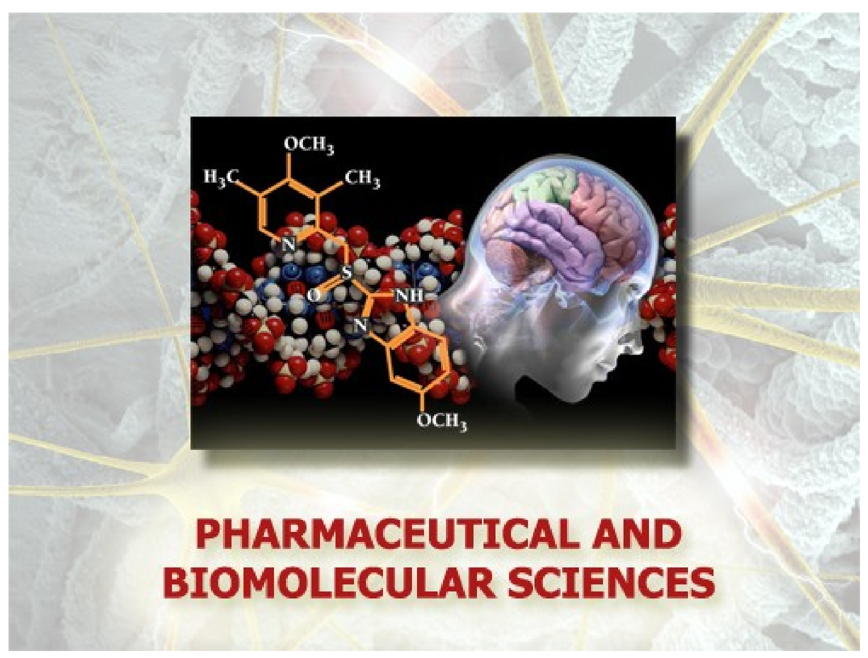
Scuola di Dottorato in
Scienze della Natura e Tecnologie Innovative

Dottorato in Scienze Farmaceutiche e Biomolecolari (XXXV ciclo)



Air Pollution, Pregnancy and DNA mutations.

Candidato: Enrico Cocchi

Tutor: Prof. Roberto Bono

UNIVERSITÀ
DI TORINO

Dottorato in
Scienze Farmaceutiche e Biomolecolari

Tesi svolta presso il
Dipartimento di Scienze della Sanità Pubblica e Pediatriche

CICLO: XXXV

Air Pollution, Pregnancy and DNA mutations.

TESI PRESENTATA DA:
Enrico Cocchi

TUTOR:
Prof. Roberto Bono

COORDINATORE DEL DOTTORATO:
Prof. Roberta Cavalli

ANNI ACCADEMICI:
2019-2023

SETTORE SCIENTIFICO-DISCIPLINARE DI AFFERENZA:
MED/42 Igiene generale e applicata

Acknowledgments

These 3 years of intensive work allowed with Prof. Bono research group allowed me to pave my way in the field of research. The amount of knowledge and skills that I acquired thanks to the wonderful people I met in this awesome experience is unbelievably high to possibly express it through word, as it is my gratitude for each of them. I have to deeply thank, from the very bottom of my hearth, Dr. Gharavi group at Columbia University, were I spent an important part of this wonderful research journey. To the best to come, to all of you.

Abstract

My project was specifically focused on trying to investigate how air pollution plays its detrimental role on pregnancy and newborn outcomes. This subject of research has been extensively investigated in past years, but models normally utilized for this specific kind of study are inherited from historical epidemiological models developed to take into account the chronic and cumulative effects of exposure, rather than using specifically developed models able to consider the peculiar characteristics of pregnancy and its related outcomes. In fact, while lot is known regarding the detrimental effects that air pollution is able to play if exposure is continued throughout the pregnancy, basically no data is available on how these molecules can play a specific detrimental effect in an acute manner, as example inducing delivery in later stages. I also tried to further elucidate molecular mechanisms that could be associated with those kind of effects, focusing specifically on acute inflammation processes, which have been extensively associated as the underpinning mechanisms of air pollution detrimental effects. Such pro-inflammatory trigger, elicited by air pollution exposure, may explain an important part of the pregnancy affection played by these molecules, paving the way to both new potential pharmacological targets and public health measures aimed to counter-play such effects in such sensitive population. As inflammation is strongly related to DNA mutations, I also tried to elucidate how air pollution exposure could related to such mutations, investigating somatic mutations, a specific kind of mutation not inherited by parents but expressly associated to environmental exposure.
The main aspects that I tried to elucidate in the present doctoral project were:

1. acute effect that pollutants can play on specific phases of pregnancy, such as later stages and delivery
2. pregnancy-related risk factors associated with air pollution and poor neonatal outcomes
3. molecular mechanisms underpinning air pollution detrimental effects on pregnancy and newborns
4. air pollution exposure and somatic DNA mutations potential association

Specifically, I investigated air pollution acute effect on pregnancy outcomes through specific Poisson regression model able to specifically identify such relationships. I then investigated habits and substances that mothers were exposed to during pregnancy, in order to further elucidate which could constitute risk factors for poor neonatal outcomes. Finally, focusing on acute inflammation processes, that appeared to be the stronger link between air pollution and the detrimental effect observed, I developed specific models to identify how these molecules could be related to DNA somatic mutations. Overall, my doctorate project tried to shade some light on the need of specific models and analysis approach aimed to investigate peculiar characteristics of such sensitive period of life and possible harmful effects played by different environmental exposure sources, widening our analysis on molecular mechanisms and DNA alterations as well. The research plan of my doctorate allowed bringing out several important limitations of current approaches to study focusing on pregnancy and newborn health, especially regarding air pollution exposure effects on preterm birth during latest stage of pregnancy. I further deepened this analysis investigating potential molecular mechanisms underlying such relationship, revealing a central role played by inflammation on the detrimental effects of pregnancy related exposure and maternal habit. I also detected how some hospital routine procedure expose newborns to harmful endocrine disruptors, such as bisphenols, and their subsequent effects on health. Finally, I was able to provide the first literature evidence of how air pollution exposure can be an important determinant of DNA somatic mutations. I hope that this work can be a starting point to realize limitations of our current research approaches to pregnancy and newborn related studies, and increase the awareness of how urgent the need for further public health measures and pharmacological therapies able to address these problems.

Table of Contents

Acknowledgments

Abstract

1. Introduction

1.1. Air pollution

1.1.1. Definition and Characteristics

1.2. Air Pollution and Human Health

1.3. Air Pollution and Pregnancy

1.3.1. Detrimental Effects
1.3.2. Underlying Pathological Mechanism
1.3.3. Chronic vs. Acute Exposure

1.4. Air Pollution and DNA Mutations

1.4.1. DNA, cancer and inherited mutations
1.4.2. Underlying Pathological Mechanism

1.5. Air Pollution Measurement

1.5.1. Single Pollutant Measurement
1.5.1.1 Carbon Monoxide (CO)
1.5.1.2 Nitrogen Dioxide (NO_2)
1.5.1.3 Particulate Matter (PM)
1.5.1.4 Ozone (O_3)
1.5.1.5 Aeroallergens (AA)
1.5.2. Air Quality Index

2. Experimental Projects

2.2.1. Study Lines
2.2.1.1. Air Pollution and Acute Inflammation in Newborns and Pregnancy
2.2.1.1.1. Study Line 1: Air Pollution as an Acute Preterm Birth Trigger
2.2.1.1.1.1. Aim and Rationale
2.2.1.1.1.2. Research Plan
2.2.1.1.1.3. Air Pollutants and Metereological Data
2.2.1.1.1.4. Cohort
2.2.1.1.1.5. Results
2.2.1.1.1.6. Discussion
2.2.1.1.2. Study Line 2: Air Pollution and Pregnancy Acute Inflammation
2.2.1.1.2.1. Aim and Rationale
2.2.1.1.2.2. Research Plan
2.2.1.1.2.3. Epidemiological Sample and Recruitment
2.2.1.1.2.4. Results
2.2.1.1.2.5. Discussion
2.2.1.2. Bisphenols as Endocrine Disruptors in Newborns
2.2.1.2.1. Study Line 3: *Bisphenols detection in Newborns*
2.2.1.2.1.1. Aim and Rationale
2.2.1.2.1.2. Research Plan
2.2.1.2.1.3. Epidemiological Sample and Recruitment
2.2.1.2.1.4. Results
2.2.1.2.1.5. Discussion

2.2.1.3. Air Pollution and DNA mutations
2.2.1.3.1. Study Line 4: Air Pollution and Somatic Mutations
2.2.1.3.1.1. Aim and Rationale
2.2.1.3.1.2. Research Plan
2.2.1.3.1.3. Epidemiological Sample and and CHIP Variants Calling
2.2.1.3.1.4. Results
2.2.1.3.1.5. Discussion
3. Project Conclusion
4. In the meantime
4.1. Formaldehyde exposure and Health
4.2. Other Papers
5. References
6. Appendix

Air Pollution, Pregnancy and DNA mutations

1. Introduction

1.1. Air Pollution

1.1.1. Definition and Characteristics

Air Pollution, accordingly to the World Health Organization (WHO), is defined as "*contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere*"[1]. Pollutants derive from many different sources, among which combustion devices, motor vehicles, and industrial facilities represent some of the most common sources of air pollution. Presence of those molecules in the atmosphere is detectable independently by human processes, but we consider those as pollutants as they reach to such levels identified as harmful to human health. Even if detectable independently by human action, nowadays air pollution is mainly related to both industrial production and transportation, thus it represents a continuously growing global problem. The huge number of industrial processes and machines generating pollutants has brought the number of detectable pollutants in every different media (water, soil, air, etc.) to such a level that makes it very difficult to be monitored. In western countries, the main pollutants that have been identified and thoroughly studied are molecules related to combustion of fossil fuels. These pollutants are released into the air, where those remain for long period of time. In developed countries levels of such pollutants are periodically or continuously monitored, among those the most represented are carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and particulate matter (PM_{10} and $PM_{2.5}$). Each of these pollutants has specific characteristics and processes that increase their concentration in the atmosphere. Alongside those important pollutants, a special role is played by aeroallergens (AA), which is any airborne substance able to trigger an allergic reaction. In western countries, AA is mainly represented by the pollen of specific seasonal plants, which are capable of eliciting allergic reactions in susceptible subjects. AA is an air pollutant, which is mostly unstudied and neglected in historical approach to epidemiological air pollution studies, but a constantly increasing number of studies on the matter is revealing its importance for human health and air pollution related reactions. A detailed description of each pollutant characteristics is beyond the scope of the current dissertation, but I provide here below the main concerns and molecular mechanisms that those pollutants play on human health, and especially during pregnancy and childhood, which has been the main focus of my doctoral project.

1.2. Air Pollution and Human Health

Air particles are defined as "*pollutants*" when their concentration reach certain levels that have been identified as harmful for human health, mostly due to cancerogenesis and exposure related risks. Pollutants have been historically studied for their chronic exposure related risks, while an important chapter of their action, regarding acute exposure associated risks, have been completely neglected until recently. Due to their diffusion and the number of people exposed, pollutants represent a major public health concern and represent an important source of morbidity and mortality for humans. WHO data show that almost all of the global population (99%) breathe air that exceeds WHO guideline limits and contains high levels of pollutants, with low- and middle-income countries suffering from the highest exposures[1]. The WHO reported that around 7 million people die every year due to exposure to polluted air, and that ambient air pollution, especially in low/middle-income countries, caused 4.2 million deaths in 2016[1]. Short-term exposure is able to elicit various acute reactions, especially in fragile population groups, such as children, pregnant women, and the elderly[2,3]. Long-term exposure is responsible of several forms of chronic diseases[4], increasing mortality due to respiratory and cardiovascular diseases[2,5], lung cancer morbidity[6], reproductive function disorders[7] and postnatal development alternations[8]. Based on their molecular characteristics, different pollutants can exert very different action on human health, which has been extensively studied and reported in literature. A particular role, which is my project specific focus, is the role played by air pollution on pregnancy and newborn outcomes. Different roles are played by different pollutants based on the moment and duration of exposure, their concentration, and several other factors that characterize this specific developmental process in this very sensitive moment, as I will summarize in the chapters below.

1.3. Air Pollution and Pregnancy

1.3.1 Detrimental Effects

Some of the detrimental effects played by pollutants are particularly evident during pregnancy and childhood. As example, CO has been extensively linked to low birth weight[9]. Several epidemiological lines of evidence proved how PM exposure during pregnancy is associated with high risk pregnancy and negative birth outcomes, including preterm birth (PTB), low birth weight, and post neonatal infant mortality[10,11]. An important role of such exposure has been demonstrated on transition from physiological to high-risk pregnancies. High-risk pregnancies are defined as pregnancy burdened by some form of complications, and especially gestational diabetes mellitus (GDM), hypertensive disorder of pregnancy (HDP), and preterm birth PTB. Briefly, GDM is a glucose intolerance beginning during pregnancy, with a worldwide prevalence of 6–25% depending on race, ethnicity, and other demographics[12]. HDP includes gestational hypertension, preeclampsia, and eclampsia, and it occurs in approximately 5–10% of all pregnancies and in 15% of all PTB[13,14]. PTB is defined as a birth that occurs before 37 weeks of gestation, and stillbirth is the loss of the fetus after 20 weeks of gestation. PTB and stillbirth are estimated to occur in approximately 12–24% of all pregnancy and in 18.4 of 1000 births globally, respectively[11,15]. Regarding mechanisms of actions, several literature reports indicated how air pollutants are able to penetrate into the respiratory tract and cause Oxidative Stress (OS), DNA mutation, inflammation, and to permeate into the blood circulation system, leading to chronic symptoms or even cell death. Reactions that in turns lead to pregnancy complication such as HDP, GDM, PTB, and stillbirth[11]. Pregnant women are especially vulnerable to air pollutants as they are in a state of active cell proliferation, fetal organ development, as well as other metabolic changes including increased cardiac output and respiratory volume. As example, O_2 consumption in pregnant women increase by 15-20%, and their breathing volume per minute increases by 30–40%, bringing them to higher exposition to air pollution than normal adults[16]. Moreover, the extremely sensitive situation of pregnancy in which the fetuses completely relies on the mother metabolism to ensure all its needs, results in an extremely increased risk of detrimental effects related to air pollution. These effects play role through all the biological pathways presented above, and the main link is represented by pro-inflammatory mechanisms, as presented below.

1.3.2. Underlying Pathological Mechanism

Air pollution has been extensively associated to Reactive Oxygen Species (ROS) generation in human body. ROS are forms of highly reactive free radicals, especially superoxide anion (O_2^-), that are physiologically produced in mitochondria as a result of respiratory activity[17,18]. At low concentrations, ROS are physiological components of cell regulation and intracellular signaling, maintained in an oxidoreductive balance thanks to several enzymatic and non enzymatic antioxidative processes[19]. Both organic and inorganic particles can directly generate ROS as a result of particle-cell interactions, as in case of lipid peroxidation, nucleic acid damage, and structural disruption of proteins. Several particles highly represented in air pollutants, especially transition metals, act as catalysts and serve as a source of ROS, alongside organic components that induce the expression of CYP450 cytochrome (CYP1A1) and further enhances ROS production[17,20]. The elicited OS acts on biologically important molecules, such as proteins, lipids, and nucleic acids, leading to an imbalance of their functionality, until complete loss of it[17,20]. It is relevant how even small action in the functioning of the respiratory chain or antioxidant mechanisms may initiate OS, contributing to the further escalation of the process. In fact, OS is the classic auto-maintained detrimental process of which mitochondria represents both the main source of and the first target of action. Mitochondria contain several susceptible structures (such as lipid membranes, enzymatic proteins, and mitochondrial DNA) and are susceptible to oxidative damage[18,21]. OS and inflammatory processes are the main response of human body (and living beings body in general) to

air pollution, as briefly summarized in Figure 1. Pollutants are phagocytized by pulmonary macrophages within the lungs, where they activate ROS synthesis causing a change in the oxidoreductive status of macrophages and their activation as pro-inflammatory actors[19,22].

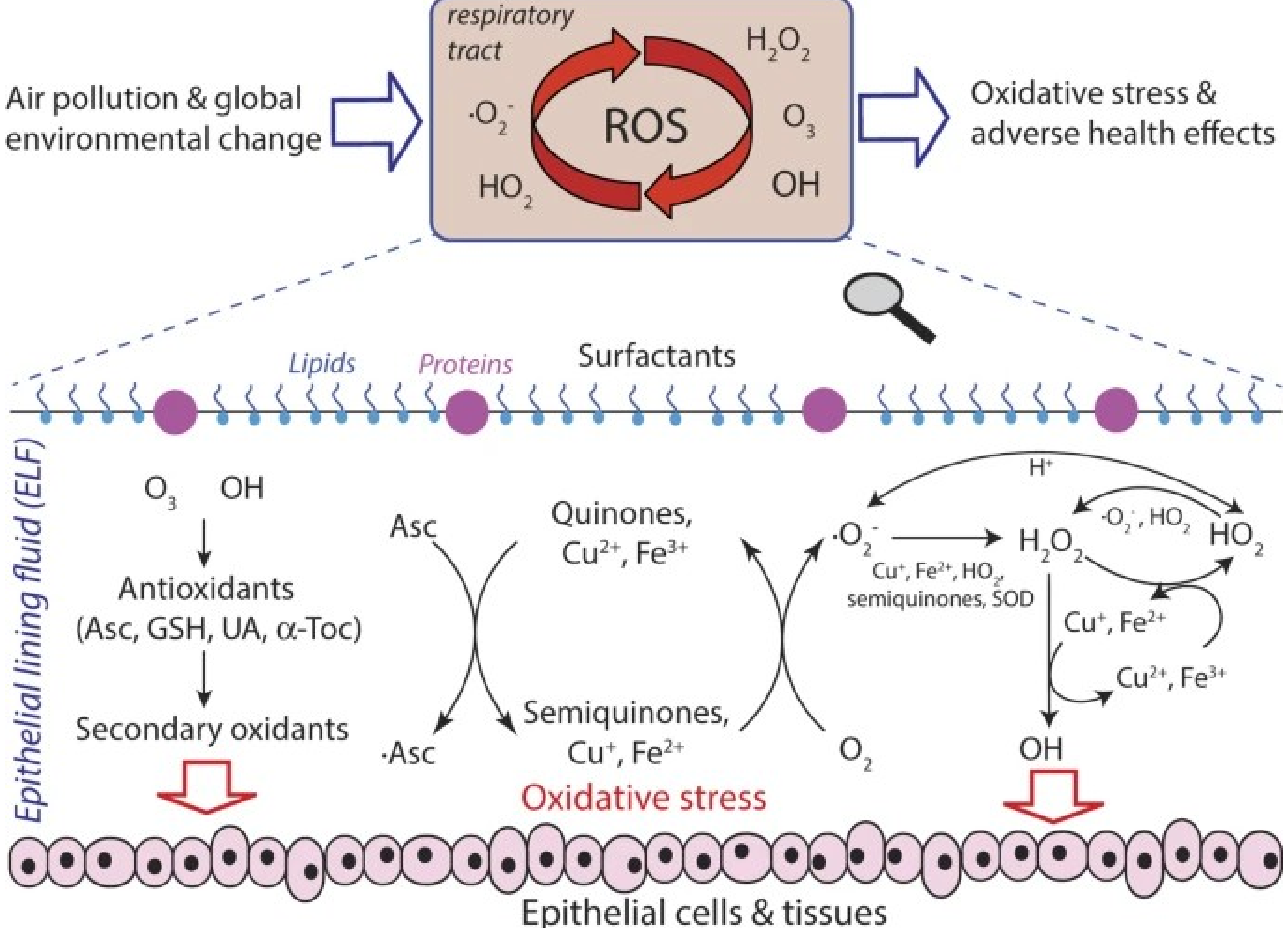


Figure . I*nteraction of air pollutants and reactive oxygen species (ROS) in the epithelial lining fluid (ELF) of the human respiratory tract. ELF can be regarded as an interface between atmospheric and physiological chemistry, through which air pollution and environmental change can induce oxidative stress and adverse health effects. Atmospheric ozone and OH radicals react with surfactants and antioxidants (ascorbate, uric acid, reduced glutathione, α-tocopherol) forming secondary organic oxidants. Redox-active components of fine particulate matter, including quinones, iron and copper ions, can trigger and sustain catalytic reaction cycles generating ROS and oxidative* stress[23].

This pathway, once triggered, induces the release of numerous inflammatory mediators that in turn activate the immune system and increase the expression of adhesion molecules in various cell types, including immune cells and endothelial blood vessels. As a result, this activation leads to unsealing the vessel barrier and the inflow of peripheral cells into several organs, eliciting the inflammatory process in loco[24]. Thus, the peripheral elicited inflammatory processes are spread throughout the whole body and organs, further contributing to ROS synthesis and their detrimental effects. Intensive ROS synthesis in the immediate vicinity of several cell types leads to a vicious cycle of oxidative stress, cellular energy metabolism modulation, impairment of their function, and death[25], as summarized in Figure 2.

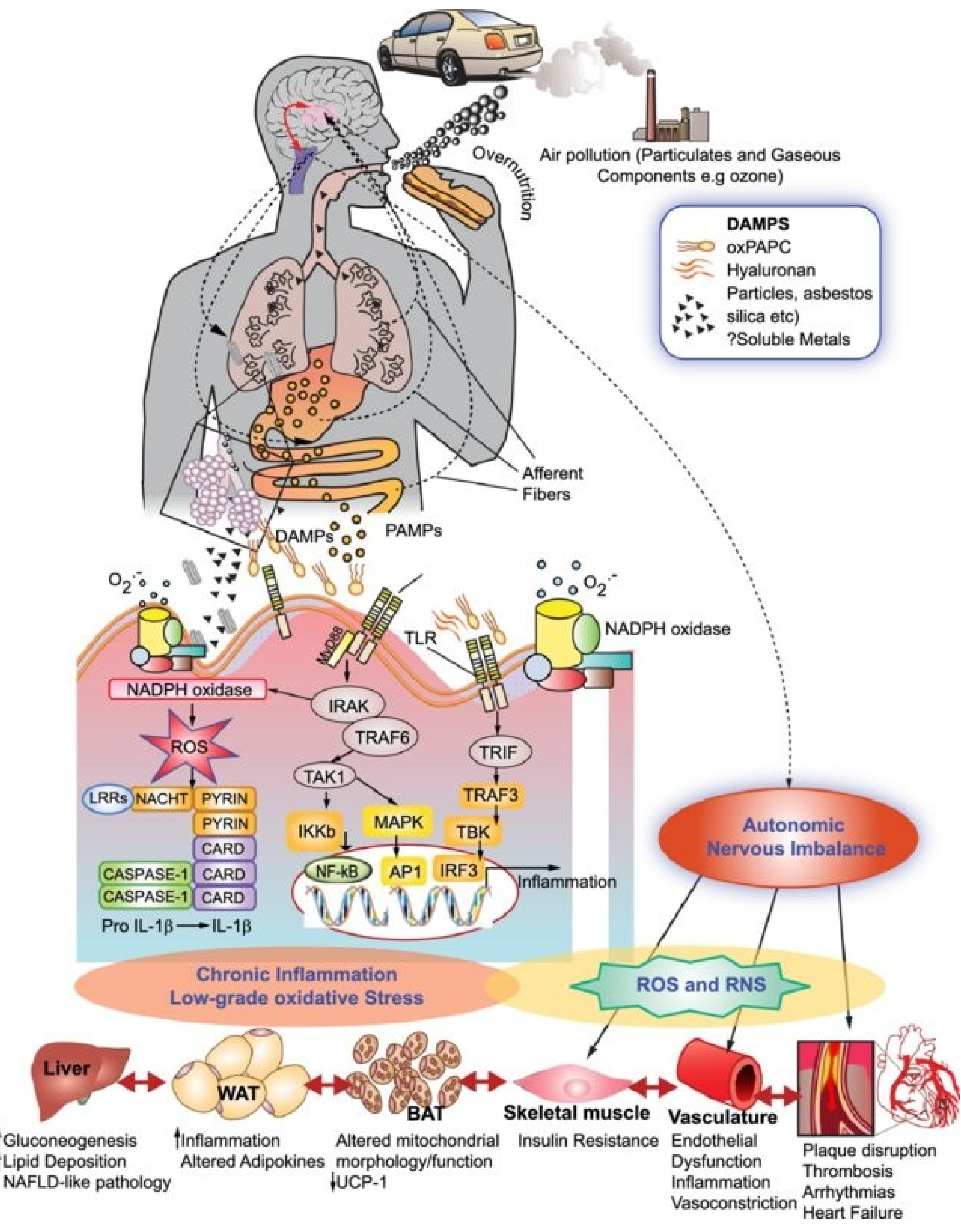


Figure 2. Hypothesized mechanisms of air pollution-mediated disease wherein inhalational or nutritional either signals directly or via the generation of signals such as DAMPs may serve to activate innate immune mechanisms such as the TLR and NLR. AP1, activator protein 1; CARD, caspase activation and recruitment domain; IKKb, IkB kinase b; IRAK, interleukin receptor-associated kinase; IRF3, interferon regulatory factor 3; MAPK, mitogen-activated protein kinase; MyD88, myeloid differentiation primary response gene 88; NAFLD, nonalcoholic fatty liver disease; PAMP, pathogen-associated molecular pattern; PAPC, palmitoyl-arachidonyl phosphocholine; RNS, reactive nitrogen species; ROS, reactive oxygen species; TAK, transforming growth factor-b-activated kinase; TBK, TANK-binding kinase 1; TRAF, TNF receptor-associated

factor; TRIF, Toll/IL-1 receptor-domain-containing adapter-inducing interferon-b; UCP-1, uncoupling protein-1; WAT, white adipose tissue[26].

Several studies showed how an increase in several pro-inflammatory markers is detectable as a consequence of air pollution exposure, especially the historical pro-inflammatory markers, such as Interleukin (IL)-1[27,28], IL-6[29], and Tumor Necrosis Factor-α (TNF-α)[30], alongside an huge and increasing number of new markers related to inflammation[30,31].

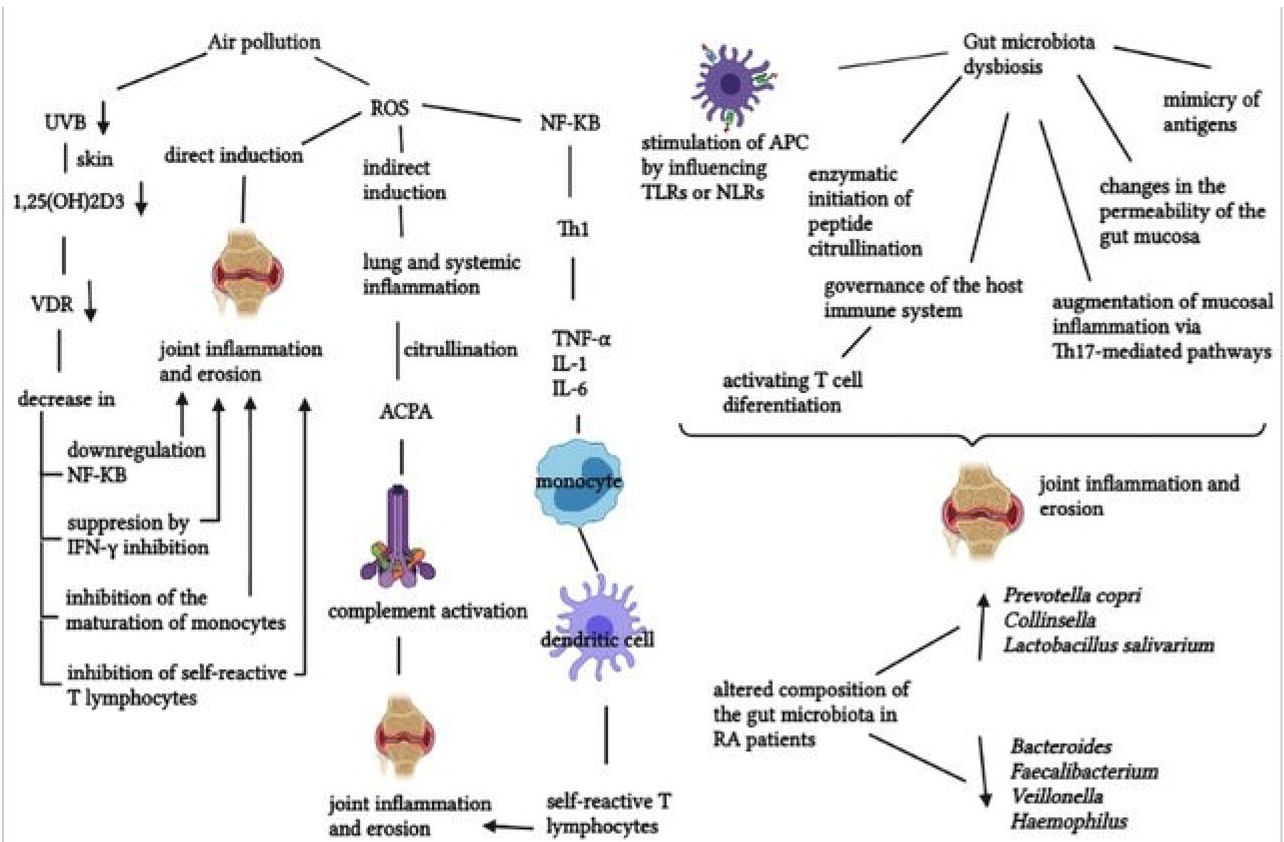


Figure 3. The involvement of air pollution and microbiota in the pathogenesis of RA. ACPA, anti-citrullinated protein antibodies; APC, antigen-presenting cells; IFN, interferon gamma; IL, interleukin; NF-KB, nuclear factor kappa-lightchain-enhancer of activated B cells; NLR, nod-like receptor; RA, rheumatoid arthritis; ROS, reactive oxygen species; TLR, toll-like receptor; TNF-α, tumor necrosis factor alpha; UVB, ultraviolet B radiation; VDR, vitamin D receptor[32].

The huge and constantly increasing lines of both experimental and clinical evidence linking air pollution exposure to pro-inflammatory effects strongly suggest that inflammation may represent the main underlying pathway linking pollutants to their detrimental effects on human health. In this regard, air pollution exposure have been historically evaluated as a chronic exposure playing its detrimental role on the long-term. Nevertheless, an increasing number of literature evidence is reporting how important acute effects are played as well. This differentiation is of central importance in the analysis of the phenomena, as it drastically changes statistical models and analysis approach to the subject. In the present dissertation, I tried to evaluate both acute and chronic exposure effects, based on the underlying disease pathophysiology. In this regard, I summarize here below some of the main evidence and differences between short- and long-term exposure effects identifiable in literature.

1.3.3. Chronic vs. Acute Exposure

Several markers of inflammation have been associated with acute short-term exposure to air pollution. As example, PM have been extensively associated with IL-6 over expression[29]. CO has been related to both an increased number of CD4 cells producers of IL-17 and IL-22[33,34]. Interestingly, IL-22 and IL-17 expression are induced by the activation of aryl hydrocarbon receptors, a transcription factor that is a target for pollution[35]. High SO_2 short-term exposure was associated with reduction in the total leucocytes count and both $PM_{2.5}$ and CO short-term exposures were associated to CD69+T regulatory cells reduction which leads to inflammation and immune-suppression[34]. All these lines of evidence testify the central pro-inflammatory role of acute air pollution exposure. This is further underlined by several clinical entities caused by acute exposure to air pollutants. As example, CO poisoning is a syndrome deriving from acute exposure to high levels of CO whose affinity to Hb is higher than O_2 affinity, displacing the latter from Hb molecules in the blood. Symptoms include headache, dizziness, chest pain, confusion, cherry red skin, vomiting, loss of consciousness, arrhythmias, seizures, and death[36], which depends on CO exposure levels, as summarized in Table 1 below:

Concentration	***Symptoms***
35 ppm (0.0035%), (0.035‰)	Headache and dizziness within six to eight hours of constant exposure
100 ppm (0.01%), (0.1‰)	Slight headache in two to three hours
200 ppm (0.02%), (0.2‰)	Slight headache within two to three hours; loss of judgment
400 ppm (0.04%), (0.4‰)	Frontal headache within one to two hours
800 ppm (0.08%), (0.8‰)	Dizziness, nausea, and convulsions within 45 min; insensible within 2 hours
1,600 ppm (0.16%), (1.6‰)	Headache, increased heart rate, dizziness, and nausea within 20 min; death in less than 2 hours
3,200 ppm (0.32%), (3.2‰)	Headache, dizziness and nausea in five to ten minutes. Death within 30 minutes.
6,400 ppm (0.64%), (6.4‰)	Headache and dizziness in one to two minutes. Convulsions, respiratory arrest, and death in less than 20 minutes.
12,800 ppm (1.28%), (12.8‰)	Unconsciousness after 2–3 breaths. Death in less than three minutes.

Table 1. Effects of different levels of CO exposure

The pathophysiology of CO poisoning is related to hypoxia, as cited above, and the strong acute inflammatory disruption deriving from oxidative stress, ROS production, etc. As summarized in the Figure 4.

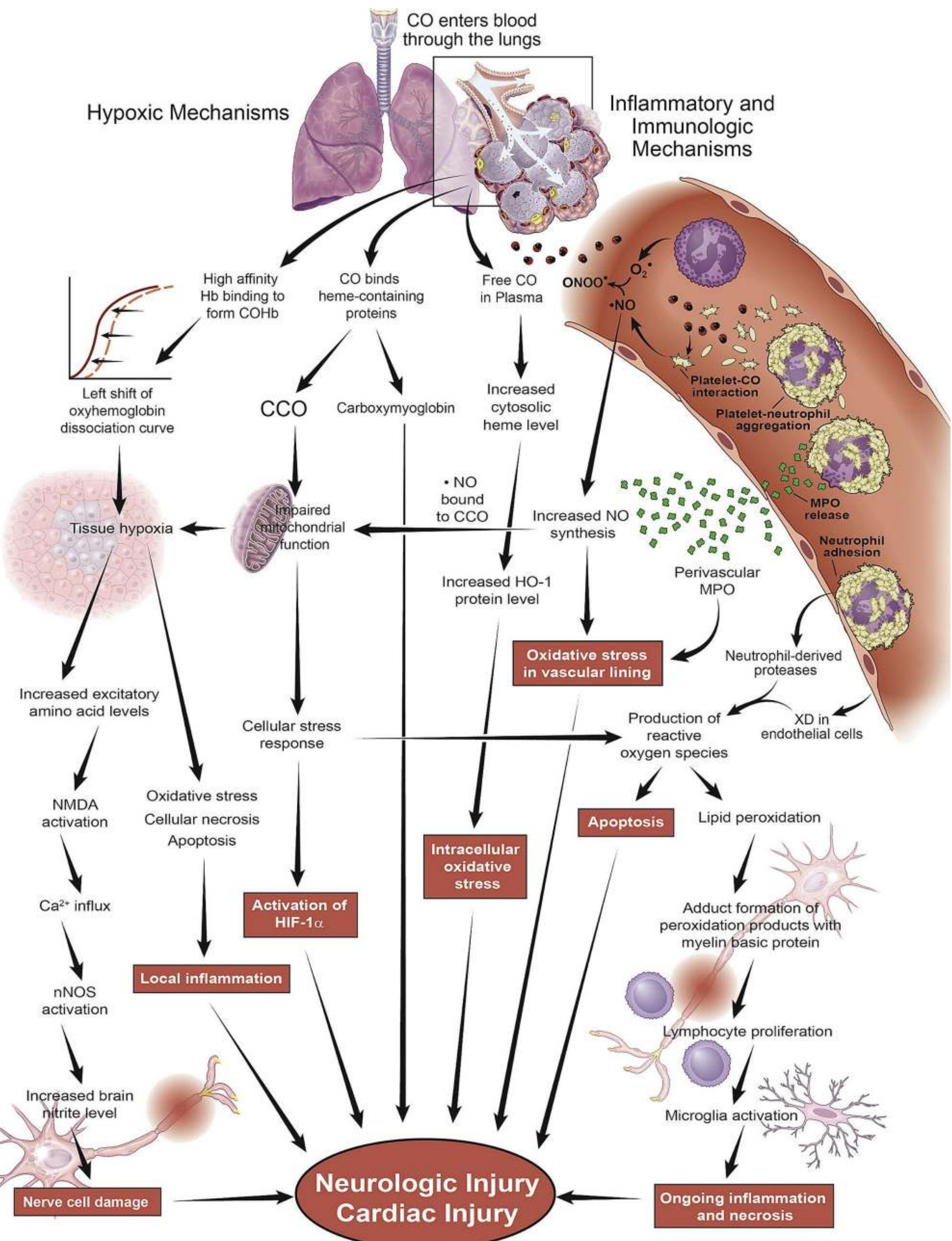


Figure 4. Pathophysiology of CO poisoning. CO diffuses rapidly into the blood after entering through the lungs. CO causes hypoxia through the formation of COHb and a leftward shift of the oxyhemoglobin dissociation curve and the binding to heme-containing proteins, particularly cytochrome c oxidase and myoglobin. CO also causes inflammation by increasing cytosolic heme levels and the heme oxygenase-1 protein, resulting in increased intracellular oxidative stress. CO binds to platelet heme protein, causing the release of NO. Excess NO produces peroxynitrite (ONOO), which in turns impairs mitochondrial function and worsens tissue hypoxia. CO induces

platelet-neutrophil aggregation and neutrophil degranulation; release of myeloperoxidase, proteases, and reactive oxygen species, which contribute to oxidative stress; lipid peroxidation; and apoptosis. The interaction of proteases with xanthine dehydrogenase in endothelial cells forms xanthine oxidase, which inhibits endogenous mechanisms against oxidative stress. Additionally, CO-induced hypoxia activates hypoxia-inducible factor 1a, which can stimulate either protective or injurious gene regulation depending on the CO dose and host factors[36].

Similarly, the rest of pollutants treated in the current work are associated to symptoms or poisoning syndromes due to acute exposure, depending on the exposure duration and pollutant concentration, as summarized in the Table 2 and Table 3.

Expose Concentration (ppm)	***Effect Concentration (ppm)***	***End Points***
200-400	/	2/4 died
>500	>500	Death in less than 2 d due to pulmonary edema
>300-400	>300	Fatal edema, bronchopneumonia
>150-200	>150	Bronchoiolitis fibrosa obliterans with death in 3-5 weeks
50-100	50	Bronciolitis and focal pneumonia with spontaneous recovery

Table 2. Effects of NO_2 acute exposure[37].

Exposure	***System Affected***	***Health Effects***
Short term	Cardiovascular	Increased rates of myocardial infarction and ischemia in those at risk
Short term	Cardiovascular	Exacerbation of cardiac failure
Short term	Cardiovascular	Increased incidence of arrhythmia
Short term	Cardiovascular	Increased incidence of deep vein thrombosis
Short term	Cardiovascular	Increased incidence of stroke
Short term	Respiratory	Increased wheeze
Short term	Respiratory	Exacerbation of asthma
Short term	Respiratory	Exacerbation of chronic obstructive pulmonary disease
Short term	Respiratory	Bronchiolitis and other respiratory infections
Long term	Cardiovascular	Increased rates of myocardial infarction
Long term	Cardiovascular	Accelerated development of atherosclerosis
Long term	Cardiovascular	Increased blood coagulability
Long term	Respiratory	Increase in systemic inflammatory markers
Long term	Respiratory	Increased incidence of pneumonia
Long term	Respiratory	Increased incidence of lung cancer
Long term	Respiratory	Impaired lung development in children
Long term	Respiratory	Development of new asthma
Long term	Reproductive	Increased incidence of preterm birth
Long term	Reproductive	Increased incidence of low birth weight
Long term	Brain	Increased risk of Alzheimer disease
Long term	Brain	Increased risk of Parkinson disease
Long term	Brain	Increased risk of neurodegenerative diseases

Table 3. Short and long-term effects of PM exposure[38].

All these differential effects depending on level and duration of exposure play different roles in pregnancy. As example, clinical evidence and experimental models demonstrated how chronic CO exposure causes a reduction in fetal growth, with an increased risk of Intrauterine Growth Restriction (IUGR), PTB, etc[39–41]. On the other side, acute exposure during pregnancy is associated to fetal hypoxia and stillbirth, with stronger effect on developing fetuses than on adults[42,43]. In this

regard, O_2 affinity of fetal blood is greater than that of maternal blood, so that its oxyhemoglobin saturation curve is normally to the left. The Hb content of fetal blood is greater than that of maternal blood, resulting in a greater O_2 capacity. Normal fetal arterial O_2 tension is relatively low compared to adult ones (about 20 mm Hg), and its O_2 content is about 12 ml per 100 ml of blood. Thus, a relatively small drop in fetal arterial O_2 tension can severely lower O_2 content. This can result from the lowered O_2 tension of maternal placental blood with which the fetal blood equilibrates. With a rise in the fetal carboxyhemoglobin level, the further shift to the left of the fetal curve exaggerates the effect of decreased O_2 tension, resulting in stronger hypoxia. Moreover, the timing of exposure further differentiates pregnancy effects during pregnancy: conception, first vs. second vs. third trimester, continuous exposure, etc. Due to the different phases of fetal development, the moment in which exposure takes place can strongly differentiate its resulting effects[44].
From an epidemiological point of view, due to all these peculiar characteristics of pregnancy and air pollution exposure relationship, it is extremely important an accurate choice of the experimental approach and statistical models. In this regard, air pollution exposure has been historically studied through long-term models that take into account primarily cumulative and chronic complications of exposure, completely neglecting both possible acute effects and precise timing of exposure during pregnancy[45,46]. My doctoral project tried to focus the attention on the latter aspects, in order to compensate for this lack and bias of available studies.

1.4. Air Pollution and DNA Mutations

1.4.1. DNA, cancer and inherited mutations

The detrimental effects further testify the strong link between air pollution and inflammation that pollutants exposure plays on DNA. Importantly, air pollution is able to cause DNA mutations in both somatic and germ cells, resulting in both increased risk of cancer and heritable mutations transmitted to the offspring, for the exposed subjects. Several lines of experimental evidence revealed how air pollutants are able to induce mutation in germ cells, which are in turn transmitted to the offspring[47,47,48]. Moreover, air pollution has been extensively linked to increased risk of cancer development, including lung, ovarian, and several other types of cancer[49,50].

1.4.2. Underlying Pathological Mechanism

An increasing number of evidence, through both clinical studies and experimental models, demonstrated how air pollutants are able to induce DNA mutation primarily through inflammatory oxidative mechanisms, as summarized in Table 4 and Figure 5.

Subjects	***Exposure assessment***	***Effects on DNA damage and repair***	***References***
98 police men and 105 controls (office clerks) in Prague (Czech Republic) and Kosice (Slovak Republic)	Concentration of polycyclic aromatic hydrocarbons in personal PM2.5 samples	Higher 8-oxodG level in lymphocytes of policemen in Kosice compared to controls, whereas no effect in policemen from Prague. Levels of 8-oxodG were very high (i.e. 53.6 lesions/106 nucleotides, corresponding to 244 lesion/106 dG)	51
41 non-smoking men (36 ± 5 y)	Urban air pollution in Cotonou and a rural village (Republic of Benin)	Higher level of oxidized DNA in taxi–moto drivers (20.5 8-oxodG/106 dG) as compared to controls in a rural village (11.1 8-oxodG/106 dG)	52
135 non-smoking men (34 ± 10 y)	Populations living in Benin, including inhabitants in a rural village (6961 particles/cm3), and three groups living in Cotonou, including suburban (19,980 particles/cm3), and dense traffic (>200,000 particles/cm3)a	Gradient in FPG sites in mononuclear blood cells as follows: 0.11 FPG sites/106 bp (rural), 0.19 FPG sites/106 bp (people living in the suburb), 0.21 FPG sites/106 bp (people living near roads with heavy traffic), and 0.27 FPG sites/106 bp (taxi–moto drivers)	53
50 male and female students (20–33 y) living in the centre of Copenhagen	Seasonal exposure to urban air pollution	Correlation between personal exposure to PM2.5 and 8-oxodG content in lymphocytes, whereas there were no correlations between PM2.5 mass concentration and FPG sites in lymphocytes or 24-h urinary excretion of 8-oxodG. No correlation between biomarkers and stationary (urban background) measurements of PM2.5	54
15 non-smoking men and women (25 ± 3 y)	Bicycling in Copenhagen (32,400 particles/cm3) or indoor (13,400 particles/cm3)a	Increased level of FPG sites after cycling in the traffic (0.08 FPG sites/106 bp) compared to cycling in the laboratory (0.02 FPG sites/106 bp)	
29 non-smoking men and women (20–40 y)	Normal air (6169–15,362 particles/cm3) or filtered air (91–542 particles/cm3)b	Filtering of air was associated with lower levels of FPG sites in mononuclear blood cells (0.38 FPG sites/106 bp) compared to inhalation of normal air (0.53 FPG sites/106 bp)	55
57 non-smoking men and women (45 ± 8 y)	Urban air pollution exposure among bus drivers in the city centre or rural/suburban area at the day of work and a day off	Bus drivers in the city centre had higher level of urinary 8-oxodG excretion as compared to bus drivers from the rural/suburban area, whereas clear differences between urinary excretions on a working day and a day off were not observed	56
Cross-sectional study of 47 female highway toll station workers (26 ± 6 y) and 27 controls (27 ± 5 y) consisting of females training to become toll station workers.	Exposure assessed as average exposure to vehicles/hour and urinary 1-hydroxypyreneglucuronide excretion	Highway toll workers had higher spot urinary 8-oxodG excretion (antibody-based detection) at the end of the workday than the controls.	57
50 bus male non-smoking drivers and controls	Stationary sampling of PM2.5 and PM10 at two locations in Prague with heavy and light traffic	Higher urinary excretion of 8-oxodG (antibody-based detection) among bus drivers compared to controls	58
95 male taxi drivers and 75 controls (community subjects) in Taipei city, Taiwan	Urinary 1-hydroxypyrene excretion	Higher spot morning urinary excretion of 8-oxodG (antibody-based) among taxi drivers compared to controls and a positive correlation between 1-hydroxypyrene and 8-oxodG	59

Table 4. Air pollution exposure and DNA mutations evidence in human studies[60].

Important mechanisms of DNA damage have been elucidated through animal models. In fact, as cited above, air pollution is a strong trigger of acute inflammation and ROS generation in humans. In this regard, one of the most important mechanism of defense against DNA mutations and is represented by the Nucleotide Excision Repair (NER), that relies on proteins called Xeroderma pigmentosum group A (XPA) and the Xeroderma pigmentosum group C (XPC)[61]. XPC-deficient mice have been showned to be more sensitive to oxidative DNA-damaging agents in the lung, compared with the XPA-deficient mice and wild-type controls[62]. NER-deficient mice have also been linked to higher incidence of tumors when exposed to air pollution derived compounds[63].101 Moreover, XPC mice exposed to chronic cigarette smoke developed lung tumors through impaired repair of oxidatively generated DNA and air pollution derived compounds can greatly inhibit NER[64]. These findings suggests that reduction of NER capacity is an important link between air pollution and DNA mutations. Recently, an important study on the matter took into account several core aspects of this relationship: inflammatory response using in vivo imaging and immunohistochemical analysis, oxidative stress by quantification of carbonylated proteins, detection of mutated DNA foci used through damage biomarker, and histopathological lesions in the lungs[31]. Results revealed the central role of XPC protein, and NER pathway, in protection against the carcinogenic potential of air pollution, further testifying the central role played by inflammation in this relationship. Authors showed how DNA is damaged both directly and indirectly (through ROS) by the various compounds detected in urban air pollution, providing data clearly linking ambient air pollution with the induction of DNA damage through inflammatory pathways. Several other lines of evidence revealed the important role of inflammation on DNA mutations[65,66]. Thus, I decided to further deepen our knowledge of air pollution and DNA mutations mechanisms in the present work, investigating somatic mutations, which are the DNA mutations primarily associated with inflammation and involved in cancer development and air pollution exposure link[67].

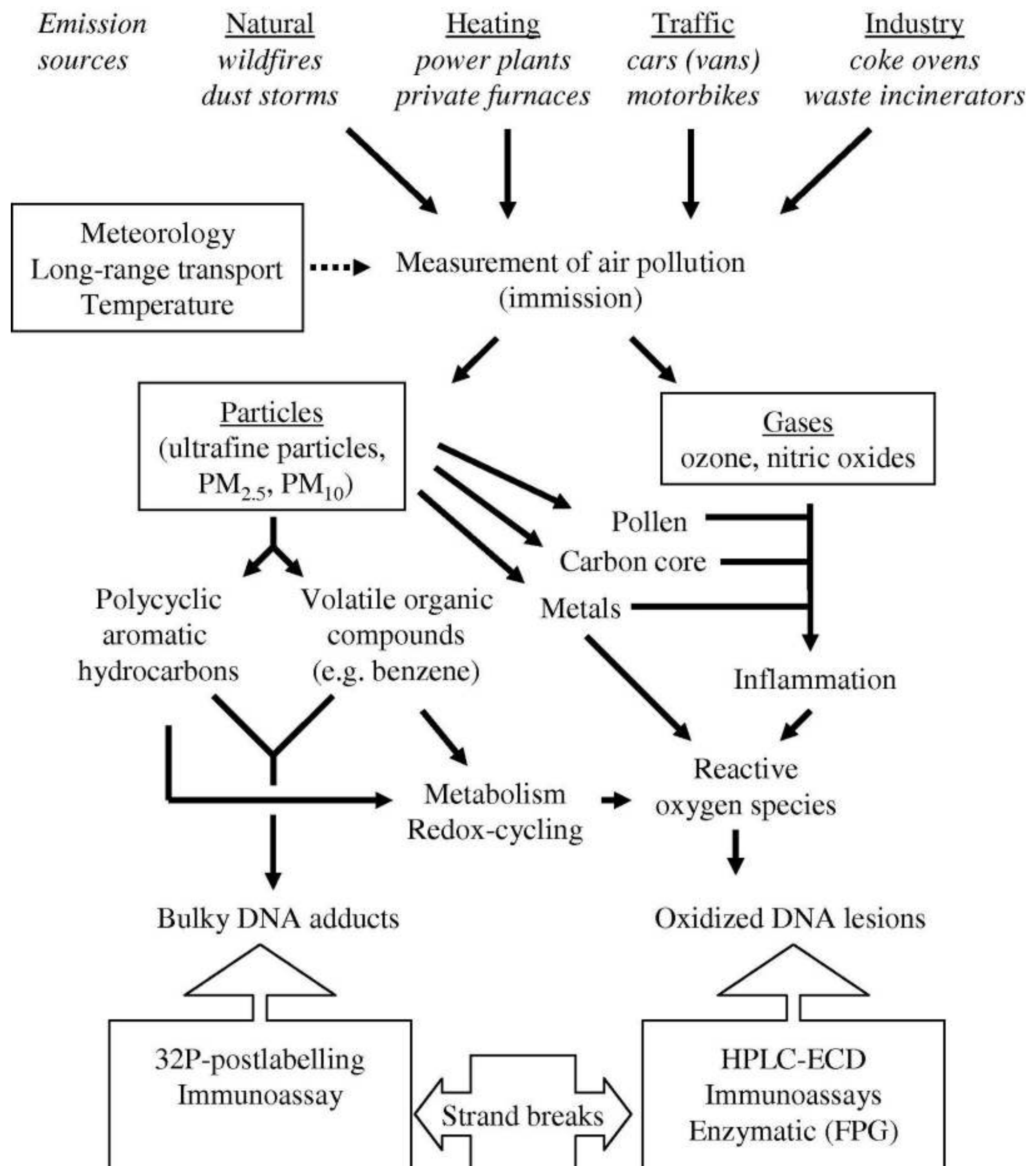


Figure 5. Relationship between exposure and the measurement of DNA damage in tissues[60].

1.5. Air Pollution Measurement

1.5.1. Single Pollutant Measurement

As cited above, particles assume the status of pollutant when their concentration exceed certain levels that have been identified as harmful for human health. Thus, air pollution measurement is the first step in identifying when those levels are exceeded and identify causes, in order to act to keep the quality of the air inside legal limits. According to the WHO, over 6000 cities in 117 countries now routinely monitor the quality of their air[68]. Pollutants measurement have extensively evolved during latest years, and it can be now mainly divided in passive vs. active measurement. Passive devices are generally simpler and cheaper. They basically work by soaking up or passively collecting air samples that is later analyzed. One of the most diffuse instrument is represented by diffusion tube. This device consists of a small and hollow plastic tube (acrylic or polypropylene), roughly 70mm long, capped at each end, as illustrated in Figure 6.

Figure 6. A typical diffusion tube attached, by a cable tie, to a drainpipe. The bottom end is open to the atmosphere. Pollution is captured by a chemical inside the red cap at the top[69].

One of the caps is either completely removed to activate the tube or contains a filter specifically permeable only to the gas under study. The other cap contains metal mesh discs coated with a chemical reagent that absorbs the gas once it entered the tube[70]. Tubes can be placed on the outside of buildings, or suitably located street furniture and after the defined period of time, the tube is taken down and analyzed to retrieve the specific gas under study concentration. Diffusion tubes are very easy to deploy and analyze, but it comes with several backwards. The main disadvantage is low accuracy (accuracy is estimated ±20% of the real concentration), and it only gives a cumulative amount over the defined period of time, making it impossible to examine shorter fluctuations. On the other end, active devices use fans to suck the air in, filter it, and either analyze it automatically or collect and store it for later analysis. Active devices are equipped with specific sensors that use physical or chemical methods to detect the gas under study. Physical methods measure an air sample without changing it, for example, by seeing how much of a certain wavelength of light it absorbs. Chemical methods change the sample in some way, through a chemical reaction, and measure that. Active measurement is nowadays mainly represented by air quality sensors, which range from small handheld devices to large-scale static monitoring stations in urban areas, depending on the area under study[71]. Due to the variety of pollutants that needs to be monitored in atmospheric air, several different methods of analysis are currently employed, including gas chromatography, spectrometry, spectroscopy, spectrophotometry, and flame photometry. A specific discussion of each method characteristics is beyond the aim of the current project, but main technologies used for each pollutant are presented below.

1.5.1.1. Carbon Monoxide (CO)

CO is generally measured by non-dispersive infrared light absorption based on the Beer-Lambert law, which exploits the attenuation of light to the properties of the material through which the light is travelling[72]. Once the air enters the analyzer, it splits into two paths: the former path represents the ambient air, pulled in directly from outside, while the latter pulls ambient air through a scrubber, which removes all of the CO. This air is now CO-free (called the background air). CO has a peak absorption at a wavelength of ~1600 nanometers, thus light emitted at this wavelength is not able to pass through CO. Inside the analyzer, the sample and background air are both exposed to UV light and are alternately pumped through different cells, both equipped with a sensor that measures the intensity of light that passes through the air, deriving CO concentration from this. A graphical representation is provided in Figure 7. CO can also be measured using electrochemical gel sensors and metal-oxide semiconductor detectors[73].

Structure of Non-Dispersive Infrared Absorption Sensor

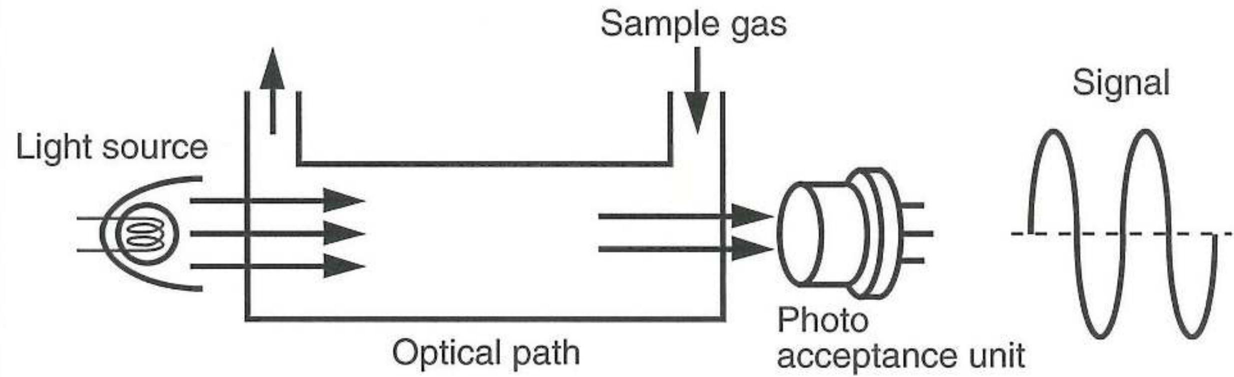


Figure 7. Graphical representation of non-dispersive infrared light absorption based devices for CO monitor[72].

1.5.1.2. Nitrogen Dioxide (NO_2)

NO_2 can be measured passively with diffusion tubes or actively through chemiluminescence analyzers, which measures NO_2 levels from the light they give off. Chemiluminescence principle is based on the emission of light (called luminescence) as the result of a chemical reaction. This principle exploits the fact that decay of excited state atoms to a lower energy level causes light emission[74]. These devices use a stabilized photodiode to measure the intensity of the light produced by the reaction of NO_2 with O_3, whose intensity is directly proportional to the concentration of NO_2.

1.5.1.3. Particulate Matter (PM)

PM are currently measured using tapered element oscillating microbalance. This device is based on a glass tube that vibrates more or less, as collected PM accumulate on it. PM particles are deposited on the filter and the added mass causes a change in its oscillation frequency, which is detected electronically. The element is periodically cycled to return it to its natural frequency75. The inlet to the device only allows particles of the desired size range to enter, as graphically resumed in Figure 8. PM can also be measured using optical photodetectors[76], which measure the light reflected from samples of light, and gravimetric analysis, in which PM are basically collected on filters and then weighed[77].

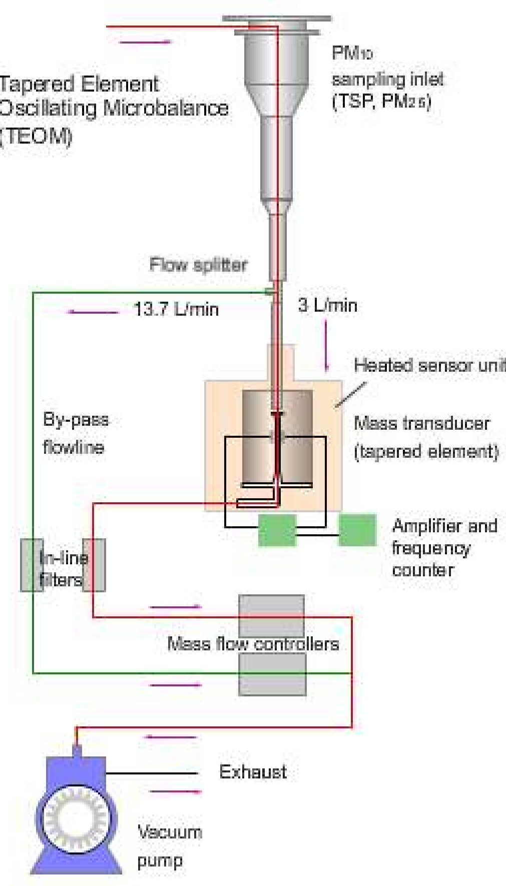


Figure 8. Graphical representation of tapered element oscillating microbalance instrument functioning[78].

1.5.1.4. Ozone (O_3)

Similarly to CO, O_3 is measured by non-dispersive infrared light absorption devices based on the Beer-Lambert law[79]. The mechanism of measurement is identical to CO one, with the difference that O_3 peak absorption is at a wavelength of ~254 nanometers, falling in the ultraviolet light region of the electromagnetic spectrum (between visible light and x-rays). That is why ozone in the stratosphere protects the earth from the sun harmful ultraviolet radiation[79].

1.5.1.5. Aeroallergens (AA)

AA is a more sophisticated and complicated analysis, thus these are typically analyzed through passive methods and HIRST sampler, which consists of three core parts: a swivel head, a suction pump and a deposition drum. The latter represents the actual sampling part. This part rotates at 2 mm/h. Collection was operated through weekly application of specific adhesive tape on the drum part of the samples. This tape is able to capture aeroallergens, ensuring no loss of rebound or natural detachment. A constant airflow is provided by the air pump part, thus converted to m3[80].

1.5.2. Air Quality Index

Single pollutant measurement represents the first step to be able to monitor air pollutants. Nevertheless, it only gives sparse information regarding this specific pollutant concentration, ignoring the total and integrative effects that different pollutants can have on air quality and subsequently on human health. This is why recently an overall index have been developed by EPA,

called the Air Quality Index (AQI), able to take into account the total effects due to different pollutants concentration at a given timepoint[81,82]. AQI is defined by the formulae[83]:

$$Ip = [IHi – ILo / BPHi – BPLo] (Cp – BPLo) + Ilo$$

Where:

- Ip = index of pollutant p
- Cp = truncated concentration of pollutant p
- BPHi = concentration breakpoint (≥Cp)
- BPLo = concentration breakpoint (≤Cp)
- IHi = AQI value corresponding to BPHi
- ILo = AQI value corresponding to BPLo

AQI is calculated on the basis of the above mentioned pollutants, where the worst sub-index drives the AQI. AQI ranges from 0 to 500, 0 being good and 500 being severe. There major pollutants that can be taken into account for AQI calculation are PM_{10} and $PM_{2.5}$, CO, O_3, NO_2. To calculate AQI, data for a minimum of three pollutants must be present, of which one should be either PM_{10} or $PM_{2.5}$. Ranges are described as below:

- 0-50: The range shows that the air quality is good and it poses no health threat.
- 51-100: This range is moderate and the quality is acceptable. Some people may experience discomfort.
- 101-150: The air quality in this range is unhealthy for sensitive groups. They experience breathing discomfort.
- 151-200: The range shows unhealthy air quality and people start to experience effects such as breathing difficulty.
- 201-300: Air quality is very unhealthy in this range and health warnings may be issued for emergency conditions. All people are likely to be affected.
- 301-500: This is the hazardous category of air quality and serious health impacts such as breathing discomfort, suffocation, airway irritation, etc. may be experienced by all.

AQI index and single pollutant breakpoints are provided in Figure 9.

O_3 (ppm)	PM_{10} (ug/m^3)	$PM_{2.5}$ (ug/m^3)	CO (ppm)	SO_2 (ppm)	NO_2 (ppm)	AQI Values	Level of Health Concern
0.000 – 0.059	0 – 54	0.0 – 15.4	0.0 – 4.4	0.000 – 0.034	–	0 – 50	Good
0.060 – 0.075	55 – 154	15.5 – 40.4	4.5 – 9.4	0.035 – 0.144	–	51 – 100	Moderate
0.076 – 0.095	155 – 254	40.5 – 65.4	9.5 – 12.4	0.145 – 0.224	–	101 – 150	Unhealthy for Sensitive Groups
0.096 – 0.115	255 – 354	65.5 – 150.4	12.5 – 15.4	0.225 – 0.304	–	151 – 200	Unhealthy
0.116 – 0.374	355 – 424	150.5 – 250.4	15.5 – 30.4	0.305 – 0.604	0.65 – 1.24	201 – 300	Very Unhealthy
–	425 – 504	250.5 – 350.4	30.5 – 40.4	0.605 – 0.804	1.25 – 1.64	301 – 400	Hazardous
–	505 – 604	350.5 – 500.4	40.5 – 50.4	0.805 – 1.004	1.65 – 2.04	401 – 500	Hazardous

Figure 9. EPA single pollutant breakpoint and AQI[84]

AQI has the main advantages of provide an overall index of air pollution level, merging different pollutants effects, providing the single pollutant that is mostly driving the index at the same time.

Nevertheless, for some specific analysis, single pollutant analysis still represents the best option, depending on the analysis aim.

2. Experimental Projects

2.2.1. Study Lines

2.2.1.1. Air Pollution and Acute Inflammation in Newborns and Pregnancy

As presented above, the main link between air pollution and its detrimental effects on human health, especially during pregnancy, is represented by acute inflammation pathway. In the present dissertation project, I decided to further investigate this association through studies focused on pregnancy outcomes, especially pregnancy complications, PTB and newborn outcomes. In order to focus on acute inflammation, and its subsequent short-term effect of air pollution exposure on pregnancy outcomes, I firstly performed a comprehensive analysis of 3 years data of pregnancies at Sant'Anna Hospital in Turin, Italy, one of the biggest child and health hospitals in Europe. Data was collected for >21'000 newborns and their mother, to precisely investigate, from an epidemiological point of view, how short term exposure to air pollution may be related to poor outcomes specifically associated with acute inflammation effects, such as PTB. I then decided to deepen my analysis and study possible molecular mechanisms underpinning this relationship, collecting samples from newborn and their mothers and analyze specific pro-inflammatory mediators and markers (IL-1, IL-6 and Isoprostane (IsoP)) possibly related to living environment and maternal habit predisposing to acute inflammation consequences. Finally, I decided to further explore air pollution and acute inflammation exposure consequences on DNA mutations, focusing specifically on somatic mutation, which are not inherited from parents but accumulated during life, and specifically linked to inflammation, as presented above. My analysis lines presented several aspects of innovation that were elaborated in order to overcome historical approaches bias to air pollution exposure studies. Firstly, I adapted a specific model in order to investigate the precise acute effect, within 7 days from delivery, of air pollution exposure on PTB, and not overall or a trimester-long exposure. Secondly, I was able to demonstrate, through pro-inflammatory markers sampling and analysis, how maternal habits and exposure induces an inflammatory state detectable not only on the mother, but on the newborn itself. Finally, I operated the first analysis on somatic mutations induction as a result of air pollution exposure. At the best of our knowledge, none of those aspects has ever been this specifically investigated.

2.2.1.1.1. Study Line 1: Air Pollution as an Acute Preterm Birth Trigger

2.2.1.1.1.1. Aim and Rationale

This study was aimed to investigate the precise acute effect of air pollution on inducing PTB on the short-term after exposure. As previously described, PTB indicates infants born before 37 weeks of gestation, thus not fully developed for extra-uterine life. PTB is a public health concern[85] affecting ~10% of total births worldwide, with higher rates in developing countries and numbers constantly increasing[86]. In developed countries it represents the main cause of infant mortality[87] and is associated with significantly impaired health outcomes[88–90] and dramatic community costs[91]. Alongside historical and well known risk factors, namely low socioeconomic status, age, ethnicity, tobacco, substance abuse, poor nutritional status, and the presence of birth defects,[92] several epidemiological studies linked PTB with chemicals air pollution[93]. While air pollution role on PTB is now known, the timing with which this effect occurs is still controversial[16]. Some studies claimed early exposure during pregnancy as responsible[94], while others indicated late exposure[95]. This

confusion is probably due to different pollution-induced reactions causing different detrimental effects on pregnancy depending on the fetal developmental stage in which they occur, their duration and outcome analyzed. In this regard, studies focusing on definite narrow temporal windows in order to better elucidate specific pollutants, co-factors and underlying mechanisms of action are lacking[92]. Moreover, PTB is a complex subject in which the long-term nature of fetal development is associated with an acute medical condition that mandates delivery, with different factors potentially playing very different roles in each scenario. PTB delivery is due to conditions that makes it impossible to continue the uterine fetal development. Those acute factors either prematurely trigger the delivery itself, or are responsible for such a strong fetal suffering that overcomes the string mortality and morbidity risks associated with PTB. Air pollution role in PTB has been historically studied through models inherited from other diseases, such as cardiovascular and neurological ones, that have a peculiar chronic nature[96]. Similarly, pollution-induced chronic inflammation during pregnancy proved to cause developmental impairments that predispose to PTB, but PTB itself is an acute medical emergency, whose main trigger is an inflammatory distress inducing uterine contractions and subsequent preterm delivery[97]. Interestingly, acute inflammation is the main elucidated mechanism contributing to both PTB and air pollutants-related detrimental effects[98,99]. Short-term exposure to air pollution proved to trigger inflammation, detectable through the increased inflammatory markers[100], and mice models revealed how this induced systemic inflammation leads to acute diseases[101]. Thus, air pollutants exposure can trigger PTB through an acute inflammatory reaction that elicits uterine contractions and subsequent preterm delivery. In this regard, it is well known how acute inflammation induces uterine contractions and PTB[102], and short term exposure to air pollutants has been demonstrated to induce an upregulation of delivery-involved inflammatory cytokines, such as IL-1, IL-6, and TNF-α with a strong established role in PTB too[102–104]. Interestingly, this air pollution exposure-induced acute inflammation proved to be reversible over days and is pharmacologically addressable[105]. All these lines of evidence advocate for both possible preventive and therapeutic approaches to pollutants exposure and PTB, reinforcing the need to better elucidate underlying mechanisms and possible pharmacological targets through studies focused on specific narrow temporal windows currently lacking. Thus, the aim of the present study is to assess the specific effect of maternal exposure to chemical and biological air pollutants in the narrow window of 7 days before delivery in both preterm and at-term births. This is in order to better elucidate the role of the air pollutants, of other co-factors, and underlying mechanisms acutely associated with preterm delivery. In fact, PTB is related to different possible reasons, but almost the totality of them is strictly related to acute uncontrolled events that brings physicians to decide that the intrauterine permanence of the fetus is far more dangerous than delivery, when delivery is not due to uncontrolled contractions. The aim of this research line was exactly to elucidate the acute effect of air pollution on those acute events. A specific generalized linear Poisson model model, which was firstly developed and used in another publication that was specifically aimed to investigate such acute effect of air pollutants on pediatric respiratory emergency room admissions[106.] was adapted. This “acute” effect focus is the innovative part and the core aim behind this work. In fact, epidemiological studies on air pollution role on PTB are based on models historically related to chronic diseases, but the complex nature of PTB, where a long-term process (fetal development) is paired with an acute one (preterm delivery) strongly biases reported relationships, as testified by the contradictory lines of evidence currently available in literature, especially regarding the timing of this relationship. In order to overcome this barrier, I developed a specific model and analysis strategy that is presented below.

2.2.1.1.1.2. Research Plan

This study serves a tailored non-stationary Poisson model correcting for seasonality and all possible confounding variables. I investigated the role of both chemical ($PM_{2.5}$, NO_2, and O_3) and biological (AA) pollutants on preterm vs. at-term birth risk. This model specifically corrected for temporal medium/long trend function, meteorological (daily temperature, daily relative humidity, cumulative daily precipitations) and non-meteorological confounding variables (day of the week, holidays and summer population decrease), that frequently bias these analyses, as testified by the significant differences in preterm vs. at-term. Relative Risk (RR) associated with each pollutant was assessed at any time lag between 0 and 7 days prior delivery in order to specifically investigate the exact timing of each pollutant in the acute prior-delivery phase. In order to focus on a narrow temporal window and correctly assess short-term effects, the association between PTB (dependent variable) and concentrations of chemical and biological air pollutants (independent variables) were analyzed using Generalized Linear Models (GLMs) fitting a non-stationary Poisson process[106,107]. The following model was used:

$$f(\lambda_t + 1) = \alpha + \sum_{i=1}^{k} \beta X_i + NS(Z_t)$$

Where:

- f: log link function
- λ_t: count of daily PTB at day t
- α: intercept constant
- β: estimated parameters vector
- X_i: matrix of k independent variables (exposure and adjustment variables)
- $NS(Z_t)$: natural spline smoothing function of calendar day Z

As in several literature reports focusing on the same analysis model, in order to take the medium/long term trend that may shape the time data under analysis into account, a natural spline smoothing function was calculated on 14 degrees of freedom (df)[107]. I restricted possible df to a maximum of 18 (corresponding to a window of ~60-days), in order to avoid overfitting[108,109]. I identified the best df as the value that minimizes the absolute values of the residuals partial autocorrelation function (PACF) sum[108]. PACF residuals were corrected for day of the week, as clarified in the point below, to remove the 7-day positive correlation, when estimating the spline smoothing function. I performed further variables adjustment in order to correct for the location under study that might otherwise bias the model:

a) Day of the week (Monday to Sunday)

b) Holidays: I considered main holidays in our zone: Christmas and Easter; ±3 days around them; other holidays; other days. This resulted in a 4-level variable that was used in the model.

c) Summer population decrease: in our specific zone, population is known to decrease during summer holidays. This variable was intended to adjust for this effect, and resulted in a categorical variable with factors considering such dates: from Saturday before Mid-August to the next Sunday,

for a total amount of 16 days per year; from July 16th to the end of August (removing the aforementioned period); all other days[110]

d) Daily average daily temperature (°C)

e) Daily average humidity (%) - relative

f) Daily precipitations (mm) - cumulative

Once the model was fitted to the actual data, I considered one chemical pollutant among $PM_{2.5}$, NO_2, O_3 or AA, alongside medium/long trend function, non-meteorological variables (day of the week, holidays and summer population decrease), and meteorological variables (daily temperature, daily relative humidity, cumulative daily precipitations). Temperature and humidity underwent natural splines with 1 and 2 df transformation, respectively. Those df was chosen through PACF criterion as explained above. Daily precipitations were binary coded: 1 if cumulative precipitation ≥1 mm and 0 otherwise. Exposure variables were included in all models at different single time lags: starting with the same day of PTBs evaluation (Lag 0) to 7 days before (Lag 7). This time frame was identified in order to focus the analysis on specific short-term potential effects of air pollutants in eliciting preterm vs. full-term births. At the best of our knowledge, this is the first literature report of such acute effect focused analysis. Thus, no previous specific time frame is available. The identification of such threshold was based on a pathophysiological rationale, as my analysis was based on investigating acute inflammation potential role in delivery and preterm vs. full-term births compensation capacities. Thus, as acute inflammation is a process known to exert its effect within few days from the exposure, I arbitrarily select this specific 7-days time frame in order to include in the analysis lag all potential acute inflammatory effects of air pollutants, from 24 hours to a week from exposure. Associations between exposure variables and preterm or full-term births are reported as relative risk (RR) with respective 95% confidence intervals (CI). RR values and their CI are calculated as exponential of GLM resulting coefficients for each specific exposure variable under analysis. Exposure variables association coefficients were calculated based on 10 μg/m3 increase in $PM_{2.5}$, NO_2, and O_3 concentrations, and 10 grains/m3 increase in case of AA.

2.2.1.1.1.3. Air Pollutants and Meteorological Data

Meteorological data: meteorological data were obtained from a station placed at 254 a.s.l., on the roof of the Department of Physics of the University of Turin, located at about 2.5 kilometers from the S. Anna Hospital. The station is permanently active and collects weather data in the urban surface layer of the city. Data is collected every 5 seconds and subsequently averaged every 5 min. Data were aggregated in a daily form for the analysis.

Chemical air pollution data: I extracted concentrations of air pollutants (NO_2, $PM_{2.5}$ and O_3) from data collected at the urban monitoring station located in Turin, viale Augusto Monti. The Local Environmental Protection Agency (ARPA Piemonte) collected and summarized data, which is coordinated by the regional air pollution service of Piedmont Region, as per the current European legislation (DIR 2008/50/ECX). Data were collected hourly and daily aggregated for the analysis.

AA data: Corylaceae, Cupressaceae, Gramineae, Urticaceae, Ambrosia, and Betula pollens were considered in this study. A station located ~12 m above the ground without surrounding obstacles, as required in such cases, on the flat roof of a building located 2.6 km from S. Anna Hospital, allowed collection of AA data. A HIRST sampler compose the station. HIRST sampler consists of

three core parts: a swivel head, a suction pump and a deposition drum. The latter represents the actual sampling part. This part rotates at 2 mm/h. Collection was operated through weekly application of specific adhesive tape on the drum part of the samples. This tape is able to capture aeroallergens, ensuring no loss of rebound or natural detachment. A constant airflow of ~10 L/min is provided by the air pump part. This represents ~14.4 m^3 daily. AA counts were performed by the Department of Life Science and System Biology of the University of Turin, and expressed in my analysis in form of grains/m^3.

2.2.1.1.1.4. Cohort

Turin, the capital of Piedmont region (North-Western Italy) is one of the most polluted European cities, located 239 m above sea level (a.s.l.); it has 886,837 inhabitants, a population density equal to 6,813 per km^2. Daily data for the period January 1st 2015 to December 31st 2017 (1,096 days) for the city of Turin were collected or derived as described below. Daily data on births were obtained from the birth registry and registration records of the main obstetric hospital in Turin, Italy (Fetal Maternal Department, Sant'Anna Obstetric Gynecological Hospital), recording maternal and fetal data during the 3-years period January 1st 2015 – December 31st 2017. Recorded data:

- Maternal age at delivery
- Gestational age (weeks of amenorrhea)
- Apgar test at 1 and 5 minutes of life
- Sex
- Twins
- Mother’s country of origin

Data were anonymized in respect to the European General Data Protection Regulation (GDPR 101/2018).

Exclusion criteria: all pregnancies resulting in stillbirth were excluded from the current analyses. At the same manner major fetal anomalies detected prenatally were excluded from the analysis as well.

2.2.1.1.1.5. Results

In the study period, 21,509 births were observed, of which 3,167 PTB (14.7%). Median gestational age was 39+1 for the whole cohort, 39+3 for term and 35+1 for PTB. The peculiar differences between preterm and term subpopulations were:

- Higher number of twins rate in PTB (preterm: 42.2% vs. term: 1.87%, $p<0.01$)
- Lower APGAR 1/5 scores in PTB (Preterm Apgar1 7.6 ± 2.1 vs Term 8.8 ± 0.9, $p<0.01$; Preterm Apgar5: 8.2 ± 1.5 vs Term 8.9 ± 0.6, $p<0.01$)
- Higher maternal age in PTB (Preterm 34.1 ± 5.5 vs Term 33.3 ± 5.4, $p<0.01$)

Cohorts characteristics are summarized in Table 5.

	All cohort (n. 21,509)				Term (n. 18,345)				Preterm (n. 3,164		
	Median (IQR)	IQR / Median	min-max	mean±SD	Median (IQR)	IQR / Median	min-max	mean±SD	Median (IQR)	IQR / Median	min-max
Maternal age at delivery (weeks)	33.7 (7.4)	0,2	14.8-48.6	33.4 ± 5.4	33.6 (7.4)	0,2	14.8-48.6	33.3 ± 5.4	34.3 (7.7)	0,2	15.6-46.9

Gestational Age (weeks + days)	39+1(2+1)	0,05	21+5-49+0	38.7 ± 2.3	39+3 (1.7)	0,04	37+0-49+0	39.4 ± 1.5	35+1(3+0)	0,09	21+5-36+6
Apgar 1	9 (0)	0	0 - 10	8.6 ± 1.2	9 (0)	0	0 - 10	8.8 ± 0.9	8 (2)	0,2	0 - 10
Apgar 5	9 (0)	0	0 - 10	8.9 ± 0.8	9 (0)	0	0 - 10	8.9 ± 0.6	9 (1)	0,11	0 - 10
Sex Prevalence (%)	Male 11,066 (51.4%) Female 10,443 (48.6%)				Male: 9,423 (51.4%) Female: 8,919 (48.6%)				Male: 1,640 (51.8%) Female: 1,524 (48.2%)		
Twins Prevalence (%)	Single: 20,005 (93.1%) Double: 1,438 (6.6%) >2: 66 (0.3%)				Single: 17,999 (98.3%) Double: 317 (1.7%)				Single: 1,977 (62.5%) Double: 1,121 (35.4%) >2: 66 (2.1%)		
Maternal country of origin Prevalence (%)	Italy: 16,032 (74.5%) Romania: 1,769 (8.2%) Morocco: 768 (3.6%) Nigeria: 440 (2%) Albany: 292 (1.4%) Peru: 251 (1.2%) Egypt: 244 (1.1%) China: 152 (0.7%) Other: 1,561 (7.3%)				Italy: 13,649 (74.4%) Romania: 1,512 (8.2%) Morocco: 651 (3.6%) Nigeria: 368 (2%) Albany: 252 (1.4%) Peru: 214 (1.2%) Egypt: 218 (1.2%) China: 136 (0.7%) Other: 1,342 (7.3%)				Italy: 2,381 (75.2%) Romania: 257 (8.1%) Morocco: 116 (3.7%) Nigeria: 72 (2.3%) Albany: 40 (1.3%) Peru: 37 (1.2%) Egypt: 26 (0.8%) China: 17 (0.5%) Other: 221 (6.9%)		

Table 5. Characteristics of the cohort under analysis.

Temporal distribution of births was analyzed in order to identify the best smoothing strategy possible. Figure 10 graphically reports the daily number of births in the period under study (Jan 1st 2015 – Dec 31st 2017) and the natural spline smoothing function of calendar day Z with 14° of freedom utilized in the models.

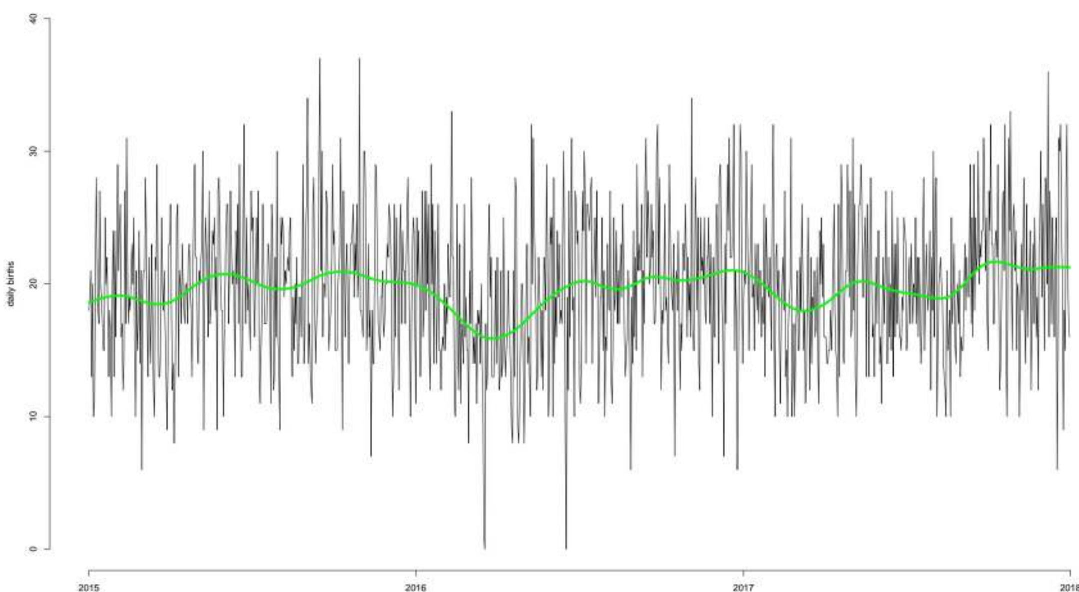


Figure 10. Daily incidence of births during the period under analysis, and the natural spline smoothing function of calendar day Z with 14° of freedom utilized in the Poisson non-stationary models (green line).

Figure 11 summarizes the average number of daily births (panel A) and PTBs (panel B) by month of the year, revealing the same unclear pattern of seasonality detectable in daily data in Figure 10 and addressed through the spline smoothing function.

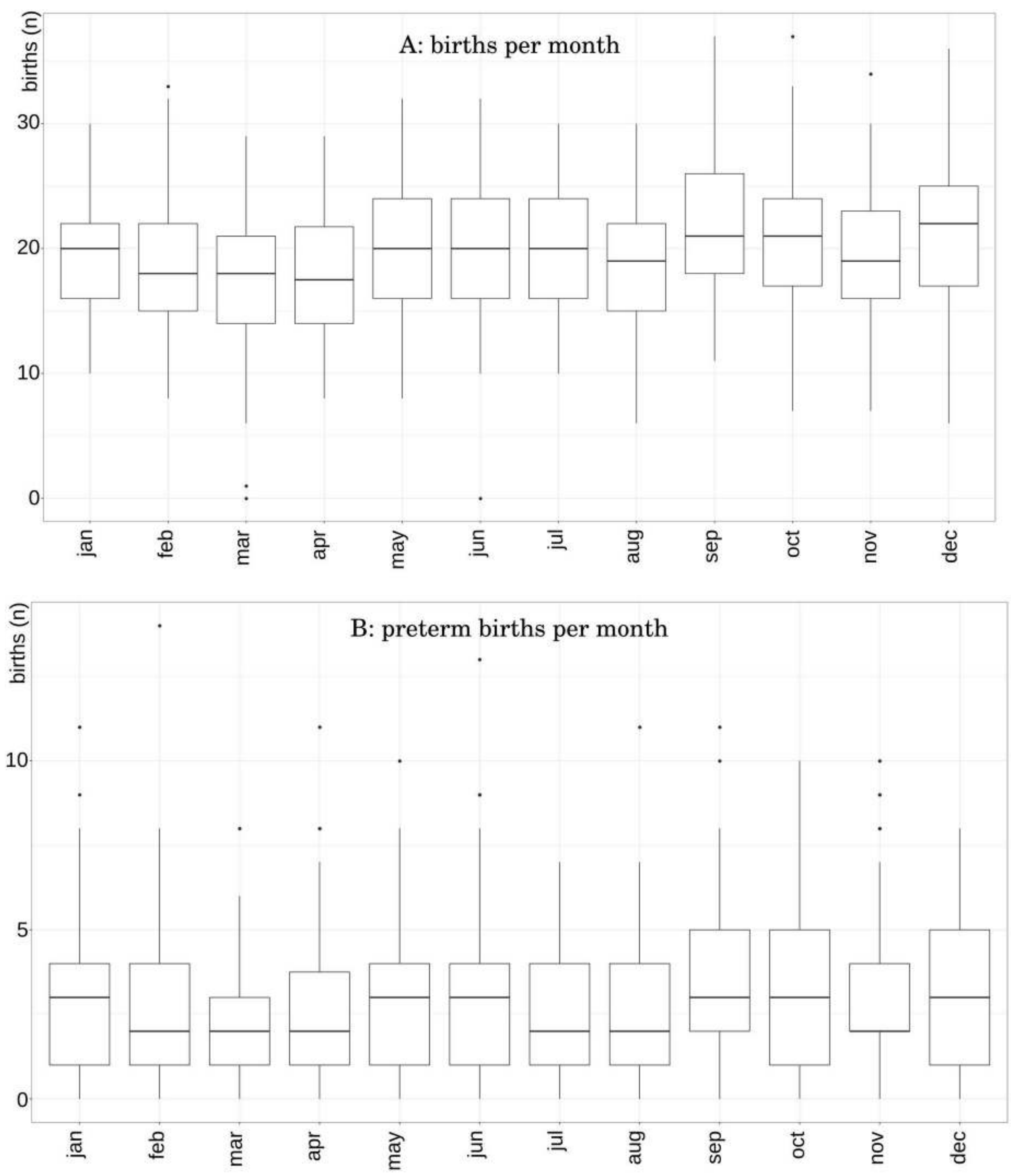


Figure 11. Average number of daily births (panel A) and preterm births (panel B) by month of the year during the 3-years period under analysis (Jan 1st 2015 – Dec 31st 2017).

Table 6 shows a general description of the median and mean concentrations of chemical and biological air pollutants in the area under study during the 3 years examined.

	Available data (days)	Median (IQR)	Interquartile ratio	min-max	Mean ± SD
PM2.5 (μg/m^3)	996	17.0 (24.0)	1.4	5.0 - 163.0	25.6 ± 21.4
NO_2 (μg/m^3)	1081	49.9 (29.6)	0.6	10.9 - 129.8	54.1 ± 21.0
O_3 (μg/m^3)	1083	19.9 (51.1)	2.6	1.7 - 262.3	41.1 ± 46.6
Aeroallergens (grains/m^3)	888	11.8 (42.1)	28.1	0.0 – 563.0	36.0 ± 61.4

Table 6. Distribution of daily concentrations of air pollution and aeroallergens during the three years examined (Jan 1st 2015 – Dec 31st 2017).

Chemical pollutants show rather high concentration levels, as well known in Turin area. Due to their natural seasonality, AA concentrations showed larger variability if compared to chemical air pollutants (even if chemicals also have a seasonality dependent on their primary and/or secondary origin). According to the origin, pollutants are categorized as primary and secondary. Primary pollutants are directly emitted from their sources, while secondary ones are either formed by the atmospheric transformation of primary pollutants or other chemical compounds. $PM_{2.5}$, and NO_2 (Figure 12a,12d) showed a prevailing maximum level during the coldest months, as expected due to its primary origin. Contrarily, O_3 shows a behavior typical of a pollutant of secondary origin (Figure 12b). Aeroallergens showed high concentrations during warm season, and virtually absent in winter (Figure 12c), as expected.

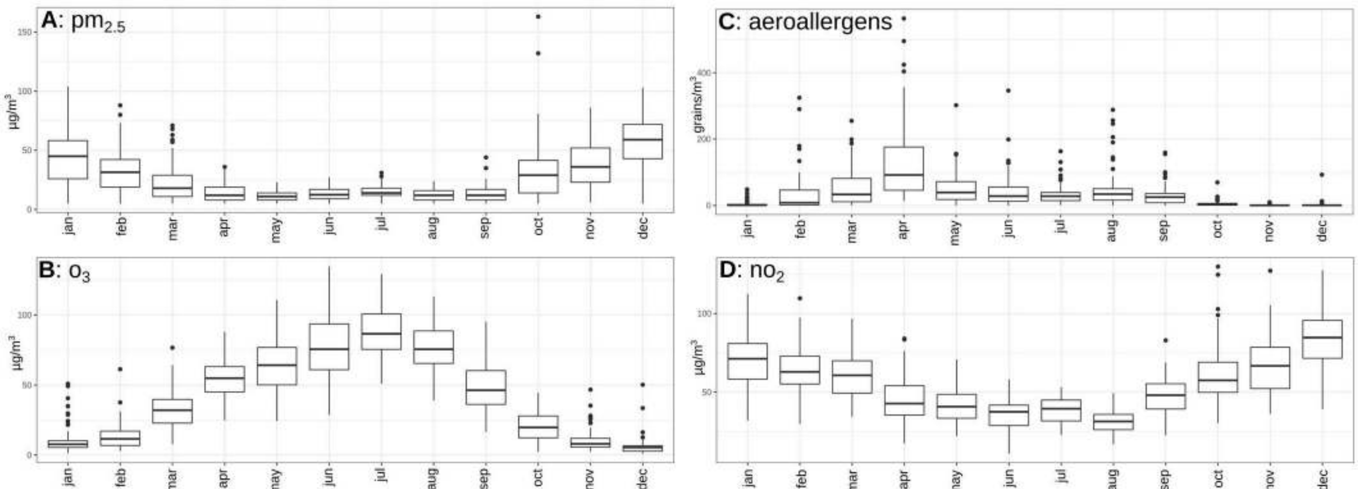


Figure 12. Mean concentration of each pollutant analyzed by month during the 3-years period under analysis.

A positive linear association between $PM_{2.5}$ and NO_2 was detected (r^2: 0.51, $p<0.01$) and negative associations of O_3 with $PM_{2.5}$ (r^2: -0.34, $p<0.01$) and NO_2 (r^2: -0.46, $p<0.01$). The associations between chemical pollutants and aeroallergens were weaker but still significant ($PM_{2.5}$ r^2: 0.06, $p<0.01$; NO_2 r^2: 0.04, $p<0.01$; O_3 r^2: 0.08, $p<0.01$). All correlation plots are reported in Figure 13.

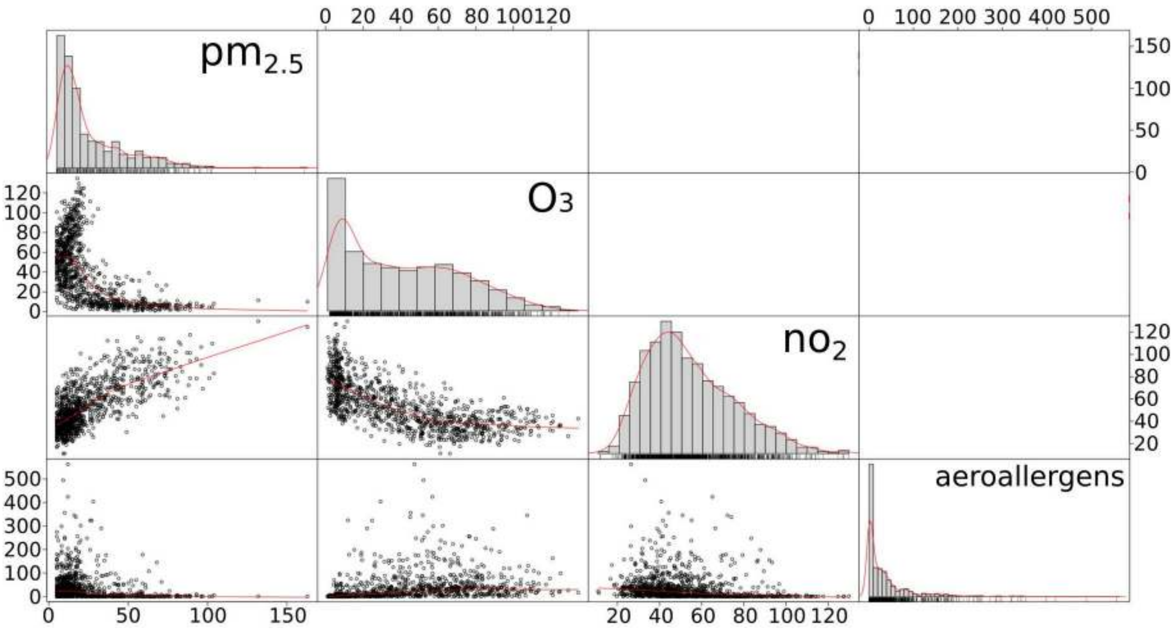


Figure 13. Correlation between pollutants under analysis.

Correlation between birth and environmental data: potential confounders and their relationship were considered before the analysis. Table 5 and Table 6 summarize the mean number of daily births, air pollution and aeroallergen concentrations according to the considered potential confounders. PTB were less frequent during weekends and holidays than during the other days. $PM_{2.5}$ and NO_2 were inversely associated with temperature, due to their secondary component nature, whose synthesis is favored by the sun. O_3 and AA resulted positively associated with temperature, as expected. Airborne pollution was lower during rainy days, as known.

Preterm: the associations between exposure variables and PTB through the Poisson model adjusted for all potential confounders (day of the week, holidays and summer population decrease, daily temperature, daily relative humidity, and cumulative daily precipitations), are summarized in Figure 14.

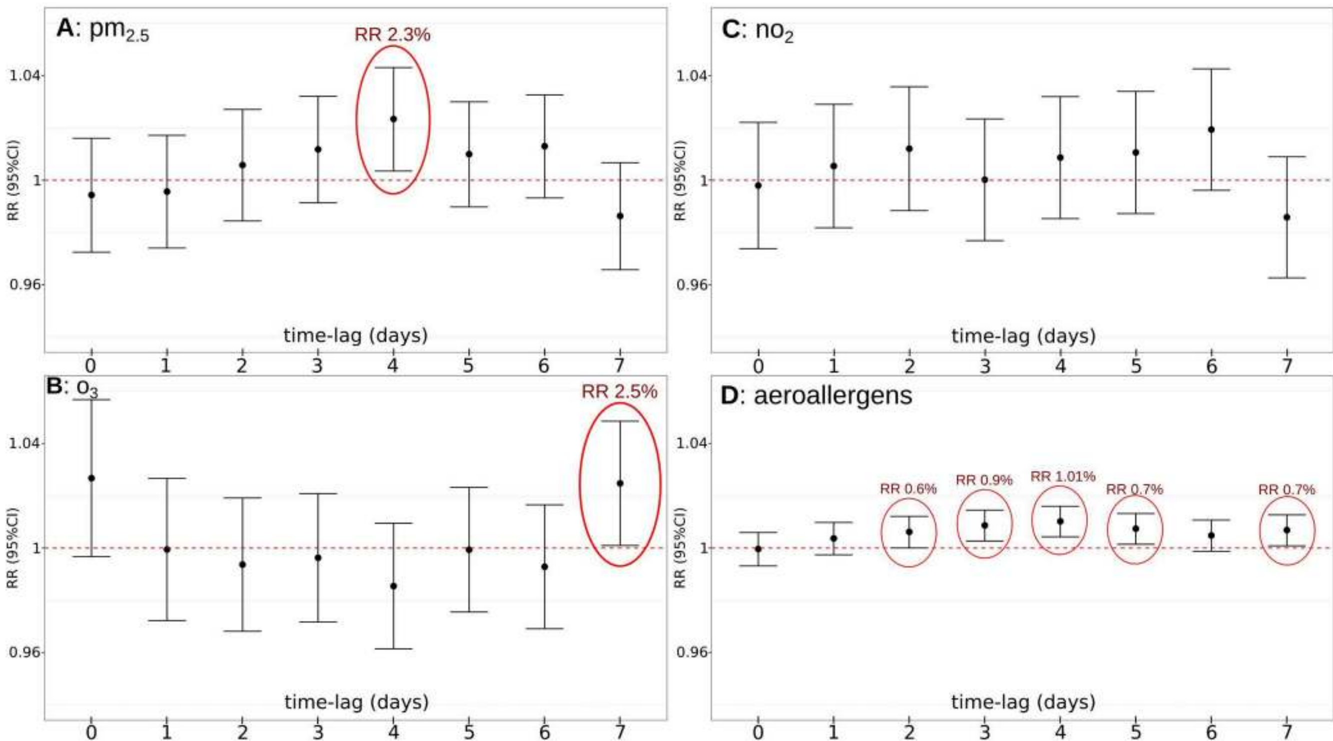


Figure 14. Graphical representation of chemical pollutants under analysis ($PM_{2.5}$, NO_2, O_3) and AA concentrations relative risk and confidence interval association with PTB as per 7-days time lags. Red circles underline significant factors with their relative lag RR. All models were adjusted for day of the week, holidays, summer population decrease, and yearly medium/long-term trend function.

With the exception of NO_2, all pollutants resulted to be significantly associated with PTB at some time lag:

1) An increase of 10 μg/m^3 of $PM_{2.5}$ is associated with a significant 1.023 (RR 95% C.I. 1.003 – 1.043, $p<0.05$) increased risk of PTB after 4 days.

2) An increase of 10 μg/m^3 of O_3 is associated with a significant 1.025 (RR 95% C.I. 1.001 – 1.048, $p<0.05$) increased risk of PTB after 7 days.

3) An increase of 10 grains/m^3 of aeroallergens is associated with an increased risk of PTB from 2 to 7 days after exposure:

- 1.006 (RR 95% C.I. 1.002 – 1.012, $p<0.05$) after 2 days

- 1.009 (RR 95% C.I. 1.002 – 1.014, $p<0.01$) after 3 days
- 1.01 (RR 95% C.I. 1.004 – 1.016, $p<0.01$) after 4 days
- 1.007 (RR 95% C.I. 1.001 – 1.013, $p<0.04$) after 5 days
- 1.007 (RR 95% C.I. 1.0008 – 1.03, $p<0.05$) after 7 days

This characteristics lag distribution effect detectable for AA: either the significant effect detectable at lag7 or the non significant detected at lag6, may be related to the specific cohort under analysis. The resulted obtained suggest that aeroallergens are likely to play a role in eliciting PTB, and this role is probably played over different days, probably underlying a stronger pro-inflammatory effect on some specific target, accordingly to their characteristics. Further analysis with larger, more different cohorts and more years of data collection are needed in order to better elucidate the real significant nature and lags of aeroallergens on PTB.

Term: correlation between at term delivery and chemical pollutants and aeroallergens was performed in the exact same way than for PTB and results are summarized in Figure 15.

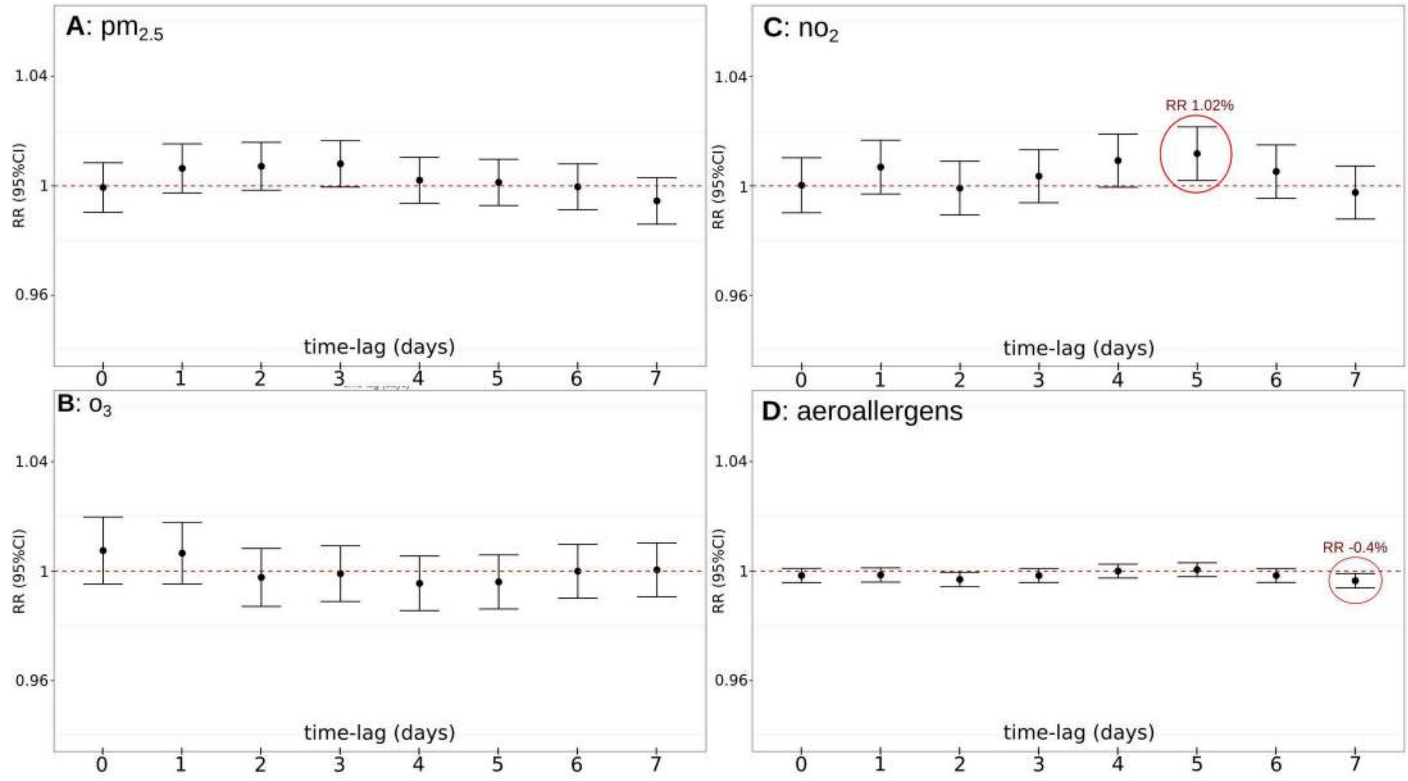


Figure 15. Graphical representation of chemical pollutants under analysis ($PM_{2.5}$, NO_2, O_3) and AA concentrations relative risk and confidence interval association with term delivery as per 7-days' time lags. Red circles underline significant factors with their relative lag RR. All models were adjusted for day of the week, holidays, summer population decrease, and yearly medium/long-term trend function.

Essentially, no pollutant resulted significantly associated with an increased risk of term delivery, with only 2 barely significant associations: an increase of 10 μg/m^3 of NO_2 resulted in an increased risk of 1.01 (RR 95% C.I. 1.002 – 1.021, $p<0.05$) after 5 days from exposure, and an increase of 10 grains/m^3 of aeroallergens resulted in a very fleble protective effect after 7 days from exposure (RR 0.996, RR 95% C.I. 0.999 – 0.993, $p<0.05$). No other air pollutants resulted associated at any time lag.

2.2.1.1.1.6. Discussion

This first line of study was specifically aimed to investigate the precise acute effect of air pollution on potentially eliciting preterm vs full-term birth. This acute focus on is the core behind model selection. Graphical analysis of the data and time distribution of daily births showed that the variable is best approximated by Poisson distribution, with higher λ in case of at-term births. This specific analysis approach allowed to show that a specific role of some specific pro-inflammatory air pollutants in eliciting preterm births is detectable. On the other hand, this acute inductive effect of air pollution, is undetectable in case of full-term births, in which I identify different effects of the various pollutants under analysis. Accordingly to this analysis aim, this difference may reflect some addictive effects of pregnancy-related pathological conditions that are brought to delivery by air pollution related stress in case of preterm births, while the same effect is likely compensated in physiological full-term births. These results show how air pollutants could play a significant role on PTB in the days immediately preceding delivery. In the 7 days preceding PTB delivery an increased exposure to airborne chemicals and aeroallergens is detectable. An exposure to $PM_{2.5}$ 5 days preceding PTB increases its risk of 2.3% for every 10 μg/m^3 $PM_{2.5}$. In a similar manner, 7 days before PTB, an exposure to O_3 increases its risk of 2.5% every 10 μg/m^3. Fascinatingly, an increased exposure of 10 grains/m3 of AA showed an almost continuous (from day 2 to day 7) increase of risk of ~1% on PTB, further testifying their strong inflammatory nature revealed in asthma and other diseases. Interestingly, none of these effects was detectable on physiological term delivery. Contrarily, AA resulted in a weak but protective effect and NO2, that showed no effect on PTB, resulted in an increased ~1% risk of delivery every 10 μg/m3 at 5 days time-lag. This fascinating difference between PTB and term births may testify the different trigger effect that these pollutants play in at risk vs physiological pregnancies, with the latter might have enough resources to cope with this pro-inflammatory noxa. This difference may be related to molecular differences in inflammatory pathways elicited by these pollutants on different physical states, whose deeper exploration is mandatory to better clarify PTB pathophysiology. This analysis is corrected for all possible environmental, seasonal and social confounding factors that frequently bias these analyses, as testified by the significant differences in PTB vs. term incidence once grouped by day of the week, holidays, population decrease, and temperature. Due to the huge portion of individuals exposed, air pollution may explain an important portion of cases, and allow preventive strategies. As example, considered an annual mean concentration of $PM_{2.5}$ of ~25 μg/m3 in Turin area versus 10 μg/m^3 in the near Switzerland, this single difference can lead to an increased ~5% risk of PTB only attributable to $PM_{2.5}$ levels. Considered ~1,050 PTB cases per year, this traduces in ~50 PTB cases only attributable to $PM_{2.5}$ exposures during days immediately preceding delivery. The focus on narrow time-frame before delivery is of core importance for both preterm delivery abrupt nature and potential preventive measures applicable. As example, strongly recommendations for at-risk women to avoid highly polluted areas in the last trimester of gestation may easily save several newborn lives and avoid PTB-related complications. Moreover, as previously explained, Il-1, IL-6, and TNF-a, among other inflammatory cytokines, exert a central role in PTB and are increasingly produced as a consequence of air pollution exposure. In this regard, pollutants-induced acute inflammation may offer the molecular explanation for this relationship, and inflammation can also be effectively and safely monitored through these biomarkers.

These findings suggest that a specific role of some specific pro-inflammatory air pollutants on delivery induction is detectable in PTB, while lower and different direction effects are detectable in full-term births. This fascinating difference advocate for extensive studies on air pollutants pro-

inflammatory pathophysiology mechanisms and their specific relationship with PTB. In this view, further identification of physiological differences underlying such relationship may bring to identify potential pharmacological targets able to further help in prevent PTB. Moreover, to the best of our knowledge, this analysis is the first analysis specifically focusing on acute eliciting effects of air pollutants on PTB vs. full-term births. I was able to detect pollutants effects that have not been previously reported, probably due to their specific pro-inflammatory short-term effect, which is undetectable in case of long-term studies, that represented the main approach to date. Thus, it is of main importance to broaden our time-frame reference when investigating such complex matters, in which an overlap of long- and short-term effects play role on the same outcome, elaborating specific models able to consider precise time frames to differently investigate chronic vs. acute effects. Finally, the detection of a precise short-term acute effect of air pollutants on PTB advocate for implementation of preventive regulations regarding air quality, especially in urban centers and for at-risk pregnancies. Considering the huge amount of pregnant women exposed, it is of central importance to further investigate this relationship and possible preventive thresholds, in order to rapidly and easily apply air pollution related recommendations for the last months of at-risk pregnancies and to further underlying the beneficial effect of policies aimed to contrast air pollution. These anti-pollution policies may prevent a huge number of PTB, alongside other countless beneficial effects on other aspects of human health, as well known and reported.

2.2.1.1.2. Study Line 2: Air Pollution and Pregnancy Acute Inflammation

2.2.1.1.2.1. Aim and Rationale

As previously described, Oxidative Stress (OS), is defined as an imbalance between pro-oxidant and antioxidant factors, it primarily derives from pro-inflammatory mechanisms action and is known to play a central role in impaired pregnancy and neonatal outcomes[111,112]. Several diseases have been linked to OS, both in pregnancy (eclampsia, miscarriage, preterm labor, and intrauterine growth restriction) and newborns (bronchopulmonary dysplasia, chronic lung disease, retinopathy, necrotizing enterocolitis, periventricular leukomalacia, and patent ductus arteriosus), but the mechanisms underlying such relationship are yet to be elucidated[111]. ROS represents the molecular basis of OS. ROS are molecules with one or more unpaired electrons, extremely unstable and highly reactive. When produced in excess in the body, they become mediators of cell and tissue damage, resulting in OS[113]. Several risk factors have been associated with OS development in pregnancy and newborns, among which maternal obesity seems to be strongly correlated[114–116]. Obesity is associated with elevated ROS generation, lipid peroxidation, decreased antioxidant levels, and subsequent OS in non-pregnant and pregnant women as too[117]. Similarly, during pregnancy, plasmatic concentration of pro-oxidants increases whereas enzymatic antioxidants activity decreases[118], possibly as a result of gestational metabolic and oxygen demands. Obesity during pregnancy may lead to an imbalance in maternal and fetal pro-oxidant/antioxidant status with greater lipid accumulation in the placenta, because of its high metabolic activity[119]. Furthermore, ROS generation in placentas increases according to maternal adiposity[120], and fetuses of obese mothers show increased production of proinflammatory adipocytokines, such as leptin, as well as increased insulin resistance and OS[116]. In turn, these factors are associated with a greater risk of developing certain diseases in later stages of life, including hypertension, type 2 diabetes mellitus (T2DM), and obesity[121]. Tobacco smoking is another factor strongly related to OS[122–124] and an historical risk factor during pregnancy. Cigarette smoking during pregnancy is one of the leading environmental factors that can adversely affect the health of mother and newborn. In this vulnerable phase, active and passive exposure to tobacco smoke have been shown to have deleterious effects on the development process and may result in permanent damage[125,126]. Maternal smoking during pregnancy strongly increases risk of preterm birth, intrauterine growth retardation, placental abruption, abortion, placenta praevia and several other pregnancy and neonatal complications[127–129]. Urbanization and traffic-air pollution represents one of the main environmental and public health challenges nowadays. In this regard, pregnancy and neonatal age is a particularly important window of susceptibility. Traffic-air exposure can affect both mother and the developing fetus. Growing evidence links exposure traffic-air pollutants during early life to adverse pregnancy outcomes, including preterm birth[16], reduced lung function, low birth weight, intrauterine growth retardation, impaired neurodevelopment and susceptibility to later metabolic diseases[130–132]. Moreover, exposure to air pollutants in early life can show up later consequences in childhood or in adulthood with chronic or lifelong conditions. The precise molecular mechanism that led up to air pollution-associated health outcomes has not been clearly elucidated but several studies attribute adverse birth and health outcomes to OS. Pro-inflammatory cytokines, such as IL-1β and IL-6, are directly correlated with OS levels, playing a central role in human pathology, including COVID-19, neurological, kidney, cardiac, infectious and several other diseases[132–136]. At the same manner, Isoprostanes are well known to be OS markers and mediators[137,138]. In this regard, the aim of the present study was to identify maternal and environmental risk factors directly associated with OS. I served maternal and neonatal pro-inflammatory biomarkers, namely Isoprostane and interleukins, in

order to first identify subjects with higher levels of OS. I then analyzed habits and exposure of such subjects in order to identify which of those could be causative related to higher levels of OS detected in some of the patients.

2.2.1.1.2.2. Research Plan

In this cohort study, the levels of OS markers listed below were detected in each mother-newborn couple, in order to identify subjects with higher levels of OS, through non-parametric comparative tests. The habits and exposure characteristics of such subjects were then analyzed and compared to OS levels detected, in order to recognize which of those factor may constitute risk factors for OS levels during pregnancy ad neonatal life. The study methods are specified here below

Questionnaire: data on individual, socio-demographic and lifestyle habits (e.g. clinical characteristics, scholarization, smoking, diet, working information, living environment) were obtained from mother during birth-hospitalization through a standardized questionnaire (PRAMS questionnaire), as described in a previous work from my group[139] which is presented in the next research line. Infant data were filled in by neonatologists.

Biological analysis: a pool of fresh urine was collected from each mother before the discharge and before the third day of the newborn's life. Infants' urine samples were collected by means of a specific polypropylene bag (Urinocol® Pediatric, BRAUN) placed inside each newborn diaper. Urine samples stored at -80°C until analysis.

Urinary biomarkers:

- *IL-1β*: is a master regulator of inflammation via controlling a variety of innate immune processes. The urinary IL-1β concentration was measured by a high-competitive enzyme-linked immunoassay (ELISA) performed with a specific microplate kit (Oxford, MI, USA), according to manufacturer's instructions.
- *IL-6*: is another historical inflammatory marker and a target biomarker of inflammatory diseases. The urinary IL-6 concentration was measured by a high-competitive enzyme-linked immunoassay (ELISA) performed with a specific microplate kit (Oxford, MI, USA), according to manufacturer's instructions.
- *IsoP*: is a specific indicator of lipid peroxidation, both in vitro and in vivo. The urinary 15-F2t-IsoP concentration was measured by a competitive enzyme-linked immunoassay (ELISA) performed with a specific microplate kit (Oxford, MI, USA), according to manufacturer's instructions.
- *Creatinine*: urinary creatinine was determined to normalize the excretion rate of the all the aforementioned urinary biomarkers as previously described.[179]

Statistical analysis: due to non-normal distribution, non-parametric correlation tests (Spearman correlation) were performed between maternal and neonatal inflammation marker levels and a group of independent variables: living environment (i.e. suburban vs high urbanization), life style habits (i.e. active and passive smoking exposure), anthropometric characteristics (i.e., weight and BMI during pregnancy), age and level of education. Data were expressed as mean ± SD or counts and percentages. BMI was used both as continuous and categorized variable (following WHO categorization: overweight or obese = OwO vs mother not overweight or obese = not OwO)[180], in order to perform multinomial logistic regression. In addition, IsoP concentrations were used both as

continuous and categorized variable (tertiles), in order to perform multinomial logistic regression. As the statistical distribution of the quantitative parameters was found to be non-Gaussian (Kolmogorov–Smirnov test), non-parametric tests were used to assess between group differences (Mann–Whitney U-test, Spearman correlations test). Two-sided P value < 0.05 was considered to indicate statistical significance. The Multinomial Logistic Regression model was used to assess the association between neonatal IsoP concentrations (tertiles), as dependent variable, and maternal OwO and urbanization level, included in the model as categorical variables, controlled for active and passive smoke. To account for intragroup correlation, the Huber-White standard error estimate for cluster sampling was applied. Results were reported as odds ratios (ORs) with 95% confidence intervals (CIs). All analyses were carried out using the software STATA 16.1 (StataCorp LLC: College Station, TX, USA).

2.2.1.1.2.3. Epidemiological Sample and Recruitment

The epidemiological sample was selected to be representative of the Turin newborn population (Turin, Piedmont Region). As I already presented in my previous article[139], subjects were recruited from July 2016 to October 2017 by consulting the register of births of the Sant'Anna Gynecological Hospital. Each mother and their relative newborn were enrolled and informed about the aim of this study and signed a written informed consent for both of them. The local Ethics Committee of "A.O.U. Città della Salute e della Scienza" of Turin (October 22nd, 2015, file No. CS/709) approved the study protocol and, subsequently, sensitive data were replaced by an anonymous identification code to ensure full privacy of data.

2.2.1.1.2.4. Results

A total number of 125 mother-newborn couples were enrolled in the analysis. Table 7 describes the characteristics of the study population.

PART A	**NEWBORN GROUP (N = 125)**	
Sex (Mean ± SD)	**male**	80 (64%)
	female	45 (36%)
Height at birth (Mean ± ND)	50.1 ± 6.7	
Weight at birth (Mean ± SD)	3.3 ± 0.4	
Cranial circumference (Mean ± SD)	33.9 ± 1.16	
Respiratory procedures N (%)	**NO**	110 (88%)
	YES	15 (12%)
General medical check-up N (%)	**NO**	45 (36%)
	YES	80 (64%)
Diseases at birth N (%)	**NO**	80 (64%)
	YES	45 (36%)
PART B	**MOTHER GROUP (N = 125)**	
Age (Mean ± SD)	34.2 ± 4.6	
Height (Mean ± SD)	164.1 ± 7.3	
Weight (Mean ± SD)	**Pre-pregnancy**	64.2 ± 13.8
	End of pregnancy	76.8 ± 13.7

BMI (Mean ± SD)	**Δ**	12.7 ± 4.6	
	Pre-pregnancy	23.8 ± 4.9	
	End of pregnancy	28.5 ± 4.8	
Educational level N (%)	**Low level**	15 (11.9%)	
	Medium level	49 (38.9 %)	
	High level	62 (49.2%)	
Occupation N (%)	**Yes**	112 (88.9%)	
	No	5 (4%)	
	Others	9 (7.1%)	
Living place N (%)	**Suburban**	69 (54.8%)	
	Urban	57 (45.2%)	
Smoking Habits N (%)	**No**	101 (80.8%)	
	Passive [Average exposure time: 3 hours]	16 (12.8%)	
	Yes	8 (6.4%)	
		< 10 sig/die[a]	4 (50%)
		> 10 sig/die	4 (50 %)

Table 7. (part A) Physical characteristics of the newborn sample at birth (i.e, sex, height, weight and cranial circumference), general information; Characteristics of mother population according to age, height, weight and BMI pre-pregnancy, weight and BMI at the end of pregnancy, educational level, occupation, living place, active and passive smoking (part B).

Table 8 reports the inflammation marker concentrations in the study population, split up for newborn and mother group.

	Inflammatory markers concentrations in the study population	
mean ± SD / median [Q1-Q3]	**Mother**	**Newborn**
IsoP	2.6 ± 2.7 / 1.9 [0.9-3.5]	6.6 ± 4.7 / 5.9 [3.2-8.9]
IL1	30.9 ± 75.7 / 3.8 [3.8-15.5]	/
IL6	77.4 ± 146 / 9.3 [9.3-86]	/

Table 8. Mean, standard deviation, median and Q1-Q3 of inflammatory marker levels in newborns and mothers (IL1 and IL6 were not screened in the newborn samples).

Among marker levels, significant associations of maternal and newborn IsoP levels ($p<0.01$) and maternal IL-1β and IL-6 levels were detected (figure 16). No other significant correlation was identified among maternal/newborn oxidative stress biomarkers (figure 16).

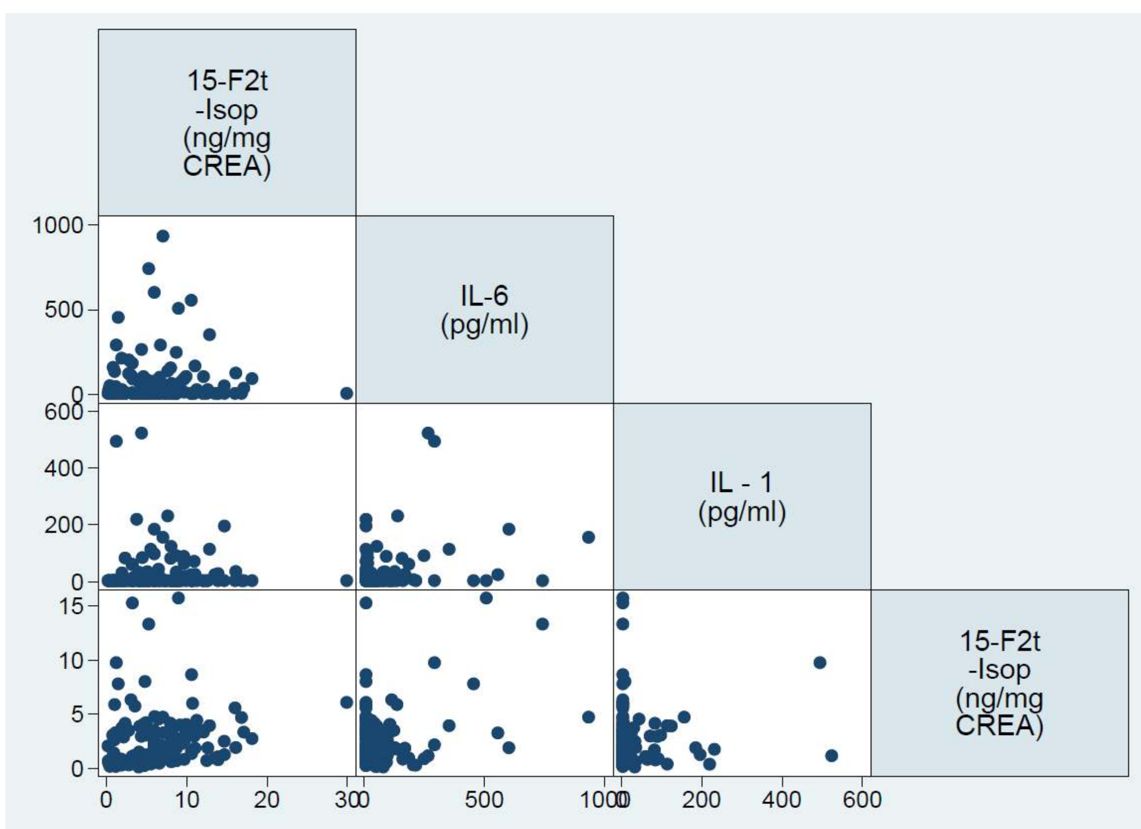


Figure 16. correlation matrix of maternal/newborn biomarker levels

Maternal BMI, living environment and tobacco smoke exposure (both active and passive) showed significant correlations with inflammatory markers. Maternal IL-1 levels resulted associated with active (cigarette/die, p<0.01 and smoke yes/no p: 0.01) and passive (exposition hours, p: 0.02 and exposition yes/no p: 0.03) smoke. No significant association was found between tobacco smoke exposure and inflammatory levels in newborns. Maternal BMI resulted to significantly correlate with maternal IL-1 (p: 0.01), IL-6 (p: 0.01), and neonatal Isop (p: 0.04) levels (figure 17).

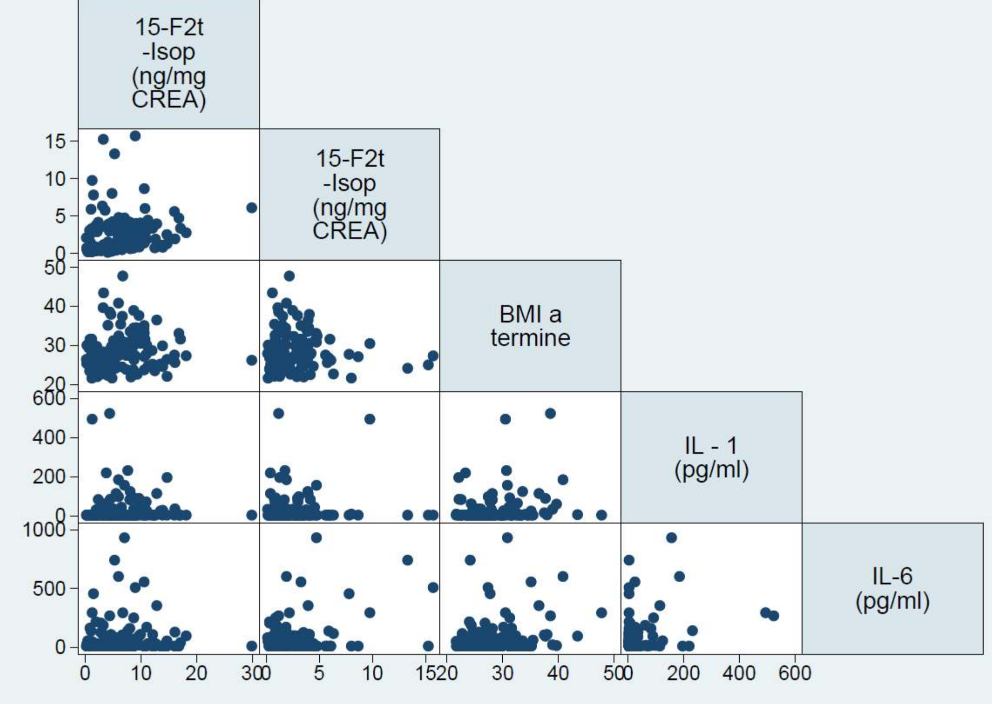


Figure 17. Correlation matrix between BMI at the term of pregnancy and maternal and newborn markers.

Moreover, higher urbanization level proved to be significantly associated with higher OS in newborns (Mann-Whitney = p: 0.02). Finally, I performed a Multinomial Logistic Regression between neonatal IsoP (categorized in tertiles) and maternal OwO (Obese/overweight vs Not Obese/overweight), Urbanization (High vs Low), active and passive smoke exposure (figure 18). The model proved that newborn with higher maternal OwO showed and increased likelihood of having higher oxidative stress and inflammation level (OR: 3.1, 95% CI 1.14–6.3, p = 0.04) while urbanization, even not statistically associated (OR: 1.74, 95% CI 1.00–1.89, p = 0.06), show a trend of decrease of OS levels in lower urbanized newborns (high urbanization was the level of reference used in the analysis, thus Figure 18 shows a protective effect for the “low urbanization group”). No other significant associations with active and passive tobacco smoke exposure were found.

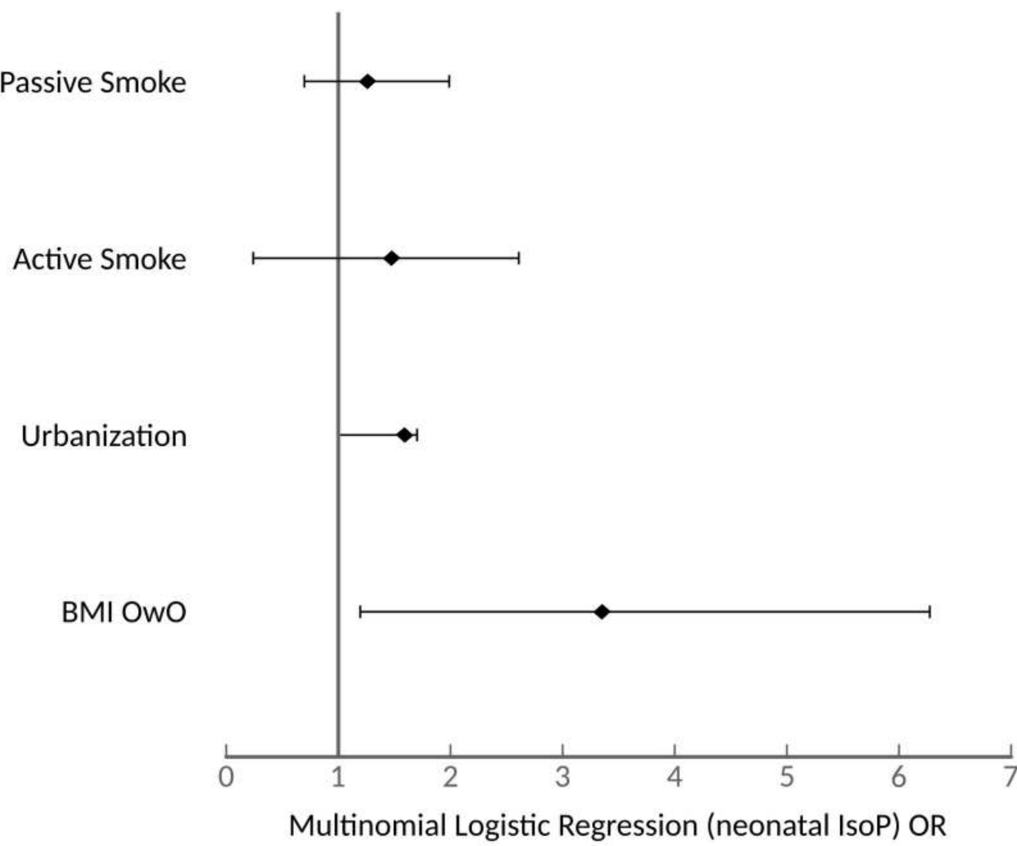


Figure 18. Multinomial Logistic Regression between neonatal IsoP (categorized in tertiles) and maternal OwO and urbanization, controlled for active and passive smoke exposure.

2.2.1.1.2.5. Discussion

This comprehensive analysis of maternal clinical characteristics and lifestyle habits allowed identification of 3 main factors strongly associated with OS: maternal BMI, urbanization levels and tobacco smoke exposure. Importantly, a direct and significant association between maternal and neonatal IsoP levels was detected, proving a direct relationship and contribution between mother and newborns in terms of OS. These results indicates how strongly maternal OS is able to influence not only foetus development but directly increase neonatal OS levels, advocating increased focused to preventive strategies and health education campaigns aimed to control and reduce maternal OS

levels during pregnancy. In this regard, my analysis showed some important factors that could be controlled in order to reduce maternal OS levels. Maternal overweight/obesity proved to significantly increase OS levels in newborns. This is an important risk factors for impaired pregnancy and neonatal outcomes[140,141]. First, subclinical metabolic dysfunctions in obese women are proved to be associate with adverse pregnancy outcomes, such as GDM and pre-eclampsia[142]. Besides, obese women have an increased risk of spontaneous preterm births and perinatal mortality[143]. Neonates born to obese women have an increased risk of overgrowth, greater insulin resistance and they are more susceptible to develop chronic diseases during adolescence and lifespan[141,144]. In particular, newborns of obese mothers have higher concentrations pro-oxidant and pro-inflammatory species (e.g.IL-6) than infants of lean mothers[140,141]. In these perspectives, my results are in line with the previous literature and proved a direct link between maternal and neonatal OS levels. Moreover, this direct link may help explaining underlying mechanisms of different neonatal chronic diseases associated with maternal overweight/obesity, such as bronchopulmonary dysplasia, chronic lung disease, retinopathy, necrotizing enterocolitis, periventricular leukomalacia or patent ductus arteriosus. Similarly, I also showed a trend, even not statistically significant, between maternal urban environmental exposure and increased neonatal OS levels. Air pollution is gaining evidence as a detrimental factor in pregnancy and neonatal outcomes and most of its consequences on human health are associated with inflammation and OS levels. In general, the urban environment has a number of features that could have adverse effects, mainly on children's respiratory health, especially during the first few years of life when the immune system are rapidly developing[145]. Prenatal and maternal exposures can affect immune development and it is evident that the variability of newborn immune responses is not random: in fact, the development of the fetal immune system during the perinatal and postnatal periods appears to be responsive to maternal characteristics and experiences[145,146]. These results of direct association with neonatal OS levels show the important association between maternal and perinatal urban environmental exposure and the immune and inflammatory response, underlying the strongly connection between environment and health outcome. This could be an important evidence on which develop preventive measures aimed to reduce risks associated with air pollution exposure in this delicate life phase. Finally, smoking in pregnancy constitutes the largest remediable risk factor for maternal and child health, such as increased risks of obstetric complications, higher rates of spontaneous abortions, ectopic pregnancies, placental abruption, premature labour, and preterm birth[147,148]. Besides, maternal tobacco smoking is a source of active oxidizing agents that release free radicals harmful for health[149]. Confirming the literature, these results showed that tobacco smoke, both in active and passive form, proved a strong influence on increased maternal OS and inflammation levels, as proved by its significant association with IL-1/IL-6. Overall, this study confirmed literature evidence of maternal overweight/obesity, air pollution exposure and tobacco smoking influence on OS levels. Moreover, the direct association identified between neonatal and maternal OS levels further advocate for public health preventive measures that, reducing risks connected to lifestyle habits, can improve pregnancy and neonatal outcomes.

2.2.1.2. Bisphenols as Endocrine Disruptors in Newborns

2.2.1.2.1. Study Line 3: Bisphenols detection in Newborns

2.2.1.2.1.1. Aim and Rationale

Bisphenols (BPs) have long been used for the production of polycarbonate plastics and epoxy resins[150]. BPs are well-known as harmful substances for human health, and they are usually ingested mainly through the diet, but BP intake can also occur by inhalation or dermal contamination. Once they have entered the body, BPs exert estrogenic and/or oxidant activity[140,141], which is particularly harmful during pre- and neonatal life[152,153]. BPs are detoxified to inactive forms in the liver, primarily through conjugation and glucuronidation, then excreted in the urine within 2 to 6 h. Thus, the detection of conjugated BP forms is critical for determining BP risk exposure. In the glucuronidated form, BPs—particularly Bisphenol A (BPA) (their main representative member)—are inactive, whereas their free (unconjugated) forms, albeit basically unstable, promote bio toxicity through mild estrogenic activity[154]. BPs are prevalently found in urine (75–90%), but they are also traceable, albeit at lower concentrations, in other body fluids such as blood, breast milk, semen, cord blood, fetal serum, and placental tissue[155–158]. In this scenario, gestation appears to be a critical window of fetal exposure to BPs, as they trigger cellular responses even at very low doses[149], influencing sex- and gender-differentiation morphology or leading to immune hyper-responsiveness[160]. Infants appear to be particularly susceptible to the harmful effects of BPs[151], and prior studies have demonstrated that exposure to these chemicals can occur through both breastfeeding and skin contact with plastic devices (e.g., polycarbonate feeding bottles and pacifiers) routinely employed in neonatal intensive care units[159,160,162]. This harmful exposure is further exacerbated by the inefficient UDP-glucuronosyltransferase system in newborns, which is required for BP detoxification and not completely developed until ~2–3 months of age[162–164]. In 2015, following a public consultation on the harmful effects of BPA on human health[227], the European Food Safety Authority (EFSA) reduced the temporary tolerable daily intake (t-TDI) threshold of BPA from 50 to 4 µg/kg bw/day. Since then, plastics manufacturers have been striving to replace BPA with alternative compounds, including bisphenol S (BPS), its closest chemical relative. However, BPS has also recently been shown to exert genotoxic and biological activity similar to that of BPA[165,166]. These lines of evidence indicate that these BPA alternative compounds may also become a serious public health concern in the near future due to lack of regulatory limits[167,168]. Given these premises, the aim of the present study was to assess the modulation and possible health effects of BP exposure on perinatal and neonatal life. The final aim of this study was to investigate direct and indirect pathways and effects of BP exposure during pregnancy and the first days of life. Toward this end, I determined both BPA and BPS concentration levels in a group of newborn babies and their mothers. I particularly intended to focus my study on the identification of factors, devices, and behaviors responsible for high BP uptake in this sensitive population in the very first days of life. Finally, I evaluated wether higher BP concentrations are correlated with specific hospital procedures or babies health status immediately after birth.

2.2.1.2.1.2. Research Plan

In this cohort study, BPs levels were tested in the urine of each mother-newborn couple. Such levels were tested both in form of free and conjugated in both component of the dyad. Characteristics and maneuvers that these newborns underwent to, both in terms of delivery and postnatally, were then reviewed in order to identify risk factors associated with higher levels of BPs in the cohort, mainly through non parametric association tests.

Questionnaire: a standardized questionnaire (PRAMS questionnaire[169]) was administered to the mothers during hospitalization. Questions regarded individual, socio-demographic, and clinical

characteristics such as smoking, diet, the mother's working and lifestyle habits, especially in the last month of pregnancy, and education. Infant data, filled in by neonatologists, included gestational age and weight, sex, head circumference, length, the baby's health status at the time of birth, and any medical check-ups.

Biological Analysis: a pool of fresh urine was collected from each subject, for both babies and their mothers. Infants' urine samples were collected by means of a specific BP-free polypropylene bag (Urinocol Pediatric, B BRAUN, Milan, Italy) placed inside each newborn's diaper from their birth to their third day of life, during hospitalization. Mother urine samples were collected in glass tubes (Pyrex, CORNING, Corning, NY, USA) during hospitalization. All urine samples (infants and mothers) were then stored in glass mini-jars, pre-treated with methanol to reduce the risk of environmental contamination, and stored at −80°C until analysis.

Free BPA and BPS collection: 2 milliliters of urine was transferred in a tube pre-treated with methanol, vortexed, and acidified with HCl at pH 1. Subsequently, NaCl was added (salting-out process) together with 30 μL of a standard solution containing BPS-d8 and BPA-d16 with a final concentration of 0.10 and 0.05 mg/L, respectively. Next, 750 μL of chloroform and 500 μL of acetone were added for liquid–liquid extraction (LLE). Each sample was vortexed and then sonicated for 1 min. Finally, the sample was centrifuged at 1250× g for 20 min at room temperature, and the resulting supernatant was collected and transferred to a new pre-treated vial. LLE was repeated by adding 800 μL of chloroform. All of the supernatant was then brought to dryness by means of a nitrogen stream at room temperature and then resuspended in 100 μL of a solution composed of 5 mM ammonium acetate in ultrapure water (70%) and 5 mM ammonium acetate in acetonitrile (30%).

Total (Free + Conjugated) BPA and BPS collection: 4 milliliters of fresh urine was thawed, vortexed, and aliquoted in two vials: 2 mL was used for the determination of free BP, while the other 2 mL was used for the determination of bisphenol-glucuronides after 12 h of incubation with 20 units of β-glucuronidase/arylsulfatase.

Conjugated BPA and BPS collection: After 12 h of incubation with the enzyme β-glucuronidase/arylsulfatase, the same extraction procedure was performed as described above.

BPA and BPS Analysis: The instrumental analysis for both the free and conjugated procedures was carried out with a UPLC Shimadzu Nexera X2 System interfaced through an ESI source (Turbo Ion Spray™) to a Sciex 5500 QTrap mass spectrometer. The analytes were detected in negative ion mode. Concerning the LC set-up, the chromatographic column consisted of a Phenomenex (Bologna, Italy) Luna Omega C18 (1.6 μm, 100 mm × 2.1 mm), and the mobile-phase solvents for reverse-phase analysis were 5 mM ammonium acetate in water and 5 mM ammonium acetate in acetonitrile. The final pH value of both solvents was corrected to 7.5–8.0 by adding a few drops of 33% ammonia solution to ensure a more significant presence of BPA and BPS in deprotonated form. Concerning chromatographic gradient, the flow rate was set at 0.35 mL min^{-1}. Mobile phases, initially consisting of 30% of 5 mM ammonium acetate in acetonitrile and 70% of 5 mM ammonium acetate in water, were held for 2 min then increased linearly to 100% organic solvent over 12 min, held there for 3 min, and finally brought back to the initial condition in 0.1 min. Moreover, 7 min of re-equilibration was necessary between samples. Moving to MS parameters, the drying gas (nitrogen) was set at 325 °C, 20.0 psi, and 10 L min^{-1}; capillary voltage was set at 2000 V. Data acquisition was made in multiple reaction monitoring (MRM) mode by monitoring the

transitions of deprotonated ions [M-H]−. For each analyte, two transitions were monitored: one for quantification and the other for confirmation. All the MS/MS parameters are described in Table 1. Procedural blank samples with ultrapure water in place of urine were collected, extracted, and analyzed by HPLC-MS/MS following the same protocol. In all processed blanks, I did not observe BPA contamination above the method limit of detection (LOD). All solvents and reagents (i.e., acetonitrile, acetone, ammonium acetate, and chloroform) were from VWR International (Radnor, PA, USA). All aqueous solutions were prepared with ultrapure water, Millipore Milli-QTM. Analytical standard compounds were purchased from Sigma-Aldrich (Milan, Italy). The method was based upon previously published methods with slight modifications[158]. To check method performances, I validated my procedure by verifying quality parameters according to Eurachem guidelines[160]. I checked for selectivity, sensitivity, linearity, accuracy, and repeatability. In particular, method sensitivity LOD was 0.0035 ng mL−1 for BPA and 0.0030 ng mL−1 for BPS.

Creatinine: urinary creatinine was determined to normalize the excretion rate of all the aforementioned urinary biomarkers as previously described[125].

2.2.1.2.1.3. Epidemiological Sample and Recruitment

The epidemiological sample was recruited from July 2016 to October 2017 by consulting the register of births of the Sant'Anna Gynecological Hospital (Turin, Piedmont Region), following these selection criteria: (1) full-term pregnancy (>37 GA-W-); (2) physiological pregnancy conditions; (3) no drugs or pharmacological treatment during pregnancy; (4) single babies (no twins) with Apgar scores > 5; and (5) healthy babies at birth (not admitted to the neonatal intensive care unit or in life-threatening conditions). Each adult subject was informed about the aim of this study and signed their written informed consent. Newborns were enrolled in the study upon written authorization from both parents. Sensitive data were replaced by anonymous identification codes to ensure full privacy of data. The local Ethics Committee of "A.O.U. Città della Salute e della Scienza" of Turin (22 October 2015, file No. CS/709) approved the study protocol.

2.2.1.2.1.4. Results

Table 9 describes the characteristics of the study population (newborns and mothers). 200 mothers with their babies were enrolled, but only 134 subjects followed all the selection criteria and were eligible. In fact, in addition to the selection criteria, about 33% of enrolled mothers/babies were excluded for three other main reasons:

- Not-fully completed information (questionnaire and infant data);
- Insufficient urine samples (mothers: <30 mL/babies: <5 mL);
- Withdrawal of informed consent.

PART A	NEWBORN GROUP (N = 125)	
Sex (Mean ± SD)	**male**	84 (62.7%)
	female	50 (37.3%)
Height at birth (Mean ± ND)	50.1 ± 6.7	
Weight at birth (Mean ± SD)	3.3 ± 0.4	
Cranial circumference (Mean ± SD)	38.3 ± 1.6	
Breast-fed N (%)	111 (82.8%)	

Infant formula dispensed N (%)	36 (26.9%)		
Glucose dispensed N (%)	38 (28.4%)		
Pacifier dispensed N (%)	42 (31.3%)		
Respiratory procedures N (%)	16 (11.9%)		
General medical check-up N (%)	85 (63.4%)		
Diseases at birth N (%)	32 (23.9%)		
PART B	**MOTHER GROUP (N = 125)**		
Age (Mean ± SD)	33.9 ± 4.7		
Height (Mean ± SD)	164.1 ± 7.1		
Weight (Mean ± SD)	**Pre-pregnancy**	64.4 ± 13.9	
	End of pregnancy	76.8 ± 13.7	
	Δ	12.3 ± 4.8	
BMI (Mean ± SD)	**Pre-pregnancy**	23.9 ± 4.9	
	End of pregnancy	28.5 ± 4.8	
Educational level N (%)	**Low level**	20 (14.9%)	
	Medium level	51 (38%)	
	High level	63 (47.1%)	
Occupation N (%)	**Yes**	114 (85.1%)	
	No	6 (4.5%)	
	Others	14 (10.4%)	
Living place N (%)	**Rural**	40 (29.9%)	
	Suburban	51 (38.1%)	
	Urban	43 (32.1%)	
Smoking Habits N (%)	**No**	98 (73.1%)	
	Passive [Average exposure time: 3 hours]	21 (15.7%)	
	Yes	15 (11.2%)	
		< 10 sig/die[a]	8 (53.3%)
		> 10 sig/die	7 (46.7%)
Delivery N (%)	Vaginal	81 (60.4%)	
	Caesarian	46 (34.3%)	
	Vacuum	7 (5.2%)	

Table 9. Physical characteristics of the newborn sample at birth concentrations (part A); Characteristics of mother population (part B).

The newborn sample was homogeneous in terms of height, weight, and cranial circumference (Table 9, part A). About 80% of the babies were breastfed, with only 27% of babies (n = 36) receiving infant formula as extra integration before being discharged from the hospital. Moreover, about 30% of newborns were administered oral glucose as extra integration and/or pacifiers. Sixteen babies (12%) required respiratory procedures, while 24% of newborns (n = 32) were diagnosed with potentially pathological conditions after general medical check-up. The sample of mothers was homogenous according to age, height, weight, and BMI variables. Among mothers (n = 134), 94% were European while 6% were classified as other nationalities (non-European). The education level was low (i.e., primary and middle school) in 14.9% of subjects, while it was medium and high

(secondary school and above) in 38% and 47.1% of subjects, respectively. Fifteen mothers (11%) reported being active smokers during pregnancy, whereas twenty-one (15.7%) were passive smokers. Table 10 reports the BP concentrations in the study population, split up for newborn and mother groups. On average, newborns had BP levels twice as high as mothers, except for free BP forms (both for free BPA and free BPS). The determination of conjugated BPs was performed by an indirect procedure based on the enzymatic cleavage of the glucuronic acid/sulfuric acid moiety from the phenol group(s) of BPs. A recent report[171] evidenced the possibility of underestimating the real quantity of bound BPs because of different yields of enzymatic reactions together with the chance of the formation of intermediates. In order to evaluate reaction yields, I performed the enzymatic incubation for different times (from 4 to 12 h), confirming that a stable result was reached at 12 h. I also checked for the presence of partially cleaved BP metabolites by high-resolution mass-spectrometry analysis (data not shown), excluding this likelihood.

BP Levels (ng/mL) (Mean ± SD; C.I. Range)	**Newborn**	**Mother**	**BP Levels (ng/mg CREA) (Mean ± SD; I.C. Range)**	**Newborn**	**Mother**
Total BPA	0.13 ± 0.3 [0.02/0.74]	0.15 ± 0.23 [<LOD/0.62]	Total BPA	0.48 ± 1.13 [0.02/2.5]	0.24 ± 0.43 [<LOD/0.9]
Conjugated BPA	0.11 ± 0.3 [0.01/0.6]	0.13 ± 0.2 [<LOD/0.53]	Conjugated BPA	0.41 ± 1.05 [>0.004/2.3]	0.2 ± 0.35 [<LOD/0.75]
Free BPA	0.01 ± 0.37 [<LOD/0.62]	0.02 ± 0.06 [<LOD/0.13]	Free BPA	0.07 ± 0.09 [<0.01/0.25]	0.03 ± 0.16 [<LOD/0.1]
Total BPS	0.09 ± 0.2 [<LOQ/0.6]	0.01 ± 0.03 [<LOQ/0.06]	Total BPS	0.2 ± 0.53 [<LOD/0.1]	0.04 ± 0.15 [<LOD/0.1]
Conjugated BPS	0.08 ± 0.17 [<LOD/0.5]	0.005 ± 0.04 [<LOD/0.05]	Conjugated BPS	0.15 ± 0.5 [<LOD/0.9]	0.02 ± 0.14 [<LOD/0.04]
Free BPS	0.01 ± 0.04 [<LOD/0.05]	0.004 ± 0.002 [<LOD/0.02]	Free BPS	0.02 ± 0.07 [<LOD/0.01]	0.02 ± 0.06 [<LOD/0.03]
Creatinine (CREA) (mg/L)	MOTHER		0.66 ± 0.5 [0.1/1.9]		
	NEWBORN		0.8 ± 0.4 [0.1/1.5]		

Table 10. BP levels (BPA and BPS) in the newborn and mother groups. Data is expressed in terms of ng/mL (column 2 and 3) and ng/mg$_{CREA}$ (column 5 and 6) for both newborns and mothers.

Biological variables had non-Gaussian distributions, thus non-parametric tests (Mann–Whitney test and Spearman's rho correlations) were performed. In the mother group, no BP differences were detected in terms of anthropometric characteristics (i.e., weight and BMI composition pre-pregnancy and at the end of pregnancy), age, level of education, and nationality. Besides, the analysis showed no significant correlation between maternal and neonatal BP concentration levels. In the newborn group, BP analysis revealed no significant differences (Mann–Whitney U-test) between mothers' lifestyle habits, such as daily consumption of food and beverages in plastic packaging (BPA total $p = 0.6$; BPS total $p = 0.2$), daily use of a microwave and dishwasher (BPA total $p = 0.5$; BPS total $p = 0.8$), or make-up use (BPA total $p = 0.7$; BPS total $p = 0.08$). In newborns, the most important correlation was found between pacifier use and BPS concentrations, in terms of both total BPS (Mann–Whitney U-test, $p = 0.004$; Spearman's rho = 0.182, $p = 0.035$

Figure 19B) and free BPS concentrations (Mann–Whitney U-test, $p = 0.003$; Spearman's rho = 0.190, $p = 0.03$ Figure 19A).

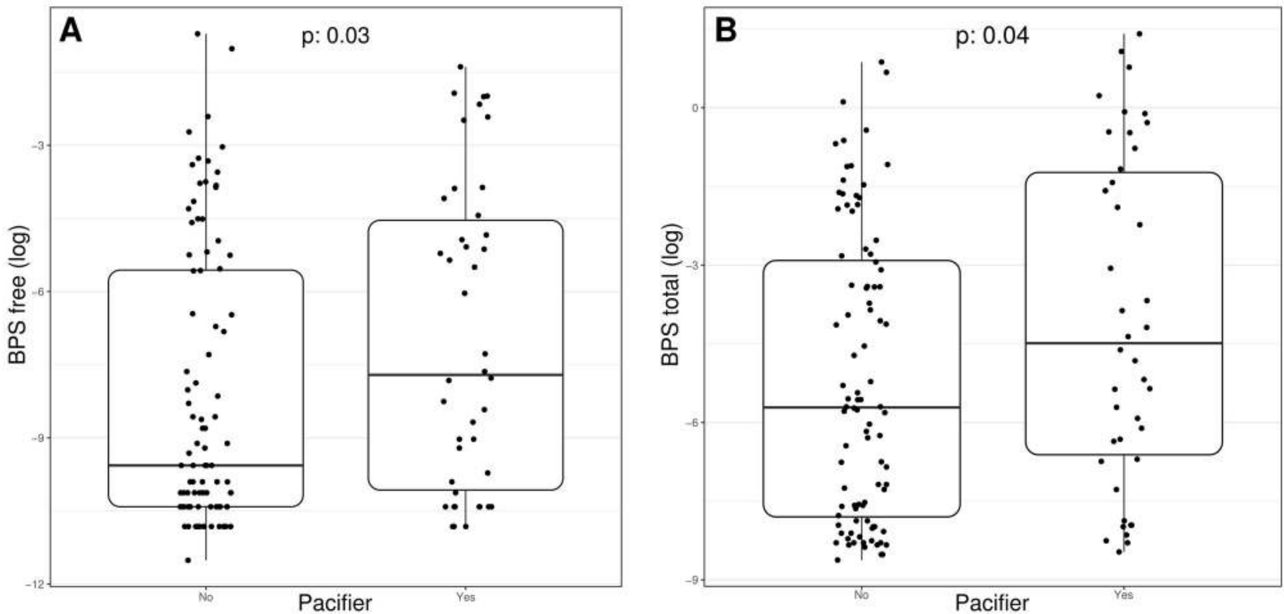


Figure 19. Non-parametric correlations between pacifier use and BPS total (A) and free BPS (B) concentrations in newborns.

A significant correlation was also found between oral glucose administration and concentration levels of free BPA (Mann–Whitney U-test, $p = 0.003$; Spearman's rho = 0.182, $p = 0.035$ Figure 20).

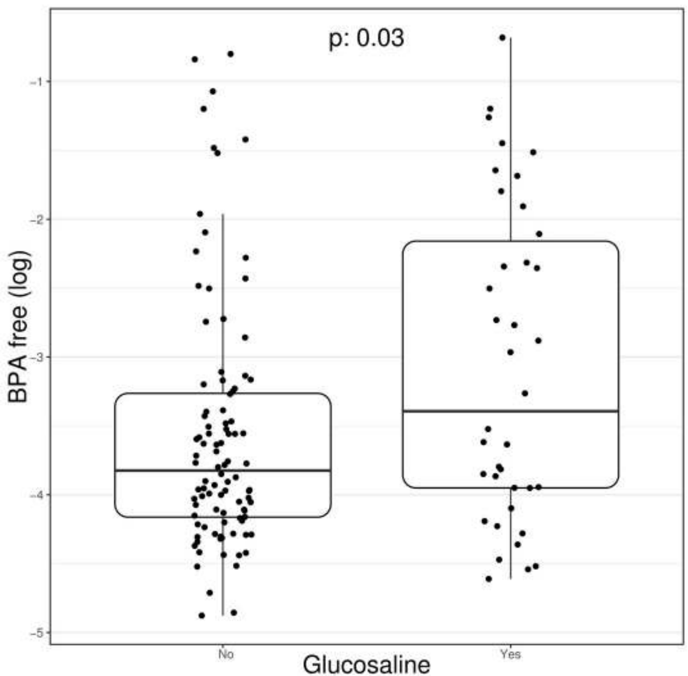


Figure 20. Non-parametric correlations between oral glucose administration and free BPA levels in newborns.

Subsequently, positive correlations were found between drugs administered to the mother during delivery and total BPA (Mann–Whitney U-test, $p = 0.05$; Spearman's rho = 0.146, $p = 0.04$), total BPS (Mann–Whitney U-test, $p = 0.001$; Spearman's rho = 0.198, $p = 0.001$), and free BPS (Mann–

Whitney U-test, p = 0.05; Spearman's rho = 0.170, p = 0.05) concentrations. In further analyses, despite the exclusion of pre-term births from the population study, positive correlations were found between BPA total (Mann–Whitney U-test p = 0.001; Spearman's rho = 0.172, p = 0.001) and free BPA (Mann–Whitney U-test p = 0.001; Spearman's rho = 0.256, p = 0.003) concentrations and neonatal glycemic control after birth.

2.2.1.2.1.5. Discussion

Neonatal routine procedures may expose newborns to BPs. The observation that newborns receiving oral glucose after birth displayed high BPA levels, especially in its toxic free form, suggests that oral glucose administration through non-BPA-free containers and syringes should be avoided. Besides, drug administration to the mother via non-BPA-free containers and syringes during delivery is also positively correlated with higher levels of both total BPS and BPA concentrations in newborns. Using syringes to administer dextrose solutions or other parenteral products to newborns is a widespread practice, and these results show the risk associated with such a practice[172]. Thus, these findings advance, once again, the need for extending EU regulations to non-alimentary or medical sources of BP contamination. They also call for further studies on BP exposure in newborns in order to identify which devices and procedures are best suited to minimize BP intake in this extremely vulnerable population. In fact, medical devices are a specific product category in which BP may be present, and preventive actions must be adopted to reduce or even eliminate this exposure source[173]. Newborn exposure to BPS can still occur through pacifiers. The observation that the use of pacifiers was associated with higher levels of both total and free BPS indicates the urgent need of in-depth analyses and regulations of BPA alternatives such as BPS. The vast majority of regulations have been directed towards BPA-containing plastics, while they have failed to take into account other plastic products containing alternative BPs. In fact, BPS was regarded as a "safe" alternative to BPA because of its greater stability against high temperatures and resistance to sunlight compared to BPA. No previous epidemiological study has explored the developmental effects of BPS, but in vitro studies have shown that BPS can bind to estrogenic receptors (ERs) and drive estrogen-induced gene transcription[174,175]. Furthermore, studies in zebrafish have shown that BPS exposure can alter the homeostasis of sex steroid hormones and disrupt reproduction or development, miming BPA exposure[176,177]. Current knowledge on the impact of BPS exposure is limited, but, on these bases, further studies are advocated to deepen knowledge of human BPS intake and exposure, focusing more on newborn health effects[178]. This analysis showed that the lifestyle habits of pregnant women in the last month of pregnancy did not cause significant changes in newborns' BP levels. In fact, no stratification effect was detected due to lifestyle, food habits, education level, or work. This could be explained by analyzing some pregnancy-related factors, such as more attention in the diet habits of pregnant women (no consumption of plastic-free or precooked food) or lifestyle habits. BP levels could exacerbate health conditions at birth and increase the duration of hospitalization. This analysis also showed that newborns with higher levels of free BPA/BPS had higher risk of diseases at birth and higher duration of hospitalization. Further investigation is needed to elucidate if specific conditions are related to this association and how they are related. Nevertheless, BPs' toxic effects may play a role in worsening a newborn's clinical status and recovery time through hormonal/metabolic impairment, resulting in a longer hospital stay.

This is the first report of newborn BP contamination due to widely employed non-alimentary products designed for newborn care. Moreover, this study points to a central role of hospitalization

procedures and neonatal devices as primary sources of BP exposure. These results indicate that lifestyle and alimentary habits, especially in the last month of pregnancy, did not influence BP risk exposure. Despite this, in general, harmful behaviors could negatively impact newborn health conditions, and it is necessary to improve preventive strategies to counteract this trend. Further studies are advocated in order to clarify both the impact of other BP forms on humans and the potential sources and consequences of BP exposure during neonatal and childhood life.

2.2.1.3. Air Pollution and DNA mutations

2.2.1.3.1. Study Line 4: Air Pollution and Somatic Mutations

2.2.1.3.1.1. Aim and Rationale

Somatic mutations are an important aspect of genetically driven disease. Genetic testing is historically focused on germline mutations: mutations that are inherited from parents or determined in germ cells, before zygote formation. Germline mutations are present in all cell types and thus technically easy to differentiate from artifacts or sequencing errors[179]. On the other hand, somatic mutations occur after conception (post-zygotic) and are thus shared only by a fraction of the daughter cells, resulting in Post-Zygotic Mosaicism (PZM). Depending on both the tissue and fraction of cells affected, PZM mutations can lead to disease but, due to their nature, such mutations are difficult to detect and differentiate from technical artifacts. Somatic mutations, and the deriving cancerogenesis, are strongly linked to inflammation, which is the link with air pollution under study in this project. This technical challenge is mainly due to their low Variant Allele Frequency (VAF) in bio samples assayed (typically peripheral blood). VAF is the percentage of sequence reads observed matching a specific DNA variant divided by the overall coverage at that locus[180–182]. In this regard, new pipelines have been recently developed to identify somatic mutations in sequencing data, among which the well-known and probably most widely used Broad Institute GATK Mutect2[165,166]. Through several orientation and contamination filters, this pipeline allows to differentiate somatic mutations with low VAF from technical artifacts, and has been successfully utilized for several publications on the subject[185,186]. Both above mentioned air pollutants and somatic mutations have been proved to represent independent risk factors for cardiovascular disease and other inflammation-driven disease. In this regard, specific types of PZM, namely Clonal Hematopoiesis of Indeterminate Potential (CHIP) has been strongly associated with disease. CHIP is defined as the presence of a clonally expanded hematopoietic stem cell caused by leukemogenic mutation in individuals without evidence of hematologic malignancy[169]. CHIP has been associated with a two-fold increase in cardiovascular risk, independent of traditional risk factors[187–189]. Nonetheless, the exact mechanisms and possible interactions between CHIP and air pollution have never been studied yet. The specific aim of this study line was to identify whether an increased exposure to air pollution is correlated to increased levels of CHIP somatic mutations, in two geographically and ethnically different cohorts.

2.2.1.3.1.2. Research Plan

In this cohort study, samples were analyzed to identify somatic CHIP mutations in the cohort. Air pollution data was retrieved for each sample county of residence (median county surface ~150 km^2) in form of 10 years median AQI prior to sampling, and this data was then compared to the number and allele frequency of somatic variants identified for each sample. Regression models served to reveal possible associations among air pollution data and somatic variants identified, once weighted for all possible confounders.

We collected AQI data for each patient home residency for 10 years before sampling, both as countywide in case of US samples and Turin for local samples from Regina Margherita Children Hospital. Based on AQI, I elaborated a generalized Poisson regression model, based on the known formulae:

$$f(y) = \beta 0 + \Sigma\beta_j * X_{ij} + \varepsilon$$

where f is the link log function, y is the conditional expectation ($E(y=y_i|X=x_i)$) for any given value of the independent variable (X). β_0 is the intercept, and β is the regression coefficient for each independent variable x. ε is the error of the estimate.

This model dependent variable is represented by the number of CHIP mutations for each patient, that is a typically Poisson distributed variable, as I will show below. I built the model considering median AQI of the last 10 years, as indicator of the county pollution state, as independent variable. Covariates considered in the model included the 10-years median number of days in which a specific pollutant was driving the overall AQI, thus representing the main pollutant in the county air. Such parameter are considered for all the above mentioned pollutants: CO, NO_2, O_3 and PM_{10}, that I was able to retrieve for each location. Demographic variables are also taken into account for the sequenced patients, in order to avoid any possible patient-driven bias in the association: age at sampling, gender, ethnicity and race are also included as model covariates. Age is expected to exert the main influence on CHIP mutations, confirming literature reports. The overall result of such model give us an idea of both the overall and specific pollutants influence on CHIP mutations for each patient, avoiding age and other possible demographic biases, representing a totally innovative analysis on air pollution influence on somatic mutations and human disease.

2.2.1.3.1.3. Epidemiological Sample and CHIP Variants Calling

This analysis focuses on two different cohorts of kidney patients, already sequenced through Next Generation Sequencing Exome/Genome Sequencing (ES/GS). 436 US patients spanned all over the country, and 44 Italian patients represent the population, all residing within Turin region. Genomic DNA was isolated from peripheral blood mononuclear cell (PBMC) and used to construct a Kapa HyperPrep PCR-free DNA library for genome sequencing. ES/GS was performed with NovaSeq 6000 platform (Illumina). Reads were processed using the pipeline at the Institute for Genomic Medicine at Columbia University. Reads were mapped against the Genome Reference Consortium Human Build 37 (GRCh37/hg19)[190]. Starting from sequenced data paired-end FASTQ files, I first used BWA-MEM 0.7.17[191] to align sequenced reads and biobambam2[192] to remove duplicates. GATK best practices guidelines were followed with a re-evaluation step of sequenced bases quality (BaseRecalibrator[193], ApplyBQSR[194] and HaplotypeCaller[195]), to generate SNPs and indels gVCF files. I also used FastQC[196] to assess the quality of sequenced reads and all reads showed a mean coverage >29x across the genome. I served GATK CollectAlignmentSummaryMetrics[197] and Seven Bridges Genome Coverage tools to further investigate alignment and coverage results. Consistent with GATK best practices guidelines I further refined my gVCF files through VariantRecalibrator[198] (for both SNPs and indels), ApplyVQSR[199], and CollectVariantCallingMetrics[200]. I finally annotated my variants through VEP v103[201]. I served Peddy[202] to investigate ancestry, sex and compare familial-relationship in this cohort. GATK 4.1.0.0 version was used for the analysis. I then called somatic variants using GATK Mutect2 best practice guidelines, considering only PASS variants as per FilterMutectCalls[203] post-calling step and further filtering remaining variants based on the following criteria:

- coding regions only
- read depth >=20
- altered allele depth >=3
- altered allele F1R2 >=1

- altered allele F2R1 >=1

Somatic variants were called only on non-blacklisted and coding regions. I then served vcfeval and investigated presence/absence vs homozygous/heteroygous variant mismatch. CHIP variants were identified limiting my examination to variants that had been described previously in the literature in 160 recurrently mutated candidate genes in myeloid and lymphoid cancers, accordingly to previous publications of the Broad Institute group that I collaborated with for this project[204]. I removed potential false positive variants by using variant-calling algorithms that had filters for known artifacts such as strand-bias and clustered reads, as well as by performing additional filtering for rare error modes using a “panel of normals” (PON: sequence data from a panel of normal persons)[205,206]. The lower limit of detection for variants depended on the depth of coverage. I run Mutect2 in tumor-only mode with PON and germline resources. Next, I run FilterMutectCalls[203] and annotate the variants. For inclusion, I used the following thresholds. DP >= 20, altered allele depth >= 3, altered F1R2 >= 1, altered F2R1 >= 1, VAF >= 0.02, gnomAD popmax Allele Frequency <= 0.0001. I also filter out any variant that is more frequent than DNMT3A_R882H or JAK2_V617F. Indels in ASXL1 and PPM1D are kept irrespective of their frequencies. I then select mutations that pass CHIP rules and curate those manually. Manual curation was performed by the group of CHIP experts at Broad Institute that I collaborated with, aimed to identify whether the variants have been known to be clonal drivers. Finally, mutations were reviewed in IGV[207] to make sure those are true mutations and not artifacts.

2.2.1.3.1.4. Results

A total of 436 subjects were enrolled, of which 44 (10.1%) samples from the Italian cohort, and 392 (89.9%) from the US cohort. The median age at sampling was 45 years (Q1 31 – Q3 56) and the categorical variables considered in the analysis are summarized below.

Variable	***Category***	***N (%)***
gender	male	266 (61%)
	female	170 (49%)
race	white	286 (65.6%)
	black	89 (20.4%)
	asian	32 (7.3%)
	native	1 (0.02%)
	multi	28 (6.4%)
ethnicity	hispanic	99 (22.7%)
	other	337 (77.3%)

Table 11: categorical variables values in the cohort under analysis.

The 10 years median AQI for each county of residence resulted to be extremely variable in the cohort, with a median value of 40 and ranging from 16 to 77, as showed in the Figure 21.

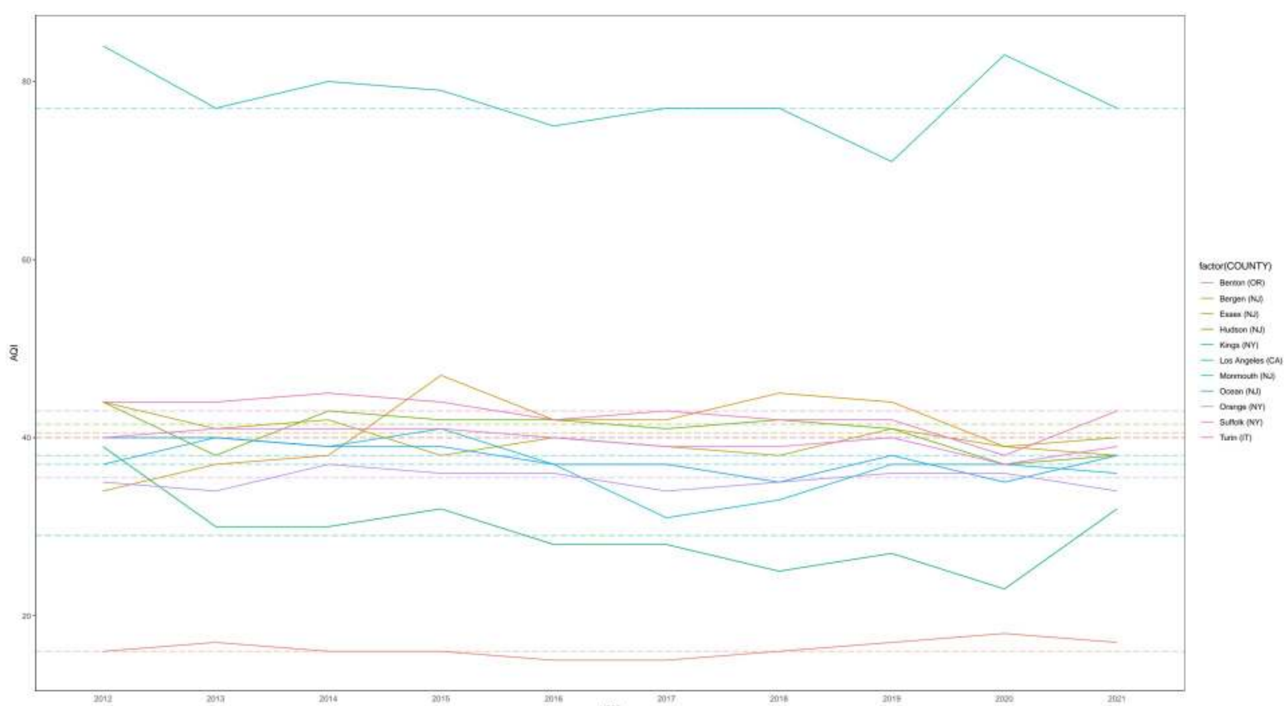


Figure 21: 10 years AQI and relative median values for each county. 10th quantiles are reported in the picture.

As summarized in Figure 22, 10 years median AQI values spanned very widely from the median value in the different counties under analysis.

Figure 22: 10 years median AQI exposure value for each sample relative to the overall median.

CHIP somatic mutations were identified in the cohort with the methods described above. As expected for this kind of mutations, CHIP distribution showed a classical Poisson distribution with low λ across the cohort, as graphically represented in the Figure 23.

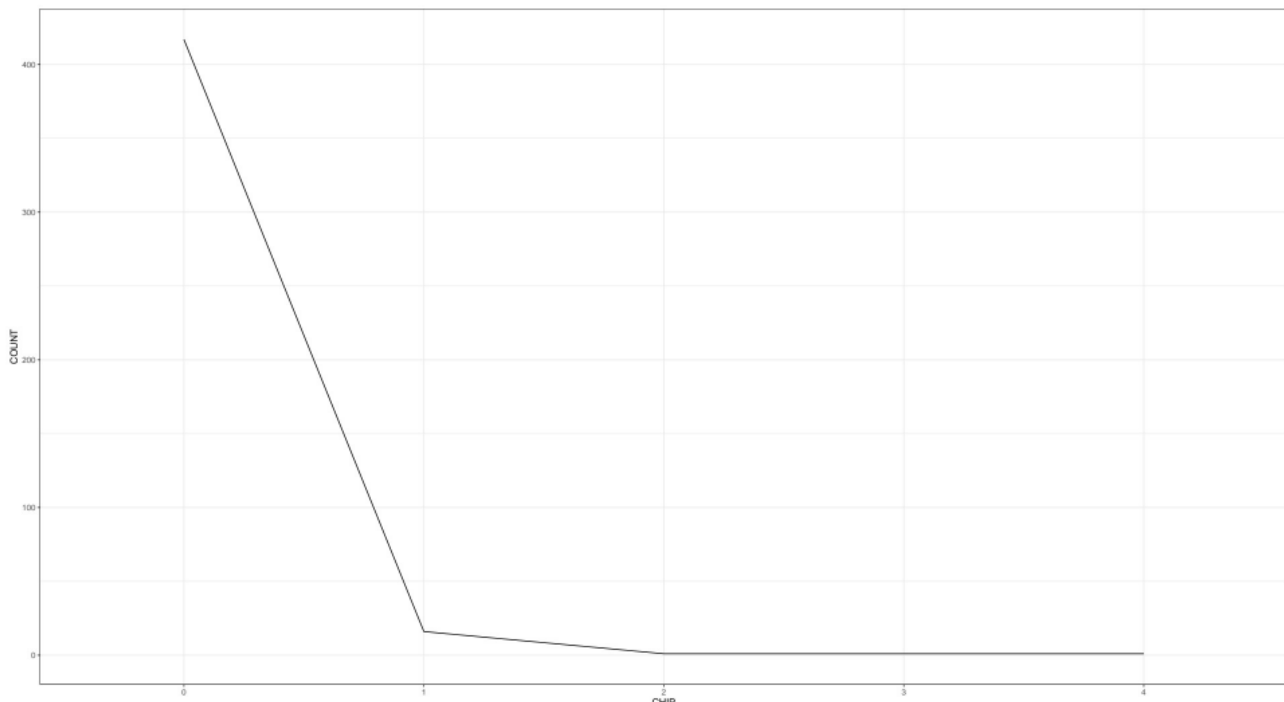


Figure 23: CHIP somatic mutations distribution across the cohort.

In order to ensure the goodness of my calls, I analyzed identified CHIP mutations, both in terms of their overall number and their VAF, compared to age distribution of the samples. As previously cited, CHIP mutations, and somatic mutations in general, are accumulated during life, thus a clear correlation with age of the samples analyzed is expected. A clear correlation was identifiable for both of those values: CHIP number ~ age (p: 0.003, R^2 0.14) and CHIP VAF ~ age (p: 0.0001, R^2: 0.18), as graphically represented in Figure 24 and Figure 25.

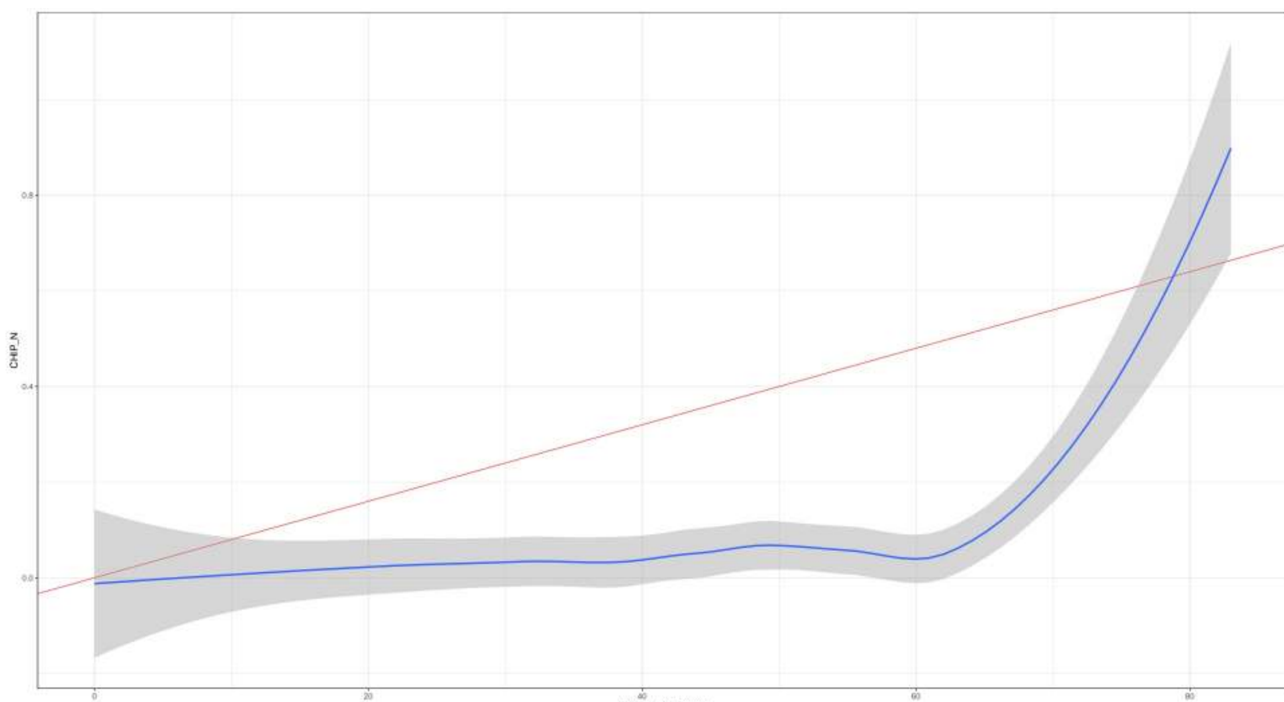


Figure 24: correlation of CHIP N with age of the samples.

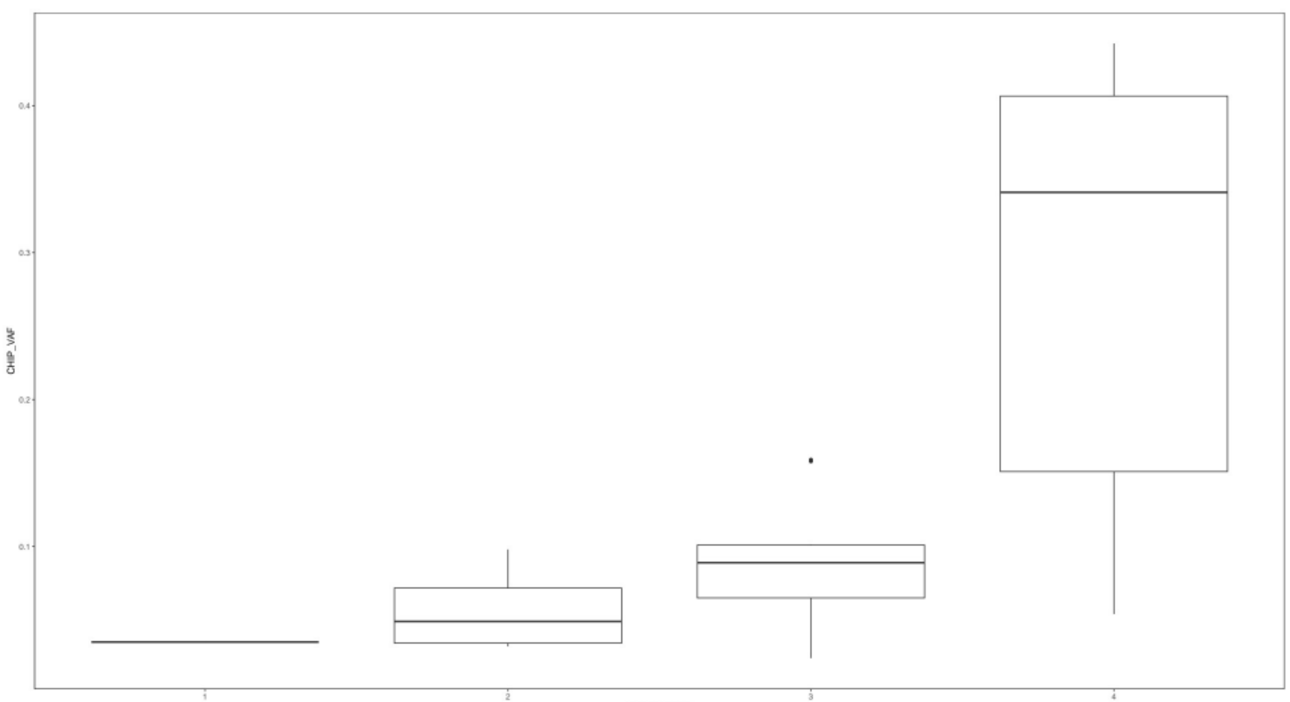


Figure 25: correlation of CHIP VAF with age of the samples. Age groups: 1: <30 years old, 2: 31-40 years old, 3: 41-50 years old, 4: >50 years old

I built the generalized regression model with a Poisson distribution of the dependent variable as summarized below. For each sample, I considered in the formula the number of identified CHIP somatic mutations, correcting for all possible demographic confounders such as age at sampling, gender, race, ethnicity. I considered, as independent variable, the 10 years AQI median value, and for each pollutant under analysis (CO, NO_2, O_3, $PM_{2.5}$, and PM_{10}) the median number of days in which this specific pollutant represented the main driver of AQI, as explained in the AQI section of the present dissertation. Through this model, I was able to identify an independent role of CHIP mutations driver for age at sampling and PM_{10}. Specifically, age at sampling resulted the main independent driver of association, as expected, with a p value of 0.00008. PM_{10} as the main determinant of AQI resulted to have a risk p value of 0.01. Basically, age at sampling resulted in an increased RR of 6% [RR 1.03 – 1.09], PM10 in an increased RR of 3.4% [RR 1.006 – 1.063].

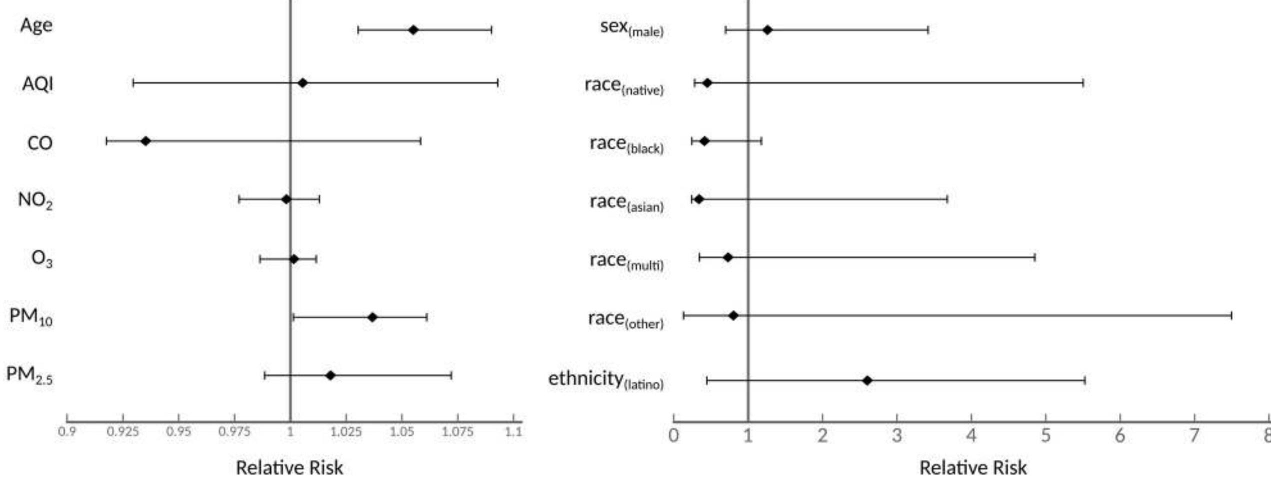


Figure 26: Relative Risk for CHIP number of each variable under analysis in the Poisson regression model

2.2.1.3.1.5. Discussion

As presented above, DNA somatic mutations primarily originates from disruptive inflammatory processes acting on DNA replication and repair mechanisms. In this doctoral project, I wanted to deepen this analysis on possible molecular consequences of air pollution exposure investigating its potential role as DNA somatic mutations determinant. At the best of our knowledge, this is the first literature report of such strong association identifiable with a specific pollutant exposure, such as PM_{10}, and DNA somatic mutations. My focus on CHIP mutations was primarily driven by the necessity of overcoming current barriers to somatic mutations identification in blood originated DNA samples. Thanks to the collaboration with the Broad Institute group at the Massachussets Institute of Technology, I was able to ensure maximum reliability of my CHIP somatic calls. The identification of such strong link between air pollution, and particularly this specific pollutant, with DNA somatic mutation, pave the way to multiple lines of investigation on the reasons and mechanisms underpinning such relationship. Firstly, it is a further clue of the role of inflammation on health detrimental effects of air pollution exposure, as inflammatory mechanisms have been identified as the main driver of DNA somatic mutations in humans and animal models, as presented above. Moreover, it focus the attention on a specific pollutant role, such as PM_{10}. Further investigation of this pollutant characteristic and molecular mechanisms have the potentiality of giving us deeper insight on both its role on human health and the molecular mechanisms behind it. This evidence is also a potential explanation of some of the detrimental effects played by air pollution exposure on fetal development, where the accumulation of somatic mutations, both in the mother and the developing child, can play an important role on pregnancy and newborn outcomes. Further investigation of underlying mechanisms is mandatory in order to better elucidate molecular actors of this relationship, and identifying both possible pharmacological approaches to counter-play this effect and implementation of public health preventive measures in this sense.

3. Project Conclusion

The different lines of research of my doctoral project allowed to bring out several important limitations of current approaches to study focusing on pregnancy and newborn health. At the best of our knowledge, this is the first literature report of the acute effect that air pollution exposure is able to exert on PTB during latest stage of pregnancy. These innovative results pave the way to potential easily implementable public health measures that could increase protection of pregnant women and reduce the burden of PTB in developed countries. I further deepened this analysis investigating potential molecular mechanisms underlying such relationship, revealing a central role played by inflammation, easily detectable through its markers, on the detrimental effects of pregnancy related exposure and maternal habit. This result confirms several lines of literature evidence in this direction, and reinforce the urge for potential pharmacological treatments that could help to counter-play and prevent such detrimental effects. I also detected how some hospital routine procedure expose newborns to harmful endocrine disruptors, such as bisphenols. This is a further demonstration of how current measures of monitor and prevention might be lacking and how further public health measures are necessary in order to increase protection of this sensitive phase of our lives. Finally, I was able to provide the first literature evidence of how air pollution exposure can be an important determinant of DNA somatic mutations, and subsequent cancerogenesis and health associated problems. This is an integrative step in understanding how inflammation might represent the main link between environmental exposure to pollutants and their detrimental effects on human health, further underlying the urge of investigation on both public health preventive measures and pharmacological treatments able to address this specific molecular link. I hope that my work can be a starting point to realize limitations of our current research approaches to pregnancy and newborn related studies, and increase the awareness of how urgent the need for further public health measures and pharmacological therapies able to address these problems in such sensitive phase of life.

4. In the meantime

4.1. Formaldehyde exposure and Health

Formaldehyde (FA) is a ubiquitous toxic chemical employed worldwide due to its disinfectant and preservative properties. Despite being classified as a human carcinogen of Group 1 by the International Agency for Research on Cancer (IARC), FA is still employed as formalin in pathology wards as standard fixative[208]. Nonetheless, the current lack of an effective alternative chemical makes formalin still very used, with the subsequent necessity of measures to control exposure, as in case of isolation of formalin related activities or adoption of new standard procedures aimed to reduce its use. The possible mechanisms underlying the FA-induced long-term effects include inflammation, OS, and apoptosis[209,210]. As side project of my doctorate, still focused on genetics consequence of different environmental exposure, in collaboration with the rest of Prof. Bono research group, I decided to further analyze the role of FA in human mutations. This work resulted in two publications to date[211,212]. As a background, several literature reports evidenced how FA exposure is strongly associated with a wide variety of toxic effects, depending on length and level of exposure. In fact, while low exposure levels (~0.1 ppm) act as irritant stimulus for exposed body parts, and especially eyes, nose and upper respiratory airways, high concentrations brings to pulmonary impairment and related disease, i.e. asthma[213,214]. Importantly, duration of exposure is a main determinant of the detrimental effects, as example, prolonged FA exposure have been extensively associated with cancerogenesis, as in case of nasopharyngeal cancer[215,216] and leukaemia[217,218]. Despite these lines of literature evidence, FA genotoxic effect is still debated. Importantly, cytogenetic outcomes, such as increased number of chromosomal aberrations (CAs) and micronucleated cells (MNc), were reported in several biomonitoring studies on chronic exposures[219,220], with contradictive conclusions in literature[221,222]. The specific focus that I wanted to explore was related to the individual susceptibility role in determining such effects. The heritable variability of polymorphisms related to FA metabolic pathways may be associated with an altered efficiency of the processes in which they are involved. Among these, a core role is played by the glutathione S-transferase (GST) and cytochrome P450 (CYP) families, that act in a two-step detoxification process of a wide spectrum of environmental xenobiotics, including FA[223,224]. At the same manner, genome integrity is guarded by several mechanisms of DNA repair, according to the type of damage. In this regard, the Base Excision Repair (BER) and the Nucleotide Excision Repair (NER) pathways are responsible for correcting DNA small base changes and bulky adducts[225]. These can be viewed as DNA-repair genes, which are involved in the protection mechanism against cancer development[226]. Several literature evidences underlined how variations in these DNA repair mechanisms could strongly reduce FA tolerance.
The main aims of the two studies conducted on this topic were:

1. Assessment of FA exposure effect on detectable CAs (assessed on Peripheral Blood Lymphocytes) and the role, both in terms of risk or protective effect, played by several genetic polymorphisms. I conducted this research on a cohort of 57 exposed pathologists vs. 48 non-exposed controls. All subjects were genotyped for the most common cancer-associated gene polymorphisms possibly related with CAs: CYP1A1 exon 7 (A > G), CYP1A1*2A (T > C), CYP2C19*2 (G > A), GSTT1 (Positive/Null), GSTM1 (Positive/null), GSTP1 (A > G), XRCC1 (G399A), XRCC1 (C194T), XRCC1 (A280G), XPD (A751C), XPC exon 15 (A939C), XPC exon 9 (C499T), TNFα − 308 (G > A), IL10 − 1082 (G > A), IL10 − 819 (C > T) and IL6 − 174 (G > C)[212]
2. Evaluation of FA exposure on SCEs, detectable on PBLs, on the same population presented above, and the relative role played by the most common cancer-associated gene polymorphisms, namely CYP1A1 exon 7 (A > G), CYP1A1*2A (T > C), CYP2C19*2 (G > A), GSTT1 (presence/absence), GSTM1 (presence/absence), GSTP1 (A > G), XRCC1

(G399A), XRCC1 (C194T), XRCC1 (A280G), XPC exon 15 (A939C), XPC exon 9 (C499T), TNFα − 308 G > A), IL10 – 1082 (G > A), and IL6 − 174 (G > C)[211]

I was able to detect how exposed subjects showed a significantly higher SCEs frequency than controls (mean exposure to air-FA resulted to be 55.2 μg/m^3 vs. 18.4 μg/m^3 respectively). Air-FA was directly correlated with SCEs frequency ($p < 0.01$) and inversely with the replication index ($p < 0.05$). Regression models revealed how FA exposure was a significant predictor of SCEs developement, highlighting no specific role for the selected polymorphisms[212].
At the same manner, positive correlations were found between CAs frequency and air-FA concentration ($p < 0.001$). Moreover, significant associations was detected between CAs frequency and the mutated allele for CYP1A1 exon 7 (A>G) ($\beta = 0.353$; $p = 0.019$), CYP2C19*2 (G>A) ($\beta = 0.504$; $p = 0.007$), GSTT1-positive ($\beta = -0.447$; $p = 0.004$), GSTM1-positive ($\beta = -0.533$; $p = 0.001$) and XRCC1 (399, G>A) ($\beta = -0.331$; $p = 0.044$) genotypes, as summarized in Figure 26 below[211].

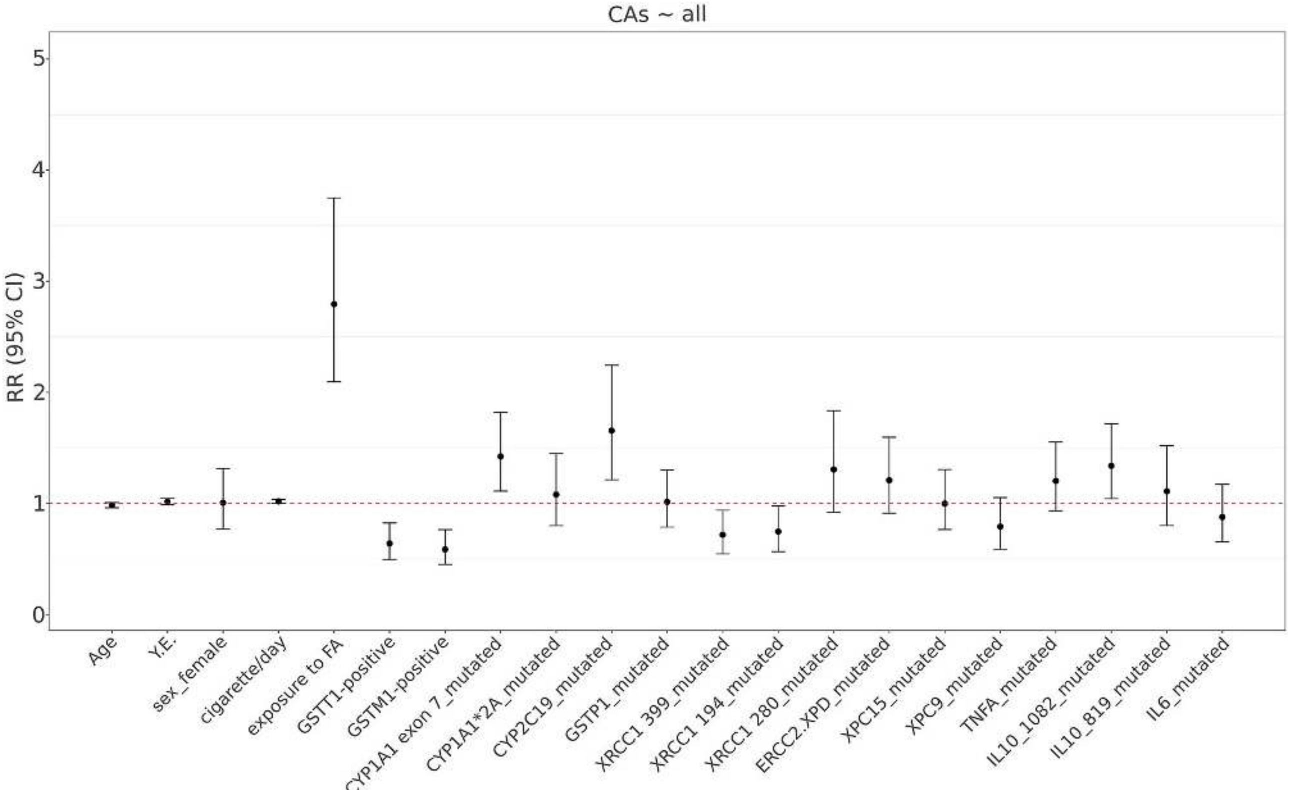


Figure 26: RR (95% CI) of developing CAs according to the various mutated allelic variants of genes considered, FA exposure and confounding factors as age, years of exposure, sex and cigarette/day.

Overall, these lines of evidence reinforce the role of FA as genotoxicity inductor, especially in workers chronically exposed to low air-FA levels. It also highlight the importance of the role played by some genetic polymorphisms in this association, underlining the importance of individual susceptibility biomarkers assessment in occupational health studies and the urge to define safer exposure limits for exposed workers. This represented another important step in research of underlying mechanisms responsible of detrimental effects on human health of toxic exposure.

4.2. Other Papers

During the three years of my doctorate, I published several other studies than the 4 studies cited above (for full text of those 4 please see Appendix A), especially with the Columbia University group of Dr. Gharavi laboratory, where I spent 2 years partially overlapping with my doctorate. Here below the list of the published studies, in the Appendix A full text of those. For a full list of my publications (37 at the moment of writing), please visit: https://www.researchgate.net/profile/Enrico-Cocchi-2 or my Scopus WOS profile: https://www.scopus.com/authid/detail.uri?authorId=56017406800

1. Rare Single Nucleotide and Copy Number Variants and the Etiology of Congenital Obstructive Uropathy: Implications for Genetic Diagnosis. March 2023. Journal of the American Society of Nephrology
2. Oxidative Stress in the Early Neonatal Period as a Possible Effect of BMI, Smoking Habits, and Level of Urbanization of the Mother. February 2023. Journal of Biological Regulators and Homeostatic Agents
3. Clinical and Genetic Characteristics of CKD Patients with High-risk APOL1 Genotypes. February 2023. Journal of the American Society of Nephrology
4. The spectrum of epilepsy with eyelid myoclonia: delineation of disease subtypes from a large multicenter study. October 2022. Epilepsia
5. Electroclinical Features and Long-term Seizure Outcome in Patients With Eyelid Myoclonia With Absences. March 2022. Neurology
6. Diagnostic sequencing to support genetically stratified medicine in a tertiary care setting. January 2022. Genetics in medicine
7. Reduced mortality in COVID-19 patients treated with colchicine: Results from a retrospective, observational study. March 2021. Plos One
8. Clinical exome sequencing is a powerful tool in the diagnostic flow of monogenic kidney diseases: an Italian experience. December 2020. Journal of Nephrology
9. The potential impact of enhanced hygienic measures during the COVID-19 outbreak on hospital-acquired infections: A pragmatic study in neurological units. August 2020. Journal of the Neurological Sciences
10. Clinical Genetic Screening in Adult Patients with Kidney Disease. July 2020. Clinical Journal of the American Society of Nephrology
11. Molecular and functional characterization of urine-derived podocytes from patients with Alport syndrome. July 2020. The Journal of Pathology
12. The switch from proteasome to immunoproteasome is increased in circulating cells of patients with fast progressive immunoglobulin A nephropathy and associated with defective CD46 expression. June 2020. Nephrology Dialysis Transplantation
13. Pre-discharge Cardiorespiratory Monitoring in Preterm Infants. the CORE Study. June 2020. Frontiers in Pediatrics
14. Valproate impact and sex-dependent seizure remission in patients with idiopathic generalized epilepsy. May 2020. Journal of the Neurological Sciences
15. Acute and chronic glomerular damage is associated with reduced CD133 expression in urinary extracellular vesicles. December 2019. American journal of physiology. Renal physiology
16. Doing without valproate in women of childbearing potential with idiopathic generalized epilepsy: Implications on seizure outcome. December 2019. Epilepsia
17. Post-transplant recurrence of steroid resistant nephrotic syndrome in children: the Italian experience. October 2019. Journal of Nephrology

5. References

1. Air pollution. https://www.who.int/health-topics/air-pollution.
2. Anderson, J. O., Thundiyil, J. G. & Stolbach, A. Clearing the air: a review of the effects of particulate matter air pollution on human health. *J Med Toxicol* **8**, 166–175 (2012).
3. Mannucci, P. M., Harari, S., Martinelli, I. & Franchini, M. Effects on health of air pollution: a narrative review. *Intern Emerg Med* **10**, 657–662 (2015).
4. Chen, H., Goldberg, M. S. & Villeneuve, P. J. A systematic review of the relation between long-term exposure to ambient air pollution and chronic diseases. *Rev Environ Health* **23**, 243–297 (2008).
5. Kurt, O. K., Zhang, J. & Pinkerton, K. E. Pulmonary health effects of air pollution. *Curr Opin Pulm Med* **22**, 138–143 (2016).
6. Li, R., Zhou, R. & Zhang, J. Function of PM2.5 in the pathogenesis of lung cancer and chronic airway inflammatory diseases. *Oncol Lett* **15**, 7506–7514 (2018).
7. Carré, J., Gatimel, N., Moreau, J., Parinaud, J. & Léandri, R. Does air pollution play a role in infertility?: a systematic review. *Environ Health* **16**, 82 (2017).
8. Sram, R. J., Veleminsky, M., Veleminsky, M. & Stejskalová, J. The impact of air pollution to central nervous system in children and adults. *Neuro Endocrinol Lett* **38**, 389–396 (2017).
9. Liu, Y., Xu, J., Chen, D., Sun, P. & Ma, X. The association between air pollution and preterm birth and low birth weight in Guangdong, China. *BMC Public Health* **19**, 3 (2019).
10. Lee, J.-T. Review of epidemiological studies on air pollution and health effects in children. *Clin Exp Pediatr* **64**, 3–11 (2021).
11. Kim, J. H., Choi, Y. Y., Yoo, S.-I. & Kang, D. R. Association between ambient air pollution and high-risk pregnancy: A 2015-2018 national population-based cohort study in Korea. *Environ Res* **197**, 110965 (2021).
12. Hu, C.-Y. *et al.* Human epidemiological evidence about the association between air pollution exposure and gestational diabetes mellitus: Systematic review and meta-analysis. *Environmental Research* **180**, 108843 (2020).

13. Assibey-Mensah, V. *et al.* Ambient wintertime particulate air pollution and hypertensive disorders of pregnancy in Monroe County, New York. *Environmental Research* **168**, 25–31 (2019).
14. Mobasher, Z. *et al.* Associations between ambient air pollution and Hypertensive Disorders of Pregnancy. *Environmental Research* **123**, 9–16 (2013).
15. Ambient air pollution the risk of stillbirth: A prospective birth cohort study in Wuhan, China - ScienceDirect. https://www-sciencedirect-com.ezproxy.cul.columbia.edu/science/article/pii/S143846391730531X.
16. Klepac, P., Locatelli, I., Korošec, S., Künzli, N. & Kukec, A. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environ. Res.* **167**, 144–159 (2018).
17. Pavlin, M., Repič, M., Vianello, R. & Mavri, J. The Chemistry of Neurodegeneration: Kinetic Data and Their Implications. *Mol Neurobiol* **53**, 3400–3415 (2016).
18. Angelova, P. R. & Abramov, A. Y. Role of mitochondrial ROS in the brain: from physiology to neurodegeneration. *FEBS Lett* **592**, 692–702 (2018).
19. Sarniak, A., Lipińska, J., Tytman, K. & Lipińska, S. Endogenous mechanisms of reactive oxygen species (ROS) generation. *Postepy Hig Med Dosw (Online)* **70**, 1150–1165 (2016).
20. Wang, Y., Xiong, L. & Tang, M. Toxicity of inhaled particulate matter on the central nervous system: neuroinflammation, neuropsychological effects and neurodegenerative disease. *J Appl Toxicol* **37**, 644–667 (2017).
21. Nissanka, N. & Moraes, C. T. Mitochondrial DNA damage and reactive oxygen species in neurodegenerative disease. *FEBS Lett* **592**, 728–742 (2018).
22. Tan, H.-Y. *et al.* The Reactive Oxygen Species in Macrophage Polarization: Reflecting Its Dual Role in Progression and Treatment of Human Diseases. *Oxid Med Cell Longev* **2016**, 2795090 (2016).
23. Lakey, P. S. J. *et al.* Chemical exposure-response relationship between air pollutants and reactive oxygen species in the human respiratory tract. *Sci Rep* **6**, 32916 (2016).
24. Chen, W.-W., Zhang, X. & Huang, W.-J. Role of neuroinflammation in neurodegenerative diseases (Review). *Mol Med Rep* **13**, 3391–3396 (2016).

25. Peixoto, M. S., de Oliveira Galvão, M. F. & Batistuzzo de Medeiros, S. R. Cell death pathways of particulate matter toxicity. *Chemosphere* **188**, 32–48 (2017).
26. Rajagopalan, S. & Brook, R. D. Air Pollution and Type 2 Diabetes: Mechanistic Insights. *Diabetes* **61**, 3037–3045 (2012).
27. Xu, F. *et al.* Effects on IL-1β signaling activation induced by water and organic extracts of fine particulate matter (PM2.5) in vitro. *Environ Pollut* **237**, 592–600 (2018).
28. Calderon-Garciduenas, L. *et al.* Brain immune interactions and air pollution: macrophage inhibitory factor (MIF), prion cellular protein (PrPC), Interleukin-6 (IL-6), interleukin 1 receptor antagonist (IL-1Ra), and interleukin-2 (IL-2) in cerebrospinal fluid and MIF in serum differentiate urban children exposed to severe vs. low air pollution. *Frontiers in Neuroscience* **7**, (2013).
29. Puett, R. C. *et al.* Inflammation and acute traffic-related air pollution exposures among a cohort of youth with type 1 diabetes. *Environ Int* **132**, 105064 (2019).
30. Zeng, X. *et al.* Air pollution associated acute respiratory inflammation and modification by GSTM1 and GSTT1 gene polymorphisms: a panel study of healthy undergraduates. *Environmental Health* **22**, 14 (2023).
31. de Oliveira Alves, N. *et al.* Inflammation response, oxidative stress and DNA damage caused by urban air pollution exposure increase in the lack of DNA repair XPC protein. *Environment International* **145**, 106150 (2020).
32. Radu, A.-F. & Bungau, S. G. Management of Rheumatoid Arthritis: An Overview. *Cells* **10**, 2857 (2021).
33. Brüske, I. *et al.* Impact of ambient air pollution on the differential white blood cell count in patients with chronic pulmonary disease. *Inhalation Toxicology* **22**, 245–252 (2010).
34. Cecconi, A. *et al.* Influence of air pollutants on circulating inflammatory cells and microRNA expression in acute myocardial infarction. *Sci Rep* **12**, 5350 (2022).
35. Vogel, C. F. A., Van Winkle, L. S., Esser, C. & Haarmann-Stemmann, T. The aryl hydrocarbon receptor as a target of environmental stressors – Implications for pollution mediated stress and inflammatory responses. *Redox Biology* **34**, 101530 (2020).
36. Guzman, J. A. Carbon Monoxide Poisoning. *Critical Care Clinics* **28**, 537–548 (2012).

37. Toxicants, N. R. C. (US) S. on R.-E. *ACUTE TOXICITY OF NITROGEN DIOXIDE*. *Assessment of Exposure-Response Functions for Rocket-Emission Toxicants* (National Academies Press (US), 1998).
38. Thangavel, P., Park, D. & Lee, Y.-C. Recent Insights into Particulate Matter (PM2.5)-Mediated Toxicity in Humans: An Overview. *Int J Environ Res Public Health* **19**, 7511 (2022).
39. Delomenie, M. *et al.* Carbon Monoxide Poisoning during Pregnancy: Presentation of a Rare Severe Case with Fetal Bladder Complications. *Case Rep Obstet Gynecol* **2015**, 687975 (2015).
40. Venditti, C. C., Casselman, R. & Smith, G. N. Effects of chronic carbon monoxide exposure on fetal growth and development in mice. *BMC Pregnancy Childbirth* **11**, 101 (2011).
41. Koren, G. *et al.* A multicenter, prospective study of fetal outcome following accidental carbon monoxide poisoning in pregnancy. *Reprod Toxicol* **5**, 397–403 (1991).
42. Margulies, J. L. Acute carbon monoxide poisoning during pregnancy. *Am J Emerg Med* **4**, 516–519 (1986).
43. Caravati, E. M., Adams, C. J., Joyce, S. M. & Schafer, N. C. Fetal toxicity associated with maternal carbon monoxide poisoning. *Ann Emerg Med* **17**, 714–717 (1988).
44. Salam, M. T. *et al.* Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: results from the Children's Health Study. *Environ Health Perspect* **113**, 1638–1644 (2005).
45. Badida, P., Krishnamurthy, A. & Jayaprakash, J. Meta analysis of health effects of ambient air pollution exposure in low- and middle-income countries. *Environ Res* **216**, 114604 (2023).
46. Kim, Y. & Radoias, V. Severe Air Pollution Exposure and Long-Term Health Outcomes. *Int J Environ Res Public Health* **19**, 14019 (2022).
47. Ledford, H. Air pollution causes sperm mutations. *Nature* (2008) doi:10.1038/news.2008.439.
48. Somers, C. M., McCarry, B. E., Malek, F. & Quinn, J. S. Reduction of Particulate Air Pollution Lowers the Risk of Heritable Mutations in Mice. *Science* **304**, 1008–1010 (2004).
49. Jiang, C.-L. *et al.* Air pollution and DNA methylation alterations in lung cancer: A systematic and comparative study. *Oncotarget* **8**, 1369–1391 (2017).
50. Niehoff, N. M. *et al.* Air Pollution and Breast Cancer: An Examination of Modification By Underlying Familial Breast Cancer Risk. *Cancer Epidemiol Biomarkers Prev* **31**, 422–429 (2022).

51. Singh, R. *et al.* Effects of environmental air pollution on endogenous oxidative DNA damage in humans. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **620**, 71–82 (2007).
52. Ayi Fanou, L. *et al.* Survey of air pollution in Cotonou, Benin--air monitoring and biomarkers. *Sci Total Environ* **358**, 85–96 (2006).
53. Avogbe, P. H. *et al.* Ultrafine particulate matter and high-level benzene urban air pollution in relation to oxidative DNA damage. *Carcinogenesis* **26**, 613–620 (2005).
54. Sørensen, M. *et al.* Personal exposure to PM2.5 and biomarkers of DNA damage. *Cancer Epidemiol Biomarkers Prev* **12**, 191–196 (2003).
55. Bräuner, E. V. *et al.* Exposure to ultrafine particles from ambient air and oxidative stress-induced DNA damage. *Environ Health Perspect* **115**, 1177–1182 (2007).
56. Loft, S., Poulsen, H. E., Vistisen, K. & Knudsen, L. E. Increased urinary excretion of 8-oxo-2'-deoxyguanosine, a biomarker of oxidative DNA damage, in urban bus drivers. *Mutat Res* **441**, 11–19 (1999).
57. Lai, C.-H. *et al.* Exposure to traffic exhausts and oxidative DNA damage. *Occup Environ Med* **62**, 216–222 (2005).
58. Rossner, P. *et al.* Oxidative and nitrosative stress markers in bus drivers. *Mutat Res* **617**, 23–32 (2007).
59. Chuang, C.-Y., Lee, C.-C., Chang, Y.-K. & Sung, F.-C. Oxidative DNA damage estimated by urinary 8-hydroxydeoxyguanosine: influence of taxi driving, smoking and areca chewing. *Chemosphere* **52**, 1163–1171 (2003).
60. Møller, P. *et al.* Air pollution, oxidative damage to DNA, and carcinogenesis. *Cancer Lett* **266**, 84–97 (2008).
61. Schärer, O. D. Nucleotide Excision Repair in Eukaryotes. *Cold Spring Harb Perspect Biol* **5**, a012609 (2013).
62. Melis, J. P. M. *et al.* Mouse Models for Xeroderma Pigmentosum Group A and Group C Show Divergent Cancer Phenotypes. *Cancer Research* **68**, 1347–1353 (2008).
63. John, K. *et al.* Benzo[a]pyrene (BP) DNA adduct formation in DNA repair–deficient p53 haploinsufficient [Xpa(− / −)p53(+/−)] and wild-type mice fed BP and BP plus chlorophyllin for 28 days. *Carcinogenesis* **33**, 2236–2241 (2012).

64. Mehta, M., Chen, L.-C., Gordon, T., Rom, W. & Tang, M. Particulate matter inhibits DNA repair and enhances mutagenesis. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **657**, 116–121 (2008).
65. Kay, J., Thadhani, E., Samson, L. & Engelward, B. Inflammation-induced DNA damage, mutations and cancer. *DNA Repair (Amst)* **83**, 102673 (2019).
66. Ferguson, L. R. Chronic inflammation and mutagenesis. *Mutat Res* **690**, 3–11 (2010).
67. Luzzatto, L. Somatic mutations in cancer development. *Environ Health* **10 Suppl 1**, S12 (2011).
68. Coules, C. Over 6,000 cities now monitor air quality, WHO reveals. *AirQualityNews* https://airqualitynews.com/2022/04/04/over-6000-cities-now-monitor-air-quality-who-reveals/ (2022).
69. Diffusion tube. *Wikipedia* (2023).
70. Diffusion Tube Monitoring. https://www.cheshireeast.gov.uk/business/environmental_health/local_air_quality/what_is_pollution_like_near_me/diffusion_tube_monitoring/diffusion_tube_monitoring.aspx.
71. Air pollution measurement. *Wikipedia* (2023).
72. Jha, R. K. Non-Dispersive Infrared Gas Sensing Technology: A Review. *IEEE Sensors Journal* **22**, 6–15 (2022).
73. Fine, G. F., Cavanagh, L. M., Afonja, A. & Binions, R. Metal Oxide Semi-Conductor Gas Sensors in Environmental Monitoring. *Sensors (Basel)* **10**, 5469–5502 (2010).
74. Vacher, M. *et al.* Chemi- and Bioluminescence of Cyclic Peroxides. *Chem. Rev.* **118**, 6927–6974 (2018).
75. Tapered element oscillating microbalance. *Wikipedia* (2022).
76. Masic, A. *et al.* Evaluation of optical particulate matter sensors under realistic conditions of strong and mild urban pollution. *Atmospheric Measurement Techniques* **13**, 6427–6443 (2020).
77. reserved, M.-T. I. I. all rights. Emissions Testing – Gravimetric Determination of Particulate Matter. https://www.mt.com/es/en/home/applications/Laboratory_weighing/emissions-testing-particulate-matter.html.
78. Queensland;, c=AU; o=The S. of. Tapered element oscillating microbalance | Measuring air quality. https://www.qld.gov.au/environment/management/monitoring/air/air-monitoring/measuring/oscillating-microbalance.

79. How We Measure Ozone - Air (U.S. National Park Service). https://www.nps.gov/subjects/air/howwemeasure-ozone.htm.
80. Cocchi, E. *et al.* Air Pollution and Aeroallergens as Possible Triggers in Preterm Birth Delivery. *Int J Environ Res Public Health* **20**, 1610 (2023).
81. Tan, X. *et al.* A review of current air quality indexes and improvements under the multi-contaminant air pollution exposure. *J Environ Manage* **279**, 111681 (2021).
82. K, P. & Kumar, P. A critical evaluation of air quality index models (1960-2021). *Environ Monit Assess* **194**, 324 (2022).
83. US EPA, O. How is the AQI calculated? https://www.epa.gov/outdoor-air-quality-data/how-aqi-calculated (2021).
84. Saad, S. M. *et al.* Development of indoor environmental index: Air quality index and thermal comfort index. *AIP Conference Proceedings* **1808**, 020043 (2017).
85. Blencowe, H. *et al.* Born Too Soon: The global epidemiology of 15 million preterm births. *Reprod Health* **10**, S2 (2013).
86. Martin, J. A. & Osterman, M. J. K. Describing the Increase in Preterm Births in the United States, 2014-2016. *NCHS Data Brief* 1–8 (2018).
87. Macdorman, M. F. & Mathews, T. J. Recent trends in infant mortality in the United States. *NCHS Data Brief* 1–8 (2008).
88. Raju, T. N. K. *et al.* Long-Term Healthcare Outcomes of Preterm Birth: An Executive Summary of a Conference Sponsored by the National Institutes of Health. *J Pediatr* **181**, 309-318.e1 (2017).
89. Cresi, F. *et al.* Cardiorespiratory events in infants with gastroesophageal reflux symptoms: Is there any association? *Neurogastroenterology & Motility* **30**, e13278 (2018).
90. Squillacioti, G., Bellisario, V., Levra, S., Piccioni, P. & Bono, R. Greenness Availability and Respiratory Health in a Population of Urbanised Children in North-Western Italy. *Int J Environ Res Public Health* **17**, (2020).
91. Hall, E. S. & Greenberg, J. M. Estimating community-level costs of preterm birth. *Public Health* **141**, 222–228 (2016).
92. Goldenberg, R. L., Culhane, J. F., Iams, J. D. & Romero, R. Epidemiology and causes of preterm birth. *Lancet* **371**, 75–84 (2008).

93. Alman, B. L. *et al.* Associations between PM2.5 and risk of preterm birth among liveborn infants. *Ann Epidemiol* **39**, 46-53.e2 (2019).
94. Mendola, P. *et al.* Air Pollution and Preterm Birth: Do Air Pollution Changes over Time Influence Risk in Consecutive Pregnancies among Low-Risk Women? *Int J Environ Res Public Health* **16**, (2019).
95. Rappazzo, K. M., Daniels, J. L., Messer, L. C., Poole, C. & Lobdell, D. T. Exposure to fine particulate matter during pregnancy and risk of preterm birth among women in New Jersey, Ohio, and Pennsylvania, 2000-2005. *Environ Health Perspect* **122**, 992–997 (2014).
96. Shi, L. *et al.* Long-term effects of PM2·5 on neurological disorders in the American Medicare population: a longitudinal cohort study. *The Lancet Planetary Health* **4**, e557–e565 (2020).
97. Taguchi, A. *et al.* Recent Progress in Therapeutics for Inflammation-Associated Preterm Birth: A Review. *Reprod Sci* **24**, 7–18 (2017).
98. Faro, J. *et al.* Intra-amniotic inflammation induces preterm birth by activating the NLRP3 inflammasome†. *Biol Reprod* **100**, 1290–1305 (2019).
99. Rider, C. F. & Carlsten, C. Air pollution and DNA methylation: effects of exposure in humans. *Clin Epigenetics* **11**, 131 (2019).
100. Brook, R. D. *et al.* Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **121**, 2331–2378 (2010).
101. Yue, W. *et al.* Short term Pm2.5 exposure caused a robust lung inflammation, vascular remodeling, and exacerbated transition from left ventricular failure to right ventricular hypertrophy. *Redox Biol* **22**, 101161 (2019).
102. Boyle, A. K., Rinaldi, S. F., Rossi, A. G., Saunders, P. T. K. & Norman, J. E. Repurposing simvastatin as a therapy for preterm labor: evidence from preclinical models. *FASEB J* **33**, 2743–2758 (2019).
103. Pope, C. A. *et al.* Exposure to Fine Particulate Air Pollution Is Associated with Endothelial Injury and Systemic Inflammation. *Circ Res* **119**, 1204–1214 (2016).
104. Belousova, V. S., Svitich, O. A., Timokhina, E. V., Strizhakov, A. N. & Bogomazova, I. M. Polymorphism of the IL-1β, TNF, IL-1RA and IL-4 Cytokine Genes Significantly Increases the Risk of Preterm Birth. *Biochemistry (Mosc)* **84**, 1040–1046 (2019).

105. Muñoz-Pérez, V. M. *et al.* Preterm Birth, Inflammation and Infection: New Alternative Strategies for their Prevention. *Curr Pharm Biotechnol* **20**, 354–365 (2019).
106. Bono, R. *et al.* Air pollution, aeroallergens and admissions to pediatric emergency room for respiratory reasons in Turin, northwestern Italy. *BMC Public Health* **16**, (2016).
107. Samoli, E. *et al.* Acute effects of ambient particulate matter on mortality in Europe and North America: results from the APHENA study. *Environ Health Perspect* **116**, 1480–1486 (2008).
108. Peng, R. D., Dominici, F. & Louis, T. A. Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **169**, 179–203 (2006).
109. Stafoggia, M., Colais, P., Serinelli, M., & Gruppo collaborativo EpiAir. [Methods of statistical analysis to evaluate the short term effects of air pollution in the EpiAir Project]. *Epidemiol Prev* **33**, 53–63 (2009).
110. Marcon, A. *et al.* Association between PM10 concentrations and school absences in proximity of a cement plant in northern Italy. *Int J Hyg Environ Health* **217**, 386–391 (2014).
111. Simon-Szabo, Z. *et al.* Oxidative stress and peripartum outcomes (Review). *Exp Ther Med* **22**, 771 (2021).
112. Joo, E. H. *et al.* Effect of Endogenic and Exogenic Oxidative Stress Triggers on Adverse Pregnancy Outcomes: Preeclampsia, Fetal Growth Restriction, Gestational Diabetes Mellitus and Preterm Birth. *Int J Mol Sci* **22**, 10122 (2021).
113. Liguori, I. *et al.* Oxidative stress, aging, and diseases. *Clin Interv Aging* **13**, 757–772 (2018).
114. Negro, S. *et al.* Maternal obesity and perinatal oxidative stress: the strength of the association. *J Biol Regul Homeost Agents* **31**, 221–227 (2017).
115. Alcala, M. *et al.* Antioxidants and Oxidative Stress: Focus in Obese Pregnancies. *Front Physiol* **9**, 1569 (2018).
116. Malti, N. *et al.* Oxidative stress and maternal obesity: Feto-placental unit interaction. *Placenta* **35**, 411–416 (2014).
117. Fernández-Sánchez, A. *et al.* Inflammation, Oxidative Stress, and Obesity. *Int J Mol Sci* **12**, 3117–3132 (2011).

118. Ballesteros-Guzmán, A. K. *et al.* Prepregnancy Obesity, Maternal Dietary Intake, and Oxidative Stress Biomarkers in the Fetomaternal Unit. *Biomed Res Int* **2019**, 5070453 (2019).

119. Myatt, L. & Maloyan, A. Obesity and Placental Function. *Semin Reprod Med* **34**, 42–49 (2016).

120. Saben, J. *et al.* Maternal Obesity is Associated with a Lipotoxic Placental Environment. *Placenta* **35**, 171–177 (2014).

121. Thompson, L. P. & Al-Hasan, Y. Impact of oxidative stress in fetal programming. *J Pregnancy* **2012**, 582748 (2012).

122. Kumar, S. N. *et al.* Structural changes, increased hypoxia, and oxidative DNA damage in placenta due to maternal smokeless tobacco use. *Birth Defects Res* **113**, 1198–1214 (2021).

123. Caliri, A. W., Tommasi, S. & Besaratinia, A. Relationships among smoking, oxidative stress, inflammation, macromolecular damage, and cancer. *Mutat Res Rev Mutat Res* **787**, 108365 (2021).

124. Bono, R. *et al.* Bisphenol A, Tobacco Smoke, and Age as Predictors of Oxidative Stress in Children and Adolescents. *Int J Environ Res Public Health* **16**, 2025 (2019).

125. Alberg, A. J., Shopland, D. R. & Cummings, K. M. The 2014 Surgeon General's report: commemorating the 50th Anniversary of the 1964 Report of the Advisory Committee to the US Surgeon General and updating the evidence on the health consequences of cigarette smoking. *Am J Epidemiol* **179**, 403–412 (2014).

126. CDCTobaccoFree. Surgeon General's Reports on Smoking and Health. *Centers for Disease Control and Prevention* https://www.cdc.gov/tobacco/sgr/index.htm (2022).

127. Napierala, M. *et al.* The effect of tobacco smoke on oxytocin concentrations and selected oxidative stress parameters in plasma during pregnancy and post-partum - an experimental model. *Hum Exp Toxicol* **36**, 135–145 (2017).

128. Delpisheh, A., Brabin, L. & Brabin, B. J. Pregnancy, smoking and birth outcomes. *Womens Health (Lond)* **2**, 389–403 (2006).

129. Fazel, N. *et al.* Environmental tobacco smoke exposure during pregnancy affects complications and birth outcomes in women with and without asthma. *BMC Pregnancy Childbirth* **20**, 314 (2020).

130. Gómez-Roig, M. D. *et al.* Environmental Exposure during Pregnancy: Influence on Prenatal Development and Early Life: A Comprehensive Review. *FDT* **48**, 245–257 (2021).

131. Nagiah, S. *et al.* Oxidative stress and air pollution exposure during pregnancy: A molecular assessment. *Hum Exp Toxicol* **34**, 838–847 (2015).

132. Babić Leko, M. *et al.* IL-1β, IL-6, IL-10, and TNFα Single Nucleotide Polymorphisms in Human Influence the Susceptibility to Alzheimer's Disease Pathology. *J Alzheimers Dis* **75**, 1029–1047 (2020).

133. Bodro, M. *et al.* Increased CSF levels of IL-1β, IL-6, and ACE in SARS-CoV-2–associated encephalitis. *Neurology - Neuroimmunology Neuroinflammation* **7**, (2020).

134. Sverrisson, K., Axelsson, J., Rippe, A., Asgeirsson, D. & Rippe, B. Acute reactive oxygen species (ROS)-dependent effects of IL-1β, TNF-α, and IL-6 on the glomerular filtration barrier (GFB) in vivo. *Am J Physiol Renal Physiol* **309**, F800-806 (2015).

135. Maass, D. L., White, J. & Horton, J. W. IL-1beta and IL-6 act synergistically with TNF-alpha to alter cardiac contractile function after burn trauma. *Shock* **18**, 360–366 (2002).

136. Kaneko, N., Kurata, M., Yamamoto, T., Morikawa, S. & Masumoto, J. The role of interleukin-1 in general pathology. *Inflamm Regen* **39**, 12 (2019).

137. Montuschi, P., Barnes, P. J. & Roberts, L. J. Isoprostanes: markers and mediators of oxidative stress. *FASEB J* **18**, 1791–1800 (2004).

138. Czerska, M., Zieliński, M. & Gromadzińska, J. Isoprostanes - A novel major group of oxidative stress markers. *Int J Occup Med Environ Health* **29**, 179–190 (2016).

139. Bellisario, V. *et al.* Bisphenol A and S in the Urine of Newborns: Plastic for Non-Food Use Still without Rules. *Biology (Basel)* **10**, 188 (2021).

140. Bittner, G. D., Denison, M. S., Yang, C. Z., Stoner, M. A. & He, G. Chemicals having estrogenic activity can be released from some bisphenol A-free, hard and clear, thermoplastic resins. *Environ Health* **13**, 103 (2014).

141. García-Córcoles, M. T. *et al.* Determination of bisphenols with estrogenic activity in plastic packaged baby food samples using solid-liquid extraction and clean-up with dispersive sorbents followed by gas chromatography tandem mass spectrometry analysis. *Talanta* **178**, 441–448 (2018).

142. Duty, S. M. *et al.* Potential sources of bisphenol A in the neonatal intensive care unit. *Pediatrics* **131**, 483–489 (2013).

143. Mercogliano, R. & Santonicola, S. Investigation on bisphenol A levels in human milk and dairy supply chain: A review. *Food Chem Toxicol* **114**, 98–107 (2018).

144. Talsness, C. E., Andrade, A. J. M., Kuriyama, S. N., Taylor, J. A. & vom Saal, F. S. Components of plastic: experimental studies in animals and relevance for human health. *Philos Trans R Soc Lond B Biol Sci* **364**, 2079–2096 (2009).

145. Geens, T., Roosens, L., Neels, H. & Covaci, A. Assessment of human exposure to Bisphenol-A, Triclosan and Tetrabromobisphenol-A through indoor dust intake in Belgium. *Chemosphere* **76**, 755–760 (2009).

146. Vandenberg, L. N. *et al.* A round robin approach to the analysis of bisphenol A (BPA) in human blood samples. *Environ Health* **13**, 25 (2014).

147. Wilson, N. K., Chuang, J. C., Morgan, M. K., Lordo, R. A. & Sheldon, L. S. An observational study of the potential exposures of preschool children to pentachlorophenol, bisphenol-A, and nonylphenol at home and daycare. *Environ Res* **103**, 9–20 (2007).

148. Pirard, C., Sagot, C., Deville, M., Dubois, N. & Charlier, C. Urinary levels of bisphenol A, triclosan and 4-nonylphenol in a general Belgian population. *Environ Int* **48**, 78–83 (2012).

149. Welshons, W. V., Nagel, S. C. & vom Saal, F. S. Large effects from small exposures. III. Endocrine mechanisms mediating effects of bisphenol A at levels of human exposure. *Endocrinology* **147**, S56-69 (2006).

150. Golub, M. S. *et al.* Bisphenol A: developmental toxicity from early prenatal exposure. *Birth Defects Res B Dev Reprod Toxicol* **89**, 441–466 (2010).

151. Nachman, R. M. *et al.* Serial Free Bisphenol A and Bisphenol A Glucuronide Concentrations in Neonates. *J Pediatr* **167**, 64–69 (2015).

152. Edginton, A. N. & Ritter, L. Predicting plasma concentrations of bisphenol A in children younger than 2 years of age after typical feeding schedules, using a physiologically based toxicokinetic model. *Environ Health Perspect* **117**, 645–652 (2009).

153. Ackerman, L. K. *et al.* Determination of bisphenol A in U.S. infant formulas: updated methods and concentrations. *J Agric Food Chem* **58**, 2307–2313 (2010).

154. Mielke, H. & Gundert-Remy, U. Bisphenol A levels in blood depend on age and exposure. *Toxicol Lett* **190**, 32–40 (2009).
155. Moreman, J. *et al.* Acute Toxicity, Teratogenic, and Estrogenic Effects of Bisphenol A and Its Alternative Replacements Bisphenol S, Bisphenol F, and Bisphenol AF in Zebrafish Embryo-Larvae. *Environ Sci Technol* **51**, 12796–12805 (2017).
156. Feng, Y. *et al.* Bisphenol AF may cause testosterone reduction by directly affecting testis function in adult male rats. *Toxicol Lett* **211**, 201–209 (2012).
157. Rochester, J. R. & Bolden, A. L. Bisphenol S and F: A Systematic Review and Comparison of the Hormonal Activity of Bisphenol A Substitutes. *Environ Health Perspect* **123**, 643–650 (2015).
158. Skledar, D. G. *et al.* Influence of metabolism on endocrine activities of bisphenol S. *Chemosphere* **157**, 152–159 (2016).
159. CDC - PRAMS Questionnaires - Pregnancy Risk Assessment Monitoring System - Reproductive Health. https://www.cdc.gov/prams/questionnaire.htm (2022).
160. Welcome to Eurachem. https://www.eurachem.org/.
161. Griffiths, A. J., Miller, J. H., Suzuki, D. T., Lewontin, R. C. & Gelbart, W. M. Somatic versus germinal mutation. *An Introduction to Genetic Analysis. 7th edition* (2000).
162. Strom, S. P. Current practices and guidelines for clinical next-generation sequencing oncology testing. *Cancer Biol Med* **13**, 3–11 (2016).
163. Ma, J. *et al.* Mutation allele frequency threshold does not affect prognostic analysis using next-generation sequencing in oral squamous cell carcinoma. *BMC Cancer* **18**, 758 (2018).
164. Kim, J. *et al.* The use of technical replication for detection of low-level somatic mutations in next-generation sequencing. *Nature Communications* **10**, 1047 (2019).
165. (How to) Call somatic mutations using GATK4 Mutect2. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360035531132.
166. Chen, Z. *et al.* Systematic comparison of somatic variant calling performance among different sequencing depth and mutation frequency. *Sci Rep* **10**, 3501 (2020).
167. Coudray, A., Battenhouse, A. M., Bucher, P. & Iyer, V. R. Detection and benchmarking of somatic mutations in cancer genomes using RNA-seq data. *PeerJ* **6**, e5362 (2018).

168. Van Horebeek, L. *et al.* A robust pipeline with high replication rate for detection of somatic variants in the adaptive immune system as a source of common genetic variation in autoimmune disease. *Hum Mol Genet* **28**, 1369–1380 (2019).

169. Marnell, C. S., Bick, A. & Natarajan, P. Clonal hematopoiesis of indeterminate potential (CHIP): Linking somatic mutations, hematopoiesis, chronic inflammation and cardiovascular disease. *J Mol Cell Cardiol* **161**, 98–105 (2021).

170. Senguttuvan, N. B., Subramanian, V., Venkatesan, V., Muralidharan, T. R. & Sankaranarayanan, K. Clonal hematopoiesis of indeterminate potential (CHIP) and cardiovascular diseases-an updated systematic review. *J Genet Eng Biotechnol* **19**, 105 (2021).

171. Kumar, P., Kopecky, S. L., Yang, E. H. & Oren, O. Clonal Hematopoiesis of Indeterminate Potential and Cardiovascular Disease. *Curr Oncol Rep* **22**, 87 (2020).

172. GRCh37 - hg19 - Genome - Assembly - NCBI. https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/.

173. bwa.1. https://bio-bwa.sourceforge.net/bwa.shtml.

174. Tischler, G. biobambam2. (2023).

175. BaseRecalibrator. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360036898312-BaseRecalibrator (2020).

176. ApplyBQSR. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360037055712-ApplyBQSR (2021).

177. HaplotypeCaller. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360037225632-HaplotypeCaller (2023).

178. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

179. CollectAlignmentSummaryMetrics (Picard). *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360036883111-CollectAlignmentSummaryMetrics-Picard- (2019).

180. VariantRecalibrator. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360036510892-VariantRecalibrator (2022).

181. ApplyVQSR. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360037056912-ApplyVQSR (2020).

182. CollectVariantCallingMetrics (Picard). *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360037057132-CollectVariantCallingMetrics-Picard- (2020).

183. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122 (2016).

184. Pedersen, B. S. & Quinlan, A. R. Who's Who? Detecting and Resolving Sample Anomalies in Human DNA Sequencing Studies with Peddy. *Am J Hum Genet* **100**, 406–413 (2017).

185. FilterMutectCalls. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360036856831-FilterMutectCalls.

186. Jaiswal, S. & Ebert, B. L. Clonal hematopoiesis in human aging and disease. *Science* **366**, eaan4673 (2019).

187. Panel of Normals (PON). *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360035890631-Panel-of-Normals-PON-.

188. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213–219 (2013).

189. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178–192 (2013).

190. Gerona, R., Vom Saal, F. S. & Hunt, P. A. BPA: have flawed analytical techniques compromised risk assessments? *Lancet Diabetes Endocrinol* **8**, 11–13 (2020).

191. Catalano, P. M. & Shankar, K. Obesity and pregnancy: mechanisms of short term and long term adverse consequences for mother and child. *BMJ* **356**, j1 (2017).

192. Waters, T. P., Huston-Presley, L. & Catalano, P. M. Neonatal body composition according to the revised institute of medicine recommendations for maternal weight gain. *J Clin Endocrinol Metab* **97**, 3648–3654 (2012).

193. Chu, S. Y. *et al.* Maternal obesity and risk of gestational diabetes mellitus. *Diabetes Care* **30**, 2070–2076 (2007).

194. Cnattingius, S. *et al.* Maternal obesity and risk of preterm delivery. *JAMA* **309**, 2362–2370 (2013).

195. Starling, A. P. *et al.* Associations of maternal BMI and gestational weight gain with neonatal adiposity in the Healthy Start study. *Am J Clin Nutr* **101**, 302–309 (2015).

196. Gern, J. E. The Urban Environment and Childhood Asthma study. *J Allergy Clin Immunol* **125**, 545–549 (2010).

197. Ly, N. P. *et al.* Characterization of regulatory T cells in urban newborns. *Clinical and Molecular Allergy* **7**, 8 (2009).

198. Gould, G. S., Havard, A., Lim, L. L., The Psanz Smoking In Pregnancy Expert Group, null & Kumar, R. Exposure to Tobacco, Environmental Tobacco Smoke and Nicotine in Pregnancy: A Pragmatic Overview of Reviews of Maternal and Child Outcomes, Effectiveness of Interventions and Barriers and Facilitators to Quitting. *Int J Environ Res Public Health* **17**, 2034 (2020).

199. Rogers, J. M. Tobacco and pregnancy. *Reprod Toxicol* **28**, 152–160 (2009).

200. Murin, S., Rafii, R. & Bilello, K. Smoking and smoking cessation in pregnancy. *Clin Chest Med* **32**, 75–91 (2011).

201. Leng, H. *et al.* [Effects of different types and concentration of oral sweet solution on reducing neonatal pain during heel lance procedures]. *Zhonghua Er Ke Za Zhi* **51**, 654–658 (2013).

202. Bisphenol A : Part 1. Facts and figures on human and environmental health issues and regulatory perspectives | RIVM. https://www.rivm.nl/publicaties/bisphenol-a-part-1-facts-and-figures-on-human-and-environmental-health-issues-and.

203. Grignard, E., Lapenna, S. & Bremer, S. Weak estrogenic transcriptional activities of Bisphenol A and Bisphenol S. *Toxicol In Vitro* **26**, 727–731 (2012).

204. Hashimoto, Y. *et al.* Measurement of estrogenic activity of chemicals for the development of new dental polymers. *Toxicol In Vitro* **15**, 421–425 (2001).

205. Ji, K., Hong, S., Kho, Y. & Choi, K. Effects of bisphenol s exposure on endocrine functions and reproduction of zebrafish. *Environ Sci Technol* **47**, 8793–8800 (2013).

206. Qiu, W. *et al.* Actions of Bisphenol A and Bisphenol S on the Reproductive Neuroendocrine System During Early Development in Zebrafish. *Endocrinology* **157**, 636–647 (2016).

207. Wan, Y. *et al.* Relationship between maternal exposure to bisphenol S and pregnancy duration. *Environ Pollut* **238**, 717–724 (2018).

208. Protano, C. *et al.* The Carcinogenic Effects of Formaldehyde Occupational Exposure: A Systematic Review. *Cancers (Basel)* **14**, 165 (2021).

209. Augenreich, M. *et al.* Vascular dysfunction and oxidative stress caused by acute formaldehyde exposure in female adults. *Am J Physiol Heart Circ Physiol* **319**, H1369–H1379 (2020).

210. Matsuoka, T., Takaki, A., Ohtaki, H. & Shioda, S. Early changes to oxidative stress levels following exposure to formaldehyde in ICR mice. *J Toxicol Sci* **35**, 721–730 (2010).

211. Ghelli, F. *et al.* The formation of SCEs as an effect of occupational exposure to formaldehyde. *Arch Toxicol* **96**, 1101–1108 (2022).

212. Ghelli, F. *et al.* The role of phase I, phase II, and DNA-repair gene polymorphisms in the damage induced by formaldehyde in pathologists. *Sci Rep* **11**, 10507 (2021).

213. Ezratty, V. *et al.* Effect of Formaldehyde on Asthmatic Response to Inhaled Allergen Challenge. *Environ Health Perspect* **115**, 210–214 (2007).

214. Gerald McGwin, J., Lienert, J. & John I. Kennedy, J. Formaldehyde Exposure and Asthma in Children: A Systematic Review. *Environmental Health Perspectives* **118**, 313 (2010).

215. Vaughan, T. *et al.* Occupational exposure to formaldehyde and wood dust and nasopharyngeal carcinoma. *Occup Environ Med* **57**, 376–384 (2000).

216. Kwon, S.-C., Kim, I., Song, J. & Park, J. Does formaldehyde have a causal association with nasopharyngeal cancer and leukaemia? *Ann Occup Environ Med* **30**, 5 (2018).

217. Zhang, L., Steinmaus, C., Eastmond, D. A., Xin, X. K. & Smith, M. T. Formaldehyde exposure and leukemia: a new meta-analysis and potential mechanisms. *Mutat Res* **681**, 150–168 (2009).

218. Collins, J. J. & Lineker, G. A. A review and meta-analysis of formaldehyde exposure and leukemia. *Regul Toxicol Pharmacol* **40**, 81–91 (2004).

219. Santovito, A. *et al.* Combined analysis of chromosomal aberrations and glutathione S-transferase M1 and T1 polymorphisms in pathologists occupationally exposed to formaldehyde. *Arch Toxicol* **85**, 1295–1302 (2011).

220. Souza, A. D. & Devi, R. Cytokinesis blocked micronucleus assay of peripheral lymphocytes revealing the genotoxic effect of formaldehyde exposure. *Clin Anat* **27**, 308–312 (2014).

221. Pala, M. *et al.* Occupational exposure to formaldehyde and biological monitoring of Research Institute workers. *Cancer Detect Prev* **32**, 121–126 (2008).

222. Peteffi, G. P. *et al.* Environmental and biological monitoring of occupational formaldehyde exposure resulting from the use of products for hair straightening. *Environ Sci Pollut Res Int* **23**, 908–917 (2016).

223. Kural, C. *et al.* Glutathione S-Transferases and Cytochrome P450 Enzyme Expression in Patients with Intracranial Tumors: Preliminary Report of 55 Patients. *Med Princ Pract* **28**, 56–62 (2019).

224. Umansky, C. *et al.* Endogenous formaldehyde scavenges cellular glutathione resulting in redox disruption and cytotoxicity. *Nat Commun* **13**, 745 (2022).

225. Luch, A., Frey, F. C. C., Meier, R., Fei, J. & Naegeli, H. Low-Dose Formaldehyde Delays DNA Damage Recognition and DNA Excision Repair in Human Cells. *PLoS One* **9**, e94149 (2014).

226. de Graaf, B., Clore, A. & McCullough, A. K. Cellular Pathways for DNA Repair and Damage Tolerance of Formaldehyde-Induced DNA-Protein Crosslinks. *DNA Repair (Amst)* **8**, 1207–1214 (2009).

Appendix A: full text of the 21 papers published studies during my doctorate

International Journal of *Environmental Research and Public Health*

MDPI

Article

Air Pollution and Aeroallergens as Possible Triggers in Preterm Birth Delivery

Enrico Cocchi [1,2,3,†], Valeria Bellisario [1,4,†], Francesco Cresi [5], Claudio Plazzotta [5], Claudio Cassardo [6], Consolata Siniscalco [7], Licia Peruzzi [2,‡] and Roberto Bono [1,*,‡]

[1] Department of Public Health and Pediatrics, University of Turin, 10126 Turin, Italy
[2] Pediatric Nephrology Unit, Regina Margherita Children's Hospital, 10126 Turin, Italy
[3] Pediatric Residency School, University of Turin, 10126 Turin, Italy
[4] Biostatistics Residency School, University of Turin, 10126 Turin, Italy
[5] Neonatal Intensive Care Unit, Sant'Anna Obstetric Gynecological Hospital, 10126 Turin, Italy
[6] Physics Department, University of Turin, 10125 Turin, Italy
[7] Department of Life Sciences and Systems Biology, University of Turin, 10123 Turin, Italy
* Correspondence: roberto.bono@unito.it; Tel.: +39-011-6705818
† These authors contributed equally to this work.
‡ These authors contributed equally to this work.

Abstract: Preterm birth (PTB) identifies infants prematurely born <37 weeks/gestation and is one of the main causes of infant mortality. PTB has been linked to air pollution exposure, but its timing is still unclear and neglects the acute nature of delivery and its association with short-term effects. We analyzed 3 years of birth data (2015–2017) in Turin (Italy) and the relationships with proinflammatory chemicals (PM2.5, O_3, and NO_2) and biological (aeroallergens) pollutants on PTB vs. at-term birth, in the narrow window of a week before delivery. A tailored non-stationary Poisson model correcting for seasonality and possible confounding variables was applied. Relative risk associated with each pollutant was assessed at any time lag between 0 and 7 days prior to delivery. PTB risk was significantly associated with increased levels of both chemical (PM2.5, RR = 1.023 (1.003–1.043), O_3, 1.025 (1.001–1.048)) and biological (aeroallergens, RR ~ 1.01 (1.0002–1.016)) pollutants in the week prior to delivery. None of these, except for NO_2 (RR = 1.01 (1.002–1.021)), appeared to play any role on at-term delivery. Pollutant-induced acute inflammation eliciting delivery in at-risk pregnancies may represent the pathophysiological link between air pollution and PTB, as testified by the different effects played on PTB revealed. Further studies are needed to better elucidate a possible exposure threshold to prevent PTB.

Keywords: public health; preterm birth; newborn; acute inflammation; delivery; air pollution

Citation: Cocchi, E.; Bellisario, V.; Cresi, F.; Plazzotta, C.; Cassardo, C.; Siniscalco, C.; Peruzzi, L.; Bono, R. Air Pollution and Aeroallergens as Possible Triggers in Preterm Birth Delivery. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1610. https://doi.org/10.3390/ijerph20021610

Academic Editor: Paul B. Tchounwou

Received: 15 November 2022
Revised: 9 January 2023
Accepted: 10 January 2023
Published: 16 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Preterm birth (PTB) indicates infants born before 37 weeks of gestation, thus, they are not fully developed for extra-uterine life. PTB is a public health concern [1] affecting ~10% of total births worldwide, with higher rates in developing countries and numbers that are constantly increasing [2]. In developed countries, PTB represents the main cause of infant mortality [3] and has been associated with significantly impaired health outcomes [4–6] and dramatic community costs [7].

Alongside historical and well known risk factors, namely low socioeconomic status, age, ethnicity, tobacco, substance abuse, poor nutritional status, and the presence of birth defects [8], recent epidemiological studies have linked PTB with chemical air pollution [8,9]. While the effects of air pollution on PTB are now known, the timing with which the effects occur is still controversial [10]. Some studies have claimed that early exposure during pregnancy is responsible [11], while other studies have indicated that late exposure is responsible [12,13]. This confusion is probably due to different pollution-induced reactions

causing different detrimental effects on pregnancy depending on the fetal developmental stage in which they occur, their duration, and analyzed outcome. In this regard, studies that have focused on definite narrow temporal windows in order to better elucidate specific pollutants, co-factors, and underlying mechanisms of action are lacking [8]. Moreover, PTB is a complex subject in which the long-term nature of fetal development is associated with an acute medical condition that mandates delivery, with different factors potentially playing very different roles in each scenario. PTB delivery is due to conditions that make it impossible to continue the uterine fetal development. The acute factors either prematurely trigger the delivery itself, or are responsible for such a strong fetal suffering that it overcomes the string mortality and morbidity risks associated with PTB. The role of air pollution in PTB has been historically studied through models inherited from other diseases, such as cardiovascular and neurological diseases, that have a peculiar chronic nature [14]. Similarly, pollution-induced chronic inflammation during pregnancy has been proven to cause developmental impairments that predispose to PTB; however, PTB itself is an acute medical emergency, whose main trigger is an inflammatory distress that induces uterine contractions and subsequent preterm delivery [15]. Interestingly, acute inflammation is the main elucidated mechanism contributing to both PTB and air pollutant-related detrimental effects [16,17]. Short-term exposure to air pollution proved to trigger inflammation, detectable through increased inflammatory markers [18], and mice models have revealed how this induced systemic inflammation could lead to acute diseases [19]. Thus, air pollutant exposure can trigger PTB through an acute inflammatory reaction that elicits uterine contractions and subsequent preterm delivery. In this regard, it is well known how acute inflammation induces uterine contractions and PTB [20], and short-term exposure to air pollutants has been demonstrated to induce an upregulation of delivery-involved inflammatory cytokines, such as interleukins (IL-1 and IL-6), and tumor necrosis factor α (TNF-α), with a strong established role in PTB as well [18,21,22]. Interestingly, this air pollution exposure-induced acute inflammation has been proven to be reversible over days and is pharmacologically addressable [23]. All these lines of evidence advocate for both possible preventive and therapeutic approaches to pollutant exposure and PTB, and reinforce the need to better elucidate underlying mechanisms and possible pharmacological targets through studies focused on specific narrow temporal windows, which are currently lacking.

Thus, the aim of the present study is to assess the specific effects of maternal exposure to chemical and biological air pollutants in the narrow window of 7 days before delivery in both preterm and at-term births. This is in order to better elucidate the role of the air pollutants, other co-factors, and the underlying mechanisms acutely associated with preterm delivery.

2. Materials and Methods

2.1. Data Collection Area

Turin, the capital of the Piedmont region (North-Western Italy), is one of the most polluted European cities, located 239 m above sea level (a.s.l.); it has 886,837 inhabitants, a population density equal to 6813 per km^2. Daily data for the period from 1 January 2015 to 31 December 2017 (1096 days) for the city of Turin were collected or derived as described below. The locations of the data sources are shown in Supplementary Materials.

2.2. Study Population

Daily data on births were obtained from the birth registry and registration records of the main obstetric hospital in Turin, Italy (Fetal Maternal Department, Sant'Anna Obstetric Gynecological Hospital), recording maternal and fetal data during the 3-year period from 1 January 2015 to 31 December 2017. Recorded data included:

1. Maternal age at delivery;
2. Gestational age (weeks of amenorrhea);
3. Apgar 1/5;
4. Sex;

5. Twins;
6. Mother's country of origin.

Data were anonymized with respect to the European General Data Protection Regulation (GDPR 101/2018).

2.3. Meteorological Data

Meteorological data were obtained from a station placed at 254 a.s.l., on the roof of the Department of Physics of the University of Turin, located at about 2.5 km from the S. Anna Hospital. The station is permanently active and collects weather data in the urban surface layer of the city. Data were collected every 5 s and subsequently averaged every 5 min. Data were aggregated in a daily form for the analysis.

2.4. Chemical Air Pollution Data

We extracted concentrations of air pollutants (NO_2, PM2.5, and O_3) from data collected at the urban background monitoring station located in Turin, viale Augusto Monti. The data collection and summary was performed by the Local Environmental Protection Agency (ARPA Piemonte), which is coordinated by the regional air pollution service of Piedmont Region, as per the current European legislation (DIR 2008/50/ECX). Data were collected hourly and daily aggregated for the analysis.

2.5. Aeroallergen Data

Corylaceae, Cupressaceae, Gramineae, Urticaceae, Ambrosia, and Betula pollens were considered in this study. The measurements of pollen daily data were conducted in a station located ~12 m above the ground without surrounding obstacles, as required in such cases, on the flat roof of a building located 2.6 km from S. Anna Hospital. The sampling station consisted of a HIRST sampler. The HIRST sampler consists of three core parts: a swivel head, a suction pump, and a deposition drum. The latter represents the actual sampling part. This part rotates at 2 mm/h. Collection involved a weekly application of specific adhesive tape on the drum part of the samples. This tape was able to capture aeroallergens, ensuring no loss of rebound or natural detachment. A constant airflow of ~10 L/min was provided by using an air pump, which represented ~14.4 m^3 daily. Aeroallergen counts were performed by the Department of Life Science and System Biology of the University of Turin, and expressed in our analysis in the form of grains/m^3.

2.6. Statistical Analysis

We summarized quantitative variables as means $\pm$ SD, medians $\pm$ interquartile ranges (IQR), showing minimum and maximum values as well. To allow comparison of variability among different variables, the interquartile ratio (as IQR/median ratio) was also computed. The normality of data was checked through the Kolmogorov–Smirnov test. Quantitative parameters were found to be non-normally distributed, thus, non-parametric tests, such as the Mann–Whitney U-test, served to assess between-group differences for those variables. Group comparisons were computed through chi-square or Fisher's exact test, as appropriate, based on categorical data under analysis. A two-sided p-value < 0.05 was considered to be significant. Pearson's r coefficient was used to investigate correlations among exposure variables. In order to focus on a narrow temporal window and to correctly assess short-term effects, the associations between PTB (the dependent variable) and concentrations of chemical and biological air pollutants (independent variables) were analyzed using generalized linear models (GLMs) fitting a non-stationary Poisson process [24,25]. We used the following model:

$$f(\lambda_t + 1) = \alpha + \sum_{i=1}^{k} \beta Xi + NS(Z_t)$$

where

- f: log link function
- λ_t: count of daily PTB at day t
- α: intercept constant
- β: estimated parameters vector
- X_i: matrix of k independent variables (exposure and adjustment variables)
- $NS(Z_t)$: natural spline smoothing function of calendar day Z

As in several literature reports that have focused on the same analysis model, in order to take the medium-/long-term trend that may shape the time data under analysis into account, a natural spline-smoothing function was calculated on 14 degrees of freedom (df) [25]. We restricted possible df to a maximum of 18 (corresponding to a window of ~60-days) in order to avoid overfitting [26,27]. We identified the best df as the value that minimized the absolute values of the residuals partial autocorrelation function (PACF) sum [26]. PACF residuals were corrected for the day of the week, as clarified in the point below, to remove the 7-day positive correlation, when estimating the spline-smoothing function.

We performed further variable adjustments in order to take into consideration characteristics of the location under study that might otherwise bias the model under analysis:

(a) Day of the week (from Monday to Sunday);
(b) Holidays, we considered main holidays in our zone, i.e., Christmas and Easter, ± 3 days around them, other holidays, and other days. This resulted in a 4-level variable that was used in the model.
(c) Summer population decrease, in our specific zone, population is known to decrease during summer holidays and this variable was intended to adjust for this effect, resulting in a categorical variable with factors considering such dates, i.e., from Saturday before Mid-August to the next Sunday (for a total amount of 16 days per year), from 16 July to the end of August (removing the aforementioned period), and all other days [28];
(d) Daily average daily temperature (°C);
(e) Daily average humidity (%)—relative;
(f) Daily precipitations (mm)—cumulative.

Once the model was fitted to the actual data, we considered one chemical pollutant among PM2.5, NO_2, O_3, or aeroallergens, alongside medium-/long-trend function, non-meteorological variables (day of the week, holidays, and summer population decrease), and meteorological variables (daily temperature, daily relative humidity, and cumulative daily precipitations). Temperature and humidity underwent natural splines with 1 and 2 df transformation, respectively. The df was chosen through PACF criterion as explained above. Daily precipitations were binary coded, i.e., 1 if cumulative precipitation ≥ 1 mm and 0 otherwise.

Exposure variables were included in all the models at different single time lags: starting from the same day of the PTB evaluation (Lag 0) to 7 days before (Lag 7). We identified such a time frame in order to focus the analysis on specific short-term potential effects of air pollutants in eliciting preterm vs. full-term births. To the best of our knowledge, this is the first literature report of such an acute effect-focused analysis. Thus, no previous specific time frame is available. The identification of such a threshold was based on a pathophysiological rationale, as our analysis was based on investigating acute inflammation potential roles in delivery and preterm vs. full-term birth compensation capacities. Thus, as acute inflammation is a process known to exert its effect within a few days from the exposure, we arbitrarily selected this specific 7-day time frame in order to include in the analysis lag all potential acute inflammatory effects of air pollutants, from 24 h to a week from exposure. Associations between exposure variables and preterm or full-term births are reported as relative risk (RR) with respective 95% confidence intervals (CIs). RR values and their CIs are calculated as exponential of GLM resulting coefficients for each specific exposure variable under analysis. Exposure variable association coefficients were calculated based on 10 $\mu g/m^3$ increase in PM2.5, NO_2, and O_3 concentrations, and a 10 grains/m^3 increase in the case of aeroallergens.

3. Results

3.1. Population Enrolled

In the study period, 21,509 births were observed, of which 3167 PTBs (14.7%). Median gestational age was 39 + 1 for the whole cohort, 39 + 3 for term, and 35 + 1 for PTB. The peculiar differences between preterm and term subpopulations were:

- Higher twins prevalence in PTB (preterm 42.2% vs. term 1.87%, $p < 0.01$);
- Lower Apgar 1/5 scores in PTB (preterm Apgar 1 score 7.6 ± 2.1 vs. term 8.8 ± 0.9, $p < 0.01$; preterm Apgar 5 score 8.2 ± 1.5 vs. term 8.9 ± 0.6, $p < 0.01$);
- Higher maternal age in PTB (preterm 34.1 ± 5.5 vs. term 33.3 ± 5.4, $p < 0.01$).

Cohort characteristics are summarized in Table 1.

3.2. Temporal Distribution of Births

The temporal distribution of births was analyzed in order to identify the best smoothing strategy possible. Figure 1 graphically reports the daily number of births in the period under study (1 January 2015–31 December 2017) and the natural spline smoothing function of calendar day Z with 14° of freedom utilized in the models.

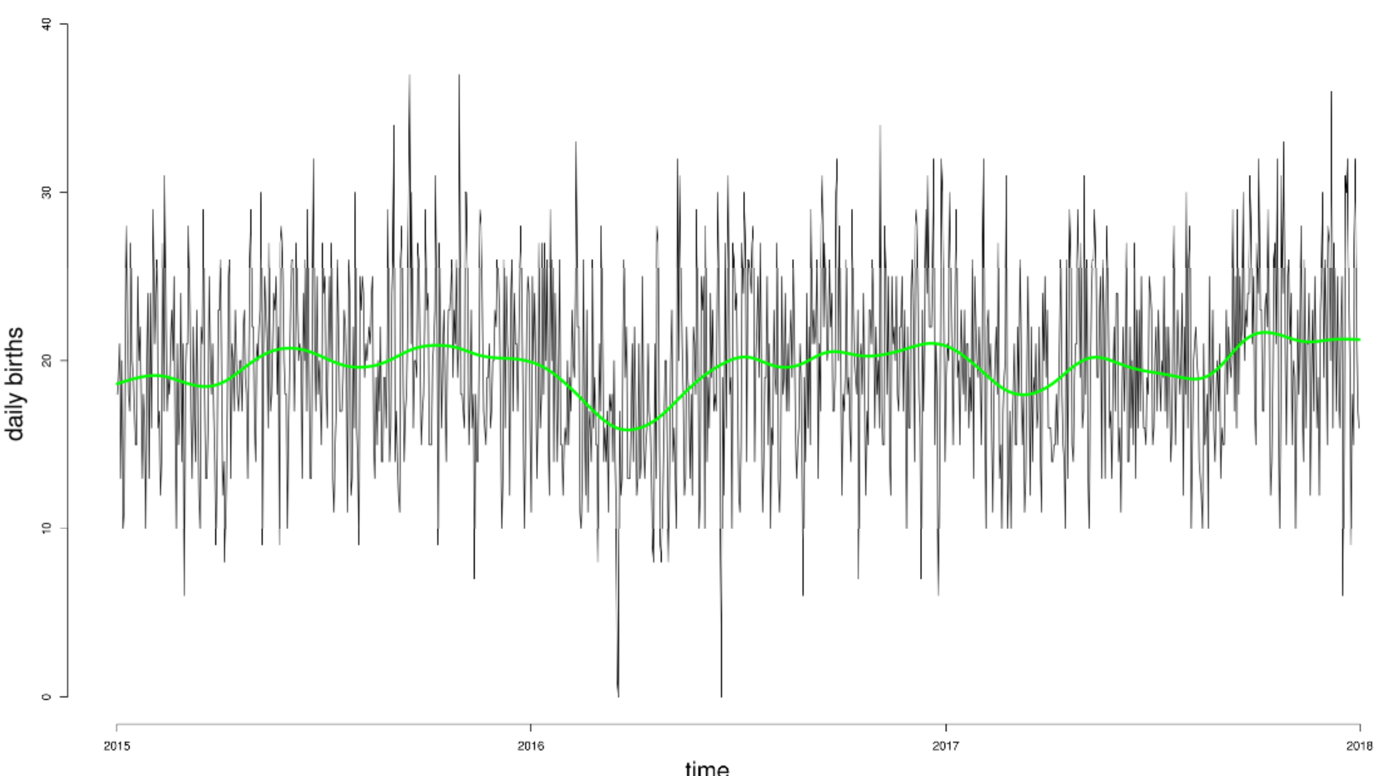


Figure 1. Daily incidence of births during the period under analysis, and the natural spline smoothing function of calendar day Z with 14° of freedom utilized in the Poisson non-stationary models (green line).

Figure 2 summarizes the average number of daily births (Figure 2A) and PTBs (Figure 2B) by month of the year, revealing the same unclear pattern of seasonality detectable in daily data in Figure 1 and addressed through the spline smoothing function.

Table 1. Characteristics of the cohort under analysis.

	All Cohort (*n* = 21,509)				**Term (*n* = 18,345)**				**Preterm (*n* = 3164)**			
	Median (IQR)	**IQR/Median**	**Min–Max**	**Mean ± SD**	**Median (IQR)**	**IQR/Median**	**Min–Max**	**Mean ± SD**	**Median (IQR)**	**IQR/Median**	**Min–Max**	**Mean ± SD**
Maternal age at delivery (weeks)	33.7 (7.4)	0.2	14.8–48.6	33.4 ± 5.4	33.6 (7.4)	0.2	14.8–48.6	33.3 ± 5.4	34.3 (7.7)	0.2	15.6–46.9	34.1 ± 5.5
Gestational age (weeks + days)	39 + 1 (2 + 1)	0.05	21 + 5–49 + 0	38.7 ± 2.3	39 + 3 (1.7)	0.04	37 + 0–49 + 0	39.4 ± 1.5	35 + 1 (3 + 0)	0.09	21 + 5–36 + 6	34.3 ± 2.7
Apgar 1	9 (0)	0	0–10	8.6 ± 1.2	9 (0)	0	0–10	8.8 ± 0.9	8 (2)	0.2	0–10	7.6 ± 2.1
Apgar 5	9 (0)	0	0–10	8.9 ± 0.8	9 (0)	0	0–10	8.9 ± 0.6	9 (1)	0.11	0–10	8.2 ± 1.5
Sex prevalence (%)	Male 11,066 (51.4%) Female 10,443 (48.6%)				Male 9423 (51.4%) Female 8919 (48.6%)				Male 1640 (51.8%) Female 1524 (48.2%)			
Twins prevalence (%)	Single 20,005 (93.1%) Double 1438 (6.6%) >2: 66 (0.3%)				Single 17,999 (98.3%) Double 317 (1.7%)				Single 1977 (62.5%) Double 1121 (35.4%) >2: 66 (2.1%)			
Maternal country of origin prevalence (%)	Italy 16,032 (74.5%) Romania 1769 (8.2%) Morocco 768 (3.6%) Nigeria 440 (2%) Albany 292 (1.4%) Peru 251 (1.2%) Egypt 244 (1.1%) China 152 (0.7%) Other 1561 (7.3%)				Italy 13,649 (74.4%) Romania 1512 (8.2%) Morocco 651 (3.6%) Nigeria 368 (2%) Albany 252 (1.4%) Peru 214 (1.2%) Egypt 218 (1.2%) China 136 (0.7%) Other 1342 (7.3%)				Italy 2381 (75.2%) Romania 257 (8.1%) Morocco 116 (3.7%) Nigeria 72 (2.3%) Albany 40 (1.3%) Peru 37 (1.2%) Egypt 26 (0.8%) China 17 (0.5%) Other 221 (6.9%)			

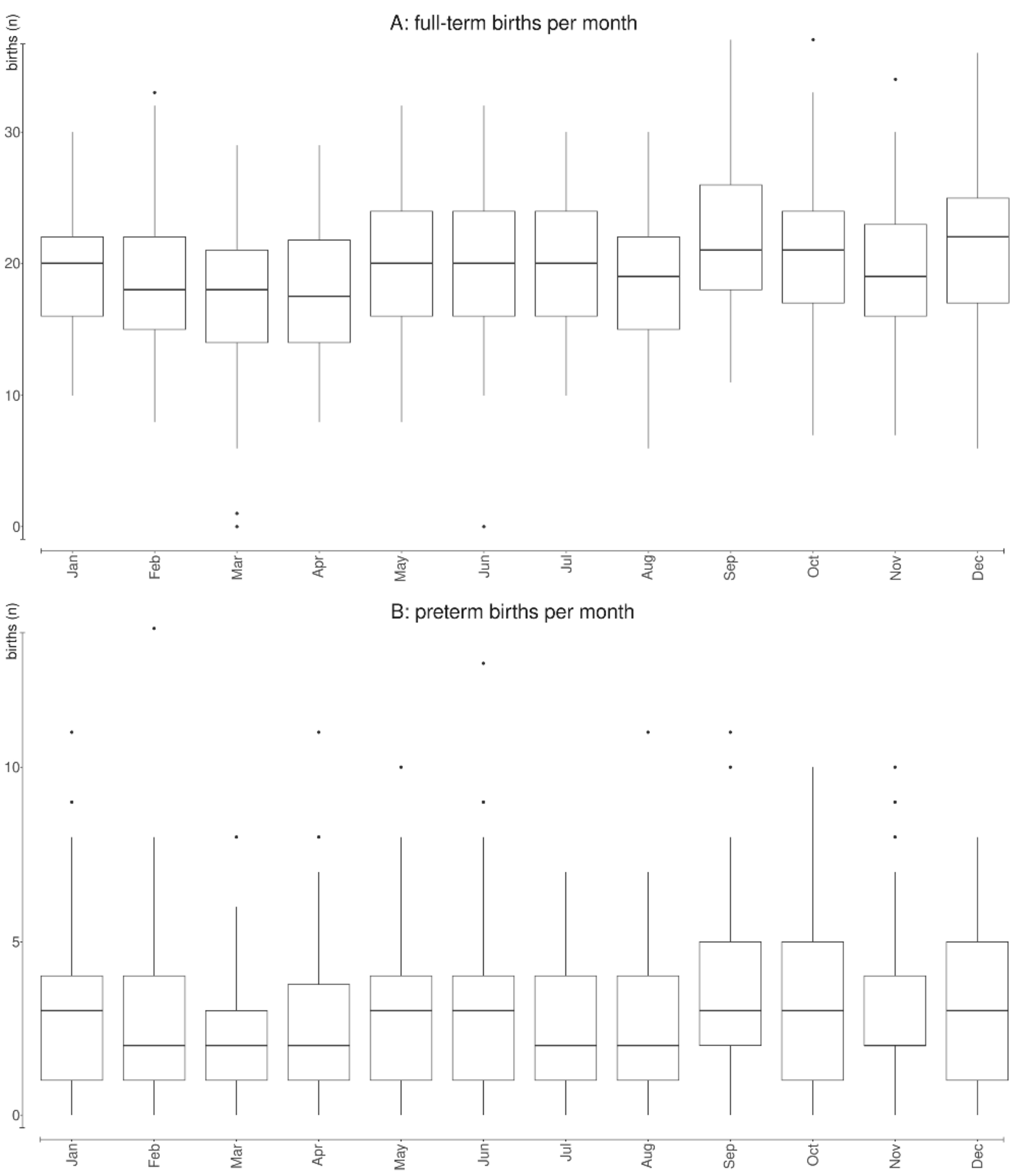


Figure 2. Average number of daily births (**A**) and preterm births (**B**) by month of the year during the 3-year period under analysis (1 January 2015–31 December 2017).

3.3. Environmental Pollutants and Allergens

Table 2 shows a general description of the daily concentrations of chemical and biological air pollutants in the area under study during the 3 years examined.

Table 2. Distribution of daily concentrations of air pollution and aeroallergens during the three years examined (1 January 2015–31 December 2017).

	Available Data (Days)	Median (IQR)	Interquartile Ratio	Min–Max	Mean ± SD
PM2.5 (μg/m^3)	996	17.0 (24.0)	1.4	5.0–163.0	25.6 ± 21.4
NO_2 (μg/m^3)	1081	49.9 (29.6)	0.6	10.9–129.8	54.1 ± 21.0
O_3 (μg/m^3)	1083	19.9 (51.1)	2.6	1.7–262.3	41.1 ± 46.6
Aeroallergens (grains/m^3)	888	11.8 (42.1)	28.1	0.0–563.0	36.0 ± 61.4

Chemical pollutants show rather high concentration levels, in accordance with the situation already studied in this area [24,28]. Due to their natural seasonality, aeroallergen concentrations showed larger variability as compared with chemical air pollutants (even if chemicals also have a seasonality dependent on their primary and/or secondary origin). According to the origin, pollutants are categorized as primary and secondary. Primary pollutants are directly emitted from their sources, while secondary pollutants are either formed by the atmospheric transformation of primary pollutants or other chemical compounds. PM2.5, and NO_2 (Figure 3A,D) showed a prevailing maximum level during the coldest months, as expected due to its primary origin. Contrarily, O_3 showed a behavior typical of a pollutant of secondary origin (Figure 3B). Aeroallergens showed high concentrations during warm season, and were virtually absent in winter (Figure 3C), as expected.

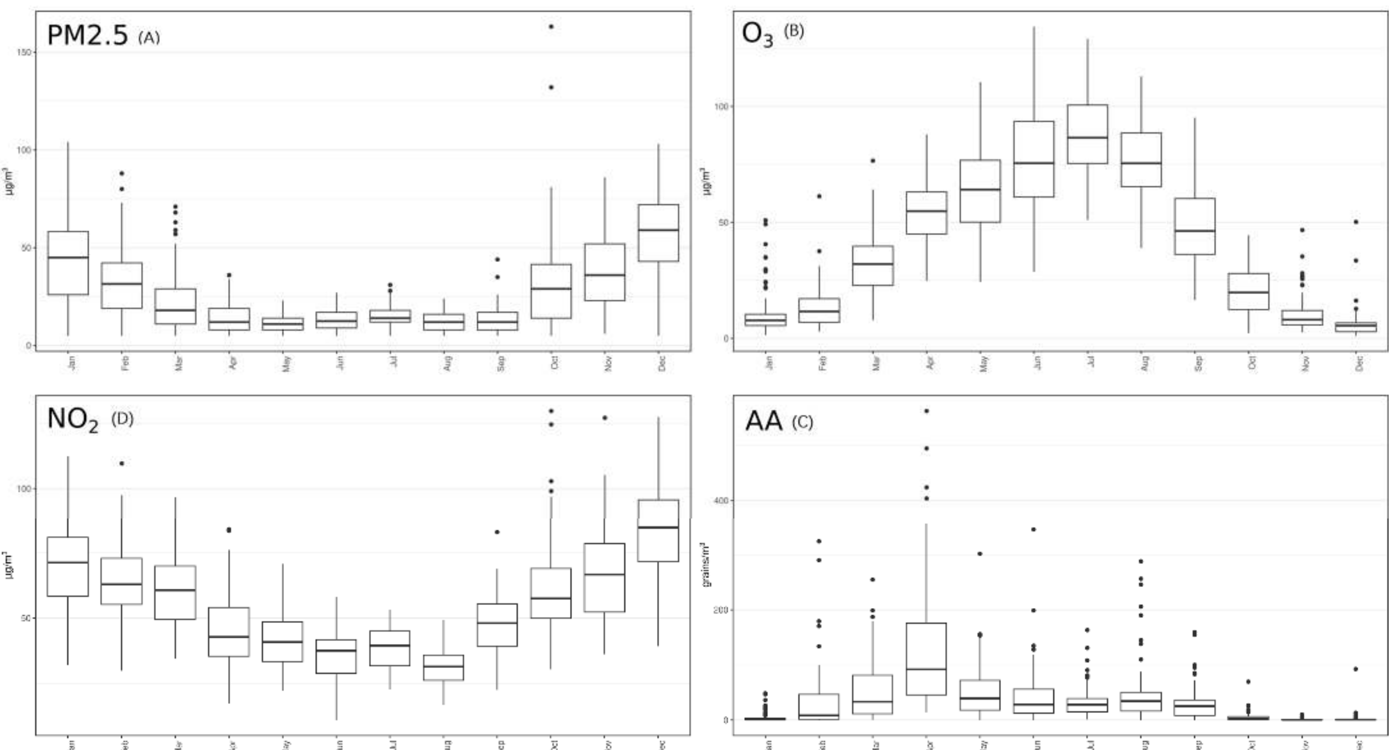


Figure 3. Mean concentration of each pollutant analyzed by month during the 3-years period under analysis. AA: aeroallergens.

A positive linear association between PM2.5 and NO_2 was detected (RR = 0.51, $p < 0.01$) and negative associations of O_3 with PM2.5 (RR = −0.34, $p < 0.01$) and NO_2 (RR = −0.46, $p < 0.01$). The associations between chemical pollutants and aeroallergens were weaker but still significant (PM2.5, RR = 0.06, $p < 0.01$; NO_2, RR = 0.04, $p < 0.01$; O_3 RR = 0.08, $p < 0.01$). All correlation plots are reported in Figure 4.

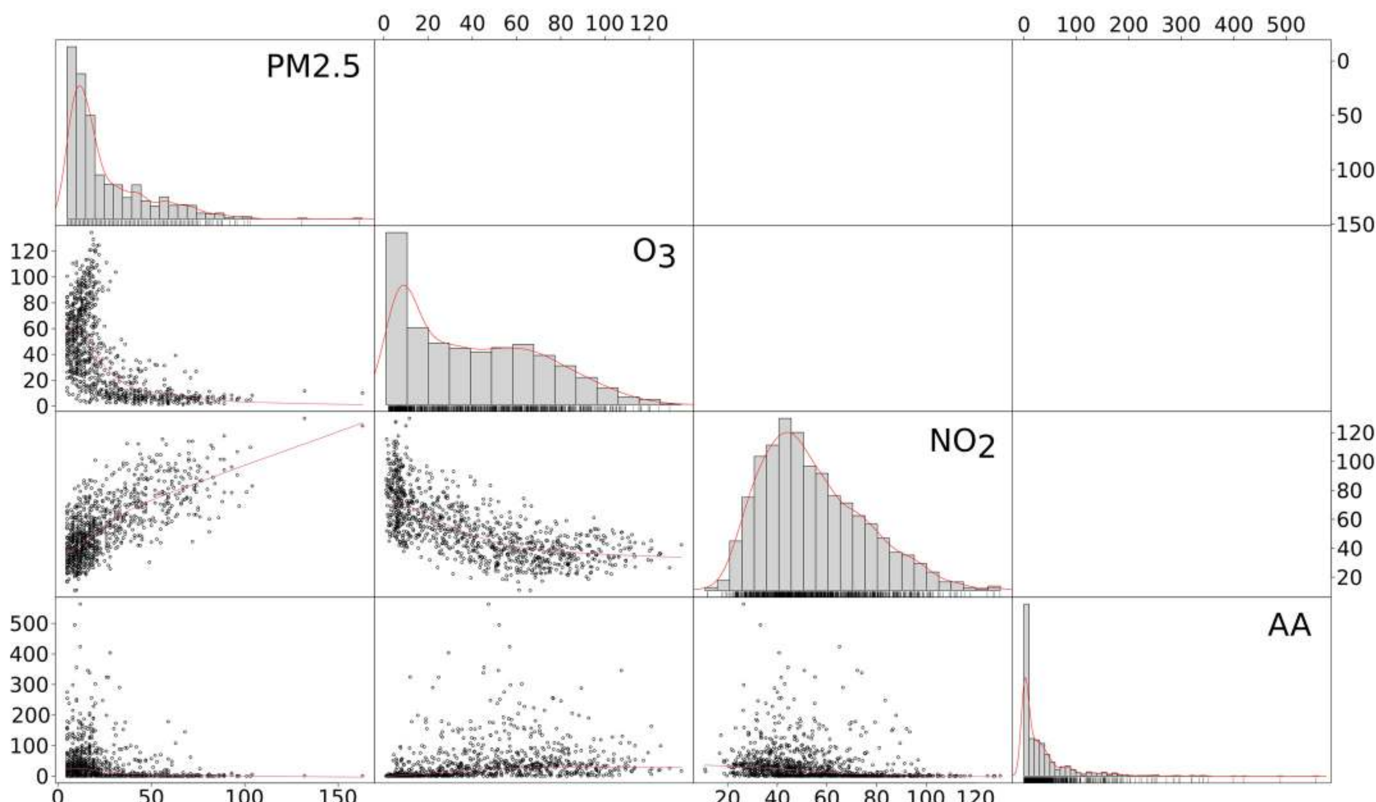


Figure 4. Matrix correlation plot among pollutants under analysis. The red lines over histograms represent the density function of that variable distribution. The red straight lines in scatter plots represents the correlation r 2 splope between the pollutants plotted in each scatter plot. AA: aeroallergens.

3.4. Correlations between Birth and Environmental Data

Potential confounders and their relationships were considered before the analysis. Tables 1 and 2 summarize the mean number of daily births, air pollution concentrations, and aeroallergen concentration according to the considered potential confounders. PTBs were less frequent during weekends and holidays than during the other days. PM2.5 and NO_2 were inversely associated with temperature, due to their secondary component nature, whose synthesis is favored by the sun. O_3 and aeroallergens were positively associated with temperature, as expected. Airborne pollution was lower during rainy days, as expected.

3.4.1. Preterm

The associations between exposure variables and PTB through the Poisson model adjusted for all potential confounders (day of the week, holidays, summer population decrease, daily temperature, daily relative humidity, and cumulative daily precipitations), are summarized in Figure 5.

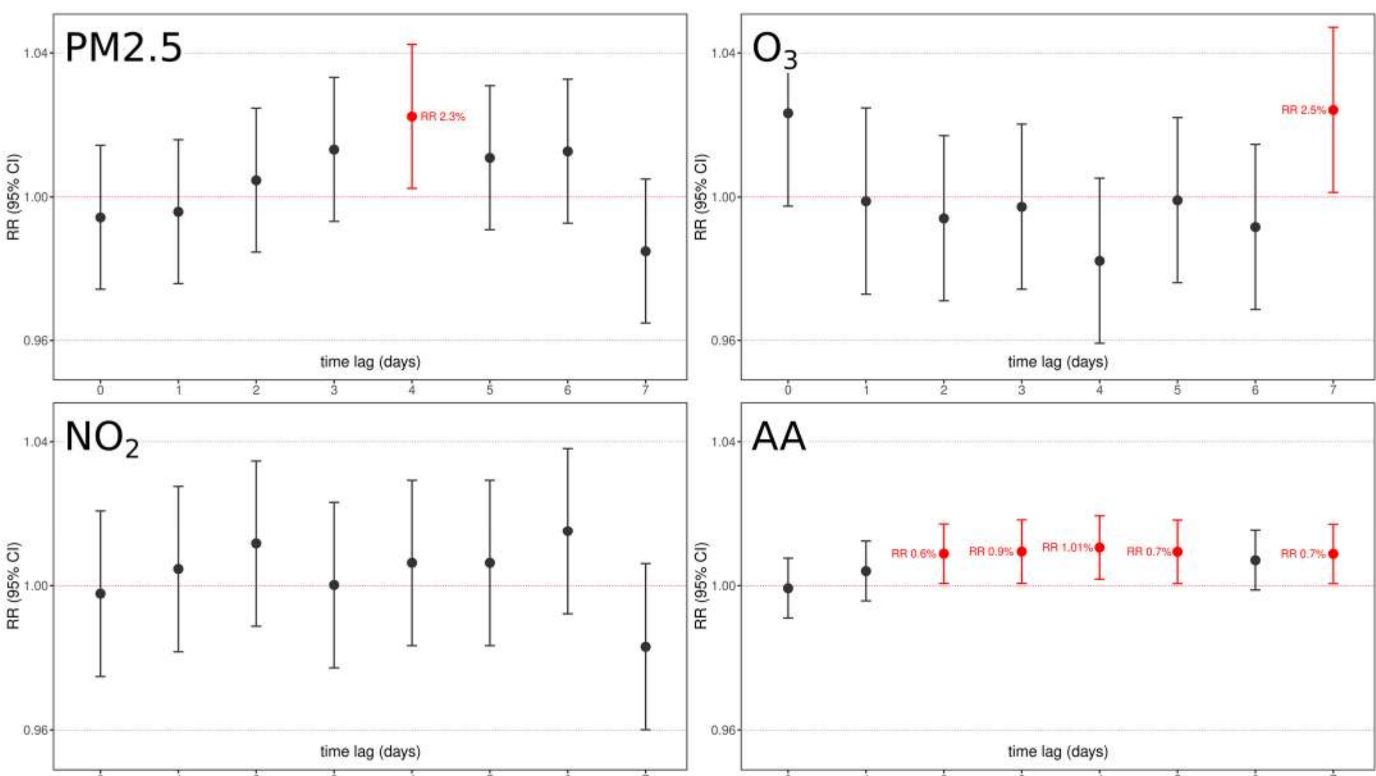


Figure 5. Graphical representation of chemical pollutants under analysis (PM2.5, NO_2, O_3) and aeroallergen (AA) concentrations relative risk and confidence interval association with PTB as per 7-days time lags. Red boxes underline significant factors with their relative lag RR. All models were adjusted for day of the week, holidays, summer population decrease, and yearly medium/long-term trend function.

With the exception of NO_2, all pollutants were significantly associated with PTB at some time lag:

(1) An increase of 10 μg/m^3 of PM2.5 is associated with a significant 1.023 (RR 95% C.I. 1.003–1.043, $p < 0.05$) increased risk of PTB after 4 days.

(2) An increase of 10 μg/m^3 of O_3 is associated with a significant 1.025 (RR 95% C.I. 1.001–1.048, $p < 0.05$) increased risk of PTB after 7 days.

(3) An increase of 10 grains/m^3 of aeroallergens is associated with an increased risk of PTB from 2 to 7 days after exposure:

- 1.006 (RR 95% C.I. 1.002–1.012, $p < 0.05$) after 2 days;
- 1.009 (RR 95% C.I. 1.002–1.014, $p < 0.01$) after 3 days;
- 1.01 (RR 95% C.I. 1.004–1.016, $p < 0.01$) after 4 days;
- 1.007 (RR 95% C.I. 1.001–1.013, $p < 0.04$) after 5 days;
- 1.007 (RR 95% C.I. 1.0008–1.03, $p < 0.05$) after 7 days.

The characteristic lag distribution effects that are detectable for aeroallergens, i.e., either the significant effect detectable at lag7 or the not significant detectable at lag6, may be related to the specific cohort under analysis. The results obtained suggest that aeroallergens are likely to play a role in eliciting PTB, and this role is probably played over different days, probably underlying a stronger proinflammatory effect on some specific target, according to their characteristics. Further analysis with larger, more different cohorts, and more years of data collection are needed in order to better elucidate the real significant nature and lags of aeroallergens on PTB.

3.4.2. Term

The correlation analysis between at-term delivery and chemical pollutants and aeroallergens was performed in the exact same way as for PTB and the results are summarized in Figure 6.

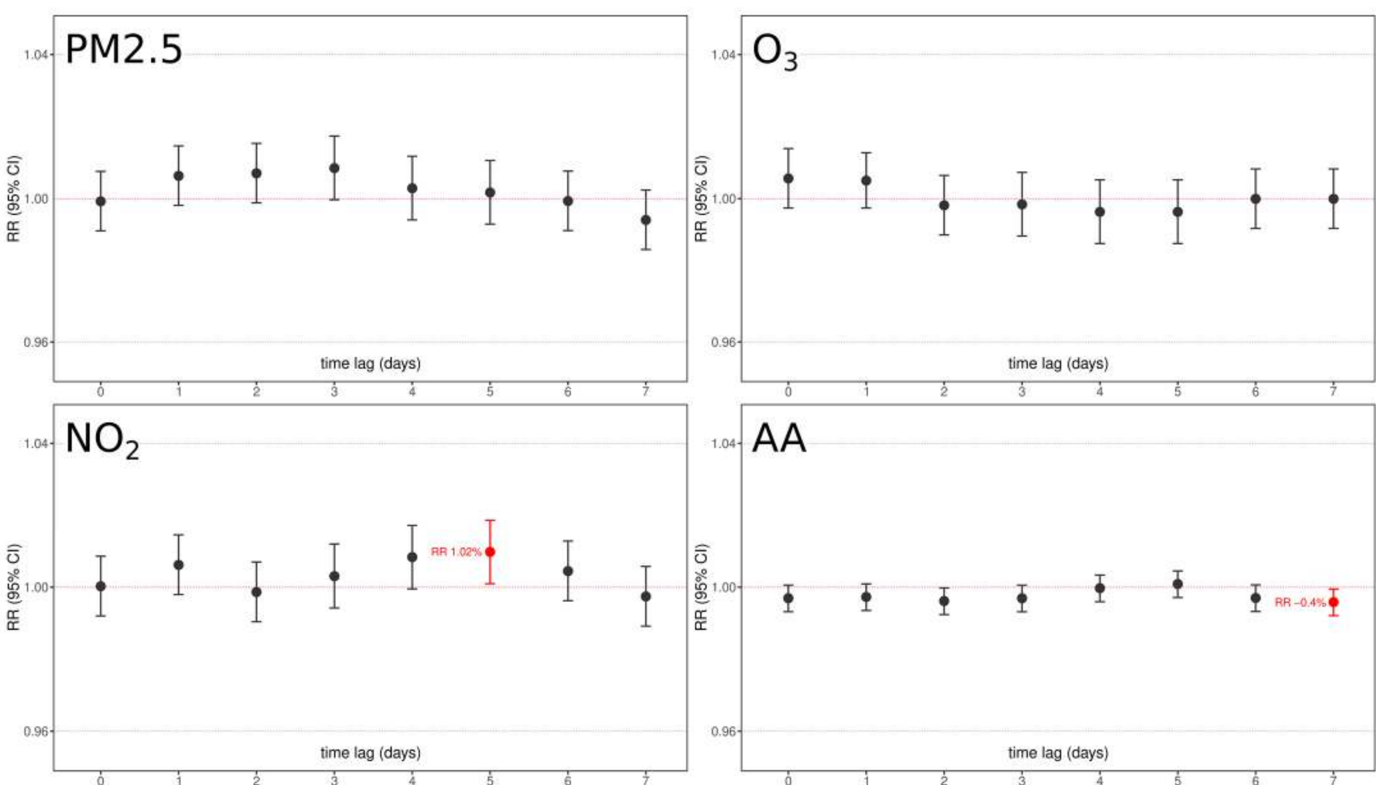


Figure 6. Graphical representation of chemical pollutants under analysis (PM2.5, NO_2, O_3) and aeroallergen (AA) concentrations relative risk and confidence interval association with full-term births as per 7-days time lags. Red boxes underline significant factors with their relative lag RR. All models were adjusted for day of the week, holidays, summer population decrease, and yearly medium/long-term trend function.

Essentially, the results showed that no pollutant was significantly associated with an increased risk of term delivery, with only two barely significant associations: an increase of 10 µg/m^3 of NO_2 resulted in an increased risk of 1.01 (RR 95% C.I. 1.002–1.021, $p < 0.05$) after 5 days from exposure, and an increase of 10 grains/m^3 of aeroallergens resulted in a very fleble protective effect after 7 days from exposure (RR 0.996, RR 95% C.I. 0.999–0.993, $p < 0.05$). No other air pollutants were associated at any time lag.

4. Discussion

Air pollution impact on PTB is well known and has been variously assessed, but specific time-frame evaluation is still lacking. The peculiar and complex nature of PTB, where a long-term process (fetal development) is paired with an acute one (preterm delivery) needs tailored models that are focused on specific outcomes and confounders based on the period considered. While PTB has normally been assessed with models historically developed for chronic diseases, no comprehensive model has been applied for correctly assessing the short-term effects of air pollutants in a narrow temporal window before delivery. The complex nature of PTB may lead to confusion in the interpretations of the different risk effects identified through epidemiological studies, as reported. In fact, while the assessment of long-term effects is of certain importance in order to elucidate the influence of prolonged exposures in fetal development and subsequent predisposition to PTB, the assessment of

short-term effects associated with actual delivery may pave the way to easier and more effective interventions immediately applicable to prevent PTB. Long-term changes would require radical and prolonged lifestyle changes, while short-term exposure preventive measures could be performed more easily and immediately. Moreover, air pollution levels are monitored widely and continuously in developed countries and the data are easily retrievable, providing important information that may help in identifying acute exposure thresholds [29–31], especially for at-risk pregnancies.

Regarding the regression model selection, as presented above, our analysis was specifically aimed at investigating the precise acute effects of air pollution on potentially eliciting preterm vs. full-term birth. This acute focus is the core behind model selection. The graphical analysis of our data and the time distribution of daily births showed that the variable is best approximated by Poisson distribution, with a higher λ in the case of at-term births. Thus, a generalized linear Poisson model was considered to be the best model for the analysis, according to previous literature reports that had focused on the same acute effects under investigation [24–26]. On the one hand, the results of this specific analysis approach show that a specific role of some specific proinflammatory air pollutants in eliciting preterm births is detectable. On the other hand, this acute inductive effect of air pollution is undetectable in the case of full-term births, in which we identify different effects of the various pollutants under analysis. According to the aim of our analysis, this difference may reflect some addictive effects of pregnancy-related pathological conditions that are brought to delivery by air pollution-related stress in the case of preterm births, while the same effects are likely compensated for in physiological full-term births.

Our results show that air pollutants could play significant roles on PTB in the days immediately preceding delivery. In the 7 days preceding PTB delivery, an increased exposure to airborne chemicals and aeroallergens is detectable. An exposure to PM2.5, 5 days preceding PTB, increases its risk by 2.3% for every 10 $\mu g/m^3$ PM2.5. In a similar manner, 7 days before PTB, an exposure to O_3 increases its risk by 2.5% every 10 $\mu g/m^3$. Fascinatingly, an increased exposure of 10 grains/m^3 of aeroallergens showed an almost continuous (from day 2 to day 7) increase of risk of ~1% on PTB, further testifying their strong inflammatory nature that has been revealed in asthma and other diseases. Interestingly, none of these effects were detectable on physiological term delivery. Contrarily, aeroallergens resulted in a weak but protective effect, and NO_2 that showed no effect on PTB resulted in an increased ~1% risk of delivery every 10 $\mu g/m^3$ at 5 days time lag. This fascinating difference between PTB and term births may testify the different trigger effects that these pollutants play in at-risk vs. physiological pregnancies, where the latter might have enough resources to cope with this proinflammatory noxa. This difference may be related to molecular differences in inflammatory pathways elicited by these pollutants on different physical states, whose deeper exploration is mandatory to better clarify PTB pathophysiology. Our analysis is corrected for all possible environmental, seasonal, and social confounding factors that frequently bias these analyses, as testified by the significant differences in PTB vs. term incidence once grouped by day of the week, holidays, population decrease, and temperature.

Due to the huge number of individuals exposed, air pollution may explain an important portion of cases, and allow preventive strategies. For example, consider an annual mean concentration of PM2.5 of ~25 $\mu g/m^3$ in the Turin area [32] versus 10 $\mu g/m^3$ in nearby Switzerland, this single difference can lead to an increased ~5% risk of PTB only attributable to PM2.5 levels. Considering ~1050 PTB cases per year, this represents ~50 PTB cases only attributable to PM2.5 exposures during days immediately preceding delivery.

The focus on a narrow time frame before delivery is of core importance for both the abrupt nature of preterm delivery and applicable potential preventive measures. For example, a strong recommendation for at-risk women to avoid highly polluted areas in the last trimester of gestation may easily save several newborn lives and may avoid PTB-related complications. Moreover, as previously explained, Il-1, IL-6, and TNF-a, among other inflammatory cytokines, exert a central role in PTB and are increasingly produced as a consequence of air pollution exposure. In this regard, pollutant-induced acute inflammation

may offer the molecular explanation for this relationship, and inflammation can also be effectively and safely monitored through these biomarkers.

Study limitations: The limitations of the current studies are mainly related to the single-center nature of the study and the highly polluted area considered. In fact, the area under analysis is a particularly polluted area, and thus, a good candidate in order to elucidate the possible role of air pollution exposure on PTB, but RR and exposure effects may vary in other areas. The single-center nature of the study is also associated with the absence of spatial resolution, as only one monitor served for all patients under analysis. Moreover, air pollutants have been recorded in a single site relatively close to the hospital, which is the only one available with such precise measurement. We considered such measurements as indicators for the overall area under analysis, even if a measurement of each woman's living place would lead to a much more accurate analysis. Another main limitation of the current study is the absence of laboratory testing for inflammation markers in the cohort under analysis, as this was intended as an exploratory analysis in order to confirm/exclude possible roles of such molecules on PTB. Inflammation marker testing is currently being performed on a subcohort of the presented one, as a continuation of the current work. The multiple comparison of lag times could result in a type I error; to be consistent with the same model presented in our previous work [24] we did not correct for Bonferroni, in order to avoid exclusion of interesting trends. Nonetheless, the Bonferroni threshold over the 7-day lag comparison would set a 0.007 p-value threshold.

5. Conclusions

Our findings suggest that a specific role of some specific proinflammatory air pollutants on delivery induction is detectable in PTB, while lower and different direction effects are detectable in full-term births. This fascinating difference advocates for extensive studies on air pollutant proinflammatory pathophysiology mechanisms and their specific relationships with PTB. In this view, further identification of physiological differences underlying such relationships may help to identify potential pharmacological targets that would further help to prevent PTB. Moreover, to the best of our knowledge, our analysis is the first analysis that has specifically focused on acute eliciting effects of air pollutants on PTB vs. full-term births. We were able to detect pollutant effects that have not been previously reported, probably due to their specific proinflammatory short-term effects, which were undetectable in the case of long-term studies that have represented the main approach to date. Thus, in particular, it is important to broaden our time-frame reference when investigating such complex matters, in which an overlap of long- and short-term effects plays a role on the same outcome, elaborating specific models that are able to consider precise time frames to differently investigate chronic vs. acute effects. Finally, the detection of precise short-term acute effects of air pollutants on PTB advocates for the implementation of preventive regulations regarding air quality, especially in urban centers and for at-risk pregnancies. Considering the huge number of pregnant women exposed, it is of central importance to further investigate this relationship and possible preventive thresholds, in order to rapidly and easily apply air pollution-related recommendations for the last months of pregnancies and to further reinforce the beneficial effects of policies aimed at reducing air pollution. These anti-pollution policies may prevent a huge number of PTBs, alongside other countless beneficial effects on other well known and reported aspects of human health.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijerph20021610/s1, Figure S1: The figure geographically illustrates Turin position inside Piedmont region, and Piedmont position in Italy.

Author Contributions: Conceptualization, E.C., V.B., L.P. and R.B.; methodology, E.C., V.B., L.P. and R.B.; software, E.C.; formal analysis, E.C. and V.B.; data curation, F.C., C.C. and C.S.; writing—original draft preparation, E.C., V.B., L.P. and R.B.; writing—review and editing, E.C., V.B., F.C., C.P., C.C., C.S., L.P. and R.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The original database was built for a previous research objective and was approved by the competent Ethics Committee of the hospital "City of Health and Science of Turin" (Prot. no. 0000064, CS21071 of 2 January 2019), in compliance with the Declaration of Helsinki. This work, which was derived from the original database, was purely epidemiological in nature and did not involve, in any way, the subjects included in the database. Furthermore, the data were provided by the Hospital in anonymized form, thus, not requiring further opinions from the Ethics Committee.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the raw data collected are available, following a justified request to the corresponding author.

Acknowledgments: Thanks are due to ARPA Piedmont, which made the air pollution data available.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Blencowe, H.; Cousens, S.; Chou, D.; Oestergaard, M.; Say, L.; Moller, A.-B.; Kinney, M.; Lawn, J.; the Born Too Soon Preterm Birth Action Group. Born Too Soon: The global epidemiology of 15 million preterm births. *Reprod. Health* **2013**, *10* (Suppl. S1), S2. [CrossRef]
2. Martin, J.A.; Osterman, M.J.K. *Describing the Increase in Preterm Births in the United States, 2014–2016*; NCHS Data Brief No. 312; U.S. Department Of Health & Human Services: Washington, DC, USA, 2018; pp. 1–8.
3. Macdorman, M.F.; Mathews, T.J. *Recent Trends in Infant Mortality in the United States*; NCHS Data Brief No. 9; U.S. Department of Health & Human Services: Washington, DC, USA, 2008; pp. 1–8.
4. Raju, T.N.; Pemberton, V.L.; Saigal, S.; Blaisdell, C.J.; Moxey-Mims, M.; Buist, S. Long-Term Healthcare Outcomes of Preterm Birth: An Executive Summary of a Conference Sponsored by the National Institutes of Health. *J. Pediatr.* **2016**, *181*, 309–318.e1. [CrossRef]
5. Cresi, F.; Cocchi, E.; Maggiora, E.; Pirra, A.; Logrippo, F.; Ariotti, M.C.; Peila, C.; Bertino, E.; Coscia, A. Pre-discharge Cardiorespiratory Monitoring in Preterm Infants. the CORE Study. *Front. Pediatr.* **2020**, *8*, 234. [CrossRef] [PubMed]
6. Squillacioti, G.; Bellisario, V.; Levra, S.; Piccioni, P.; Bono, R. Greenness Availability and Respiratory Health in a Population of Urbanised Children in North-Western Italy. *Int. J. Environ. Res. Public Health* **2019**, *17*, 108. [CrossRef] [PubMed]
7. Hall, E.; Greenberg, J. Estimating community-level costs of preterm birth. *Public Health* **2016**, *141*, 222–228. [CrossRef]
8. Goldenberg, R.L.; Culhane, J.F.; Iams, J.D.; Romero, R. Epidemiology and causes of preterm birth. *Lancet* **2008**, *371*, 75–84. [CrossRef] [PubMed]
9. Alman, B.L.; Stingone, J.A.; Yazdy, M.; Botto, L.D.; Desrosiers, T.A.; Pruitt, S.; Herring, A.H.; Langlois, P.H.; Nembhard, W.N.; Shaw, G.M.; et al. Associations between PM2.5 and risk of preterm birth among liveborn infants. *Ann. Epidemiol.* **2019**, *39*, 46–53.e2. [CrossRef]
10. Klepac, P.; Locatelli, I.; Korošec, S.; Künzli, N.; Kukec, A. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environ. Res.* **2018**, *167*, 144–159. [CrossRef]
11. Mendola, P.; Nobles, C.; Williams, A.; Sherman, S.; Kanner, J.; Seeni, I.; Grantz, K. Air Pollution and Preterm Birth: Do Air Pollution Changes over Time Influence Risk in Consecutive Pregnancies among Low-Risk Women? *Int. J. Environ. Res. Public Health* **2019**, *16*, 3365. [CrossRef]
12. Rappazzo, K.M.; Daniels, J.L.; Messer, L.C.; Poole, C.; Lobdell, D. Exposure to Fine Particulate Matter during Pregnancy and Risk of Preterm Birth among Women in New Jersey, Ohio, and Pennsylvania, 2000–2005. *Environ. Health Perspect.* **2014**, *122*, 992–997. [CrossRef]
13. Liu, Y.; Xu, J.; Chen, D.; Sun, P.; Ma, X. The association between air pollution and preterm birth and low birth weight in Guangdong, China. *BMC Public Health* **2019**, *19*, 3. [CrossRef] [PubMed]
14. Shi, L.; Shi, L.; Wu, X.; Wu, X.; Yazdi, M.D.; Yazdi, M.D.; Braun, D.; Braun, D.; Abu Awad, Y.; Abu Awad, Y.; et al. Long-term effects of PM2·5 on neurological disorders in the American Medicare population: A longitudinal cohort study. *Lancet Planet. Health* **2020**, *4*, e557–e565. [CrossRef]
15. Taguchi, A.; Yamashita, A.; Kawana, K.; Nagamatsu, T.; Furuya, H.; Inoue, E.; Osuga, Y.; Fujii, T. Recent Progress in Therapeutics for Inflammation-Associated Preterm Birth: A Review. *Reprod. Sci.* **2015**, *24*, 7–18. [CrossRef]
16. Faro, J.; Romero, R.; Schwenkel, G.; Garcia-Flores, V.; Arenas-Hernandez, M.; Leng, Y.; Xu, Y.; Miller, D.; Hassan, S.S.; Gomez-Lopez, N. Intra-amniotic inflammation induces preterm birth by activating the NLRP3 inflammasome. *Biol. Reprod.* **2018**, *100*, 1290–1305. [CrossRef]
17. Rider, C.F.; Carlsten, C. Air pollution and DNA methylation: Effects of exposure in humans. *Clin. Epigenet.* **2019**, *11*, 131. [CrossRef]

18. Pope, C.A.; Bhatnagar, A.; McCracken, J.P.; Abplanalp, W.; Conklin, D.J.; O'Toole, T. Exposure to Fine Particulate Air Pollution Is Associated With Endothelial Injury and Systemic Inflammation. *Circ. Res.* **2016**, *119*, 1204–1214. [CrossRef]
19. Yue, W.; Tong, L.; Liu, X.; Weng, X.; Chen, X.; Wang, D.; Dudley, S.C.; Weir, E.K.; Ding, W.; Lu, Z.; et al. Short term PM2.5 exposure caused a robust lung inflammation, vascular remodeling, and exacerbated transition from left ventricular failure to right ventricular hypertrophy. *Redox Biol.* **2019**, *22*, 101161. [CrossRef]
20. Boyle, A.K.; Rinaldi, S.F.; Rossi, A.G.; Saunders, P.T.K.; Norman, J.E. Repurposing simvastatin as a therapy for preterm labor: Evidence from preclinical models. *FASEB J.* **2018**, *33*, 2743–2758. [CrossRef] [PubMed]
21. Belousova, V.S.; Svitich, O.A.; Timokhina, E.V.; Strizhakov, A.N.; Bogomazova, I.M. Polymorphism of the IL-1β, TNF, IL-1RA and IL-4 Cytokine Genes Significantly Increases the Risk of Preterm Birth. *Biochemistry* **2019**, *84*, 1040–1046. [CrossRef] [PubMed]
22. Aung, M.T.; Yu, Y.; Ferguson, K.K.; Cantonwine, D.E.; Zeng, L.; McElrath, T.F.; Pennathur, S.; Mukherjee, B.; Meeker, J.D. Prediction and associations of preterm birth and its subtypes with eicosanoid enzymatic pathways and inflammatory markers. *Sci. Rep.* **2019**, *9*, 17049. [CrossRef] [PubMed]
23. Munoz-Perez, V.M.; Ortiz, M.I.; Carino-Cortes, R.; Fernandez-Martinez, E.; Rocha-Zavaleta, L.; Bautista-Avila, M. Preterm Birth, Inflammation and Infection: New Alternative Strategies for their Prevention. *Curr. Pharm. Biotechnol.* **2019**, *20*, 354–365. [CrossRef]
24. Bono, R.; Romanazzi, V.; Bellisario, V.; Tassinari, R.; Trucco, G.; Urbino, A.; Cassardo, C.; Siniscalco, C.; Marchetti, P.; Marcon, A. Air pollution, aeroallergens and admissions to pediatric emergency room for respiratory reasons in Turin, northwestern Italy. *BMC Public Health* **2016**, *16*, 722. [CrossRef] [PubMed]
25. Samoli, E.; Peng, R.; Ramsay, T.; Pipikou, M.; Touloumi, G.; Dominici, F.; Burnett, R.; Cohen, A.; Krewski, D.; Samet, J.; et al. Acute Effects of Ambient Particulate Matter on Mortality in Europe and North America: Results from the APHENA Study. *Environ. Health Perspect.* **2008**, *116*, 1480–1486. [CrossRef]
26. Peng, R.D.; Dominici, F.; Louis, T.A. Model choice in time series studies of air pollution and mortality. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2006**, *169*, 179–203. [CrossRef]
27. Stafoggia, M.; Colais, P.; Serinelli, M.; Gruppo collaborativo EpiAir. Methods of statistical analysis to evaluate the short term effects of air pollution in the EpiAir Project. *Epidemiol. Prev.* **2009**, *33*, 53–63.
28. Marchetti, P.; Pesce, G.; Villani, S.; Antonicelli, L.; Ariano, R.; Attena, F.; Bono, R.; Bellisario, V.; Fois, A.; Gibelli, N.; et al. Pollen concentrations and prevalence of asthma and allergic rhinitis in Italy: Evidence from the GEIRD study. *Sci. Total. Environ.* **2017**, *584–585*, 1093–1099. [CrossRef] [PubMed]
29. Squillacioti, G.; Bellisario, V.; Grignani, E.; Mengozzi, G.; Bardaglio, G.; Dalmasso, P.; Bono, R. The Asti Study: The Induction of Oxidative Stress in A Population of Children According to Their Body Composition and Passive Tobacco Smoking Exposure. *Int. J. Environ. Res. Public Health* **2019**, *16*, 490. [CrossRef]
30. Bellisario, V.; Comoretto, R.I.; Berchialla, P.; Koumantakis, E.; Squillacioti, G.; Borraccino, A.; Bono, R.; Lemma, P.; Charrier, L.; Dalmasso, P. The Association between Greenness and Urbanization Level with Weight Status among Adolescents: New Evidence from the HBSC 2018 Italian Survey. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5897. [CrossRef]
31. Squillacioti, G.; Carsin, A.-E.; Bellisario, V.; Bono, R.; Garcia-Aymerich, J. Multisite greenness exposure and oxidative stress in children. The potential mediating role of physical activity. *Environ. Res.* **2022**, *209*, 112857. [CrossRef]
32. Sustainability and Recognition—Piedmont Office Realty Trust. Available online: https://www.piedmontreit.com/about/sustainability-and-recognition/ (accessed on 7 April 2021).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

scientific reports

OPEN

The role of phase I, phase II, and DNA-repair gene polymorphisms in the damage induced by formaldehyde in pathologists

Federica Ghelli[1], Enrico Cocchi[1], Martina Buglisi[1], Giulia Squillacioti[1], Valeria Bellisario[1], Roberto Bono[1,3]✉ & Alfredo Santovito[2,3]

Formaldehyde (FA) is a human carcinogen used as formalin in hospital laboratories. We evaluated its association with human chromosomal aberrations (CAs) and the risk/protective role played by several genetic polymorphisms in this relationship, on a cohort of 57 exposed pathologists *vs* 48 controls. All subjects were assessed for CAs on peripheral blood lymphocytes and genotyped for the most common cancer-associated gene polymorphisms which could be related with the genotoxic outcome: *CYP1A1 exon 7 (A>G)*, *CYP1A1*2A (T>C)*, *CYP2C19*2 (G>A)*, *GSTT1* (Positive/Null), *GSTM1* (Positive/null), *GSTP1 (A>G)*, *XRCC1 (G399A)*, *XRCC1 (C194T)*, *XRCC1 (A280G)*, *XPD (A751C)*, *XPC exon 15 (A939C)*, *XPC exon 9 (C499T)*, *TNFα – 308 (G>A)*, *IL10 – 1082 (G>A)*, *IL10 – 819 (C>T)* and *IL6 – 174 (G>C)*. Air-FA concentration was assessed through personal samplers. The comparison between pathologists and controls showed a significantly higher CAs frequency in pathologists. Significant positive correlations were found between CAs frequency and air-FA concentration while significant associations were found between variation in CAs frequency and the mutated allele for *CYP1A1 exon 7 (A>G)*, *CYP2C19*2 (G>A)*, *GSTT1*-positive, *GSTM1*-positive and *XRCC1 (G399A)*. Our study confirms the role of FA as genotoxicity inductor, even in workers chronically exposed to low air-FA levels and reveals the role played by some genetic polymorphisms in this association, highlighting the importance of individual susceptibility biomarkers assessment in occupational health studies.

Formaldehyde (FA) is a compound produced worldwide and employed in an extremely wide variety of industrial and medical processes[1], resulting in a widespread exposure in both environmental and occupational contexts[2]. As it is well-known, FA is responsible of several biological effects, even at lower concentrations than those recommended by the American Conference of Governmental Industrial Hygienists (ACGIH)[3,4]. Workers exposed to FA are at increased risk of cancer, especially nasopharyngeal cancer and myeloid leukaemia[5]. Due to these effects, FA is classified as a group I human carcinogen by the International Agency for Research on Cancer (IARC) since 2006[5–7]; nevertheless, considerable discrepancies remain among guidelines suggested for occupational exposure to FA. The ACGIH recommended a Threshold Limit Value-Ceiling (TLV-C) of 0.3 ppm until 2016; the value was then dropped to a Time Weighted Average (TLV-TWA) of 0.1 ppm (0.120 mg/m^3) and a Short Term Exposure Limit (TLV-STEL) of 0.3 ppm (0.370 mg/m^3). Conversely, the European Scientific Committee on Occupational Exposure Limits recently suggested a FA-related TWA of 0.3 ppm, but a STEL of 0.6 ppm (0.740 mg/m^3)[8].

Formalin is an aqueous solution usually containing 37–40% by weight of dissolved FA: its easy preparation and low cost make this compound the main cytological fixative in pathology laboratories worldwide[9,10]. Despite these advantages, the health and safety risks associated with formalin use are currently a matter of concern and FA toxicity is nowadays the main issue for its abolition in pathology laboratories[3,10,11]. Moreover, chronic exposures to FA, such as those present in workplaces, are suspected to be related to genotoxic effects[9].

[1]Department of Public Health and Pediatrics, University of Turin, Via Santena 5 bis, 10126 Turin, Italy. [2]Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin, Italy. [3]These authors jointly supervised this work: Roberto Bono and Alfredo Santovito. ✉email: roberto.bono@unito.it

The FA genotoxic effect in occupationally exposed workers is still debated. Cytogenetic outcomes, such as increased chromosomal aberrations (CAs) and micronucleated cells (MNc), were reported in some biomonitoring studies[12,13] on chronic exposures, while this evidence was lacking in other published reports[14,15].

However, the genomic damage level due to occupational exposure to xenobiotics depends also on the individual susceptibility. From the genetic point of view, this is due to polymorphisms in a battery of genes, mainly involved in metabolic and DNA-repair pathways[16].

Phase I metabolic enzymes mostly consist of the cytochrome P450 (CYP) superfamily of microsomal enzymes[17] catalysing oxidative reactions[18], while phase II enzymes, such as glutathione S-transferases (GSTs), role is to increase the hydrophilicity of the xenobiotic compounds through conjugation reactions[17]. FA is quickly detoxified in the nasal tissues by oxidative reactions catalysed by glutathione-dependent and independent dehydrogenases, primarily the alcohol dehydrogenase 5[19].

In order to safeguard the genome's integrity and to prevent the potentially mutagenic consequences of DNA modifications, the cells evolved several mechanisms of DNA repair, according to the type of damage. The Base Excision Repair (BER) and the Nucleotide Excision Repair (NER) correct DNA small base changes (oxidation or alkylation) and bulky adducts, pyrimidine dimers and inter-strand cross-links, respectively[20]. These DNA-repair genes, which are involved in the protection mechanism against cancer development, are polymorphic[20]. Several evidences reported that defects in these DNA repair mechanisms could reduce FA tolerance at cellular level[21].

Finally, several lines of evidence recently showed the FA role as oxidative stress inductor[3,22]. This imbalance between the production of Reactive Oxygen Species (ROS) and the capacity of the antioxidant system to counteract them, leads to biomolecular damages triggering inflammation, testified by massive proinflammatory cytokine release[23], which is in turn related to carcinogenesis. Some cytokine gene polymorphisms, moreover, were found to modulate the amount of genomic damage associated with inflammatory and cancer diseases. As example, *TNF-α*, *IL-2*, *IL-6,* and *TGF-β1* polymorphisms have been showed to influence CAs level in cultured human peripheral blood lymphocytes (PBL)[24–26]. Despite of all these lines of evidence, the role of cytokine gene polymorphisms in modulating the FA exposure associated damage has not been completely clarified yet[26].

In order to better elucidate the chronic FA genotoxic effect, we evaluated CAs frequency in PBL of pathologists chronically exposed to low air-FA concentration. This allows the detection of cells carrying unstable aberrations (*i.e.* chromosome and chromatid breaks, fragments) leading, in turn, to cell death during proliferation[27]. An increased CAs frequency in PBL is, thus, a powerful predictor of cancer risk significantly associated with the early events of carcinogenesis, as confirmed by previous studies in literature[28]. In order to evaluate the individual susceptibility role, sampled subjects were assessed for phase I, phase II, and DNA-repair gene polymorphisms, involved in the biotransformation, inactivation, and the DNA-repair processes, respectively. We analysed the most studied cancer-associated gene polymorphisms[29,30], namely Cytochrome P450 1A1 (*CYP1A1*) exon 7 (A>G) *CYP1A1 2A* (T>C), *CYP2C19*2* (G>A), *GSTT1*, *GSTM1*, *GSTP*, X-ray repair cross-complementing group 1 (*XRCC1*) 399 (G>A), 194 (C>T), 280 (A>G), Xeroderma pigmentosum complementation group C (*XPC*) exon 15 (A>C), XPC exon 9 (C>T) and Xeroderma pigmentosum complementation group D (*XPD*) (A>C). Finally, since cytokines play a fundamental role in the inflammatory process leading to genomic damage[26], we assessed polymorphisms in *TNF-α* (− 308, G>A), *IL-10* − 1082 (G>A), *IL-10* (− 819, C>T), *IL-6* (− 174, G>C) as well.

The aim of the present study is thus to evaluate the role of chronic occupational FA exposure risk levels and the role of some genetic polymorphisms as possible modulators of genotoxic effects, in workers chronically exposed to low air-FA concentrations.

Results

The epidemiologic sample includes 57 pathologists and 48 controls. In Table 1 are reported the demographic characteristics of the study population and the measured air-FA concentration on the sampling day. As expected, pathologists turned out to be exposed to an air-FA concentration significantly higher than controls ($p < 0.001$). No significant differences were found, instead, between the two groups concerning confounding factors such as sex, age, smoking habits and years of employment.

In Table 2 the level of genotoxic damage in the two sample groups is reported. We found three types of aberrations: chromatid break, chromosome break and acentric fragment. As can be seen, when compared to the control group, pathologists showed higher CAs and Ab.C frequencies ($p < 0.001$). The difference in CAs frequency between exposed and controls subjects is shown in Fig. 1.

The analyses of correlations performed on the whole sample showed a significant positive correlation between age and years of employment ($r = 0.83$, $p < 0.001$), CAs and Ab.C frequencies ($r = 1.00$, $p < 0.001$), CAs frequency and air-FA concentration ($r = 0.33$, $p < 0.001$) and, lastly, Ab.C frequency and air-FA concentration ($r = 0.33$, $p < 0.001$).

Multiple linear regressions were carried out to investigate the influence of the genetic profile in CAs frequency. The model (Model M0) includes all genetic polymorphisms (wt vs carriers of at least one mutated allele) and confounding factors. There was a significant relationship between CAs frequency and exposure to air-FA ($\beta = 1.027$; $p < 0.001$), *CYP1A1 exon 7 (A>G)* ($\beta = 0.353$; $p = 0.019$), *CYP2C19*2 (G>A)* ($\beta = 0.504$; $p = 0.007$), *GSTT1*-positive ($\beta = -0.447$; $p = 0.004$), *GSTM1*-positive ($\beta = -0.533$; $p = 0.001$) and *XRCC1 (399, G>A)* ($\beta = -0.331$; $p = 0.044$) genotypes. A tendency in increasing CAs frequency, albeit not significant, was found for *IL-10 1082 (G>A)* genotype ($\beta = 0.292$; $p = 0.054$). Figure 2 shows the Relative Risk (RR) of developing CAs according to the presence of mutated allelic variants of the gene considered.

Discussion

Despite the growing awareness regarding the harmful effects of air-FA exposure, FA is currently employed in hospital pathology laboratories raising concerns about pathologists safety[31,32].

	Pathologists (n = 57)	Controls (n = 48)
Sex (n)		
Males	29	25
Females	28	23
Age (years)		
Mean ± S.D.	42.632 ± 8.778	40.208 ± 9.711
Range	25–58	25–70
Smokers (n)	14	9
Number of cigarette/day (mean ± S.D.)	12.857 ± 10.939	13.667 ± 4.272
Years of smoking habit (mean ± S.D.)	19.643 ± 8.554	16.111 ± 7.688
Non-smokers (n)	43	39
Years of employment (years)		
Mean ± S.D.	11.246 ± 7.886	12.125 ± 7.482
Range	1–33	2–32
Air-FA (μg/m^3)		
Mean ± S.D.	64.197 ± 32.385*	19.065 ± 5.173

Table 1. Demographic characteristics and air-FA exposure level of subjects belonging to the studied groups. *n* number of analysed subjects, *S.D.* standard deviation. * P < 0.001, Kruskal–Wallis, Significantly higher with respect to Controls.

Groups	N	NSM	B′	B″	AF	Total CAs	Total Ab.C	CAs/NSM % mean ± S.D	Ab.C/NSM % mean ± S.D
Pathologists	57	11,400	125	29	28	182	181	0.016 ± 0.012*	0.015 ± 0.011*
Males	29	5800	77	15	19	111	110	0.019 ± 0.013 A	0.019 ± 0.013 A
Females	28	5600	48	14	9	71	71	0.013 ± 0.010	0.013 ± 0.010
Smokers	14	2800	38	5	3	46	46	0.016 ± 0.015	0.016 ± 0.015
Non-smokers	43	8600	87	24	25	136	135	0.016 ± 0.011	0.016 ± 0.011
Controls	48	9600	42	9	18	69	68	0.007 ± 0.006	0.007 ± 0.006
Males	25	5000	14	5	8	27	26	0.005 ± 0.006	0.005 ± 0.006
Females	23	4600	28	4	10	42	42	0.009 ± 0.005 B	0.009 ± 0.005 B
Smokers	9	1800	5	2	3	10	9	0.006 ± 0.004	0.005 ± 0.004
Non-smokers	39	7800	37	7	15	59	59	0.008 ± 0.006	0.008 ± 0.006

Table 2. Frequencies of chromosomal aberrations and cells with aberrations in metaphases of lymphocytes from studied subjects. *N* number of analysed subjects, *NSM* number of scored metaphases, *B′* chromatid breaks, *B″* chromosome breaks, *AF* acentric fragments, *CAs* chromosome aberrations, *Ab.C* cells with aberrations, *S.D.* standard deviation. *P < 0.001, Kruskal–Wallis, significantly higher with respect to Controls. A P = 0.046, Kruskal–Wallis, significantly higher with respect to Females. B P = 0.020, Kruskal–Wallis, significantly higher with respect to Males.

Many studies demonstrated the FA genotoxicity both in vitro and in vivo, considering various biological systems and endpoints[5]. Moreover, occupational and environmental exposures are often chronic and mixed, and the analysis of their outcomes should take into account also the eventuality of a cumulative genomic instability induced by chronic exposures[33].

As expected, we found significantly higher CAs and Ab.C frequencies in pathologists than in controls, in agreement with literature evidence, even though conflicting results can be found[12,34,35]. In this regard, Costa et al.[34] reported that even at an average FA concentration of 0.38 ppm (i.e., 0.47 mg/m^3) frequencies of cytogenetic parameters, such as CAs, were significantly higher in pathologists than in controls. In order to explain the mechanisms leading to FA-induced genotoxicity, several hypotheses have been proposed: DNA–protein cross-links, damage to proteins required for the mitotic process and reduced expression of paxillin, an essential component of the abscission machinery required to complete cytokinesis, which may lead, respectively, to DNA replication stress and DNA breaks, chromosome malsegregation during nuclear division and cytokinesis failure leading to micronucleus formation and aneuploidy[7]. Moreover, the inflammatory process, due to the activation of neutrophils and eosinophils and/or the altered redox balance in the bone-marrow could play a central role in DNA strand breaks induced by ROS[7].

The harmful effects of xenobiotics exposure is extremely shaped by individual susceptibility. The analysis of metabolic and DNA-repair gene polymorphisms in risk assessment of hazardous chemicals assume thus particular importance[5,26].

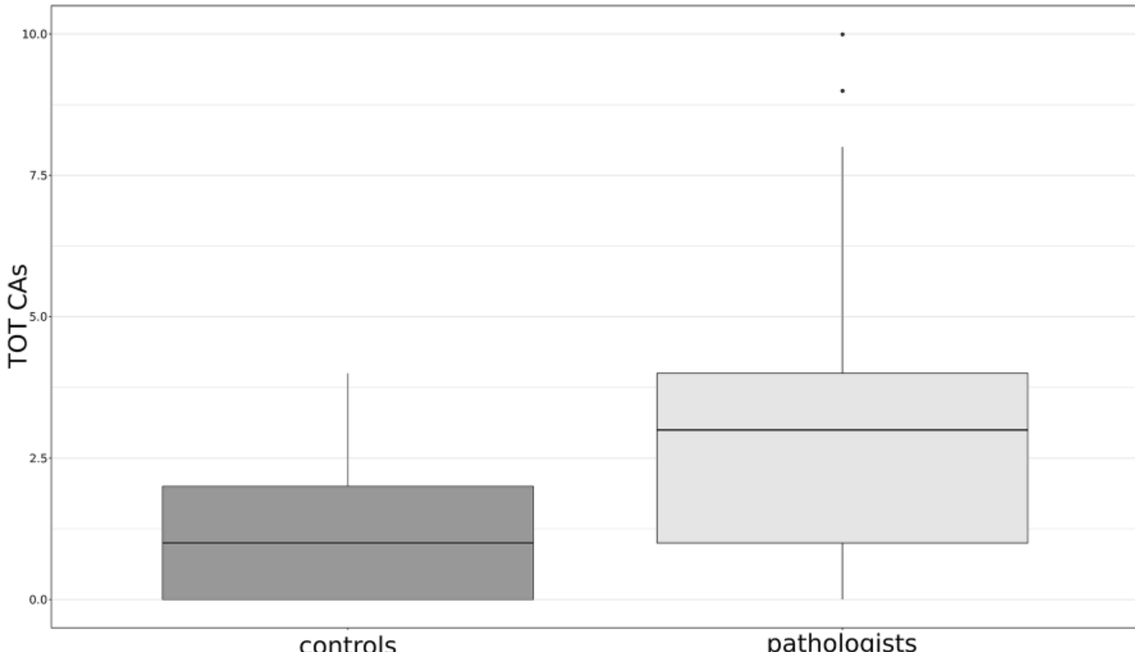


Figure 1. CAs frequency in exposed and control groups.

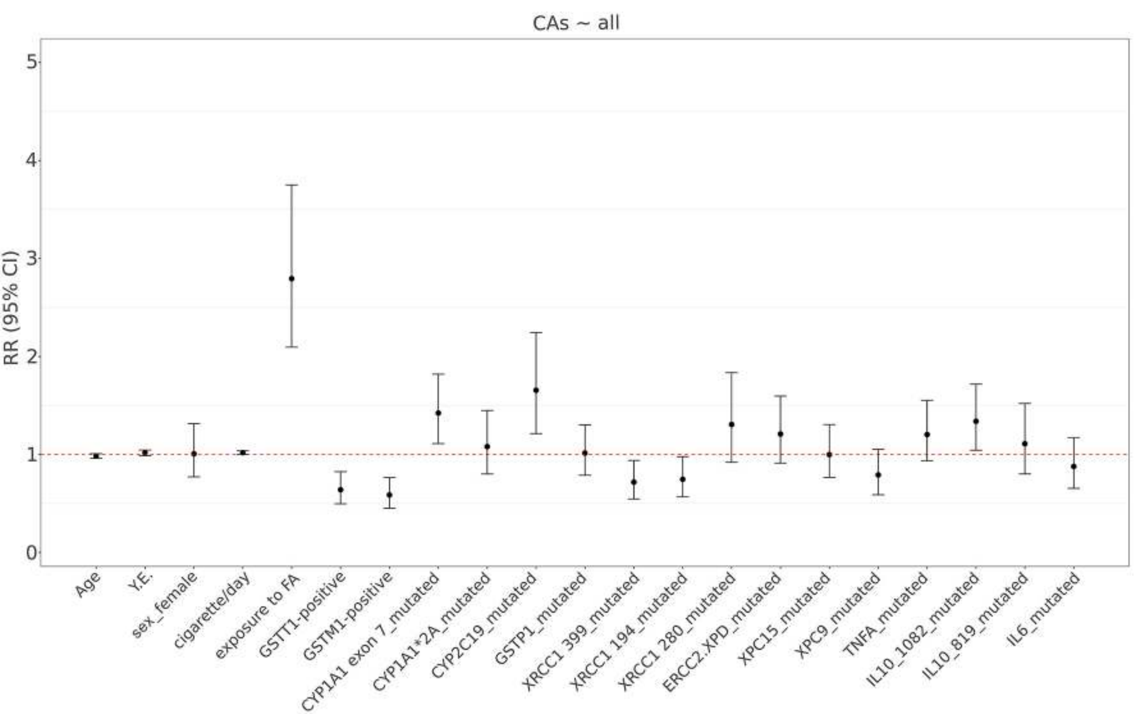


Figure 2. RR (95% CI) of developing CAs according to the various mutated allelic variants of genes considered in the present study, FA exposure and confounding factors as age, years of exposure, sex and cigarette/day (Model M0).

In this context, we focused on the role of several genetic polymorphisms in modulating CAs frequency in subjects occupationally exposed to air-FA compared to a control group.

We found a significant effect of some genetic polymorphisms in CAs frequency modulation, even among genes coding for enzymes not directly involved in FA metabolism.

Specifically, among genes coding for phase I metabolism enzymes, we found a significant increase in CAs frequency in carriers of *CYP1A1* exon 7 (A>G) and *CYP2C19*2* (G>A) polymorphisms. These genes are members of the cytochrome P450 superfamily of enzymes, mixed-function mono-oxygenases responsible for metabolizing,

mainly via oxidative reactions, several exogenous and endogenous compounds, including steroids, fatty acids, retinoid, drugs, vitamins, procarcinogens/promutagens, and environmental compounds[36,37].

The *CYP1A1* Ile462Val substitution in the heme-binding domain of exon 7, leads to a concurrent increase in the catalytic activity of the protein and was associated with lung cancer risk[38,39]. The Ile/Ile genotype was also found to be associated with an increase of aberrant cells, and to be a CAs predictor[40]. Contrary, other studies did not find any association[41].

Common variants of the *CYP2C19* gene are associated with impaired drug metabolism. *CYP2C19*2* results from a guanine (G) to adenine (A) transition at position 681 in exon 5, producing an aberrant splicing site and encoding enzymes with decreased activity[42]. This polymorphism was related to genotoxicity in a previous study of Santovito et al., where CYP2C19 A/A subjects turned out to show a frequency of sister chromatid exchanges (SCEs) significantly higher with respect to the CYP2C19 G/G homozygote genotypes[43].

In phase II enzymes, we found a significant CAs frequency decrease in *GSTT1*-positive and *GSTM1*-positive subjects. Accordingly, the higher frequency of genotoxic damage in carriers of the null-allele could be explained considering the role of these genes and their mutations on metabolism. The glutathione S-transferases represent an important group of enzymes, which detoxify both endogenous and exogenous compounds, included pharmaceuticals and environmental pollutants[44]. The *GSTM1* and *GSTT1* polymorphisms consist both in the deletion of a part of the gene, leading, in homozygous individuals, to a lack of the enzyme activity[45]. In literature, the *GSTM1*-null and *GSTT1*-null genotypes have been related to increased risk for several cancers, such as lung and colorectal cancer[45]. These are enzymes directly involved in the FA metabolism. Due to its high-water solubility and reactivity, indeed, airborne FA is absorbed mainly (~ 90%) in the upper respiratory tract, where it quickly forms intermolecular and intramolecular cross-links within proteins and nucleic acids at the site of contact. It is also rapidly metabolized to formate by FA-dehydrogenase requiring glutathione: the depletion of this compound in the absorbing tissues results in more FA bound to DNA within cells[46,47]. While some reports showed no effect of these polymorphisms in modulating the level of genomic damage[12,34,45,46], others found a significant association. As example, Santovito et al.[43] observed higher frequencies of SCEs, CAs, and Ab.C among pathologists with *GSTT1*-null genotypes than in the reference group. Several other epidemiological studies evaluating exposure to organic solvents, reported the *GSTM1*-null genotype associated with an increase in cytogenetic biomarkers, probably due to the absence of detoxification activity that may affect the amount of DNA damage[48,49].

Since DNA damage is a key step in the carcinogenic process[50], we also considered polymorphisms in both BER and NER pathways. Unexpectedly, we found a significant relationship only for *XRCC1 (399, G>A)* polymorphism, which turned out to be related to a reduction in CAs frequency. The X-ray cross-complementing group 1 (*XRCC1*) is a major DNA repair gene involved in base BER, which is able to fix DNA base damage and single-strand breaks through interacting with DNA components at the damage site. The polymorphisms related to this gene have been linked to the development of several types of cancer[51,52]. Specifically, the *XRCC1* Arg-399Gln polymorphism has been reported to reduce the oxidative damage repair activity and the 399Gln allele has been shown to be related to higher mutagen sensitivity and higher levels of DNA adducts[53]. Therefore, our result appears to be inconsistent with the gene function. Nevertheless, in literature contrasting results can be found. In workers exposed to organic solvents, Hoyos-Giraldo et al. reported that the *XRCC1* Arg399Gln polymorphism carriers did not have a significant CAs frequency increase compared to the wild type genotype carriers. Contrary, in benzene-exposed workers, a significant CAs frequency increase related to by *XRCC1* Arg399Gln variant was reported[48]. As well, in a study on active and passive smokers, Gln/Gln carriers reveal a significantly higher number of aberrations than the Arg/Gln and Arg/Arg genotypes in both the controls and exposed subjects[54].

No significant relationship was found between CAs frequency and *XRCC1* (280, A>G) polymorphism, even though in literature this association has been reported[48].

Finally, inflammation and oxidative stress are knowingly interdependent pathophysiological processes[55]. Since these are two possible mechanisms through which FA could explain its harmful effects, we evaluated the role of polymorphisms on both pro- and anti-inflammatory cytokines genes. No significant result was found, according to the study of Santovito et al.[26], which found no association between cytogenetic damage and *TNFα* − 308 (G>A), *IL10* − 1082 (G>A) and *IL10* − 819 (C>T) gene polymorphisms, with the only exception of homozygous genotypes for *IL-6* G allele, that showed a significant decrease in the frequency of SCEs compared to heterozygous subjects.

These results, however, should be considered cautiously, as we did not consider the eventual effect of all the possible confounding factors that could modulate the studied outcomes, such as the plethora of chemicals to whom pathologists could be exposed in various degrees in laboratories.

Conclusions

Our study confirms FA genotoxic effect, even in workers chronically exposed to low FA levels. Several genetic polymorphisms in metabolism and DNA-repair pathways seem to have an influence in modulating the effect of FA exposure. These findings further highlight the importance of individual susceptibility biomarkers assessment in occupational studies. Due to the extreme FA widespread presence in environmental and occupational settings, studies on both harmful effects related to FA exposure and modulators are crucial to elaborate effective Public Health preventive strategies.

Materials and methods

Epidemiological sample. The epidemiological sample consist of 57 workers occupationally exposed to FA enrolled in two pathology wards of Turin (Italy) and forty-eight hospital workers not exposed to FA recruited in the same two hospitals as control group. Each volunteer signed an informed consent form. The sampling was performed on Wednesday of each sampling week, engaging five to eight subjects every time. Since, routinely, the

exposure to FA in pathology wards occurs mainly via inhalation, each participant wore a personal passive sampler for the measurement of air-FA concentration during the sampling-day work shift. At 4 p.m. of the same day, each subject provided a venous blood sample and answered to a questionnaire administered by one interviewer. The study was approved by the Bioethical Committee of the University of Turin and was performed in accordance with the ethical standards laid down in the 2013 Declaration of Helsinki.

Questionnaire. The questionnaire was administered to each subject by an interviewer to obtained information about demographic characteristics (sex, age), personal habits (smoking) during the last year, and work characteristics (length in years of service working and type of work).

Personal air-FA collection and analysis. FA air samples were collected for working shift (8 h) on Wednesday using passive personal air samplers clipped near the breathing zone of the subject, according to Santovito et al.[12].

Blood sample collection and chromosomal aberration analysis. Blood sample collection and chromosome aberration analysis were performed according to Santovito et al.[12]. DNA extraction and genotyping procedure were carried out as described in Ruberto et al.[56]. Primer sequences, melting temperatures, PCR methodologies used, and expected PCR product sizes are reported in Supplementary Table S1 online.

Statistical analyses. Statistical analysis was assessed using the SPSS software statistical package programme (version 22.0, Chicago, USA) and R (R version 4.0.2). Differences between sex, mean age and years of employment (y.e.) among and between groups were evaluated by analysis of variance. A non-parametric Kruskal–Wallis test was used to compare age, mean y.e. and CAs frequency between groups.

Multivariate general linear model, with Bonferroni's correction, was used to evaluate the influence of age and years of exposure on CAs frequency in both groups. All p-values were two tailed and the level of statistical significance was set at $p < 0.05$ for all tests.

Association between both genetic and environmental variables with the level of genomic damage was evaluated by Poisson regression model, due to a Poisson nature and distribution of the dependent variable (CAs) (Supplementary Fig. S1 online). Genetic variables were considered both in binarized (wild type vs. any mutated allele) and multiallelic (wild type vs. heterozygous vs. homozygous).

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 15 December 2020; Accepted: 30 April 2021
Published online: 18 May 2021

References

1. Costa, S. *et al.* Cytogenetic and immunological effects associated with occupational formaldehyde exposure. *J. Toxicol. Environ. Health. Part A Curr. Issues* **76**, 217–229 (2013).
2. Viegas, S. *et al.* Genotoxic effects in occupational exposure to formaldehyde: A study in anatomy and pathology laboratories and formaldehyde-resins production. *J. Occup. Med. Toxicol.* **5**, 25 (2010).
3. Bellisario, V. *et al.* Towards a formalin-free hospital. Levels of 15–F2t-isoprostane and malondialdehyde to monitor exposure to formaldehyde in nurses from operating theatres. *Toxicol. Res.* **5**, 1122–1129 (2016).
4. Romanazzi, V. *et al.* 15–F2t isoprostane as biomarker of oxidative stress induced by tobacco smoke and occupational exposure to formaldehyde in workers of plastic laminates. *Sci. Total Environ.* **442**, 20–25 (2013).
5. Costa, S. *et al.* Occupational exposure to formaldehyde and early biomarkers of cancer risk, immunotoxicity and susceptibility. *Environ. Res.* **179**, 108740 (2019).
6. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, World Health Organization, & International Agency for Research on Cancer. *Formaldehyde, 2-Butoxyethanol and 1-tert-Butoxypropan-2-ol.* (International Agency for Research on Cancer, 2006).
7. Fenech, M., Nersesyan, A. & Knasmueller, S. A systematic review of the association between occupational exposure to formaldehyde and effects on chromosomal DNA damage measured using the cytokinesis-block micronucleus assay in lymphocytes. *Mutat. Res. Rev. Mutat. Res.* **770**, 46–57 (2016).
8. Dugheri, S. *et al.* An integrated air monitoring approach for assessment of formaldehyde in the workplace. *Saf. Health Work* **9**, 479–485 (2018).
9. Ladeira, C. *et al.* Genotoxicity biomarkers in occupational exposure to formaldehyde—The case of histopathology laboratories. *Mutat. Res. Toxicol. Environ. Mutagen.* **721**, 15–20 (2011).
10. Zanini, C., Gerbaudo, E., Ercole, E., Vendramin, A. & Forni, M. Evaluation of two commercial and three home-made fixatives for the substitution of formalin: A formaldehyde-free laboratory is possible. *Environ. Health. A Glob. Access Sci. Source* **11**, 59 (2012).
11. Patil, S., Rao, R. S., Ganavi, B. S. & Majumdar, B. Natural sweeteners as fixatives in histopathology: A longitudinal study. *J. Nat. Sci. Biol. Med.* **6**, 67–70 (2015).
12. Santovito, A. *et al.* Combined analysis of chromosomal aberrations and glutathione S-transferase M1 and T1 polymorphisms in pathologists occupationally exposed to formaldehyde. *Arch. Toxicol.* **85**, 1295–1302 (2011).
13. Souza, A. D. & Devi, R. Cytokinesis blocked micronucleus assay of peripheral lymphocytes revealing the genotoxic effect of formaldehyde exposure. *Clin. Anat.* **27**, 308–312 (2014).
14. Pala, M. *et al.* Occupational exposure to formaldehyde and biological monitoring of Research Institute workers. *Cancer Detect. Prev.* **32**, 121–126 (2008).
15. Peteffi, G. P. *et al.* Environmental and biological monitoring of occupational formaldehyde exposure resulting from the use of products for hair straightening. *Environ. Sci. Pollut. Res.* **23**, 908–917 (2016).
16. Crocco, P. *et al.* Inter-individual variability in xenobiotic-metabolizing enzymes: Implications for human aging and longevity. *Genes* **10**, 403 (2019).

17. Xu, L. & Wei, H. Association between CYP1A1 2454A>G polymorphism and colorectal cancer risk: A meta-analysis. *J. Cancer Res. Ther.* **11**, 760 (2015).
18. Almazroo, O. A., Miah, M. K. & Venkataramanan, R. Drug metabolism in the liver. *Clin. Liver Dis.* **21**, 1–20 (2017).
19. Pontel, L. B. *et al.* Endogenous formaldehyde is a hematopoietic stem cell genotoxin and metabolic carcinogen. *Mol. Cell* **60**, 177–188 (2015).
20. Langie, S. A. S. *et al.* Causes of genome instability: The effect of low dose chemical exposures in modern society. *Carcinogenesis* **36**, S61–S88 (2015).
21. Reingruber, H. & Pontel, L. B. Formaldehyde metabolism and its impact on human health. *Curr. Opin. Toxicol.* **9**, 28–34 (2018).
22. Bono, R. *et al.* Formaldehyde-induced toxicity in the nasal epithelia of workers of a plastic laminate plant. *Toxicol. Res.* **5**, 752–760 (2016).
23. Hussain, T. *et al.* Oxidative stress and inflammation: What polyphenols can do for us? *Oxid. Med. Cell. Longev.* **2016**, 7432797 (2016).
24. Higashimoto, T., Panopoulos, A., Hsieh, C. L. & Zandi, E. TNFα induces chromosomal abnormalities independent of ROS through IKK, JNK, p38 and caspase pathways. *Cytokine* **34**, 39–50 (2006).
25. Westbrook, A. M. *et al.* The role of tumour necrosis factor-α and tumour necrosis factor receptor signalling in inflammation-associated systemic genotoxicity. *Mutagenesis* **27**, 77–86 (2012).
26. Santovito, A. *et al.* Relationships between cytokine (IL-6 and TGF-β 1) gene polymorphisms and chromosomal damage in hospital workers. *J. Immunotoxicol.* **13**, 314–323 (2016).
27. Garcia-Sagredo, J. M. Fifty years of cytogenetics: A parallel view of the evolution of cytogenetics and genotoxicology. *Biochim. Biophys. Acta Gene Regul. Mech.* **1779**, 363–375 (2008).
28. Bonassi, S., Znaor, A., Norppa, H. & Hagmar, L. Chromosomal aberrations and risk of cancer in humans: An epidemiologic perspective. *Cytogenet. Genome Res.* **104**, 376–382 (2004).
29. García-González, M. A. *et al.* Relevance of GSTM1, GSTT1, and GSTP1 gene polymorphisms to gastric cancer susceptibility and phenotype. *Mutagenesis* **27**, 771–777 (2012).
30. Song, X. *et al.* Variants in nucleotide excision repair core genes and susceptibility to recurrence of squamous cell carcinoma of the oropharynx. *Int. J. Cancer* **133**, 695–704 (2013).
31. Ogawa, M., Kabe, I., Terauchi, Y. & Tanaka, S. A strategy for the reduction of formaldehyde concentration in a hospital pathology laboratory. *J. Occup. Health* **61**, 135–142 (2019).
32. D'Ettorre, G., Criscuolo, M. & Mazzotta, M. Managing formaldehyde indoor pollution in anatomy pathology departments. *Work* **56**, 397–402 (2017).
33. Kirsch-Volders, M., Fenech, M. & Bolognesi, C. Validity of the lymphocyte cytokinesis-block micronucleus assay (L-CBMN) as biomarker for human exposure to chemicals with different modes of action: A synthesis of systematic reviews. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **836**, 47–52 (2018).
34. Costa, S. *et al.* Increased levels of chromosomal aberrations and DNA damage in a group of workers exposed to formaldehyde. *Mutagenesis* **30**, 463–473 (2015).
35. Jakab, M. G. *et al.* Formaldehyde-induced chromosomal aberrations and apoptosis in peripheral blood lymphocytes of personnel working in pathology departments. *Mutat. Res. Toxicol. Environ. Mutagen.* **698**, 11–17 (2010).
36. Autrup, H. Genetic polymorphisms in human xenobiotica metabolizing enzymes as susceptibility factors in toxic response. In *Mutation Research—Genetic Toxicology and Environmental Mutagenesis*. **464**, 65–76 (2000).
37. Pan, S. T. *et al.* Computational identification of the paralogs and orthologs of human cytochrome P450 superfamily and the implication in drug discovery. *Int. J. Mol. Sci.* **17**, 1020 (2016).
38. Wright, C. M. *et al.* Genetic association study of CYP1A1 polymorphisms identifies risk haplotypes in nonsmall cell lung cancer. *Eur. Respir. J.* **35**, 152–159 (2010).
39. Roszak, A., Lianeri, M., Sowińska, A. & Jagodziński, P. P. CYP1A1 Ile462Val polymorphism as a risk factor in cervical cancer development in the polish population. *Mol. Diagn. Ther.* **18**, 445–450 (2014).
40. Sram, R. J. *et al.* Chromosomal aberrations in environmentally exposed population in relation to metabolic and DNA repair genes polymorphisms. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **620**, 22–33 (2007).
41. Motykiewicz, G. *et al.* A molecular epidemiology study in women from Upper Silesia, Poland. *Toxicol. Lett.* **96–97**, 195–202 (1998).
42. Dehbozorgi, M. *et al.* Prevalence of the CYP2C19*2 (681 G>A), *3 (636 G>A) and *17 (-806 C>T) alleles among an Iranian population of different ethnicities. *Mol. Med. Rep.* **17**, 4195–4202 (2018).
43. Santovito, A., Gendusa, C. & Cervella, P. Evaluation of baseline frequency of sister chromatid exchanges in an Italian population according to age, sex, smoking habits, and gene polymorphisms. *Am. J. Hum. Biol.* **29**, e23034–e23041 (2017).
44. Nebert, D. W. & Vasiliou, V. Analysis of the glutathione S-transferase (GST) gene family. *Hum. Genomics* **1**, 460–464 (2004).
45. Costa, S. *et al.* Genotoxic damage in pathology anatomy laboratory workers exposed to formaldehyde. *Toxicology* **252**, 40–48 (2008).
46. Nielsen, G. D., Larsen, S. T. & Wolkoff, P. Re-evaluation of the WHO (2010) formaldehyde indoor air quality guideline for cancer risk assessment. *Arch. Toxicol.* **91**, 35–61 (2017).
47. Haffner, M. J. *et al.* Formaldehyde exposure and its effects during pregnancy: Recommendations for laboratory attendance based on available data. *Clin. Anat.* **28**, 972–979 (2015).
48. Hoyos-Giraldo, L. S., Carvajal, S., Cajas-Salazar, N., Ruíz, M. & Sánchez-Gómez, A. Chromosome aberrations in workers exposed to organic solvents: Influence of polymorphisms in xenobiotic-metabolism and DNA repair genes. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **666**, 8–15 (2009).
49. Norppa, H. Cytogenetic biomarkers and genetic polymorphisms. In *Toxicology Letters*. **149**, 309–334 (2004).
50. Sellappa, S., Balakrishnan, M., Raman, S. & Palanisamy, S. Induction of micronuclei in buccal mucosa on chewing a mixture of betel leaf, areca nut and tobacco. *J. Oral Sci.* **51**, 289–292 (2009).
51. Qin, C. J. *et al.* XRCC1 R399Q polymorphism and colorectal cancer risk in the Chinese Han population: A meta-analysis. *Tumor Biol.* **36**, 461–466 (2015).
52. Wu, Y. G. *et al.* The association of XRCC1 polymorphism with osteosarcoma risk, clinicopathologic features, and prognosis in a Chinese Han population. *Cancer Manag. Res.* **10**, 4959–4967 (2018).
53. Takeshita, H., Fujihara, J., Yasuda, T. & Kimura-Kataoka, K. Worldwide distribution of four SNPs in X-ray and repair and cross-complementing group 1 (XRCC1). *Clin. Transl. Sci.* **8**, 347–350 (2015).
54. Chandirasekar, R. *et al.* XRCC1 gene variants and possible links with chromosome aberrations and micronucleus in active and passive smokers. *Environ. Toxicol. Pharmacol.* **32**, 185–192 (2011).
55. Biswas, S. K. Does the interdependence between oxidative stress and inflammation explain the antioxidant paradox?. *Oxid. Med. Cell. Longev.* **2016**, 1–9 (2016).
56. Ruberto, S. & Santovito, A. Association of *TGFβ1* codon 10 (T>C) and *IL* -10 (G>C) cytokine gene polymorphisms with longevity in a cohort of Italian population. *Am. J. Hum. Biol.* https://doi.org/10.1002/ajhb.23491 (2020).

Author contributions

Conceptualization, R.B. and A.S.; Methodology, A.S.; Software, E.C., V.B. and G.S; Validation, A.S., V.B. and G.S.; Formal analysis, E.C., A.S. F.G., M.B.; Investigation, A.S. and F.G.; Resources, R.B.; Data curation, F.G. and M.B.; Writing—original draft preparation, F.G.; Writing—review and editing, F.G. and M.B.; Supervision, A.S. and R.B.; Project administration, A.S. and R.B.; Funding acquisition, A.S. and R.B.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-89833-w.

Correspondence and requests for materials should be addressed to R.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

Archives of Toxicology (2022) 96:1101–1108
https://doi.org/10.1007/s00204-022-03238-w

GENOTOXICITY AND CARCINOGENICITY

The formation of SCEs as an effect of occupational exposure to formaldehyde

Federica Ghelli[1] · Enrico Cocchi[1] · Valeria Bellisario[1] · Martina Buglisi[1] · Giulia Squillacioti[1] · Alfredo Santovito[2] · Roberto Bono[1]

Received: 21 October 2021 / Accepted: 27 January 2022 / Published online: 12 February 2022
© The Author(s) 2022

Abstract

Formaldehyde (FA) is a ubiquitous toxic chemical employed worldwide due to its disinfectant and preservative properties. Despite being classified as a human carcinogen, FA is still employed as formalin in pathology wards as standard fixative. We evaluated its relationship with the formation of sister-chromatid exchanges (SCEs) in cultured peripheral blood lymphocytes on 57 pathologists and 48 controls and the risk/protective role played by several genetic polymorphisms. All subjects were assessed for SCEs and genotyped for the most common cancer-associated gene polymorphisms: *CYP1A1* exon 7 (A > G), *CYP1A1**2A (T > C), *CYP2C19**2 (G > A), *GSTT1* (presence/absence), *GSTM1* (presence/absence), *GSTP1* (A > G), *XRCC1* (G399A), *XRCC1* (C194T), *XRCC1* (A280G), *XPC* exon 15 (A939C), *XPC* exon 9 (C499T), *TNFα* − 308 G > A), *IL10* − 1082 (G > A), and *IL6* − 174 (G > C). Air-FA concentration was assessed through passive personal samplers. Pathologists, exposed to 55.2 μg/m^3 of air-FA, showed a significantly higher SCEs frequency than controls, exposed, respectively, to 18.4 μg/m^3. Air-FA was directly correlated with SCEs frequency and inversely with the replication index (RI). Regression models showed FA exposure as a significant predictor in developing SCEs, while did not highlight any role of the selected polymorphisms. Our study confirms the role of low air-FA levels as genotoxicity inductor, highlighting the importance to define exposure limits that could be safer for exposed workers.

Keywords Formaldehyde · Genotoxicity · Biomonitoring · Occupational exposure

Introduction

Formaldehyde (FA) is a ubiquitous toxic highly reactive chemical with carcinogenic properties (Zhao et al. 2021; Kang et al. 2021). FA is the result of both natural and anthropogenic processes, and its industrial employment is extended worldwide. It is a well-known disinfectant and preservative to which humans are exposed mainly by respiratory way, in both life and occupational settings. Workers directly involved in the production or use of FA and FA-based compounds, are exposed to higher concentration of this pollutant than the general population (Motta et al. 2021; Kang et al. 2021). Due to its ability in preserving cell and tissue morphology, formalin, an aqueous solution containing dissolved FA, is widely used in pathology wards (Shaham et al. 2002; Motta et al. 2021). Pathologists are exposed to FA in several steps of their workflow. The most relevant exposure happens in the grossing room, when workers manipulate the anatomic samples soaked in formalin. Particularly, "sampling" step, when pathologists handle the pieces from formalin container and the "rinsing" step to avoid the formalin excess under running water before slicing are the occasion when the proximity between the operator and the formalin jar determine the higher exposure to FA vapour (Vimercati et al. 2010). FA is currently classified as human carcinogen, based on studies on nasopharyngeal cancer and leukaemia (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. et al. 2006). The mechanisms by which FA could be involved in causing cancer are not completely clear to date; several pathways have been proposed, including epigenetic effects, DNA reactivity, chromosomal breakage, and oxidative stress (Costa et al. 2015).

✉ Roberto Bono
roberto.bono@unito.it

1 Department of Public Health and Pediatrics, University of Turin, Via Santena 5 bis, 10126 Turin, Italy

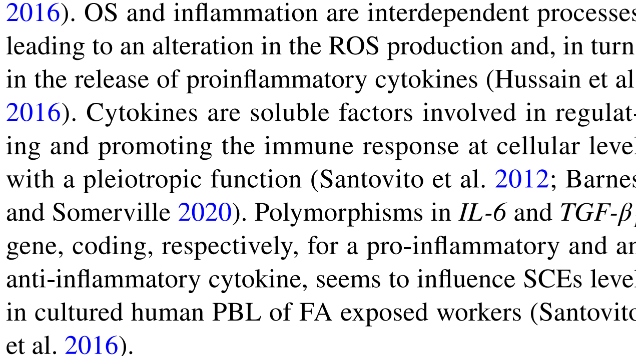
2 Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin, Italy

Cytogenetic biomarkers as Sister-Chromatid Exchanges (SCEs), assessed in cultured peripheral blood lymphocytes (PBL), are reliable tools widely employed in human biomonitoring to assess FA-related DNA damage level and, in turn, genotoxic effects (Norppa 2004; Santovito et al. 2014, 2015; Costa et al. 2019). SCEs are the result of DNA replication product interchanges between sister chromatids at apparently homologous *loci* (Norppa 2004). This phenomenon is due to molecules able to form covalent adducts to DNA or to interfere with DNA metabolic processes (Santovito et al. 2014).

Several studies to date assessed increased SCEs frequency as a consequence of FA exposure, both in vitro and in exposed worker's PBL (Shaham et al. 2002; Speit et al. 2007; Costa et al. 2008, 2019; Bernardini et al. 2020).

In this scenario, the individual genetic background is a key factor in shaping the response to FA exposure (Bernardini et al. 2020), with several polymorphisms able to influence its genotoxic effect, that can be used as susceptibility biomarkers (Šrám and Binková 2000). As example, any polymorphism involved in the xenobiotic metabolism or in the DNA damage repair pathways may affect the individual susceptibility to genotoxic agents (Norppa 2004).

Enzymes involved in xenobiotic biotransformation are generally classified into Phase I and Phase II enzymes, according to the reactions they catalyse (Jancova et al. 2010). Phase I enzymes include mostly enzymes belonging to the cytochrome P (CYP) 450 (CYP450) family, which is a key factor in oxidative reactions, resulting in transformation of a parent compound to more polar metabolite(s) (Swinney et al. 2006; Jancova et al. 2010). The metabolites obtained are often electrophilic molecules able to bind proteins and nucleic acids, leading thus to an alteration of cellular functions (Aladesanmi et al. 2017). Conversely, phase II enzymes are mostly transferases responsible of conjugation reaction, aiming at the transformation of xenobiotic or phase I enzyme products into easily excretable molecules (Swinney et al. 2006; Jancova et al. 2010). One of its major components are glutathione S-transferases (GSTs), which catalyse the conjugation of the glutathione reduced form (GSH) (Tan et al. 2017). In this regard, subjects exposed to FA carrying the *CYP2E1* intronic polymorphism (rs6413432) seems to have higher levels of genetic damage (Costa et al. 2015). Similarly, *GSTM1* and *GSTT1* null-alleles seem to be related to an enhanced genotoxic effect in FA-exposed workers (Costa et al. 2008; Ghelli et al. 2021). On the other side, DNA repair mechanisms are crucial to counteract DNA damage due to xenobiotic and carcinogens exposure. Polymorphisms in genes involved in the Nucleotide Excision Repair (NER) and in the Base Excision Repair (BER) pathways have been associated with increased risk for several types of cancer, including head, neck, and lung cancer (Costa et al. 2008). Finally, FA is known to be a powerful oxidative stress (OS) inductor (Bellisario et al. 2016; Bono et al.


2016). OS and inflammation are interdependent processes leading to an alteration in the ROS production and, in turn, in the release of proinflammatory cytokines (Hussain et al. 2016). Cytokines are soluble factors involved in regulating and promoting the immune response at cellular level with a pleiotropic function (Santovito et al. 2012; Barnes and Somerville 2020). Polymorphisms in *IL-6* and *TGF-*β_1 gene, coding, respectively, for a pro-inflammatory and an anti-inflammatory cytokine, seems to influence SCEs level in cultured human PBL of FA exposed workers (Santovito et al. 2016).

The possible role of phase I, phase II, DNA-repair, and cytokine gene polymorphisms in modulating the genotoxic effect of FA exposure in workers chronically exposed has to be better understand. Thus, we assessed the SCEs frequency in PBL of professionals working in pathology wards and genotyped them for a battery of gene polymorphisms. These last are: Cytochrome P450 1A1 (*CYP1A1*) exon 7 (A > G), *CYP1A1 2A* (T > C), *CYP2C19*2* (G > A), *GSTT1*, *GSTM1*, *GSTP*, X-ray repair cross-complementing group 1 (*XRCC1*) 399 (G > A), 194 (C > T), 280 (A > G), *Xeroderma pigmentosum* complementation group C (*XPC*) exon 15 (A > C), exon 9 (C > T), *TNF-α, IL-6, IL-10* − 1082 (G > A).

Materials and methods

Epidemiological sample and questionnaire

The epidemiological sample for this study included 105 workers, 57 pathologists recruited in two pathological wards of two Turin hospitals (Northwestern Italy) and 48 healthcare workers, recruited in the same medical facilities. The control group included mainly nurses and support personnel operating in different wards and not involved in tasks requiring the manipulation or the exposure to FA-based chemicals, such as formalin, or other carcinogenic compounds. Each subject accepted voluntarily to participate and signed an informed consent form prior to their inclusion in the study. The study was approved by the University of Turin bioethical committee (Ethics Committee of Azienda Ospedaliera Città della Salute e della Scienza of Torino—protocol code 0071900, 25 June 2013 and protocol code 0094007, 9 May 2013) and was performed in accordance with ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All the subjects working since at least 1 year who agreed to participate were enrolled.

Each subject enrolled, both belonging to the pathologists' and the controls' group, filled out a questionnaire, provided a blood specimen, and wore for an 8-h working shift an air-FA diffusive personal sampler.

An interviewer administered a questionnaire aiming to collect information about individual characteristics (sex,

age), habits referring to the last year (smoking), and working characteristics (working years and task and Personal Protective Equipment (PPE) use). Workers exposed to FA reported to adopt the prevention and protection measures, both collective and individual, established by the regulation in force (Legislative Decree No 81/2008).

Formaldehyde exposure assessment

Each volunteer wore a personal radial symmetry diffusive air sampler (Radiello®, ICS Maugeri SpA, Pavia, Italy) (https://radiello.com/, accessed on 23 December 2021). This device was clipped to the collar to assess the air-FA concentration in the worker's breathing zone during a standard working day. The samplers were equipped with a special sorbent cartridge containing a 35–50 Florisil mesh coated with 2,4-dinitrophenylhydrazine (DNPH). The quantification was performed by HPLC, as previously described (Squillacioti et al. 2020). As recommended by the Radiello® user manual, we kept at least two cartridges of each lot as a blank. The average FA concentration over exposure time is calculated as the mass of the analyte in the cartridge over time according to the following expression: C = (mass of the analyte [μg]/(Q_k [ml·min^{-1}]·exposure time [min]))·1,000,000. Temperature was assumed to be constantly 25 °C and results are expressed in μg·m^{-3}.

Blood sampling and SCEs assay

5–10 mL venous blood specimens for SCEs assay obtained were collected into heparinised Vacutainer® (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Blood samples were coded and kept refrigerated (+4 °C) until analysis and processed within 2 h, as previously described. A 0.3 mL aliquot of each sample was cultured in a 25 cm^2 flask with 6 mL of RPMI-1640 (Merck, Milan, Italy) with the addition of 20% fetal calf serum (FCS), 2% mitogenic agent phytohemagglutinin-M (Difco, 0.2 mL), L-glutamine (2 mM), antibiotics (100 μg/mL streptomycin and 100 IU/mL penicillin). Cultures were then incubated at 37 °C for 72 h, in a 5% CO_2 atmosphere. During the last 2 h of culture 0.06 μg/mL of colchicine (Merck, Milan, Italy, 0.25 μg/mL) was added in order to arrest cells in mitosis.

During the chromosome preparation step, after a 10 min centrifugation at ≤ 800 rpm, cells were slowly resuspended in 10 mL of pre-warmed hypotonic solution (0.075 M KCl, 37 °C), and subsequently incubated in water bath 15 min at 37 °C. After a new centrifugation, cells were than fixed with a 20 min incubation in cold methanol:acetic acid (3:1) at room temperature. The fixation treatment was repeated three times. After discarding the supernatant, the remaining pellet, still dissolved in a residual fixative volume, was seeded on slides.

Bromodeoxyuridine (BrdUrd, 10 μg/mL), a thymidine analogue that during replication can be efficiently incorporated into the newly synthesized DNA strands, was added at 24 h to measure SCEs in second division metaphases. After two cell cycles in BrdUrd medium, the two sister chromatids revealed a different appearance due to the amount of BrdUrd incorporated: the lighter chromatid is the one with more BrdUrd ("bleaching" effect).

In order to allow sister chromatid differentiation, cells were thereafter stained with a 20 min incubation with 10 μg/mL fluorescence dye Hoechst 33258 (Merck, Milan, Italy) in the dark at room temperature. Cells were then irradiated 30 min with an 8-W UV lamp (254 nm) at approximately 20 cm. Subsequently, after an 1 h incubation in 2 × Standard Saline Concentration (SSC) at 60 °C, and slides were then stained with a 10 min incubation in Sörensen buffer with 5% Giemsa (Merck, Milan, Italy). Microscopic analyses were performed by a light microscope (CX40, Olympus, Tokyo, Japan) at 1000 × magnification.

The determination of the SCE/cell number for each subject was performed scoring 50 well-spread second-division metaphases containing 46 chromosomes. The replication index (RI) evaluation was performed scoring a total of 100 cells from each donor and calculated according to the following formula: RI = (M1 + 2 M2 + 3 M3)/N, where M1, M2 and M3 represent the number of cells undergoing first second and third mitosis and N is the total number of scored metaphases (NSM) (Santovito et al. 2014, 2017).

DNA extraction and genotyping

An aliquot of all blood specimens had been stored at − 20 °C until the assay. The DNA extraction and genotypization analysis had been performed as previously described (Ruberto and Santovito 2021).

Primer sequences, melting temperatures, PCR methodologies used, and expected PCR product sizes are reported in Online Resource 1.

Statistical analyses

The Shapiro–Wilk test and density visualization was used to test the normal distribution. The variables were described as frequency (percentage) for categorical variables, median and interquartile range (IQR) for non-normal distributed variables, and mean and standard deviation (± SD) for endpoints with normal distribution. Continuous non-normally distributed variables were compared across relevant groups using the Mann–Whitney *U* test or Wilcoxon signed-rank test and the *T*-test was used for normal-distributed ones. Fisher's exact test served for group comparison of categorical ones. The Spearman rank correlation test was used to investigate

correlations. Group tests were two-sided with $p < 0.05$ considered statistically significant.

Multivariate regression analyses were performed to calculate associations between FA exposure and genetic and clinical variables. Several regression model were created: M0 was performed to evaluate the role of exposure and smoking in the SCEs frequency alteration, 5 different models were performed in order to study the possible role of susceptibility polymorphisms, grouped by the molecular pathway they belong to (phase I (M1), phase II (M2), BER (M3), NER (M4), CK (M5), (Online Resources 2–6), in shaping the frequency of SCEs once covariated for all clinical and exposure variables. Analyses were performed using R 4.4.1 (R Project for Statistical Computing, Vienna, Austria) and SPSS Statistics (IBM SPSS Statistics, Version 27.0. Armonk, NY: IBM Corp).

Results

The final sample includes 105 volunteers, 57 pathologists occupationally exposed to FA and 48 unexposed healthcare professionals considered as control group. The general characteristics of the subjects recruited are shown in Table 1.

The comparison between exposed and non-exposed subjects reveals a significantly higher SCEs frequency in pathologists ($p = 0.009$), as displayed in Table 2.

Concerning the different genotype makeup, the comparison between the *wt* and the group of who carries at least one mutated allele, reveal, in workers exposed to FA, a significantly higher SCEs frequency in the mutation carrier group for *CYP1A1* exon 7 ($p = 0.010$) and *XPC 9* ($p = 0.040$). Among controls, the *XRCC1 194 wt* group showed a higher RI compared to the mutation carrier group ($p = 0.019$).

Table 2 Comparison between exposed and control groups in terms of genotoxic outcomes

Groups	N	SCEs	SCEs/cell ± SD	RI ± S.D
Pathologists	**57**	**15,102**	**5.30 ± 1.33***	**1.69 ± 0.31****
Males	29	7877	5.43 ± 1.18	1.71 ± 0.33
Females	28	7225	5.16 ± 1.49	1.66 ± 0.28
Smokers	14	3841	5.49 ± 0.99	1.82 ± 0.28[a]
Non-smokers	43	11,261	5.24 ± 1.43	1.65 ± 0.31
Controls	**48**	**10,598**	**4.42 ± 1.50**	**1.92 ± 0.19**
Males	25	5952	4.78 ± 1.32	1.90 ± 0.15
Females	23	4646	4.03 ± 1.61	1.95 ± 0.22
Smokers	9	2566	5.77 ± 1.13[b]	1.74 ± 0.15[c]
Non-smokers	39	8032	4.11 ± 1.41	1.97 ± 0.17

Non-normally distributed continuous variables were compared by Mann–Whitney *U* test or Wilcoxon signed-rank test, while normally distributed variables were compared by *t*-test

N Number of analysed subjects, *SCEs* Sister Chromatid Exchanges, *RI* (Replication Index), *S.D.* Standard Deviation

*Pathologists vs controls $p < 0.05$

**Pathologists vs controls $p < 0.001$

[a]smokers vs non-smokers $p < 0.05$ (Mann–Whitney test)

[b]smokers vs non-smokers $p < 0.005$ (*t*-test)

[c]smokers vs non-smokers $p < 0.005$ (Mann–Whitney test)

Table 1 Demographic characteristics and air-FA exposure level of subjects belonging to the studied groups

	Pathologists ($n = 57$)	Controls ($n = 48$)	*p*-value
Sex			> 0.05
Males (%)	50.9	52.1	
Females (%)	49.1	47.9	
Age (years)			> 0.05
Median [IQR]	43 [12]	38 [12]	
min–max	25–60	25 – 70	
Years of employment (years)			> 0.05
Median [IQR]	10 [13]	10 [5]	
min–max	1–33	2–32	
Smoking habit			> 0.05
Non-smokers (%)	75.4%	81.3%	
Smokers (%)	24.6%	18.8%	
Cigarettes/die			> 0.05
Median [IQR]	7.50 [12]	13 [7]	
min–max	5–40	10–22	
Air-FA (μg/m^3)			< 0.001
Median [IQR]	55.2 [22.3]	18.5 [5.4]	
(ppm conversion)	(0.045 [0.018])	(0.015 [0.004])	
min–max	30.0–169.3	9.2–36.3	

Continuous variables were compared by Mann–Whitney *U* test or Wilcoxon signed-rank test, while categorical variables were compared by Fisher's exact test. n = number of analysed subjects

The correlation analysis performed on the whole sample revealed that the SCEs frequency was significantly directly correlated with the age ($\rho = 0.30$, $p < 0.01$) and the smoking habit (cigarette/die and years of smoking, respectively $\rho = 0.27$, $p < 0.01$ and $\rho = 0.28$, $p < 0.01$), air-FA concentration ($\rho = 0.29$, $p < 0.01$) and negatively correlated with the RI ($r = -0.21$, $p < 0.05$). The RI was found to be also negatively correlated with the air-FA concentration ($\rho = -0.39$, $p < 0.001$).

M0 regression model was performed to evaluate the role of exposure and smoking in the SCEs frequency alteration. As reported in Table 3, the exposure to Air-FA influence significantly the SCEs frequency ($\beta = 38.29$, $p = 0.008$). Conversely, in the same model, age, sex, smoking habit and year of exposure did not show any significant role.

The regression models performed separating the studied genes according to the molecular pathways they belong to (Online Resources 2–6), revealed a significant role of exposure in all models, except in M2 (phase 2 metabolic pathway). In this model, the similar variance in SCEs frequency due to the polymorphisms covered the effect of FA exposure as well, which significance remained borderline.

Discussion

Despite the IARC classified FA as a known human carcinogen (group 1) nearly 15 years ago, the safety of workers nowadays exposed to this chemical, and particularly those employed in the anatomy pathology wards, remains a matter of concern.

Currently, although there are alternative methods to the use of FA, such as the practice of Under-Vacuum Sealing (UVS) (Bellisario et al. 2016), FA is still considered the standard fixative for routine work. In Italy, 1726 workers employed in specialised hospital activities had been estimated to be potentially exposed to FA (Scarselli et al. 2017; Dugheri et al. 2021). To safeguard workers occupationally exposed to air-FA, various limits of exposure have been proposed, even though there is no agreement between the international agencies (Dugheri et al. 2020). The current Threshold Limit Value—Time Weighted Average (TLV – TWA) recommended by the American Conference of Governmental Industrial Hygienists (ACGIH) is 120 µg/m^3 (i.e., 0.1 ppm) and is referred to average exposure in eight-hour workdays (Dugheri et al. 2020).

Within this scenario, performing research on the various outcomes relatable to FA exposure and on other potentially influencing factors remains crucial. Specifically, the understanding of the gene-by-exposure effect could allow identifying more susceptible subjects and, eventually, defining health-based recommended limits that could be safe for all exposed workers (Faruque et al. 2020). Our aim was to deepen the possible association between exposure to air-FA and genotoxic damage. The eventual identification of significant SCEs levels according to the workers' genetic makeup in subject exposed to a deemed safe concentration of air-FA could provide relevant information to the continuous revising process of FA exposure limits.

The median air-FA concentration measured in the exposed workers breathing zone (0.06 mg/m^3, i.e., 0.05 ppm) is lower than most of the air-FA levels in pathology wards reported in literature. Recently, Costa et al. (2019) reported an average FA level of 0.38 ppm (i.e. 0.47 mg/m^3) in hospital anatomy-pathology laboratories in Portugal, while Jalali et al. (2021) measured even higher concentrations (0.64 mg/m^3) in Iranian pathology wards (Costa et al. 2019; Jalali et al. 2021). In Italy, Scarselli et al. (2017) reported a general average value of 0.49 mg/m^3 (i.e., 0.40 ppm) for workers employed in hospital activities (Scarselli et al. 2017).

In our previous work on the same sample, we reported a significantly higher chromosomal aberration (CAs) frequency in pathologists than in controls (Ghelli et al. 2021). As well, we found a higher SCEs frequency in the exposed group, in line with previous reports (Costa et al. 2013, 2019; Santovito et al. 2014). Moreover, the RI value in pathologists was significantly lower than in the control group, contrary to Santovito et al. (2014), and negatively correlated with SCEs frequency, a further evidence of cytotoxicity induction (Santovito et al. 2014).

The association between air-FA and genotoxic damage was confirmed by correlation analysis as well, revealing that air-FA concentration was positively related to SCEs frequency and negatively to RI.

Concerning the genetic background, the comparison between *wt* and mutation carrier subjects revealed, in the pathologist group, a significantly lower SCEs frequency in workers *wt* for *CYP1A1* exon 7 (A > G) and *XPC* exon 9 (C499T). *CYP1A1* is a member of the CYP450 family playing a fundamental role in the metabolism of both endogenous and exogenous substrates, such as nutrients,

Table 3 Multivariate regression model for personal characteristics and work-related factors in predicting the SCEs frequency

	Estimate	SE	*z*	*p*	
(Intercept)	158.415	43.904	3.608	< 0.001	***
Age	1.632	1.403	1.164	> 0.05	
Sex (F)	− 10.678	14.464	− 0.738	> 0.05	
Years of Smoking	0.776	1.272	0.61	> 0.05	
Cigarette/die	0.774	1.510	0.512	> 0.05	
Years of Exposure	− 0.202	1.670	− 0.121	> 0.05	
FA-Exposure (Pathologists)	38.291	14.048	2.726	< 0.01	**

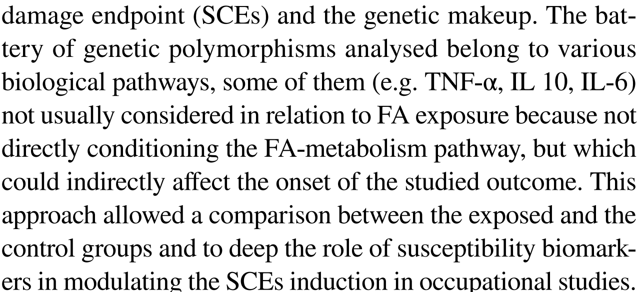
(***) < 0.001; (**) 0.001–0.01; (*) 0.01–0.05; (.) 0.05–0.1

drugs, and environmental carcinogens (Badal and Delgoda 2014). The enzyme coded by this gene, indeed, is mostly expressed in extra hepatic tissues where it is involved in the polycyclic aromatic hydrocarbons (PAHs) biotransformation, which is associated with lung cancer risk (Ezzeldin et al. 2019). *CYP1A1* expression is affected by environmental and genetic factors (Ezzeldin et al. 2019). Specifically, the Ile462Val polymorphism, causing a substitution in the enzyme heme-binding region, result in a twofold increase of the microsomal enzyme activity and is a risk factor for many types of cancer and hematopoietic malignancies, such as acute leukaemia (Zhuo et al. 2012; Roszak et al. 2014). However, previous studies did not report any association between this polymorphism and SCEs frequency (Carere et al. 2002; Santovito et al. 2017). The human *XPC* (*Xeroderma pigmentosum* complementation group C) gene encodes for a 940-amino acid protein essential within the NER (Nucleotide Excision Repair) pathway, involved in the early damage site recognition and DNA repair initiation (Dai et al. 2019). The abnormal expression of the XPC protein is related to cancer progression, and the TT genotype seems to be associated with an increased risk of bladder and breast cancers (Dai et al. 2019). To our knowledge, this is the first report highlighting a significant higher SCEs frequency in workers exposed to FA carrying the XPC exon 9 (C > T) polymorphism.

The regression analysis, however, did not evidence any significant influence of the studied polymorphisms in modulating the SCEs frequency, in line with literature (Costa et al. 2019). However, the FA exposure remains significant in almost all the models strengthening, once more, its role in the genotoxic damage induction. Interestingly, only few studies on occupational research topic analysed the influence of cytokine genetic polymorphisms, even though the key role of the inflammatory pathway in mediating the consequences of occupational exposures. In the present research we did not highlight any role of the studied cytokine polymorphisms in inducing SCEs, contrary to Santovito et al. (2015), which reported a role of IL-6 -174 (G > C). However, in that report, the authors specified that it is still unclear whether the genomic effects can be related to a direct action of these polymorphisms or, more likely, to the interaction with the cytokine network.

Despite evidence on the tobacco-smoking role in SCEs induction (DeMarini 2004), in our regression models, nor the number of cigarettes smoked every day, nor the years of smoking resulted significantly related with the selected outcome.

The novelty of our research consists in the combination of several different parameters. Firstly, the personal assessment of air-FA exposure provides a further element to the varied picture of FA exposure in the Italian scenario. Data concerning personal FA exposure were associated with both a DNA



damage endpoint (SCEs) and the genetic makeup. The battery of genetic polymorphisms analysed belong to various biological pathways, some of them (e.g. TNF-α, IL 10, IL-6) not usually considered in relation to FA exposure because not directly conditioning the FA-metabolism pathway, but which could indirectly affect the onset of the studied outcome. This approach allowed a comparison between the exposed and the control groups and to deep the role of susceptibility biomarkers in modulating the SCEs induction in occupational studies.

The results of our study, however, should be considered carefully. Indeed, the cross-sectional design of the study did not allow making causal inferences and workers employed in pathology wards are exposed to several chemical agents other than FA that could potentially have a confounding effect in the induction of SCEs.

Conclusion

The present study confirms once more the genotoxic effect of FA, highlighting the urgent need to put in place all the possible strategies to reduce and prevent the chronic exposure consequences, especially in working settings. Moreover, we focus our attention on the importance of susceptibility biomarkers, which could modulate the individual biological response to the xenobiotic exposure. This approach is of outmost importance in order to provide data for the definition of exposure limits that could be safe for exposed workers.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00204-022-03238-w.

Funding No funds, grants, or other support was received.

Declarations

Conflict of interest The authors have no financial or proprietary interests in any material discussed in this article.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the University of Turin bioethical committee (Ethics Committee of Azienda Ospedaliera Città della Salute e della Scienza of Torino—protocol code 0071900, 25 June 2013 and protocol code 0094007, 9 May 2013).

Informed consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aladesanmi OT, Agboola FK, Okonji RE (2017) Enzymes as biomarkers of environmental stress in African catfish (Clarias gariepinus) in Osun State, Nigeria. J Heal Pollut 7:71–83. https://doi.org/10.5696/2156-9614-7.14.71

Badal S, Delgoda R (2014) Role of the modulation of CYP1A1 expression and activity in chemoprevention. J Appl Toxicol 34:743–753

Barnes BJ, Somerville CC (2020) Modulating cytokine production via select packaging and secretion from extracellular vesicles. Front Immunol. https://doi.org/10.3389/FIMMU.2020.01040

Bellisario V, Mengozzi G, Grignani E et al (2016) Towards a formalin-free hospital. Levels of 15–F2t-isoprostane and malondialdehyde to monitor exposure to formaldehyde in nurses from operating theatres. Toxicol Res (camb) 5:1122–1129. https://doi.org/10.1039/c6tx00068a

Bernardini L, Barbosa E, Charão MF, Brucker N (2020) Formaldehyde toxicity reports from in vitro and in vivo studies: a review and updated data. Drug Chem Toxicol. https://doi.org/10.1080/01480545.2020.1795190

Bono R, Munnia A, Romanazzi V et al (2016) Formaldehyde-induced toxicity in the nasal epithelia of workers of a plastic laminate plant. Toxicol Res (camb) 5:752–760. https://doi.org/10.1039/c5tx00478k

Carere A, Andreoli C, Galati R et al (2002) Biomonitoring of exposure to urban air pollutants: analysis of sister chromatid exchanges and DNA lesions in peripheral lymphocytes of traffic policemen. Mutat Res Genet Toxicol Environ Mutagen 518:215–224. https://doi.org/10.1016/S1383-5718(02)00108-0

Costa S, Coelho P, Costa C et al (2008) Genotoxic damage in pathology anatomy laboratory workers exposed to formaldehyde. Toxicology 252:40–48. https://doi.org/10.1016/j.tox.2008.07.056

Costa S, García-Lestón J, Coelho M et al (2013) Cytogenetic and immunological effects associated with occupational formaldehyde exposure. J Toxicol Environ Heal Part A Curr Issues 76:217–229. https://doi.org/10.1080/15287394.2013.757212

Costa S, Carvalho S, Costa C et al (2015) Increased levels of chromosomal aberrations and DNA damage in a group of workers exposed to formaldehyde. Mutagenesis 30:463–473. https://doi.org/10.1093/mutage/gev002

Costa S, Costa C, Madureira J et al (2019) Occupational exposure to formaldehyde and early biomarkers of cancer risk, immunotoxicity and susceptibility. Environ Res 179:108740. https://doi.org/10.1016/j.envres.2019.108740

Dai Y, Song Z, Zhang J, Gao W (2019) Comprehensive assessment of the association between XPC rs2228000 and cancer susceptibility based on 26835 cancer cases and 37069 controls. Biosci Rep 39:1–16. https://doi.org/10.1042/BSR20192452

DeMarini DM (2004) Genotoxicity of tobacco smoke and tobacco smoke condensate: a review. Mutat Res Rev Mutat Res 567:447–474

Dugheri S, Massi D, Mucci N et al (2020) Formalin safety in anatomic pathology workflow and integrated air monitoring systems for the formaldehyde occupational exposure assessment. Int J Occup Med Environ Health 34:1–20. https://doi.org/10.13075/ijomeh.1896.01649

Dugheri S, Massi D, Mucci N et al (2021) Formalin safety in anatomic pathology workflow and integrated air monitoring systems for the formaldehyde occupational exposure assessment. Int J Occup Med Environ Health 34:319–338. https://doi.org/10.13075/ijomeh.1896.01649

Ezzeldin N, Abdel Galil SM, Said D et al (2019) Polyneuropathy associated with chronic hemodialysis: clinical and electrophysiological study. Int J Rheum Dis 22:826–833. https://doi.org/10.1111/1756-185X.13462

Faruque MO, De Jong K, Vonk JM et al (2020) Occupational exposures and genetic susceptibility to occupational exposures are related to sickness absence in the lifelines cohort study. Sci Rep 10:1–10. https://doi.org/10.1038/s41598-020-69372-6

Ghelli F, Cocchi E, Buglisi M et al (2021) The role of phase I, phase II, and DNA-repair gene polymorphisms in the damage induced by formaldehyde in pathologists. Sci Rep 11:10507. https://doi.org/10.1038/s41598-021-89833-w

Hussain T, Tan B, Yin Y et al (2016) Oxidative stress and inflammation: what polyphenols can do for us? Oxid Med Cell Longev. https://doi.org/10.1155/2016/7432797

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans., World Health Organization., International Agency for Research on Cancer. (2006) Formaldehyde, 2-butoxyethanol and 1-tert-butoxypropan-2-ol. International agency for research on cancer

Jalali M, Moghadam SR, Baziar M et al (2021) Occupational exposure to formaldehyde, lifetime cancer probability, and hazard quotient in pathology lab employees in Iran: a quantitative risk assessment. Environ Sci Pollut Res 28:1878–1888. https://doi.org/10.1007/s11356-020-10627-0

Jancova P, Anzenbacher P, Anzenbacherova E (2010) Phase II drug metabolizing enzymes. Biomed Pap 154:103–116. https://doi.org/10.5507/bp.2010.017

Kang DS, Kim HS, Jung JH et al (2021) Formaldehyde exposure and leukemia risk: a comprehensive review and network-based toxicogenomic approach. Genes Environ 43:1–10

Motta O, Charlier B, De Caro F et al (2021) Environmental and biological monitoring of formaldehyde inside a hospital setting: a combined approach to manage chemical risk in workplaces. J Public Health Res 10:2012. https://doi.org/10.4081/jphr.2021.2012

Norppa H (2004) Cytogenetic biomarkers and genetic polymorphisms. Toxicology letters. Elsevier, Amsterdam, pp 309–334

Roszak A, Lianeri M, Sowińska A, Jagodziński PP (2014) CYP1A1 Ile462Val polymorphism as a risk factor in cervical cancer development in the polish population. Mol Diagn Ther 18:445–450. https://doi.org/10.1007/s40291-014-0095-2

Ruberto S, Santovito A (2021) Association of TGFβ1 codon 10 (T>C) and IL-10 (G>C) cytokine gene polymorphisms with longevity in a cohort of Italian population. Am J Hum Biol. https://doi.org/10.1002/ajhb.23491

Santovito A, Cervella P, Schleicherova D, Delpero M (2012) Genotyping for cytokine polymorphisms in a Northern Ivory Coast population reveals a high frequency of the heterozygote genotypes for the TNF-α-308G/A SNP. Int J Immunogenet 39:291–295. https://doi.org/10.1111/j.1744-313X.2012.01086.x

Santovito A, Cervella P, Delpero M (2014) Increased frequency of chromosomal aberrations and sister chromatid exchanges in peripheral lymphocytes of radiology technicians chronically exposed to low levels of ionizing radiations. Environ Toxicol Pharmacol 37:396–403. https://doi.org/10.1016/j.etap.2013.12.009

Santovito A, Cervella P, Delpero M (2015) Evaluation of genomic damage in peripheral lymphocytes from occupationally exposed anesthetists: assessment of the effects of age, sex, and *GSTT1* gene

polymorphism. J Biochem Mol Toxicol 29:234–239. https://doi.org/10.1002/jbt.21689
Santovito A, Cervella P, Chiarizio M et al (2016) Relationships between cytokine (IL-6 and TGF-β 1) gene polymorphisms and chromosomal damage in hospital workers. J Immunotoxicol 13:314–323. https://doi.org/10.3109/1547691X.2015.1076547
Santovito A, Delsoglio M, Manitta E et al (2017) Association of GSTT1 null, XPD 751 CC and XPC 939 CC genotypes with increased levels of genomic damage among hospital pathologists. Biomarkers 22:557–565. https://doi.org/10.1080/1354750X.2017.1322147
Scarselli A, Corfiati M, Di MD, Iavicoli S (2017) National estimates of exposure to formaldehyde in Italian workplaces. Ann Work Expo Heal 61:33–43. https://doi.org/10.1093/annweh/wxw004
Shaham J, Gurvich R, Kaufman Z (2002) Sister chromatid exchange in pathology staff occupationally exposed to formaldehyde. Mutat Res Genet Toxicol Environ Mutagen 514:115–123. https://doi.org/10.1016/S1383-5718(01)00334-5
Speit G, Schütz P, Högel J, Schmid O (2007) Characterization of the genotoxic potential of formaldehyde in V79 cells. Mutagenesis 22:387–394. https://doi.org/10.1093/mutage/gem031
Squillacioti G, Bellisario V, Grosso A et al (2020) Formaldehyde, oxidative stress, and FeNO in traffic police officers working in two cities of northern Italy. Int J Environ Res Public Health. https://doi.org/10.3390/ijerph17051655
Šrám RJ, Binková B (2000) Molecular epidemiology studies on occupational and environmental exposure to mutagens and carcinogens, 1997–1999. Environ Health Perspect 108:57–70. https://doi.org/10.1289/ehp.108-1637778
Swinney R, Hsu S, Tomlinson G (2006) Phase I and phase II enzyme polymorphisms and childhood cancer. J Investig Med off Publ Am Fed Clin Res 54:303–320. https://doi.org/10.2310/6650.2006.05062
Tan WSD, Liao W, Zhou S, Wong WSF (2017) Is there a future for andrographolide to be an anti-inflammatory drug? Deciphering its major mechanisms of action. Biochem Pharmacol 139:71–81. https://doi.org/10.1016/j.bcp.2017.03.024
Vimercati L, Carrus A, Martino T et al (2010) Formaldehyde exposure and irritative effects on medical examiners, pathologic anatomy post-graduate students and technicians. Iran J Public Health 39:26
Zhao Y, Magaña LC, Cui H et al (2021) Formaldehyde-induced hematopoietic stem and progenitor cell toxicity in mouse lung and nose. Arch Toxicol 95:693–701. https://doi.org/10.1007/s00204-020-02932-x
Zhuo W, Zhang L, Zhu B et al (2012) Association between CYP1A1 Ile462Val Variation and acute leukemia risk: meta-analyses including 2164 cases and 4160 controls. PLoS ONE. https://doi.org/10.1371/journal.pone.0046974

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Article

Bisphenol A and S in the Urine of Newborns: Plastic for Non-Food Use Still without Rules

Valeria Bellisario [1,2,*,†], Enrico Cocchi [1,3,†], Roberta Tassinari [1], Giulia Squillacioti [1], Tiziana Musso [1], Stefano Sottemano [1], Michael Zorzi [4], Paola Dalmasso [1,2], Alessandra Coscia [1], Claudio Medana [4] and Roberto Bono [1]

1 Department of Public Health and Pediatrics, University of Turin, 10126 Turin, Italy; enrico.cocchi@unito.it (E.C.); roberta.tassinari@unito.it (R.T.); giulia.squillacioti@unito.it (G.S.); tiziana.musso@unito.it (T.M.); stefano.sottemano@unito.it (S.S.); paola.dalmasso@unito.it (P.D.); alessandra.coscia@unito.it (A.C.); roberto.bono@unito.it (R.B.)
2 Health Statistics and Biometrics Residency School, University of Turin, 10126 Turin, Italy
3 Center for Precision Medicine and Genomics, Department of Medicine, Division of Nephrology, Columbia University, New York, NY 10032, USA
4 Unit of Mass Spectrometry, Department of Molecular Biotechnology and Health Sciences, University of Turin, 10126 Turin, Italy; michael.zorzi@unito.it (M.Z.); claudio.medana@unito.it (C.M.)
* Correspondence: valeria.bellisario@unito.it; Tel.: +39-0116705876
† These authors contributed equally to this study.

Simple Summary: The aim of our study was to assess the effects of Bisphenols exposure on pregnancy and neonatal life. In this optic, we have: (a) determined Bisphenols concentration levels (Bisphenol A and Bisphenol S) in a group of newborns and their mothers, (b) identified factors, habits and devices possibly responsible for Bisphenols uptake, and (c) determined some possible health effect of Bipshenols exposure. The statistical analyses showed no significant correlations between maternal and neonatal Bisphenols concentration levels. In newborns, on the contrary, a positive correlation between pacifier use and Bisphenol S total and free concentration was detected. Beside, a significant correlation was also found between oral glucose administration and concentration levels of free Bisphenols A. Our study points to a central role of lifestyle, hospital procedures and neonatal devices in inducing Bisphenols exposure during perinatal period. This is the first report of Bisphenols contamination in newborns due to widely non-alimentary products destined for newborn care (glucose solution containers for Bisphenol A and the pacifiers for the Bisphenol S). Further studies are advocated to clarify both the impact of such other Bisphenols forms on human health and the potential Bisphenol A exposure sources during neonatal and childhood life.

Abstract: The aim of the present study was to assess the effects of bisphenol (BP) exposure on pregnancy and neonatal life. We have (a) determined BP (BPA and BPS) concentration levels in a group of newborns and their mothers; (b) identified factors, habits, and devices possibly responsible for BP uptake; and (c) determined the effect of BP exposure. No significant correlations were detected between maternal and neonatal BP concentration levels. In newborns, positive correlations between pacifier use and BPS total ($p = 0.04$) and free BPS ($p = 0.03$) concentrations were detected. A significant correlation was also found between oral glucose administration and concentration levels of free BPA ($p < 0.05$). Our study points to a central role of lifestyle, hospital procedures, and neonatal devices in inducing BP exposure, especially during the perinatal period. This is the first report of BP contamination in newborns due to widely non-alimentary products designed for newborn care, such as glucose-solution containers for BPA and pacifiers for BPS. Further studies are advocated in order to clarify both the impact of other BP forms on human health and development, as well as potential BPA exposure sources during neonatal and childhood life.

Keywords: BP neonatal exposure; BP non-alimentary contamination; human and childhood health; BP regulation

Citation: Bellisario, V.; Cocchi, E.; Tassinari, R.; Squillacioti, G.; Musso, T.; Sottemano, S.; Zorzi, M.; Dalmasso, P.; Coscia, A.; Medana, C.; et al. Bisphenol A and S in the Urine of Newborns: Plastic for Non-Food Use Still without Rules. *Biology* **2021**, *10*, 188. https://doi.org/10.3390/biology10030188

Academic Editor: Andrés Moya

Received: 7 February 2021
Accepted: 26 February 2021
Published: 3 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Bisphenols (BPs) have long been used for the production of polycarbonate plastics and epoxy resins [1]. BPs are well-known as harmful substances for human health, and they are usually ingested mainly through the diet, but BP intake can also occur by inhalation or dermal contamination.

Once they have entered the body, BPs exert estrogenic and/or oxidant activity [2,3], which is particularly harmful during pre- and neonatal life [4,5]. BPs are detoxified to inactive forms in the liver, primarily through conjugation and glucuronidation, then excreted in the urine within 2 to 6 h. Thus, the detection of conjugated BP forms is critical for determining BP risk exposure. In the glucuronidated form, BPs—particularly Bisphenol A (BPA) (their main representative member)—are inactive, whereas their free (unconjugated) forms, albeit basically unstable, promote biotoxicity through mild estrogenic activity [6]. BPs are prevalently found in urine (75–90%), but they are also traceable, albeit at lower concentrations, in other body fluids such as blood, breast milk, semen, cord blood, fetal serum, and placental tissue [7–10].

In this scenario, gestation appears to be a critical window of fetal exposure to BPs, as they trigger cellular responses even at very low doses [11], influencing sex- and gender-differentiation morphology or leading to immune hyper-responsiveness [12]. Infants appear to be particularly susceptible to the harmful effects of BPs [13], and prior studies have demonstrated that exposure to these chemicals can occur through both breastfeeding and skin contact with plastic devices (e.g., polycarbonate feeding bottles and pacifiers) routinely employed in neonatal intensive care units [11,12,14]. This harmful exposure is further exacerbated by the inefficient UDP-glucuronosyltransferase system in newborns, which is required for BP detoxification and not completely developed until ~2–3 months of age [14–16].

In 2015, following a public consultation on the harmful effects of BPA on human health [17], the European Food Safety Authority (EFSA) reduced the temporary tolerable daily intake (t-TDI) threshold of BPA from 50 to 4 µg/kg bw/day. Since then, plastics manufacturers have been striving to replace BPA with alternative compounds, including bisphenol S (BPS), its closest chemical relative. However, BPS has also recently been shown to exert genotoxic and biological activity similar to that of BPA [18,19]. These lines of evidence indicate that these BPA alternative compounds may also become a serious public health concern in the near future due to lack of regulatory limits [20,21].

Given these premises, the aim of the present study was to assess the modulation and possible health effects of BP exposure on perinatal and neonatal life. The final aim of this study was to investigate direct and indirect pathways and effects of BP exposure during pregnancy and the first days of life. Toward this end, we determined both BPA and BPS concentration levels in a group of newborn babies and their mothers. We particularly intended to focus our study on the identification of factors, devices, and behaviors responsible for high BP uptake in this sensitive population in the very first days of life. Finally, we evaluated if higher BP concentrations are correlated with specific hospital procedures or babies' health status immediately after birth.

2. Material and Methods

2.1. Epidemiological Sample

The epidemiological sample was recruited from July 2016 to October 2017 by consulting the register of births of the Sant'Anna Gynecological Hospital (Turin, Piedmont Region), following these selection criteria: (1) full-term pregnancy (>37 GA-W-); (2) physiological pregnancy conditions; (3) no drugs or pharmacological treatment during pregnancy; (4) single babies (no twins) with Apgar scores > 5; and (5) healthy babies at birth (not admitted to the neonatal intensive care unit or in life-threatening conditions). Each adult subject was informed about the aim of this study and signed their written informed consent. Newborns were enrolled in the study upon written authorization from both parents. Sensitive data were replaced by anonymous identification codes to ensure full privacy

of data. The local Ethics Committee of "A.O.U. Città della Salute e della Scienza" of Turin (22 October 2015, file No. CS/709) approved the study protocol.

2.2. Questionnaire

A standardized questionnaire (PRAMS questionnaire) [22] was administered to the mothers during hospitalization. Questions regarded individual, socio-demographic, and clinical characteristics such as smoking, diet, the mother's working and lifestyle habits, especially in the last month of pregnancy, and education. Infant data, filled in by neonatologists, included gestational age and weight, sex, head circumference, length, the baby's health status at the time of birth, and any medical check-ups.

2.3. Biological Analysis

2.3.1. Urine Sample Collection

A pool of fresh urine was collected from each subject, both for babies and their mothers. Infants' urine samples were collected by means of a specific BP-free polypropylene bag (Urinocol Pediatric, B BRAUN, Milan, Italy) placed inside each newborn's diaper from their birth to their third day of life, during hospitalization. Mothers' urine samples were collected in glass tubes (Pyrex, CORNING, Corning, NY, USA) during hospitalization. All urine samples (infants and mothers) were then stored in glass mini-jars, pre-treated with methanol to reduce the risk of environmental contamination, and stored at −80 °C until analysis.

2.3.2. Free BPA and BPS

Two milliliters of urine was transferred in a tube pre-treated with methanol, vortexed, and acidified with HCl at pH 1. Subsequently, NaCl was added (salting-out process) together with 30 µL of a standard solution containing BPS-d8 and BPA-d16 with a final concentration of 0.10 and 0.05 mg/L, respectively. Next, 750 µL of chloroform and 500 µL of acetone were added for liquid–liquid extraction (LLE). Each sample was vortexed and then sonicated for 1 min. Finally, the sample was centrifuged at 1250× *g* for 20 min at room temperature, and the resulting supernatant was collected and transferred to a new pre-treated vial. LLE was repeated by adding 800 µL of chloroform. All of the supernatant was then brought to dryness by means of a nitrogen stream at room temperature and then resuspended in 100 µL of a solution composed of 5 mM ammonium acetate in ultrapure water (70%) and 5 mM ammonium acetate in acetonitrile (30%).

2.3.3. Total (Free + Conjugated) BPA and BPS

Four milliliters of fresh urine was thawed, vortexed, and aliquoted in two vials: 2 mL was used for the determination of free BP, while the other 2 mL was used for the determination of bisphenol-glucuronides after 12 h of incubation with 20 units of β-glucuronidase/arylsulfatase.

2.3.4. Conjugated BPA and BPS

After 12 h of incubation with the enzyme β-glucuronidase/arylsulfatase, the same extraction procedure was performed as described above.

2.3.5. Analysis

The instrumental analysis for both the free and conjugated procedures was carried out with a UPLC Shimadzu Nexera X2 System interfaced through an ESI source (Turbo Ion Spray™) to a Sciex 5500 QTrap mass spectrometer. The analytes were detected in negative ion mode. Concerning the LC set-up, the chromatographic column consisted of a Phenomenex (Bologna, Italy) Luna Omega C18 (1.6 µm, 100 mm × 2.1 mm), and the mobile-phase solvents for reverse-phase analysis were 5 mM ammonium acetate in water and 5 mM ammonium acetate in acetonitrile. The final pH value of both solvents was corrected to 7.5–8.0 by adding a few drops of 33% ammonia solution to ensure a more

significant presence of BPA and BPS in deprotonated form. Concerning chromatographic gradient, the flow rate was set at 0.35 mL min^{-1}. Mobile phases, initially consisting of 30% of 5 mM ammonium acetate in acetonitrile and 70% of 5 mM ammonium acetate in water, were held for 2 min then increased linearly to 100% organic solvent over 12 min, held there for 3 min, and finally brought back to the initial condition in 0.1 min. Moreover, 7 min of re-equilibration was necessary between samples.

Moving to MS parameters, the drying gas (nitrogen) was set at 325 °C, 20.0 psi, and 10 L min^{-1}; capillary voltage was set at 2000 V. Data acquisition was made in multiple reaction monitoring (MRM) mode by monitoring the transitions of deprotonated ions $[M\text{-}H]^-$. For each analyte, two transitions were monitored: one for quantification and the other for confirmation. All the MS/MS parameters are described in Table 1. Procedural blank samples with ultrapure water in place of urine were collected, extracted, and analyzed by HPLC-MS/MS following the same protocol. In all processed blanks, we did not observe BPA contamination above the method limit of detection (LOD). All solvents and reagents (i.e., acetonitrile, acetone, ammonium acetate, and chloroform) were from VWR International (Radnor, PA, USA). All aqueous solutions were prepared with ultrapure water, Millipore Milli-QTM. Analytical standard compounds were purchased from Sigma-Aldrich (Milan, Italy). The method was based upon previously published methods with slight modifications [23]. To check method performances, we validated our procedure by verifying quality parameters according to Eurachem guidelines [24]. We checked for selectivity, sensitivity, linearity, accuracy, and repeatability. In particular, method sensitivity LOD was 0.0035 ng mL^{-1} for BPA and 0.0030 ng mL^{-1} for BPS.

Table 1. MS/MS parameters for bisphenol A (BPA) and bisphenol S (BPS) analyses.

Analyte	Molecularion (*m/z*)	Fragmention (*m/z*)	Declustering Potential (V)	Entrance Potential (V)	Collision Energy (V)	Collision Cell Exit Potential (V)
BPA	227.0	212.0	−79	−9	−23	−14
	227.0	133.0	−79	−9	−30	−14
BPA-d16	241.0	223.0	−62	−8	−25	−11
	241.0	142.0	−62	−8	−33	−18
BPS	249.0	108.0	−75	−3	−35	−13
	249.0	92.0	−75	−3	−38	−13
BPS-d8	257.0	112.0	−120	−7	−34	−14
	257.0	96.0	−120	−7	−35	−18

2.3.6. Creatinine

Urinary creatinine was determined to normalize the excretion rate of all the aforementioned urinary biomarkers as previously described [25].

2.4. Statistical Analysis

Data were expressed as mean ± SD or counts and percentages. As the statistical distribution of the quantitative parameters was found to be non-Gaussian (Kolmogorov–Smirnov test), non-parametric tests were used to assess between-group differences (Mann–Whitney U-test). For qualitative data, groups were compared using a chi-square or Fisher's exact test, as appropriate. A two-sided p-value < 0.05 was considered to indicate statistical significance. To further analyze possible sources of BP exposure, non-parametric correlations (Spearman correlations) were performed between BP levels and three groups of independent variables: (1) maternal exposure: lifestyle habits in the last months of pregnancy (i.e., daily consumption of food and beverages in plastic packaging or daily use of a microwave and dishwasher), smoking habits, anthropometric characteristics (i.e., weight and BMI composition pre-pregnancy and at the end of pregnancy), age, level of education, and nationality; (2) neonatal hospital procedures, including medical check-ups, oral glucose administration in the first three days of life, and drug administration during delivery (because both were administered from a non-BP-free container); and (3) neonatal exposure:

pacifier and formula administration in the first three days of life. All analyses were carried out including only those subjects with no missing data (i.e., complete case analysis).

3. Results

Table 2 describes the characteristics of the study population (newborns and mothers). For this pilot study, 200 mothers with their babies were enrolled, but only 134 subjects followed all the selection criteria and were eligible. In fact, in addition to the selection criteria, about 33% of enrolled mothers/babies were excluded for three other main reasons:

1. Not-fully completed information (questionnaire and infant data);
2. Insufficient urine samples (mothers: <30 mL/babies: <5 mL);
3. Withdrawal of informed consent.

Table 2. Physical characteristics of the newborn sample at birth and BP (i.e., BPA and BPS) concentrations (part A); Characteristics of mother population (part B).

PART A	NEWBORN GROUP (*n* = 134)		
Sex	male	84 (62.7%)	
	female	50 (37.3%)	
Height at birth (cm) (Mean ± SD)	52.9 ± 1.3		
Weight at birth (kg) (Mean ± SD)	3.3 ± 0.4		
Cranial circumference (cm) (Mean ± SD)	38.3 ± 1.6		
Breast-fed N (%)	111 (82.8%)		
Infant formula dispensed N (%)	36 (26.9%)		
Glucose dispensed N (%)	38 (28.4%)		
Pacifier dispensed N (%)	42 (31.3%)		
Respiratory procedures N (%)	16 (11.9%)		
General medical check-up N (%)	85 (63.4%)		
Diseases at birth N (%)	32 (23.9%)		
PART B	**MOTHER GROUP (*n* = 134)**		
Age (Mean ± SD)	33.9 ± 4.7		
Height (cm) (Mean ± SD)	164.1 ± 7.1		
Weight (kg) (Mean ± SD)	Pre-pregnancy	64.4 ± 13.9	
	End of pregnancy	76.8 ± 13.7	
	Δ	12.3 ± 4.8	
BMI (Mean ± SD)	Pre-pregnancy	23.9 ± 4.9	
	End of pregnancy	28.5 ± 4.8	
Nationality N (%)	European	126 (94%)	
	Others	8 (6%)	
	Low level	20 (14.9%)	
	Medium level	51 (38%)	
	High level	63 (47.1%)	
Occupation N (%)	Yes	114 (85.1%)	
	No	6 (4.5%)	
	Others	14 (10.4%)	
	Rural	40 (29.9%)	
	Suburban	51 (38.1%)	
	Urban	43 (32.1%)	
Smoking habits N (%)	No	98 (73.1%)	
	Passive (Average exposure time: 3 h)	21 (15.7%)	
	Yes	15 (11.2%)	
		<10 cig/die	8 (53.3%)
		>10 cig/die	7 (46.7%)
	Vaginal	81 (60.4%)	
	Caesarian	46 (34.3%)	
	Vacuum	7 (5.2%)	

The newborn sample was homogenous in terms of height, weight, and cranial circumference (Table 2, part A). About 80% of the babies were breastfed, with only 27% of babies (*n* = 36) receiving infant formula as extra integration before being discharged from the hospital. Moreover, about 30% of newborns were administered oral glucose as extra integration and/or pacifiers. Sixteen babies (12%) required respiratory procedures, while 24% of newborns (n = 32) were diagnosed with potentially pathological conditions after general medical check-up. The sample of mothers was homogenous according to age, height, weight, and BMI variables. Among mothers (*n* = 134), 94% were European while 6% were classified as other nationalities (non-European). The education level was low (i.e., primary and middle school) in 14.9% of subjects, while it was medium and high (secondary school and above) in 38% and 47.1% of subjects, respectively. Fifteen mothers (11%) reported being active smokers during pregnancy, whereas twenty-one (15.7%) were passive smokers.

Table 3 reports the BP concentrations in the study population, split up for newborn and mother groups. On average, newborns had BP levels twice as high as mothers, except for

free BP forms (both for free BPA and free BPS). The determination of conjugated BPs was performed by an indirect procedure based on the enzymatic cleavage of the glucuronic acid/sulfuric acid moiety from the phenol group(s) of BPs. A recent report [26] evidenced the possibility of underestimating the real quantity of bound BPs because of different yields of enzymatic reactions together with the chance of the formation of intermediates. In order to evaluate reaction yields, we performed the enzymatic incubation for different times (from 4 to 12 h), confirming that a stable result was reached at 12 h. We also checked for the presence of partially cleaved BP metabolites by high-resolution mass-spectrometry analysis (data not shown), excluding this likelihood.

Table 3. BP levels (BPA and BPS) in the newborn and mother groups.

BP Levels (ng/mL) (Mean ± SD; C.I. Range)	Newborn	Mother	BP Levels (ng/mg $_{CREA}$) (Mean ± SD; I.C. Range)	Newborn	Mother
Total BPA	0.13 ± 0.3 [0.02/0.74]	0.15 ± 0.23 [<LOD/0.62]	Total BPA	0.48 ± 1.13 [0.02/2.5]	0.24 ± 0.43 [<LOD/0.9]
Conjugated BPA	0.11 ± 0.3 [0.01/0.6]	0.13 ± 0.2 [<LOD/0.53]	Conjugated BPA	0.41 ± 1.05 [>0.004/2.3]	0.2 ± 0.35 [<LOD/0.75]
Free BPA	0.01 ± 0.37 [<LOD/0.62]	0.02 ± 0.06 [<LOD/0.13]	Free BPA	0.07 ± 0.09 [<0.01/0.25]	0.03 ± 0.16 [<LOD/0.1]
Total BPS	0.09 ± 0.2 [<LOQ/0.6]	0.01 ± 0.03 [<LOQ/0.06]	Total BPS	0.2 ± 0.53 [<LOD/0.1]	0.04 ± 0.15 [<LOD/0.1]
Conjugated BPS	0.08 ± 0.17 [<LOD/0.5]	0.005 ± 0.04 [<LOD/0.05]	Conjugated BPS	0.15 ± 0.5 [<LOD/0.9]	0.02 ± 0.14 [<LOD/0.04]
Free BPS	0.01 ± 0.04 [<LOD/0.05]	0.004 ± 0.002 [<LOD/0.02]	Free BPS	0.02 ± 0.07 [<LOD/0.01]	0.02 ± 0.06 [<LOD/0.03]
Creatinine (CREA) (mg/L)	MOTHER NEWBORN		0.66 ± 0.5 [0.1/1.9] 0.8 ± 0.4 [0.1/1.5]		

Because the biological variables had non-Gaussian distributions, non-parametric tests (Mann–Whitney test and Spearman's rho correlations) were performed. In the mother group, no BP differences were detected in terms of anthropometric characteristics (i.e., weight and BMI composition pre-pregnancy and at the end of pregnancy), age, level of education, and nationality. Besides, the analysis showed no significant correlation between maternal and neonatal BP concentration levels.

In the newborn group, BP analysis revealed no significant differences (Mann–Whitney U-test) between mothers' lifestyle habits, such as daily consumption of food and beverages in plastic packaging (BPA total p = 0.6; BPS total p = 0.2), daily use of a microwave and dishwasher (BPA total p = 0.5; BPS total p = 0.8), or make-up use (BPA total p = 0.7; BPS total p = 0.08). In newborns, the most important correlation was found between pacifier use and BPS concentrations, in terms of both total BPS (Mann–Whitney U-test, p = 0.004; Spearman's rho = 0.182, p = 0.035—Figure 1A) and free BPS concentrations (Mann–Whitney U-test, p = 0.003; Spearman's rho = 0.190, p = 0.03—Figure 1B).

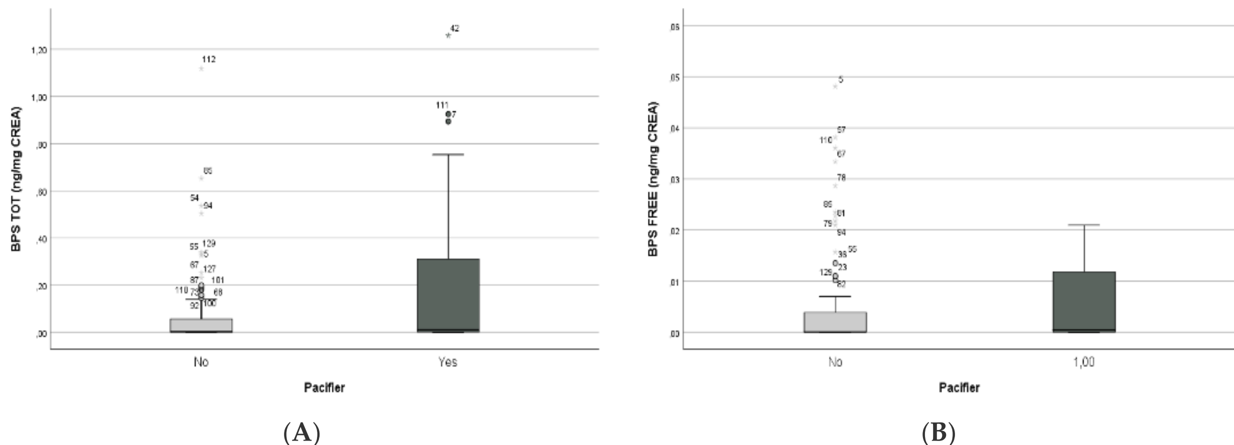


Figure 1. Non-parametric correlations between pacifier use and BPS total (**A**) and free BPS (**B**) concentrations in newborns.

A significant correlation was also found between oral glucose administration and concentration levels of free BPA (Mann–Whitney U-test, $p = 0.003$; Spearman's rho = 0.182, $p = 0.035$—Figure 2).

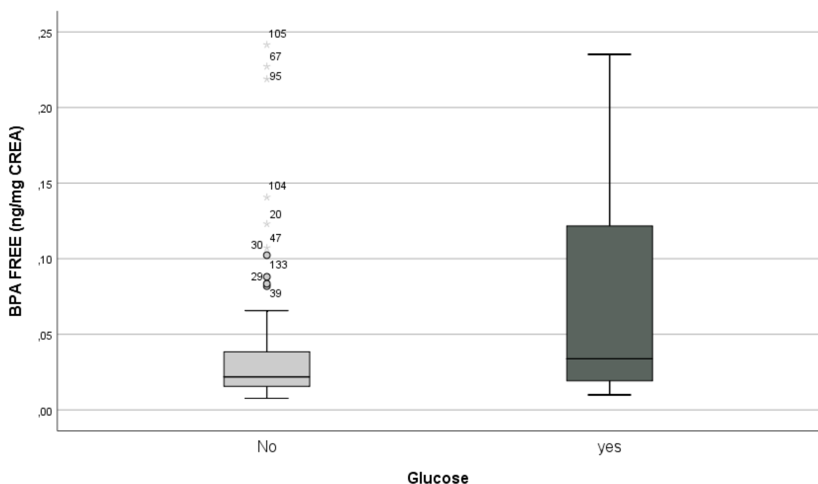
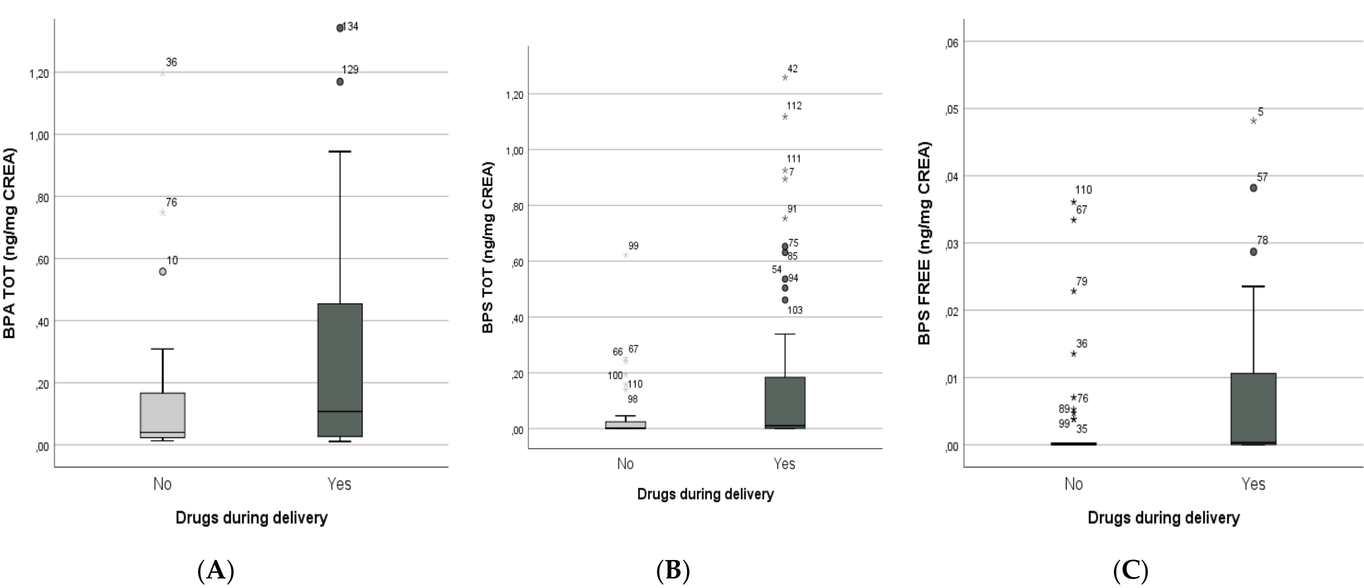


Figure 2. Non-parametric correlations between oral glucose administration and free BPA levels in newborns.

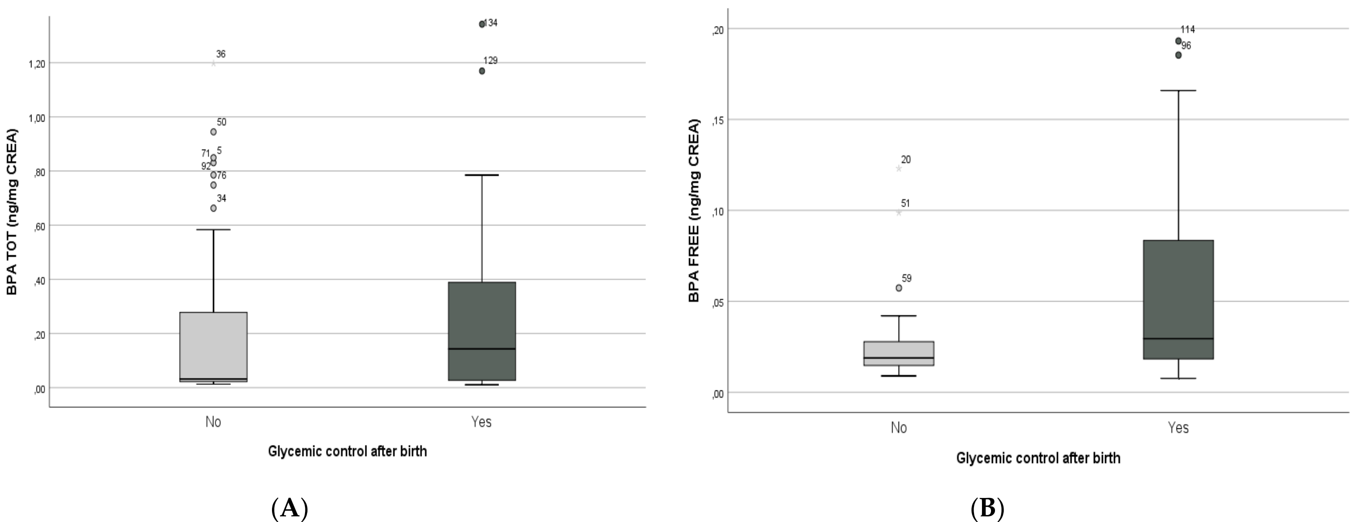
Subsequently, positive correlations were found between drugs administered during delivery and total BPA (Mann–Whitney U-test, $p = 0.05$; Spearman's rho = 0.146, $p = 0.04$—Figure 3A), total BPS (Mann–Whitney U-test, $p = 0.001$; Spearman's rho = 0.198, $p = 0.001$—Figure 3B), and free BPS (Mann–Whitney U-test, $p = 0.05$; Spearman's rho = 0.170, $p = 0.05$—Figure 3C) concentrations.

In further analyses, despite the exclusion of pre-term births from the population study, positive correlations were found between BPA total (Mann–Whitney U-test $p = 0.001$; Spearman's rho = 0.172, $p = 0.001$) and free BPA (Mann–Whitney U-test $p = 0.001$; Spearman's rho = 0.256, $p = 0.003$) concentrations and glycemic control after birth (Figure 4A,B, respectively).



(**A**) (**B**) (**C**)

Figure 3. Correlations between drugs administered during delivery and BPA total (**A**), BPS total (**B**), and free BPS (**C**) concentrations in newborns.



(**A**) (**B**)

Figure 4. Correlations between glycemic control after birth and BPA total (**A**) and free BPA (**B**) concentrations.

4. Conclusions

The toxic properties of BPs are well-known, and their widespread presence in the environment poses a substantial risk to human health. Thus, assessing BP environmental distribution and the possible presence of these compounds in newborns and their mothers' biological fluids represents an urgent need to prevent BP-related diseases. BPs are particularly dangerous for children and pregnant women, who are generally highly vulnerable [13]. In fact, because glucuronidation is less efficient in newborns, their internal dose of BPs may be higher and more persistent compared to the general population. As previously cited, the European Food Safety Authority has recently called for a reduction of BPA in food and drink packaging, recommending the replacement of BPA-containing epoxy resins with plastics containing alternative BPs, such as BPS or BPF. For these compounds, no regulation is available yet, and recent data have shown that BPS may also exert genotoxic and estrogenic activities, similar to BPA [19]. All these lines of evidence converge in the reasonable fear that BPS could rapidly become a major health concern.

In this study, we measured BP concentration levels in a group of newborns and their mothers, and assessed factors contributing to BP risk exposure during prenatal and neonatal life. Based on our findings, we can draw two main conclusions and two strong hypotheses:

Conclusion 1: Neonatal routine procedures may expose newborns to BPs. The observation that newborns receiving oral glucose after birth displayed high BPA levels, especially in its toxic free form, suggests that oral glucose administration through non-BPA-free containers and syringes should be avoided. Besides, drug administration to the mother via non-BPA-free containers and syringes during delivery is also positively correlated with higher levels of both total BPS and BPA concentrations in newborns. Using syringes to administer dextrose solutions or other parenteral products to newborns is a widespread practice, and these results show the risk associated with such a practice [27]. Thus, our findings advance, once again, the need for extending EU regulations to non-alimentary or medical sources of BP contamination. They also call for further studies on BP exposure in newborns in order to identify which devices and procedures are best suited to minimize BP intake in this extremely vulnerable population. In fact, medical devices are a specific product category in which BP may be present, and preventive actions must be adopted to reduce or even eliminate this exposure source [28].

Conclusion 2: Newborn exposure to BPS can still occur through pacifiers. The observation that the use of pacifiers was associated with higher levels of both total and free BPS indicates the urgent need of in-depth analyses and regulations of BPA alternatives such as BPS. The vast majority of regulations have been directed towards BPA-containing plastics, while they have failed to take into account other plastic products containing alternative BPs. In fact, BPS was regarded as a "safe" alternative to BPA because of its greater stability against high temperatures and resistance to sunlight compared to BPA. No previous epidemiological study has explored the developmental effects of BPS, but in vitro studies have shown that BPS can bind to estrogenic receptors (ERs) and drive estrogen-induced gene transcription [29,30]. Furthermore, studies in zebrafish have shown that BPS exposure can alter the homeostasis of sex steroid hormones and disrupt reproduction or development, miming BPA exposure [31,32]. Current knowledge on the impact of BPS exposure is limited, but, on these bases, further studies are advocated to deepen knowledge of human BPS intake and exposure, focusing more on newborn health effects [33].

Hypothesis 1: *Lifestyle habits of the mother in the last month of pregnancy do not influence BP exposure. We showed that the lifestyle habits of pregnant women in the last month of pregnancy did not cause significant changes in newborns' BP levels. In fact, no stratification effect was detected due to lifestyle, food habits, education level, or work. This could be explained by analyzing some pregnancy-related factors, such as more attention in the diet habits of pregnant women (no consumption of plastic-free or precooked food) or lifestyle habits.*

Hypothesis 2: *BP levels could exacerbate health conditions at birth and increase the duration of hospitalization. We showed that newborns with higher levels of free BPA/BPS had higher risk of diseases at birth and higher duration of hospitalization. Further investigation is needed to elucidate if specific conditions are related to this association and how they are related. Nevertheless, BPs' toxic effects may play a role in worsening a newborn's clinical status and recovery time through hormonal/metabolic impairment, resulting in a longer hospital stay.*

Strengths and limitations of this study: Strengths of our study include high-sensitivity biological analyses and the planning and sampling of the subject immediately after birth (within the first three days of the newborn's life). One of the biggest limitations is that we planned a cross-sectional study so we are only able to describe the situation. Besides, the study analyzed a small sample, and the information about lifestyle habits is limited to the last month of pregnancy.

This is the first report of newborn BP contamination due to widely employed non-alimentary products designed for newborn care. Moreover, our study points to a central

role of hospitalization procedures and neonatal devices as primary sources of BP exposure. We also show that lifestyle and alimentary habits, especially in the last month of pregnancy, did not influence BP risk exposure. Despite this, in general, harmful behaviors could negatively impact newborn health conditions, and it is necessary to improve preventive strategies to counteract this trend. Finally, we showed the potential pathological effects of BP exposure on a newborn's clinical status. Further studies are advocated in order to clarify both the impact of other BP forms on humans and the potential sources and consequences of BP exposure during neonatal and childhood life.

Author Contributions: Conception and design, analysis and interpretation of data: V.B., E.C., C.M. and R.B. Drafting the article or revising it critically: V.B., E.C., P.D., M.Z., C.M. and R.B. Approval of the final version: V.B., E.C., R.T., G.S., T.M., S.S., M.Z., P.D., A.C., C.M. and R.B. All authors have read and agreed to the published version of the manuscript.

Funding: All phases of this study were supported by the Cassa di Risparmio di Torino Foundation.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of "A.O.U. Città della Salute e della Scienza" of Turin (22 October 2015, file No. CS/709).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the participant's privacy protection.

Acknowledgments: Thank you to Enrico Bertino for his precious collaboration and for the review of the manuscript. Thank you also to the all the mothers and their babies who accepted the invitation to voluntarily participate in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zalko, D.; Soto, A.M.; Dolo, L.; Dorio, C.; Rathahao, E.; Debrauwer, L.; Faure, R.; Cravedi, J.-P. Biotransformations of bisphenol A in a mammalian model: Answers and new questions raised by low-dose metabolic fate studies in pregnant CD1 mice. *Environ. Health Perspect.* **2003**, *111*, 309–319. [CrossRef]
2. García-Córcoles, M.T.; Cipa, M.; Rodríguez-Gómez, R.; Rivas, A.; Olea-Serrano, F.; Vílchez, J.L.; Zafra-Gómez, A. Determination of bisphenols with estrogenic activity in plastic packaged baby food samples using solid-liquid extraction and clean-up with dispersive sorbents followed by gas chromatography tandem mass spectrometry analysis. *Talanta* **2018**, *178*, 441–448. [CrossRef]
3. Bittner, G.D.; Denison, M.S.; Yang, C.Z.; Stoner, M.A.; He, G. Chemicals having estrogenic activity can be released from some bisphenol a-free, hard and clear, thermoplastic resins. *Environ. Health* **2014**, *13*, 103. [CrossRef]
4. Duty, S.M.; Mendonca, K.; Hauser, R.; Calafat, A.M.; Ye, X.; Meeker, J.D.; Ackerman, R.; Cullinane, J.; Faller, J.; Ringer, S. Potential sources of bisphenol A in the neonatal intensive care unit. *Pediatrics* **2013**, *131*, 483–489. [CrossRef] [PubMed]
5. Mercogliano, R.; Santonicola, S. Investigation on bisphenol A levels in human milk and dairy supply chain: A review. *Food Chem. Toxicol.* **2018**, *114*, 98–107. [CrossRef] [PubMed]
6. Talsness, C.E.; Andrade, A.J.M.; Kuriyama, S.N.; Taylor, J.A.; vom Saal, F.S. Components of plastic: Experimental studies in animals and relevance for human health. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2009**, *364*, 2079–2096. [CrossRef]
7. Geens, T.; Roosens, L.; Neels, H.; Covaci, A. Assessment of human exposure to Bisphenol-A, Triclosan and Tetrabromobisphenol-A through indoor dust intake in Belgium. *Chemosphere* **2009**, *76*, 755–760. [CrossRef] [PubMed]
8. Vandenberg, L.N.; Gerona, R.R.; Kannan, K.; Taylor, J.A.; van Breemen, R.B.; Dickenson, C.A.; Liao, C.; Yuan, Y.; Newbold, R.R.; Padmanabhan, V.; et al. A round robin approach to the analysis of bisphenol a (BPA) in human blood samples. *Environ. Health* **2014**, *13*, 25. [CrossRef] [PubMed]
9. Wilson, N.K.; Chuang, J.C.; Morgan, M.K.; Lordo, R.A.; Sheldon, L.S. An observational study of the potential exposures of preschool children to pentachlorophenol, bisphenol-A, and nonylphenol at home and daycare. *Environ. Res.* **2007**, *103*, 9–20. [CrossRef] [PubMed]
10. Pirard, C.; Sagot, C.; Deville, M.; Dubois, N.; Charlier, C. Urinary levels of bisphenol A, triclosan and 4-nonylphenol in a general Belgian population. *Environ. Int.* **2012**, *48*, 78–83. [CrossRef]
11. Welshons, W.V.; Nagel, S.C.; vom Saal, F.S. Large effects from small exposures. III. Endocrine mechanisms mediating effects of bisphenol A at levels of human exposure. *Endocrinology* **2006**, *147*, S56–S69. [CrossRef]
12. Golub, M.S.; Wu, K.L.; Kaufman, F.L.; Li, L.-H.; Moran-Messen, F.; Zeise, L.; Alexeeff, G.V.; Donald, J.M. Bisphenol A: Developmental toxicity from early prenatal exposurea. *Birth Defects Res. Part B Dev. Reprod. Toxicol.* **2010**, *89*, 441–466. [CrossRef]

13. Nachman, R.M.; Fox, S.D.; Golden, W.C.; Sibinga, E.; Groopman, J.D.; Lees, P.S.J. Serial Free Bisphenol A and Bisphenol A Glucuronide Concentrations in Neonates. *J. Pediatr.* **2015**, *167*, 64–69. [CrossRef]
14. Edginton, A.N.; Ritter, L. Predicting plasma concentrations of bisphenol A in children younger than 2 years of age after typical feeding schedules, using a physiologically based toxicokinetic model. *Environ. Health Perspect.* **2009**, *117*, 645–652. [CrossRef] [PubMed]
15. Ackerman, L.K.; Noonan, G.O.; Heiserman, W.M.; Roach, J.A.; Limm, W.; Begley, T.H. Determination of Bisphenol A in U.S. Infant Formulas: Updated Methods and Concentrations. *J. Agric. Food Chem.* **2010**, *58*, 2307–2313. [CrossRef]
16. Mielke, H.; Gundert-Remy, U. Bisphenol A levels in blood depend on age and exposure. *Toxicol. Lett.* **2009**, *190*, 32–40. [CrossRef]
17. EFSA. Report on the two-phase public consultation on the draft EFSA scientific opinion on bisphenol A (BPA). *EFSA Support. Publ.* **2015**, *12*, 740E. [CrossRef]
18. Moreman, J.; Lee, O.; Trznadel, M.; David, A.; Kudoh, T.; Tyler, C.R. Acute Toxicity, Teratogenic, and Estrogenic Effects of Bisphenol A and Its Alternative Replacements Bisphenol S, Bisphenol F, and Bisphenol AF in Zebrafish Embryo-Larvae. *Environ. Sci. Technol.* **2017**, *51*, 12796–12805. [CrossRef] [PubMed]
19. Feng, Y.; Yin, J.; Jiao, Z.; Shi, J.; Li, M.; Shao, B. Bisphenol AF may cause testosterone reduction by directly affecting testis function in adult male rats. *Toxicol. Lett.* **2012**, *211*, 201–209. [CrossRef] [PubMed]
20. Rochester, J.R.; Bolden, A.L. Bisphenol S and F: A Systematic Review and Comparison of the Hormonal Activity of Bisphenol A Substitutes. *Environ. Health Perspect.* **2015**, *123*, 643–650. [CrossRef]
21. Skledar, D.G.; Schmidt, J.; Fic, A.; Klopčič, I.; Trontelj, J.; Dolenc, M.S.; Finel, M.; Mašič, L.P. Influence of metabolism on endocrine activities of bisphenol S. *Chemosphere* **2016**, *157*, 152–159. [CrossRef] [PubMed]
22. CDC—PRAMS Questionnaires—Pregnancy Risk Assessment Monitoring System—Reproductive Health. Available online: https://www.cdc.gov/prams/questionnaire.htm (accessed on 20 March 2020).
23. Rocha, B.A.; Da Costa, B.R.B.; De Albuquerque, N.C.P.; De Oliveira, A.R.M.; Souza, J.M.O.; Al-Tameemi, M.; Campiglia, A.D.; Barbosa, F. A fast method for bisphenol A and six analogues (S, F, Z, P, AF, AP) determination in urine samples based on dispersive liquid-liquid microextraction and liquid chromatography-tandem mass spectrometry. *Talanta* **2016**, *154*, 511–519. [CrossRef] [PubMed]
24. Magnusson, B.; Örnemark, U. (Eds.) *Eurachem Guide: The Fitness for Purpose of Analytical Methods–A Laboratory Guide to Method Validation and Related Topics*, 2nd ed.; 2014; ISBN 978-91-87461-59-0. Available online: www.eurachem.org (accessed on 20 February 2020).
25. Bono, R.; Bellisario, V.; Tassinari, R.; Squillacioti, G.; Manetta, T.; Bugiani, M.; Migliore, E.; Piccioni, P. Bisphenol a, tobacco smoke, and age as predictors of oxidative stress in children and adolescents. *Int. J. Environ. Res. Public Health* **2019**, *16*. [CrossRef] [PubMed]
26. Gerona, R.; vom Saal, F.S.; Hunt, P.A. BPA: Have flawed analytical techniques compromised risk assessments? *Lancet Diabetes Endocrinol.* **2020**, *8*, 11–13. [CrossRef]
27. Leng, H.; Zheng, X.; Yan, L.; Zhang, X.; He, H.; Xiang, M. Effects of different types and concentration of oral sweet solution on reducing neonatal pain during heel lance procedures. *Chin. J. Pediatr.* **2013**, *51*, 654–658.
28. Bisphenol A: Part 1. Facts and Figures on Human and Environmental Health Issues and Regulatory Perspectives | RIVM. Available online: https://www.rivm.nl/publicaties/bisphenol-a-part-1-facts-and-figures-on-human-and-environmental-health-issues-and (accessed on 11 February 2020).
29. Grignard, E.; Lapenna, S.; Bremer, S. Weak estrogenic transcriptional activities of Bisphenol A and Bisphenol S. *Toxicol. In Vitro* **2012**, *26*, 727–731. [CrossRef]
30. Hashimoto, Y.; Moriguchi, Y.; Oshima, H.; Kawaguchi, M.; Miyazaki, K.; Nakamura, M. Measurement of estrogenic activity of chemicals for the development of new dental polymers. *Toxicol. In Vitro* **2001**, *15*, 421–425. [CrossRef]
31. Ji, K.; Hong, S.; Kho, Y.; Choi, K. Effects of bisphenol s exposure on endocrine functions and reproduction of zebrafish. *Environ. Sci. Technol.* **2013**, *47*, 8793–8800. [CrossRef] [PubMed]
32. Qiu, W.; Zhao, Y.; Yang, M.; Farajzadeh, M.; Pan, C.; Wayne, N.L. Actions of Bisphenol A and Bisphenol S on the Reproductive Neuroendocrine System during Early Development in Zebrafish. *Endocrinology* **2016**, *157*, 636–647. [CrossRef]
33. Wan, Y.; Huo, W.; Xu, S.; Zheng, T.; Zhang, B.; Li, Y.; Zhou, A.; Zhang, Y.; Hu, J.; Zhu, Y.; et al. Relationship between maternal exposure to bisphenol S and pregnancy duration. *Environ. Pollut.* **2018**, *238*, 717–724. [CrossRef]

RESEARCH ARTICLE

Electroclinical Features and Long-term Seizure Outcome in Patients With Eyelid Myoclonia With Absences

Emanuele Cerulli Irelli, MD, Enrico Cocchi, MD, Georgia Ramantani, MD, PhD, Roberto H. Caraballo, MD, Loretta Giuliano, MD, Tulay Yilmaz, MD, Alessandra Morano, MD, Eleni Panagiotakaki, MD, PhD, Francesca F. Operto, MD, Beatriz Gonzalez Giraldez, MD, Katri Silvennoinen, MD, Sara Casciato, MD, PhD, Marion Comajuan, MD, Simona Balestrini, MD, PhD, Francesco Fortunato, MD, Antonietta Coppola, MD, PhD, Giancarlo Di Gennaro, MD, PhD, Angelo Labate, MD, PhD, Vito Sofia, MD, Gerhard J. Kluger, MD, Dorothée G.A. Kasteleijn-Nolst Trenité, MD, PhD, Antonio Gambardella, MD, PhD, Betul Baykan, MD, PhD, Sanjay M. Sisodiya, MD, PhD, Alexis Arzimanoglou, MD, PhD, Pasquale Striano, MD, PhD, and Carlo Di Bonaventura, MD, PhD, on behalf of the EMA Study Group

Correspondence
Dr. Di Bonaventura
c_dibonaventura@yahoo.it

Neurology® 2022;98:e1865-e1876. doi:10.1212/WNL.0000000000200165

Abstract

Background and Objectives

Eyelid myoclonia (EM) with absences (EMA) is a generalized epilepsy syndrome with a prognosis and clinical characteristics that are still partially undefined. We investigated electroclinical endophenotypes and long-term seizure outcome in a large cohort of patients with EMA.

Methods

In this multicenter retrospective study, patients with EMA with ≥5 years of follow-up were included. We investigated prognostic patterns and sustained terminal remission (STR), along with their prognostic factors. Moreover, a 2-step cluster analysis was used to investigate the presence of distinct EMA endophenotypes.

Results

We included 172 patients with a median age at onset of 7 years (interquartile range [IQR] 5–10 years) and a median follow-up duration of 14 years (IQR 8.25–23.75 years). Sixty-six patients (38.4%) displayed a nonremission pattern, whereas remission and relapse patterns were encountered in 56 (32.6%) and 50 (29.1%) participants. Early epilepsy onset, history of febrile seizures (FS), and EM status epilepticus significantly predicted a nonremission pattern according to multinomial logistic regression analysis. STR was achieved by 68 (39.5%) patients with a mean latency of 14.05 years (SD ±12.47 years). Early epilepsy onset, psychiatric comorbid conditions, and a history of FS and generalized tonic-clonic seizures were associated with a lower probability of achieving STR according to a Cox regression proportional hazards model. Antiseizure medication (ASM) withdrawal was attempted in 62 of 172 patients, and seizures recurred in 74.2%. Cluster analysis revealed 2 distinct clusters with 86 patients each. Cluster 2, which we defined as EMA-plus, was characterized by an earlier age at epilepsy onset, higher rate of intellectual disability, EM status epilepticus, generalized paroxysmal fast activity, self-induced seizures, FS, and poor ASM response, whereas cluster 1, the EMA-only cluster, was characterized by a higher rate of seizure remission and more favorable neuropsychiatric outcome.

From the Department of Human Neurosciences (E.C.I., A.M., C.D.B.), Sapienza, University of Rome, Italy; Department of Precision Medicine and Genomics (E.C.), Department of Medicine, Columbia University, New York, NY; Department of Neuropediatrics (G.R.), University Children's Hospital Zurich, Switzerland; Department of Neurology (R.H.C.), Hospital de Pediatría "Prof. Dr. Juan P Garrahan," Buenos Aires, Argentina; Department of Medical and Surgical Sciences and Advanced Technologies "G.F. Ingrassia" (L.G., V.S.), Section of Neurosciences, University of Catania, Italy; Departments of Neurology and Clinical Neurophysiology (T.Y., B.B.), Istanbul University, Istanbul Faculty of Medicine, Turkey; Department of Paediatric Clinical Epileptology, Sleep Disorders and Functional Neurology (E.P., M.C., A.A.), University Hospitals of Lyon (HCL), member of the ERN EpiCARE, France; Child and Adolescent Neuropsychiatry Unit (F.F.O.), Department of Medicine, Surgery and Dentistry, University of Salerno, Italy; Epilepsy Unit (B.G.G.), Neurology Service, Hospital Universitario and IIS Fundación Jiménez Díaz and CIBERER, Madrid, Spain; Department of Clinical and Experimental Epilepsy (S.B., K.S., S.M.S.), UCL Queen Square Institute of Neurology, London, UK; IRCCS NEUROMED (S.C., G.D.G.), Pozzilli, Isernia, Italy; Institute of Neurology (S.B.), UCL, London, UK; Institute of Neurology (F.F., A.L., A.G.), University Magna Graecia, Catanzaro, Italy; Department of Neuroscience and Reproductive and Odontostomatological Sciences (A.C.), Federico II University, Naples, Italy; Clinic for Neuropediatrics and Neurorehabilitation (G.J.K.), Epilepsy Center for Children and Adolescents, Schoen Clinic Vogtareuth, Germany; Department of Neurosurgery and Epilepsy (D.G.A.K.-N.T.), University Medical Center, Utrecht University, the Netherlands; and Pediatric Neurology and Muscular Diseases Unit (P.S.), IRCCS "Istituto Giannina Gaslini," Genova, Italy.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

EMA Study Group coinvestigators are listed in the Appendix 2 at the end the article.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Glossary

ASM = antiseizure medication; **ECS** = eye closure sensitivity; **EM** = eyelid myoclonia; **EMA** = EM with absences; **FS** = febrile seizures; **GGE** = genetic generalized epilepsies; **GPFA** = generalized paroxysmal fast activity; **GTCS** = generalized tonic-clonic seizures; **HR** = hazard ratio; **ID** = intellectual disability; **IQR** = interquartile range; **JME** = juvenile myoclonic epilepsy; **LEV** = levetiracetam; **LTG** = lamotrigine; **OR** = odds ratio; **PS** = photosensitivity; **PWD** = polyspike-wave discharges; **STR** = sustained terminal remission; **SWD** = spike-wave discharge; **TSCA** = 2-step cluster analysis; **VPA** = valproate.

Discussion

Early epilepsy onset was the most relevant prognostic factor for poor treatment response. A long latency between epilepsy onset and ASM response was observed, suggesting the effect of age-related brain changes in EMA remission. Last, our cluster analysis showed a clear-cut distinction of patients with EMA into an EMA-plus insidious subphenotype and an EMA-only benign cluster that strongly differed in terms of remission rates and cognitive outcomes.

In 1977, Jeavons et al.[1] originally described an epileptic condition characterized by marked photosensitivity (PS), eye closure sensitivity (ECS), and absences associated with eyelid myoclonia (EM). After this first report, several authors expanded the electroclinical description of Jeavons syndrome, also known as EM with absences (EMA).[2-4] However, EMA has not been recognized by the International League Against Epilepsy as a distinct epilepsy syndrome[5] because the features described above may be found across a range of different epilepsy syndromes, including genetic generalized epilepsies (GGE), focal genetic photosensitive epilepsies, structural epilepsies, and genetic epileptic encephalopathies.[6-10] Conversely, many authors have recognized EMA as a unique nosologic entity due to its specific electroclinical features and genetic studies reinforcing the differences between EMA and other GGE syndromes.[11-15]

As a result of these discrepancies, heterogeneous diagnostic criteria have been used across different studies to describe the electroclinical features and prognostic characteristics of EMA.[11,16] In particular, the inclusion of patients showing myoclonia in body regions other than the eyelids by many authors may have led to the inclusion of patients with juvenile myoclonic epilepsy (JME), who share several clinical features with patients with EMA, including ECS, PS, and EM.[16,17] Moreover, the majority of existing studies focusing on EMA have been conducted in small patient cohorts, leading to uncertainties regarding the true prognostic trajectories of these patients and their predictive factors. Even with strict diagnostic criteria, significant clinical heterogeneity could be observed across patients with EMA.[18] Indeed, the age at onset may vary from early infancy to early adolescence, and the presence and degree of intellectual disability (ID) vary across patients with EMA.[19,20] Others described a homogeneous group of patients with EMA showing antiseizure medication (ASM) refractoriness, high rates of EM status epilepticus, and ID, suggesting the existence of distinct subphenotypes within the EMA spectrum.[21] However, the existence of distinct EMA subgroups has not yet been investigated by modern statistical clustering approaches as applied to other neurologic diseases.

The main objective of this multicenter study was to investigate seizure outcome and prognostic factors in a large cohort of patients with well-defined EMA during a long-term follow-up. In addition, we used a cluster analysis approach to define different EMA subphenotypes corresponding to distinct prognostic trajectories.

Methods

Study Participants, Setting, and Eligibility Criteria

The study was conducted according to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines as a retrospective multicenter cohort study. Data from patients followed up from 1983 to 2020 at 14 different pediatric and adult specialized epilepsy outpatient clinics, most of them members of the European Reference Network for Rare and Complex Epilepsies, were retrospectively reviewed.

Patients were enrolled according to the following inclusion criteria: (1) history of EM with or without absences; (2) history of PS or ECS; (3) EEG generalized spike-wave discharges (SWDs) and/or polyspike-wave discharges (PWDs); (4) absence of spontaneous or provoked myoclonia in body parts other than the eyelids; (5) normal neuroimaging (when available) and neurologic examination; and (6) follow-up for at least 5 years. We excluded patients with (1) cognitive deficits other than borderline intellectual functioning and mild ID to minimize the risk of including patients with clear-cut epileptic/developmental encephalopathy[11] and (2) myoclonic jerks in body parts other than the eyelids to avoid including patients with JME.

Clinical Data Collection and EEG Assessment

Clinical charts were thoroughly reviewed for demographic data, family history of epilepsy, history of febrile seizures (FS), age at epilepsy onset, seizure types throughout the epilepsy course, occurrence of EM status epilepticus and self-induced seizures, drug regimen changes, MRI findings (when available),

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

psychiatric comorbid conditions, and follow-up duration. Follow-up information on seizure type(s), frequency, and treatment adherence was reviewed for each visit. The presence of borderline intellectual functioning or mild ID, as established by at least 1 standardized neuropsychological test, was noted for each patient.

Standard EEGs were reviewed to assess the following features: (1) background activity; (2) presence and characteristics of ECS and PS; (3) SWD and PWD occurrence and frequency; (4) presence of focal epileptiform abnormalities, defined as focal discharges confined to a single lobe; (5) asymmetry of SWDs or PWDs in both onset and amplitude; and (6) presence of focal slow waves. Sleep EEG recordings, if available, were reviewed to assess the presence of generalized paroxysmal fast activity (GPFA), defined as a generalized discharge of rhythmic polyspikes in beta frequency with a duration of at least 1 second.[22]

Clinical Outcomes

Different seizure outcome measures were assessed during follow-up in each patient. The primary endpoint was the occurrence of sustained terminal remission (STR), defined as a period of at least 4 consecutive years of freedom from all seizures at the last follow-up visit. The time from the first ASM trial to STR was also obtained for each patient, corresponding to the time period from the first ASM trial to the last seizure before STR started. The occurrence of a 2-year remission from all seizure types during clinical history was also considered. Patients who did not achieve at least a 2-year remission during their history were considered to show a nonremission pattern. When at least a 2-year remission was achieved, 2 distinct patterns of seizure control, namely a relapse and remission pattern, were distinguished according to the occurrence or absence of subsequent seizure relapses during follow-up. The time period to the first 2-year remission from the patient's history was also calculated for each patient to investigate the latency from the first ASM prescription to the initial medication response. Last, the occurrence of a 2-year remission from generalized tonic-clonic seizures (GTCS) at the last follow-up visit was evaluated.

In addition, we noted the number of ASM trials during the disease course and the number of ASMs at the last follow-up visit. The recurrence of seizures after ASM withdrawal was also investigated, considering only patients with a follow-up of at least 12 months after ASM discontinuation.

Cluster Analysis

The 2-step cluster analysis (TSCA) approach was used to investigate the presence of distinct EMA endophenotypes and to identify the electroclinical features characterizing these endophenotypes. TSCA is a hybrid cluster approach that performs group clusterization through a double-step procedure. It first separates groups with the use of a distance measure and then chooses the optimal subgroup model through a probabilistic approach. This approach provides several advantages over more traditional clustering techniques because it permits the use of both categorical and continuous variables, the handling of outliers, and the selection of the number of clusters based on statistical measurements rather than arbitrary choice; in addition, it is highly reliable and reproducible.[23,24] The following variables were used to perform TSCA: (1) the presence of mild ID or borderline intellectual functioning; (2) a family history of epilepsy in first- or second-degree relatives; (3) early-onset EMA (as defined below); (4) a history of GTCS; (5) a history of EM status epilepticus; and (6) prognostic patterns (i.e., remission, relapse, and nonremission), as defined above.

Statistical Analysis

Each variable distribution was graphically analyzed to select the appropriate statistical tests and to ensure the highest possible reliability of identified results. Among all variables, the distribution of the age at onset variable was tested and graphically analyzed, resulting in a nonnormally distributed variable. Distribution analysis showed a multimodal pattern that was further analyzed through kernel density estimation to identify underlying modes. Subsequently, the Fisher-Jenks optimization algorithm was used to confirm kernel density estimation intervals and to identify the best cutoff for categorization of the variable, which was determined to be 8.5 years. Early-onset EMA was therefore defined as EMA with an age at seizure onset ≤8 years (the main statistical analysis was also repeated with age at onset as a continuous variable, which yielded comparable results that are reported in eTable 1, links.lww.com/WNL/B882). Categorical variables were compared through the Fisher exact test, while continuous variables were compared by se of the Wilcoxon-Mann-Whitney test due to their nonnormal distribution. Group tests were 2 sided, with values of $p < 0.05$ considered statistically significant.

Kaplan-Meier estimates were performed to calculate the cumulative time-dependent probability of entering STR during follow-up. The time of entry into the analysis was the date of epilepsy diagnosis, and the time of the endpoint was the date of STR onset or the date of the last follow-up visit (depending on which occurred first), truncated at 40 years of follow-up. Cox proportional hazards model was used to investigate the association between STR occurrence and possible predictors according to previous studies. Results were presented as hazard ratios (HRs) with 95% CIs.

Multivariable multinomial logistic regression analysis was used to assess the relation between prognostic patterns (dependent variables) and their possible clinical predictors with the remission pattern used as a reference. Results were presented as odds ratios (ORs) with 95% CIs. Last, a linear regression model was used to assess the relation between the number of ASMs at the last follow-up visit (dependent variable) and its possible clinical predictors.

Data Availability

Deidentified data are available on reasonable request.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Standard Protocol Approvals, Registrations, and Patient Consents

The institutional/regional ethical committee approved the study (5970, 286/2020), and informed consent was obtained from all participants.

Results

General Clinical Features of the Study Cohort

After identification of 301 potential patients with EMA, 172 individuals (123 female, 71.5%) were included according to study criteria (the inclusion tree is presented in eFigure 1, links.lww.com/WNL/B882). The median age at epilepsy onset was 7 years (interquartile range [IQR] 5–10 years), and the median follow-up duration was 14 years (IQR 8.3–23.8 years). A history of psychiatric comorbid conditions was found in 45 patients, among whom 18 of 45 (40%) were diagnosed with mood disorders, 23 of 45 (51.1%) with behavioral disorders, and 4 (8.9%) with psychotic disorders. Descriptive statistics of the cohort with main clinical and demographic data are summarized in Table 1.

Table 1 Demographic and Clinical Characteristics

Age, median (IQR), y	22 (17–32)
Sex, female, n (%)	123 (71.5)
Age at epilepsy onset (IQR), y	7 (5–10)
Follow-up duration, median (IQR), y	14 (8.25–23.75)
History of FS, n, %	19 (11)
Family history of epilepsy in a first-degree relative, n (%)	35 (20.3)
Family history of epilepsy in a second-degree relative, n (%)	21 (12.2)
Family history of epilepsy in a third-degree relative, n (%)	13 (7.6)
Family history of FS in a first or second-degree relative, n (%)	6 (3.5)
Psychiatric comorbid conditions, n (%)	45 (26.2)
Borderline IF or mild ID, n (%)	55 (32)

Abbreviations: FS = febrile seizures; ID = intellectual disability; IF = intellectual functioning; IQR = interquartile range.

Electroclinical Characteristics

EM with or without absences was the seizure type at epilepsy onset in 131 patients (76.2%), whereas GTCS was reported as the presenting seizure type in 41 (23.8%). All patients fulfilled the criteria of either PS or ECS. A history of either PS or ECS was found in 156 (90.7%) and 134 patients (77.9%), respectively, and a history of both PS and ECS was observed in 117 (68%) patients. During follow-up, 120 (69.8%) patients experienced GTCS at least once, and 22 (12.8%) had a history of EM status epilepticus. The occurrence of self-induced seizures during history was found in 17 patients (9.9%), whereas a clear-cut catamenial worsening of EM or GTCS was reported in 15 of 123 (12.2%) female patients.

All but 6 patients showed spontaneous SWD/PWD during at least 1 standard EEG, whereas generalized discharges were provoked only by intermittent photic stimulation or eye closure in these 6 patients. SWDs were recorded in 144 (83.7%) individuals, while PWDs were recorded in 131 (76.2%). SWD/PWD frequency was ≥4 Hz in 110 (64%) patients. Focal spike or sharp waves were reported in 36 individuals (20.9%), and asymmetric/asynchronous generalized discharges were found in 10 (5.8%). A total of 159 of 172 (92.4%) patients performed at least 1 sleep EEG during follow-up, and 8 of 159 (8.8%) were found to have GPFA during sleep.

At the last year of follow-up, ECS persisted in 73 of 158 (46.2%) patients among whom this information was available, whereas PS was found in 81 of 161 (50.3%) patients.

ASM Treatment

The most common first-line ASM was valproate (VPA) in 108 of 172 patients (62.8%), followed by levetiracetam (LEV) in 19 (11%), ethosuximide in 16 (9.3%), and lamotrigine (LTG) in 8 (4.6%). During follow-up, the median number of prescribed ASMs was 3 (IQR 2–4). At the last follow-up visit, all but 16 patients were on ASMs. The median number of ASMs used at the last follow-up was 1 (IQR 1–2, range 1–5), and 78 of 172 (45.3%) patients were on a polytherapy regimen (≥2 ASMs). The most used ASM at the last follow-up visit was VPA in 95 of 172 patients (55.2%), followed by LEV in 58 (33.7%) and LTG in 36 (20.9%). The most frequently used monotherapies at the last follow-up visit were VPA in 41 patients, LEV in 17, and LTG in 13, which were associated with the following 2-year remission rates: 68.3%, 77.8%, and 46.2%, respectively. Among those on a bitherapy regimen (63 patients), the most frequently observed combination was VPA + LEV (13 of 63), which was associated with the highest 2-year remission rate (61.5% vs 36%, $p = 0.1$). ASMs used at the last follow-up visit with the respective 2-year remission rate are given in eFigure 2 (links.lww.com/WNL/B882).

Seizure Outcome and Prognostic Factors

During follow-up, 106 of 172 (61.6%) patients achieved at least a 2-year remission from all seizure types, and the mean time from epilepsy onset to the first 2-year remission was 10.45 years (SD ±10.89 years). Therefore, 66 of 172 (38.4%) individuals displayed a nonremission pattern, whereas 56 (32.6%) and 50 (29.1%) patients showed a remission and relapse pattern of seizure control, respectively. Multivariable multinomial logistic regression analysis showed that a longer follow-up duration (OR 1.04, 95% CI 1.01–1.08, $p = 0.02$), a history of GTCS (OR 3.15, 95% CI 1.05–9.43, $p = 0.04$), and a family history of epilepsy (OR 3.11, 95% CI 1.22–7.94, $p = 0.02$) were associated with a relapse pattern of seizure control, whereas early epilepsy onset (OR 4.88, 95% CI 1.82–12.98, $p = 0.002$), EM status epilepticus (OR 5.05, 95% CI 1.24–20.8, $p = 0.02$), and a history of FS (OR 9.01, 95% CI

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Table 2 Predictors of Relapse and Nonremission Patterns as Determined by Multivariable Multinomial Logistic Regression

	Relapse			No remission		
Predictor	OR	95% CI	*p* Value	OR	95% CI	*p* Value
Male sex	0.5	0.17–1.41	0.2	0.68	0.26–1.79	0.4
Early-onset epilepsy	2.42	0.91–6.41	0.08	4.88	1.82–12.98	0.002[a]
Follow-up duration	1.04	1.01–1.08	0.02	0.98	0.94–1.02	0.3
FS	3.6	0.55–23.25	0.2	9.01	1.67–47.61	0.01[a]
Family history of epilepsy in a first- or second-degree relative	3.11	1.22–7.94	0.02[a]	1.29	0.51–3.27	0.6
Borderline IF or mild ID	2.05	0.73–5.78	0.2	1.06	0.39–2.87	0.9
Psychiatric comorbid conditions	1.73	0.58–5.18	0.3	2.28	0.81–6.41	0.1
EM status epilepticus	1.86	0.37–9.43	0.5	5.05	1.24–20.8	0.02[a]
Self-induced seizures	0.72	0.11–4.48	0.7	2.39	0.55–10.41	0.2
History of GTCS	3.15	1.05–9.43	0.04[a]	2.19	0.85–5.61	0.1
History of both ECS and PS	1.71	0.64–4.57	0.3	1.65	0.64–4.22	0.3
EEG focal abnormalities	1.83	0.69–4.85	0.2	1.68	0.63–4.08	0.3

Abbreviations: ECS = eye closure sensitivity; EM = eyelid myoclonia; FS = febrile seizures; GTCS = generalized tonic-clonic seizures; ID = intellectual disability; IF = intellectual functioning; OR = odds ratio; PS = photosensitivity.
[a] Statistically significant variables ($p < 0.05$).

1.67–47.61, $p = 0.01$) significantly predicted the nonremission pattern (Table 2 shows detailed multivariable multinomial logistic regression results).

STR was achieved in 68 (39.5%) patients, and mean time from epilepsy onset to STR was 14.05 years (SD ±12.47 years). Early epilepsy onset (HR 0.41, 95% CI 0.24–0.70, $p < 0.001$), a history of GTCS (HR 0.47, 95% CI 0.27–0.82, $p = 0.008$), psychiatric comorbid conditions (HR 0.34, 95% CI 0.16–0.71), and a history of FS (HR 0.18, 95% CI 0.05–0.76) were significantly associated with a lower chance of entering STR according to the Cox proportional hazards

Table 3 Prognostic Factors for STR According to Multivariable Cox Proportional Hazards Model

Predictors	**HR**	**95% CI**	***p* Value**
Female sex	0.81	0.46–1.44	0.5
Early onset of epilepsy	0.41	0.24–0.7	<0.001[a]
History of FS	0.17	0.05–0.76	0.02[a]
Family history of epilepsy in a first- or second-degree relative	0.72	0.42–1.21	0.2
Borderline IF or mild ID	0.95	0.53–1.72	0.9
Psychiatric comorbid conditions	0.34	0.16–0.71	0.004[a]
EM status epilepticus	0.54	0.22–1.28	0.2
Self-induced seizures	0.71	0.25–2.04	0.5
GTCS during history	0.47	0.27–0.82	0.008[a]
History of both PS and ECS	1.15	0.65–2.02	0.6
EEG focal abnormalities	0.97	0.56–1.68	0.9

Abbreviations: ECS = eye closure sensitivity; EM = eyelid myoclonia; FS = febrile seizures; GTCS = generalized tonic-clonic seizures; HR = hazard ratio; ID = intellectual disability; IF = intellectual functioning; PS = photosensitivity; STR = sustained terminal remission.
[a] Statistically significant variables ($p < .05$).

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

model. Results of the Cox proportional hazards model are reported in Table 3, and the cumulative probability curves of significant prognostic factors are illustrated in Figure 1. At the last follow-up visit, 88 of 120 (73.3%) patients had achieved 2-year freedom from GTCS.

The persistence of ECS at the last medical observation was associated with significantly lower rates of 2-year remission at the last follow-up visit (25 of 73 vs 53 of 85, $p = 0.001$). In addition, the persistence of PS was associated with lower rates of 2-year remission at the last medical observation (30 of 81 vs 49 of 80, $p = 0.006$).

ASM withdrawal was attempted in 62 of 172 (36%) patients; seizure freedom at least 1 year after ASM discontinuation was observed in 16 of 62 patients (25.8%). Among those who displayed seizure relapse after at least 1 year of ASM withdrawal (46 of 62, 74.2%), GTCS relapse was observed in 28 patients (28 of 35, 80%) with a history of GTCS. Patients with seizure relapse after ASM withdrawal had significantly higher rates of GTCS during their history (36 of 47 vs 5 of 16, $p = 0.002$), whereas no significant differences were found according to other variables.

When we considered the number of ASMs at the last follow-up visit as a dependent variable, early epilepsy onset ($\beta = 0.20$, $p = 0.009$) and history of GTCS ($\beta = 0.17$, $p = 0.02$) were significantly associated with the use of a higher number of ASMs at the last follow-up visit according to a multiple linear regression model ($F = 4.5$, $p < 0.001$). The results of multiple linear regression are reported in eTable 2 (links.lww.com/WNL/B882).

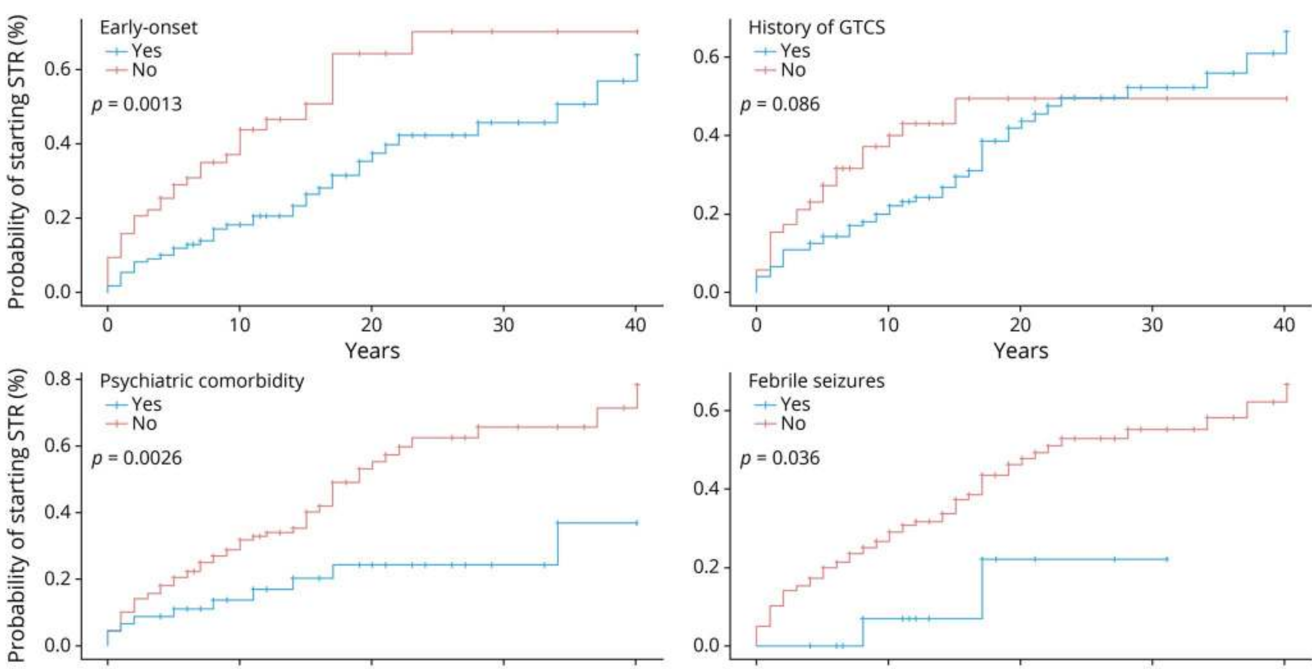
Cluster Analysis: Identification of Clinical EMA Subtypes

TSCA revealed 2 distinct clusters (86 patients per group) of patients with EMA with similar follow-up duration. The 2 clusters, hereinafter referred to as EMA-only (cluster 1) and EMA-plus (cluster 2), significantly differed in terms of age at epilepsy onset and cognitive abnormalities, with the latter showing a younger age at epilepsy onset and a higher percentage of ID/borderline intellectual functioning (16.3% vs 47.7%, $p < 0.001$). In addition, EMA-plus patients were characterized by a higher proportion of FS (5.8% vs 16.3%, $p = 0.049$), self-induced seizures (4.6% vs 15.1%, $p = 0.03$), and EM status epilepticus (4.6% vs 20.9%, $p = 0.002$).

The 2 clusters had similar rates of mood disorders (EMA-only 11.6% vs EMA-plus 10.5%, $p = 0.6$) and psychotic disorders (1.2% vs 3.5%, $p = 0.3$), whereas EMA-plus patients had higher rates of behavioral disorders (8.1% vs 17.4%, $p = 0.07$) compared with EMA-only patients.

As far as EEG characteristics, EMA-plus patients were found to have higher rates of PS persistence at the last year of follow-up (37.2% vs 57%, $p = 0.02$), as well as higher rates of ECS at

Figure 1 Prognostic Factors of STR During Follow-up



Follow-up was truncated at 40 years. Censored patients are indicated by crosses. GTCS = generalized tonic-clonic seizures; STR = sustained terminal remission.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

the last medical observation (32.6% vs 52.3%, p = 0.01), compared with EMA-only patients. In addition, GPFA during sleep was significantly more frequent among EMA-plus patients compared with EMA-only patients (3.5% vs 12.8%, p = 0.02), with a similar proportion of patients undergoing sleep EEG recordings during follow-up in the 2 clusters (95.3% vs 89.5%, p = 0.3).

In terms of seizure outcome, EMA-plus patients showed a significantly higher rate of nonremission pattern (0 vs 76.7%, $p < 0.001$), a similar rate of relapse pattern (EMA-only 31.4% vs EMA-plus 23.3%, p = 0.3), and a significantly lower rate of remission pattern (68.6% vs 0, $p < 0.001$) compared with EMA-only patients.

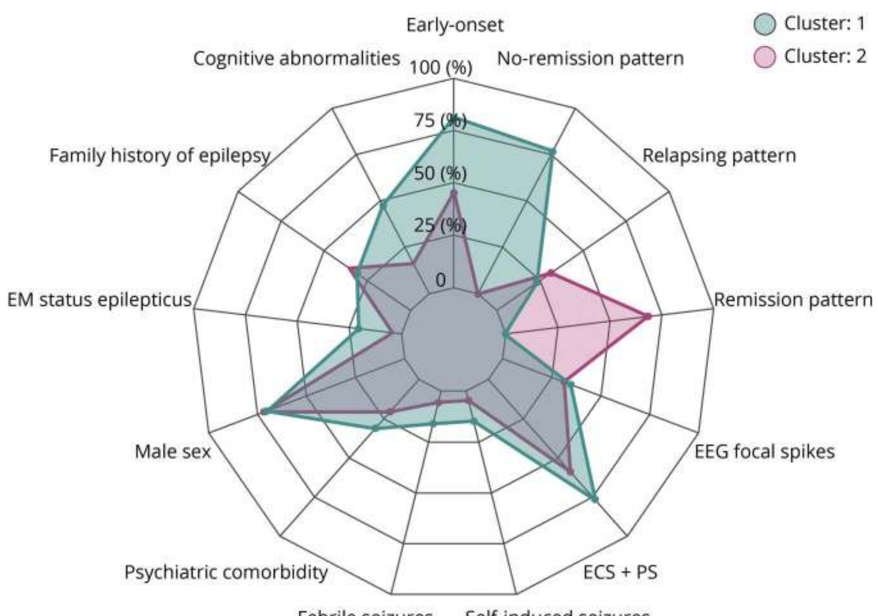
The 2 clusters did not significantly differ in terms of sex, family history of epilepsy in first- or second-degree relatives, or GTCS history. The electroclinical differences between the 2 clusters are illustrated in Figure 2, and all statistics and p values related to comparisons between clusters are reported in Table 4. Last, the 2 clusters significantly differed in terms of probability of entering STR during the follow-up (log-rank test, $p < 0.001$), as illustrated in Figure 3.

Discussion

In this multicenter study, we evaluated the electroclinical characteristics and determined the prognostic factors for distinct epilepsy evolution patterns in a cohort of 172 patients with EMA with a long-term follow-up. More than one-third of our patients displayed a nonremission pattern, whereas remission and relapse patterns were found at almost equal rates in the remaining participants.

Only 39.5% of our population achieved STR, with a median latency of 14.05 years. A similarly long delay was also observed when we considered the interval from epilepsy onset to the initial medication response (time from onset to first 2-year remission 10.45 years), suggesting a key role of age-related brain changes, as previously hypothesized for other photosensitive epilepsies.[25-27] Among the investigated prognostic factors, early epilepsy onset was the most powerful predictor in our study and was significantly associated with both failure to reach STR and a nonremission pattern of seizure control. Our observation is in line with previous findings in a much smaller subgroup (9 patients), which found treatment refractoriness in all patients with early-onset EMA.[28] A history of FS was also significantly associated both with not converting to STR and with a nonremission pattern. The negative effect of FS on long-term seizure outcome was also recently highlighted in a GGE cohort and was attributed to genetic factors that may predispose patients to both FS and ASM refractoriness.[29,30] In accordance with this hypothesis, a previous family study revealed that generalized epilepsy with FS-plus was common among relatives of patients with EMA, suggesting shared genetic determinants between these 2 syndromes.[14] Furthermore, a history of GTCS and psychiatric comorbid conditions significantly predicted failure to achieve STR, in line with previous observations across different epilepsy syndromes.[31-33] Last, a history of EM status epilepticus was associated with a 5-fold increased risk of not experiencing remission throughout the course of EMA. This observation, together with the prognostic effect of an earlier age at epilepsy

Figure 2 Electroclinical Characteristics of Clusters



Radar plot showing the electroclinical differences between clusters. Cluster 1 refers to EMA-only patients; cluster 2 refers to EMA-plus patients. ECS = eye closure sensitivity; EM = eyelid myoclonia; EMA = EM with absences; PS = photosensitivity.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Table 4 Patient Clinical Characteristics Stratified According to Cluster

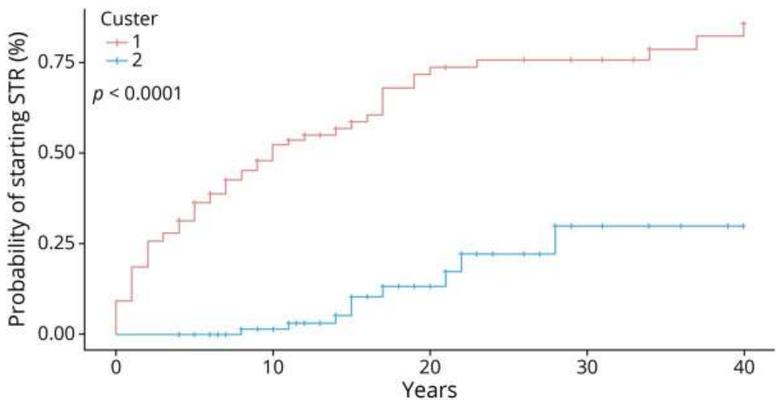
Variable	Cluster 1 (86 patients)	Cluster 2 (86 patients)	*p* Value
Sex, female, n (%)	62 (72.1)	61 (70.9)	1
Early epilepsy onset, n (%)	39 (45.3)	70 (81.4)	<0.001[a]
Age at epilepsy onset, median (IQR), y	9 (6–10.7)	6 (4–8)	<0.001[a]
Follow-up duration, median (IQR), y	15 (8–25.7)	14 (9.2–21)	0.46
History of FS, n, %	5 (5.8)	14 (16.3)	0.049[a]
Family history of epilepsy in a first- or second-degree relative, n (%)	30 (34.9)	27 (31.4)	0.7
Psychiatric comorbid conditions, n (%)	18 (20.9)	27 (31.4)	0.16
Borderline IF or mild ID, n (%)	14 (16.3)	41 (47.7)	<0.001[a]
EM status epilepticus, n (%)	4 (4.6)	18 (20.9)	0.002[a]
Self-induced seizures, n (%)	4 (4.6)	13 (15.1)	0.03[a]
History of GTCS, n (%)	61 (70.9)	59 (68.6)	0.87
History of both PS and ECS, n (%)	51 (59.3)	66 (76.7)	0.02[a]
Focal EEG abnormalities, n (%)	27 (31.4)	30 (34.9)	0.75
Remission pattern, n (%)	59 (68.6)	0	<0.001[a]
Relapsing pattern, n (%)	27 (31.4)	20 (23.3)	0.3
Nonremission pattern, n (%)	0	66 (76.7)	<0.001[a]

Abbreviations: ECS = eye closure sensitivity; EM = eyelid myoclonia; FS = febrile seizures; ID = intellectual disability; IF = intellectual functioning; IQR = interquartile range; PS = photosensitivity.
[a] Statistically significant variables ($p < 0.05$).

onset and a history of FS, may reflect shared underlying genetic components, as supported by the results of our cluster analysis.

In this study, we confirmed the existence of a subgroup of patients with EMA with an insidious phenotype, referred to as EMA-plus, and another more benign subgroup, referred to as EMA-only. Subgroups of patients with EMA characterized by a higher rate of moderate ID, status epilepticus, and ASM resistance have previously been described in small cohorts,[21,28] but these observations have not yet been corroborated in larger cohorts with modern statistical approaches. The 2 EMA patient subgroups, as delineated here, differ to a great extent in terms of both their electroclinical features and long-term seizure outcomes. EMA-plus patients were younger at epilepsy onset and had higher rates of

Figure 3 Probability of Entering STR During Follow-up Depending on the Cluster



Follow-up was truncated at 40 years. Censored patients are indicated by crosses. Cluster 1 refers to EMA-only patients, cluster 2 refers to EMA-plus patients. EMA = eyelid myoclonia with absences; STR = sustained terminal remission.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

cognitive disturbances, EM status epilepticus, FS, GPFA, and self-induction compared with EMA-only patients. In addition, EMA-plus patients showed higher rates of poor response to ASMs, whereas EMA-only patients showed a favorable long-term seizure outcome, with two-thirds of patients achieving a remission pattern of seizure control.

EMA has generally been recognized as an epilepsy syndrome with a high rate of ASM refractoriness regardless of the electroclinical characteristics of affected patients.[34,35] Using cluster analysis, we have made a clear-cut distinction between EMA patient subtypes, with EMA-plus patients having a poor response to ASM and thus a less favorable long-term seizure outcome than EMA-only patients. Further studies will clarify whether the differences between EMA subtypes may be attributed to the underlying genetic substrate, with EMA-plus patients possibly harboring pathogenic variants in genes related to EMA and EMA-like phenotypes such as *SYNGAP1*, *KIA02022*, and *CHD2*, which have been established as the most consistent genetic contributors in this setting.[36-39]

In this study, we also explored ASM withdrawal in patients with EMA. One-third of patients with EMA in our cohort discontinued ASMs during follow-up, in line with previous studies in patients with JME,[40] and one-fourth of these patients remained seizure-free after ASM discontinuation. A history of GTCS emerged as the only predictor of seizure recurrence in our study, suggesting caution when ASMs are withdrawn in these patients. However, due to the retrospective nature of our study, we were unable to quantify GTCS before ASM withdrawal. This potential limitation prevented us from determining whether a single GTCS during history could have the same prognostic significance as multiple GTCS on seizure recurrence after ASM withdrawal.

In addition, we documented an EMA onset peak during mid-childhood, as well as a female preponderance (2.51:1) and high rates of family history of epilepsy, thus providing solid evidence in support of previous findings from much smaller cohorts.[2,11,28,35,41]

As far as treatment data, the combination of VPA + LEV was most frequently associated with 2-year remission at the last follow-up visit; both of these ASMs were also associated with the highest remission rates when used as monotherapy, in line with previous literature findings.[42] VPA was also found to be the most frequently prescribed ASM during the entire follow-up duration, followed by LEV. According to our data, the decreased VPA use in female patients of childbearing age due to its well-known teratogenic adverse effects may eventually result in higher rates of seizure refractoriness in patients with EMA (as observed in other GGEs),[43] especially considering the striking female preponderance observed in this rare epilepsy syndrome.

The main limitations of our study are due to its retrospective design and the lack of systematic genetic testing in all patients, which may have contributed to the interpretation of our findings about prognostic factors and subphenotypes. However, the multicenter design, the large number of patients compared to previous cohorts, the long-term follow-up, and the strict diagnostic criteria used to define EMA support the generalizability of our results. In addition, we adopted strict criteria to define EMA to avoid the inclusion of other myoclonic syndromes such as JME, especially in patients with later onset.[17] We thus chose to exclude patients with a history of myoclonic seizures involving body regions other than the eyelids, although their classification still represents a controversial topic. Similarly, while the exclusion of patients with moderate/severe ID allowed us to minimize the risk of including patients with clear-cut developmental/epileptic encephalopathy, their exclusion prevented us from definitively characterizing the entire spectrum of EMA subphenotypes with our cluster analysis, as previously suggested.[21]

Our study reveals the clinical variables predicting the occurrence of sustained remission in patients with EMA. In particular, early age at epilepsy onset appeared to be the most relevant predictor of poor seizure outcome. Moreover, using a large database with long-term follow-up data, we outlined the distinct prognostic patterns of this rare epilepsy syndrome. Last, we identified 2 distinct EMA subphenotypes with strong implications in terms of seizure control and cognitive outcome.

Acknowledgment
The authors thank Karine Ostrowsky-Coste (University of Lyon, France), Julitta De Regnauld De Bellescize (University of Lyon, France), Pascale Keo-Kosal (University of Lyon, France), Zeynep Gokce-Samar (University of Lyon, France), Alexandra Montavont (University of Lyon, France), Joseph Toulouse (University of Lyon, France), and Clara Milleret-Pignot (University of Lyon, France), who followed up some of the patients included in this study.

Study Funding
The authors report no targeted funding.

Disclosure
The authors report no disclosures relevant to this manuscript. Go to Neurology.org/N for full disclosures.

Publication History
Received by *Neurology* September 24, 2021. Accepted in final form January 21, 2022.

Appendix 1 Authors

Name	Location	Contribution
Emanuele Cerulli Irelli, MD	Department of Human Neurosciences, Sapienza, University of Rome, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data

Continued

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Appendix 1 *(continued)*

Name	Location	Contribution
Enrico Cocchi, MD	Department of Precision Medicine and Genomics, Department of Medicine, Columbia University, New York, NY	Study concept or design; analysis or interpretation of data
Georgia Ramantani, MD, PhD	Department of Neuropediatrics, University Children's Hospital Zurich, Switzerland	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Roberto H. Caraballo, MD	Department of Neurology, Hospital de Pediatría "Prof. Dr. Juan P Garrahan," Buenos Aires, Argentina	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Loretta Giuliano, MD	Department of Medical and Surgical Sciences and Advanced Technologies "G.F. Ingrassia," Section of Neurosciences, University of Catania, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Tulay Yilmaz, MD	Departments of Neurology and Clinical Neurophysiology, Istanbul University, Istanbul Faculty of Medicine, Turkey	Major role in the acquisition of data
Alessandra Morano, MD	Department of Human Neurosciences, Sapienza, University of Rome, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Eleni Panagiotakaki, MD, PhD	Department of Paediatric Clinical Epileptology, Sleep Disorders and Functional Neurology, University Hospitals of Lyon (HCL), member of the ERN EpiCARE, France	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Francesca F. Operto, MD	Child and Adolescent Neuropsychiatry Unit, Department of Medicine, Surgery and Dentistry, University of Salerno, Italy	Major role in the acquisition of data
Beatriz Gonzalez Giraldez, MD	Epilepsy Unit, Neurology Service, Hospital Universitario and IIS Fundación Jiménez Díaz and CIBERER, Madrid, Spain	Major role in the acquisition of data
Katri Silvennoinen, MD	Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, London, UK	Major role in the acquisition of data
Sara Casciato, MD, PhD	IRCCS NEUROMED, Pozzilli, Isernia, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Marion Comajuan, MD	Department of Paediatric Clinical Epileptology, Sleep Disorders and Functional Neurology, University Hospitals of Lyon (HCL), member of the ERN EpiCARE, France	Major role in the acquisition of data

Appendix 1 *(continued)*

Name	Location	Contribution
Simona Balestrini	Institute of Neurology, UCL	Major role in the acquisition of data
Francesco Fortunato, MD	Institute of Neurology, University Magna Graecia, Catanzaro, Italy	Major role in the acquisition of data
Antonietta Coppola, MD, PhD	Department of Neuroscience and Reproductive and Odontostomatological Sciences, Federico II University, Naples, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Giancarlo Di Gennaro, MD, PhD	IRCCS NEUROMED, Pozzilli, Isernia, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Angelo Labate, MD, PhD	Institute of Neurology, University Magna Graecia, Catanzaro, Italy	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Vito Sofia, MD	Department of Medical and Surgical Sciences and Advanced Technologies "G.F. Ingrassia," Section of Neurosciences, University of Catania, Italy	Major role in the acquisition of data
Gerhard J. Kluger, MD	Clinic for Neuropediatrics and Neurorehabilitation, Epilepsy Center for Children and Adolescents, Schoen Clinic Vogtareuth, Germany	Major role in the acquisition of data
Dorothée G.A. Kasteleijn-Nolst Trenité, MD, PhD	Department of Neurosurgery and Epilepsy, University Medical Center, Utrecht University, the Netherlands	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Antonio Gambardella, MD, PhD	Institute of Neurology, University Magna Graecia, Catanzaro, Italy	Drafting/revision of the manuscript for content, including medical writing for content
Betul Baykan, MD, PhD	Departments of Neurology and Clinical Neurophysiology, Istanbul University, Istanbul Faculty of Medicine, Turkey	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Sanjay M. Sisodiya, MD, PhD	Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, London, UK	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Alexis Arzimanoglou, MD, PhD	Department of Paediatric Clinical Epileptology, Sleep Disorders and Functional Neurology, University Hospitals of Lyon (HCL), member of the ERN EpiCARE, Lyon, France	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Pasquale Striano, MD, PhD	Pediatric Neurology and Muscular Diseases Unit, IRCCS "Istituto Giannina Gaslini," Genova, Italy	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Appendix 1 (continued)

Name	Location	Contribution
Carlo Di Bonaventura, MD, PhD	Department of Human Neurosciences, Sapienza, University of Rome, Italy	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data

Appendix 2 Coinvestigators

Name	Location	Role	Contribution
Rimma Gamirova, MD	Kazan State Medical Academy, Russia	Site investigator	Data acquisition
Giacomo Fisco, MD	Sapienza, University of Rome, Italy	Site investigator	Data acquisition
Natalia Liukshina, MD	MIDEAL Medical Clinic, Russia	Site investigator	Data acquisition
Tatiana Tomenko, MD	European Medical Center UMMC-Health, Russia	Site investigator	Data acquisition
Anna T. Giallonardo, MD	Sapienza, University of Rome, Italy	Site investigator	Data acquisition
Stefano Meletti, MD, PhD	University of Modena and Reggio Emilia, Modena, Italy	Site investigator	Data acquisition
Giuseppe Gobbi, MD	IRCCS, Istituto delle Scienze Neurologiche di Bologna, Italy	Site investigator	Data acquisition
Daniela Buti, MD	Meyer Hospital, Firenze, Italy	Site investigator	Data acquisition
Susanna Casellato, MD	University Hospital of Sassari, Italy	Site investigator	Data acquisition
Salvatore Striano, MD, PhD	Federico II University, Naples, Italy	Site investigator	Data acquisition
Tullio Messana, MD	IRCCS, Istituto delle Scienze Neurologiche di Bologna, Italy	Site investigator	Data acquisition
Antonella Riva, MD	IRCCS Istituto Giannina Gaslini, Genoa, Italy	Site investigator	Data acquisition
Lucio Giordano, MD	ASST Spedali Civili of Brescia, Italy	Site investigator	Data acquisition
Edoardo Ferlazzo, MD, PhD	Magna Græcia University of Catanzaro, Italy	Site investigator	Data acquisition
Aglaia Vignoli, MD	University of Milan, Italy	Site investigator	Data acquisition
Maurizio Viri, MD	AOU Maggiore della Carità Novara, Italy	Site investigator	Data acquisition
Irene Bagnasco, MD	Marini Hospital, Torino, Italy	Site investigator	Data acquisition

References

1. Jeavons PM, Clark JE, Maheshwari MC. Treatment of generalized epilepsies of childhood and adolescence with sodium valproate ("epilim"). *Dev Med Child Neurol.* 1977;19(1):9-25.
2. Giannakodimos S, Panayiotopoulos CP. Eyelid myoclonia with absences in adults: a clinical and video-EEG study. *Epilepsia.* 1996;37(1):36-44.
3. Appleton RE, Panayiotopoulos CP, Acomb BA, Beirne M. Eyelid myoclonia with typical absences: an epilepsy syndrome. *J Neurol Neurosurg Psychiatry.* 1993;56(12):1312-1316.
4. Panayiotopoulos CP, Agathonikou A, Koutroumanidis M, Giannakodimos S, Rowlinson S, Carr CP. Eyelid myoclonia with absences: the symptoms. In: Duncan JS, Panayiotopoulos CP, editors. *Eyelid Myoclonia with Absences.* John Libby and Co Ltd; 1996:17-25.
5. Panayiotopoulos CP. Syndromes of idiopathic generalized epilepsies not recognized by the International League Against Epilepsy. *Epilepsia.* 2005;46(suppl 9):57-66.
6. Ohya T, Yamashita Y, Shibuya I, Hara M, Nagamitsu S, Matsuishi T. Eyelid myoclonia with absences occurring during the clinical course of cryptogenic myoclonic epilepsy of early childhood. *Eur J Paediatr Neurol.* 2012;16(4):399-401.
7. Ferrie CD, Agathonikou A, Parker A, Robinson RO, Panayotopoulos CP. The spectrum of childhood epilepsies with eyelid myoclonia. In: Duncan JS, Panayotopoulos CP, editors. *Eyelid Myoclonia with Absences.* John Libbey; 1996:39-48.
8. Nar Senol P, Tezer FI, Saygi S. Eyelid myoclonia seizures in adults: an alternate look at the syndrome paradox. *Epilepsy Behav.* 2015;45:265-270.
9. Oguz Akarsu E, Tekturk P, Yapici Z, Tepgec F, Uyguner ZO, Baykan B. Eyelid myoclonic status epilepticus: a rare phenotype in spinal muscular atrophy with progressive myoclonic epilepsy associated with ASAH1 gene mutation. *Seizure.* 2016;42:49-51.
10. Okazaki T, Saito Y, Hiraiwa R, et al. Pharmacoresistant epileptic eyelid twitching in a child with a mutation in SYNGAP1. *Epileptic Disord.* 2017;19(3):339-344.
11. Striano S, Striano P, Nocerino C, et al. Eyelid myoclonia with absences: an overlooked epileptic syndrome? *Neurophysiol Clin.* 2002;32(5):287-296.
12. Vaudano AE, Ruggieri A, Tondelli M, et al. The visual system in eyelid myoclonia with absences. *Ann Neurol.* 2014;76(3):412-427.
13. Adachi M, Inoue T, Tsuneishi S, Takada S, Nakamura H. Eyelid myoclonia with absences in monozygotic twins. *Pediatr Int.* 2005;47(3):343-347.
14. Sadleir LG, Vears D, Regan B, Redshaw N, Bleasel A, Scheffer IE. Family studies of individuals with eyelid myoclonia with absences. *Epilepsia.* 2012;53(12):2141-2148.
15. Viravan S, Go C, Ochi A, Akiyama T, Carter Snead O III, Otsubo H. Jeavons syndrome existing as occipital cortex initiating generalized epilepsy. *Epilepsia.* 2011;52(7):1273-1279.
16. Destina Yalçin A, Forta H, Kiliç E. Overlap cases of eyelid myoclonia with absences and juvenile myoclonic epilepsy. *Seizure.* 2006;15(6):359-365.
17. Striano S, Capovilla G, Sofia V, et al. Eyelid myoclonia with absences (Jeavons syndrome): a well-defined idiopathic generalized epilepsy syndrome or a spectrum of photosensitive conditions? *Epilepsia.* 2009;50(suppl 5):15-19.
18. Incorpora G, Sofia V, Pavone P, Biondi R, Barone B, Parano E. Clinical heterogeneity in eyelid myoclonia, with absences, and epilepsy. *Eur J Pediatr.* 2002;161(3):175-177.
19. Covanis A. Jeavons syndrome: updated review. *J Epileptol.* 2015;23:113-123.
20. Scuderi C, Musumeci SA, Ferri R, Calabrese G, Elia M. Eyelid myoclonia with absences in three subjects with mental retardation. *Neurol Sci.* 2000;21(4):247-250.
21. Capovilla G, Striano P, Gambardella A, et al. Eyelid fluttering, typical EEG pattern, and impaired intellectual function: a homogeneous epileptic condition among the patients presenting with eyelid myoclonia. *Epilepsia.* 2009;50(6):1536-1541.
22. Sun Y, Seneviratne U, Perucca P, et al. Generalized polyspike train: an EEG biomarker of drug-resistant idiopathic generalized epilepsy [published correction appears in *Neurology*. 2018;91(24):1117]. *Neurology.* 2018;91(19):e1822–e1830.
23. Benassi M, Garofalo S, Ambrosini F, et al. Using two-step cluster analysis and latent class cluster analysis to classify the cognitive heterogeneity of cross-diagnostic psychiatric inpatients. *Front Psychol.* 2020;11:1085.
24. Kent P, Jensen RK, Kongsted A. A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep Cluster analysis, Latent Gold and SNOB. *BMC Med Res Methodol.* 2014;14:113.
25. Regesta G, Tanganelli P. The evolution and prognosis of photosensitive epilepsies. In: Beaumanoir A, Gastaut H, Naquet R, editors. *Reflex Seizures and Reflex Epilepsies.* Editions Medicine and Hygiene; 1989:175-177.
26. Jeavons PM, Bishop A, Harding GF. The prognosis of photosensitivity. *Epilepsia.* 1986;27(5):569-575.
27. Panayiotopoulos CP, Michael M, Sanders S, Valeta T, Koutroumanidis M. Benign childhood focal epilepsies: assessment of established and newly recognized syndromes. *Brain.* 2008;131(pt 9):2264-2286.
28. Caraballo RH, Fontana E, Darra F, et al. A study of 63 cases with eyelid myoclonia with or without absences: type of seizure or an epileptic syndrome? *Seizure.* 2009;18(6):440-445.
29. Cerulli Irelli E, Morano A, Barone FA, et al. Persistent treatment resistance in genetic generalized epilepsy: a long-term outcome study in a tertiary epilepsy center. *Epilepsia.* 2020;61(11):2452-2460.
30. Mohanraj R, Brodie MJ. Outcomes of newly diagnosed idiopathic generalized epilepsy syndromes in a non-pediatric setting. *Acta Neurol Scand.* 2007;115(3):204-208.
31. Gomez-Ibañez A, McLachlan RS, Mirsattari SM, Diosy DC, Burneo JG. Prognostic factors in patients with refractory idiopathic generalized epilepsy. *Epilepsy Res.* 2017;130:69-73.
32. Choi H, Detyniecki K, Bazil C, et al. Development and validation of a predictive model of drug-resistant genetic generalized epilepsy. *Neurology.* 2020;95(15):e2150–e2160.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

33. Geithner J, Schneider F, Wang Z, et al. Predictors for long-term seizure outcome in juvenile myoclonic epilepsy: 25-63 years of follow-up. *Epilepsia*. 2012;53(8):1379-1386.
34. Joshi CN, Patrick J. Eyelid myoclonia with absences: routine EEG is sufficient to make a diagnosis. *Seizure*. 2007;16(3):254-260.
35. Giuliano L, Fatuzzo D, Mainieri G, et al. Eyelid myoclonia with absences: electro-clinical features and prognostic factors. *Epilepsia*. 2019;60(6):1104-1113.
36. Mayo S, Gómez-Manjón I, Fernández-Martínez FJ, Camacho A, Martínez F, Benito-León J. Candidate genes for eyelid myoclonia with absences, review of the literature. *Int J Mol Sci*. 2021;22(11):5609.
37. Vlaskamp DRM, Shaw BJ, Burgess R, et al. SYNGAP1 encephalopathy: a distinctive generalized developmental and epileptic encephalopathy [published correction appears in *Neurology*. 2019;93(20):908]. *Neurology*. 2019;92(202):e96-e107.
38. Stamberger H, Hammer TB, Gardella E, et al. NEXMIF encephalopathy: an X-linked disorder with male and female phenotypic patterns. *Genet Med*. 2021;23(2):363-373.
39. Thomas RH, Zhang LM, Carvill GL, et al. CHD2 myoclonic encephalopathy is frequently associated with self-induced seizures. *Neurology*. 2015;84(9):951-958.
40. Syvertsen M, Fløgstad I, Enger U, Landmark CJ, Koht J. Antiepileptic drug withdrawal in juvenile myoclonic epilepsy. *Acta Neurol Scand*. 2019;139(2):192-198.
41. Nilo A, Crespel A, Genton P, Macorig G, Gigli GL, Gelisse P. Epilepsy with eyelid myoclonias (Jeavons syndrome): an electro-clinical study of 40 patients from childhood to adulthood. *Seizure*. 2021;87:30-38.
42. Striano P, Sofia V, Capovilla G, et al. A pilot trial of levetiracetam in eyelid myoclonia with absences (Jeavons syndrome). *Epilepsia*. 2008;49(3):425-430.
43. Cerulli Irelli E, Morano A, Cocchi E, et al. Doing without valproate in women of childbearing potential with idiopathic generalized epilepsy: implications on seizure outcome. *Epilepsia*. 2020;61(1):107-114.

Copyright © 2022 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Genetics in Medicine (2022) 24, 862–869

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Diagnostic sequencing to support genetically stratified medicine in a tertiary care setting

Natalie Lippa[1], Louise Bier[1], Anya Revah-Politi[1,2], Halie May[1], Sulagna Kushary[1], Natalie Vena[1,3], Jessica L. Giordano[4], Hila Milo Rasouly[3], Enrico Cocchi[3], Tristan T. Sands[1,5], Ronald J. Wapner[1,4], Kwame Anyane-Yeboa[1,6], Ali G. Gharavi[1,3], David B. Goldstein[1,*]

[1]Instiitute for Genomic Medicine, Columbia University Irving Medical Center, New York, NY; [2]Precision Genomics Laboratory, Columbia University Irving Medical Center, New York, NY; [3]Division of Nephrology, Department of Medicine, Columbia University Irving Medical Center, New York, NY; [4]Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, Columbia University Irving Medical Center, New York, NY; [5]Division of Child Neurology, Department of Neurology, Columbia University Irving Medical Center, New York, NY; [6]Division of Clinical Genetics, Department of Pediatrics, Columbia University Irving Medical Center, New York, NY

ARTICLE INFO

Article history:
Received 23 September 2021
Received in revised form
14 December 2021
Accepted 16 December 2021
Available online 22 January 2022

Keywords:
Exome sequencing
Genetic diagnosis
Precision medicine
Stratified medicine

ABSTRACT

Purpose: The goal of stratified medicine is to identify subgroups of patients with similar disease mechanisms and specific responses to treatments. To prepare for stratified clinical trials, genome-wide genetic analysis should occur across clinical areas to identify undiagnosed genetic diseases and new genetic causes of disease.
Methods: To advance genetically stratified medicine, we have developed and implemented broad exome sequencing infrastructure and research protocols at Columbia University Irving Medical Center/NewYork-Presbyterian Hospital.
Results: We enrolled 4889 adult and pediatric probands and identified a primary result in 572 probands. The cohort was phenotypically and demographically heterogeneous because enrollment occurred across multiple specialty clinics (eg, epilepsy, nephrology, fetal anomaly). New gene-disease associations and phenotypic expansions were discovered across clinical specialties.
Conclusion: Our study processes have enabled the enrollment and exome sequencing/analysis of a phenotypically and demographically diverse cohort of patients within 1 tertiary care medical center. Because all genomic data are stored centrally with permission for longitudinal access to the electronic medical record, subjects can be recontacted with updated genetic diagnoses or for participation in future genotype-based clinical trials. This infrastructure has allowed for the promotion of genetically stratified clinical trial readiness within the Columbia University Irving Medical Center/NewYork-Presbyterian Hospital health care system.
© 2022 American College of Medical Genetics and Genomics.
Published by Elsevier Inc. All rights reserved.

Louise Bier and Natalie Lippa contributed equally.
[*]Correspondence and requests for materials should be addressed to David B. Goldstein at the Institute for Genomic Medicine, Hammer Health Science Building, Columbia University Irving Medical Center, 701 W. 168th St., New York, NY 10032. *E-mail address:* dbgoldstein@columbia.edu

doi: https://doi.org/10.1016/j.gim.2021.12.010
1098-3600/© 2022 American College of Medical Genetics and Genomics. Published by Elsevier Inc. All rights reserved.

Introduction

Exome sequencing (ES) has proven to be an essential diagnostic method for patients with undiagnosed genetic conditions.[1] However, beyond establishing a diagnosis, ES is also an important tool in the advancement of genetically stratified medicine, where the goal is to identify subgroups of patients with similar disease mechanisms and specific responses to medications and treatments. Next-generation sequencing can elucidate the underlying genetic mechanism of disease and identify potential biological pathways to target with medications.

Success in developing stratified medicine will need to address underlying disparities in medicine and research. Health disparities in genomics have been documented and previously described and occur both within research and clinical care.[2-4] The reference genome, population genomic databases, and clinical research have lacked inclusion of underrepresented ancestral groups.[5-7] In clinical care, public and private insurance coverage for genetic testing varies widely and is often a barrier to obtaining genomic sequencing.[8-10] Trosman et al[10] note that it is challenging to obtain comprehensive nation-wide information about genetic testing coverage and policies from Medicaid and other public payers, affecting underserved and minority groups where public payers cover a substantial number of patients. Taken together, these disparities negatively affect the ability to make genetic diagnoses in underrepresented populations. Ultimately, this leads to further inequalities by limiting the focus of precision medicine to overrepresented populations and affects the ability of underrepresented populations to participate in, and ultimately benefit from, stratified medicine clinical trials.

As genomic research has advanced, genetically stratified clinical trials are a natural first step to identify underlying disease mechanisms and develop tailored treatments. However, advancement of stratified medicine relies on widespread sequencing of cohorts of interest. Recent publications have suggested the value of centralized efforts to facilitate this work, including a unified approach for obtaining and returning clinical genomic test results, establishing overarching Institutional Review Board (IRB) protocols as is done with institutional biobanks, and establishing centralized platforms for storage and analysis of genomic data.[11-13]

Recent research has shown that genetic diagnoses often occur in clinics that have not routinely been subject to diagnostic sequencing such as kidney and ophthalmologic disorders.[14-16] Therefore, to best prepare for stratified clinical trials, genetic evaluation should occur across clinical areas to identify undiagnosed genetic diseases in patients and should not be limited only to currently established gene-disease associations. Genome-wide analysis is required both to identify new phenotypes of known genetic diseases and to identify new genetic causes of disease.

To facilitate genetically stratified medicine, we have developed and implemented broad ES infrastructure and research protocols at Columbia University Irving Medical Center/New York Presbyterian Hospital (CUIMC/NYPH), a large academic medical center and hospital system. We addressed key criteria of a successful centralized system, including ability to enroll patients from any clinical area across the medical center, without cost to participants or their insurance; central storage of data; and the ability to use phenotypic data from the electronic medical record (EMR) in genomic analyses. As our knowledge of genomic variants progresses, the centralized storage of data allows patients to be recontacted over time when new knowledge about a gene or variant emerges. In this paper, we describe the infrastructure and results of a centralized system that promotes genetically stratified clinical trial readiness. These protocols have succeeded in enrolling >4800 diverse probands and making >500 diagnoses among >300 causative genes.

Materials and Methods

Research protocols were written to allow for the enrollment of pediatric and adult subjects, including pregnant women, from across all clinical areas of the medical center. We also accepted referrals from outside of the CUIMC/NYPH system after review of medical records to ensure eligibility (ie, detailed medical records provided, including documentation of initial genetics workup with normal genetic tests where clinically indicated, and confirmation of collaboration from referring physician for sample collection and return of results). Inclusion criteria were designed to accommodate referrals from a broad range of specialties and to allow for physician judgment regarding suspicion of possible underlying genetic etiology. This protocol design accommodates both enrollment of subjects with unique phenotypes and patients from the general genetics clinic, as well as specialty cohorts of interest such as epilepsy, pediatric intensive care unit patients, Parkinson disease, ataxia, cerebral palsy (CP), immunodeficiency, chronic kidney disease (CKD), and fetal anomalies. The protocols were approved by the CUIMC IRB, and all subjects, or their parent/legal guardian, signed informed consent. Subjects were not paid for their participation, and the research funding paid the costs of all genetic testing (including clinical confirmation). Subjects enrolled were also given the option to be contacted for future research and to receive incidental findings. Figure 1 provides an overall workflow of the infrastructure of these protocols.

Subjects provided blood or buccal samples. Fetal cases used chorionic villi or amniotic fluid, if the pregnant mother was undergoing such a procedure for clinical purposes. Other tissue types (eg, skin, brain tissue, and saliva) were also obtained to evaluate for germline and/or somatic variants. Somatic variants were analyzed in cases where there

Figure 1 **Overall workflow and infrastructure of the Institute for Genomic Medicine protocols.**

was suspected mosaicism. If the subject had clinical ES in the past, sequencing data were obtained from the clinical testing laboratory in lieu of sample collection, if permissible under the clinical laboratory's policies. Phenotypic data were extracted from the medical record. If available, parental samples were obtained for trio analysis.

ES and variant filtering and analysis has been previously described and were specifically designed for areas of interest.[14,17,18] Fetal anomaly cases were analyzed as described by Petrovski et al[18]; CKD cases were analyzed per Groopman et al[14]; and all other cases were analyzed per Zhu et al.[17] The American College of Medical Genetics and Genomics (ACMG) secondary finding gene list current at the time of analysis was used if incidental findings emerged.[19,20] Candidate diagnostic variants were reviewed in multidisciplinary meetings including genetic counselors, laboratory and clinical geneticists, and physicians from the referring specialties. Local clinical champions of each subspecialty (eg, nephrology, fetal anomaly, epilepsy) provided essential feedback related to subject phenotypes and association with candidate diagnostic variants. Variants deemed causative or likely causative of the proband's phenotype by team consensus were confirmed in a Clinical Laboratory Improvement Amendments (CLIA)–certified laboratory. Variant causality was assessed using a combination of both ACMG/Association for Molecular Pathology variant classification criteria, as well as clinical context and clinical correlation from phenotype experts/referring providers.[21] CLIA-confirmed variants were uploaded to the EMR, returned to referring providers, and subsequently returned to subjects with support from centralized research genetic counselors as needed. In cases where a clinical sample could not be obtained for CLIA laboratory confirmation testing, the diagnostic variant was not returned to the subject or provider, per regulatory requirements.

Results

Demographics

In total, we enrolled 4889 probands (see Table 1). The cohort proved to be diverse with 52% of probands from an underrepresented race/ethnicity, including 27% who self-identified as Hispanic/Latino. A large proportion of our cohort was adult patients (54.5%). In addition, 33.1% of the cohort had both parents enrolled, so that a trio analysis could be conducted. Of the subspecialty cohorts, the largest was CKD, with 2187 cases (44.7%) enrolled.

Specific genetic findings

We identified a primary result (genetic finding that fully or partially explains the phenotype) in 572 probands (11.7%) with 11 probands (3 adult, 8 pediatric) having a dual genetic diagnosis from ES (see Supplemental Tables 1 and 2 Supplemental File 1). In total, 337 different genes were identified among the diagnosed probands. Diagnostic yields among subspecialty cohorts ranged from 23% in the intellectual disability/developmental delay/autism spectrum disorder cohort, to 7.6% in the CKD cohort. Table 2 shows diagnostic yields across each specialty cohort. Diagnostic yields across specialty cohorts were not necessarily directly comparable because cohorts varied in the proportion of trio/nontrio cases, inclusion criteria, and customized analyses.

The 3 most common genetic diagnoses were *COL4A5*, *COL4A3*, and *COL4A4,* with 35, 17, and 14 diagnoses, respectively, all occurring within the CKD cohort. Eighteen genetic diagnoses occurred in 5 or more probands. Nine of these diagnoses (*SCN1A, SCN2A, NF1, NSD1, CACNA1A, EYA1, HNF1A, COL4A1,* and *PTPN11*) occurred across multiple subspecialty cohorts. Table 3 shows the most common genetic diagnoses in our proband patient population.

In addition, probands received incidental findings meeting the ACMG criteria for reporting[19,20] at the time of analysis. The protocol also allowed return of incidental findings outside of the ACMG secondary gene list if they were judged to be of urgent clinical significance (Table 4). In total, 67 probands received an incidental finding including 43 adult and 19 pediatric probands. Five families from the fetal anomaly cohort also received incidental findings after delivery or at end of pregnancy. Two of these cases involved incidental findings in genes that were not included in the ACMG list. The first of these cases involved a fetus with a diagnosis of isolated omphalocele identified in the first trimester. The family opted for a chorionic villus sampling with karyotyping, microarray, Beckwith–Wiedemann Syndrome methylation studies, and

Table 1 Demographics of probands

Demographics of Probands	Total N (%)
Total number of probands	4889 (100)
Sex	
Male	2339 (47.8)
Female	2509 (51.3)
Unknown/not recorded	41 (0.9)
Race/ethnicity	
African/African American	492 (10.1)
Asian	482 (9.9)
Hispanic	1321 (27.0)
White	2302 (47.1)
Other/more than one race	244 (5.0)
Unknown	48 (1.0)

ES. ES analysis identified a homozygous variant in *F11* (NM_000128.3: c.403G>T; NP_000119.1:p.Glu135Ter), consistent with factor XI deficiency (OMIM 612416). Owing to the planned surgical correction of omphalocele after birth, the team opted to return to this result so that the surgical team could prepare for potential bleeding issues. The second case involved a fetus with an isolated complex cardiac finding noted on ultrasound at 33 weeks' gestation. The family declined invasive prenatal testing and consented to cord blood sample collection with microarray, karyotyping, and research ES. Postnatal genetics evaluation showed a female child with minor dysmorphic features. The microarray results and karyotype were normal, but research ES identified a homozygous pathogenic missense variant in *HEXB*, associated with Sandhoff disease (NM_000521.4:c.1597C>T; NP_000512.2:p.Arg533Cys; OMIM 268800). Given the seriousness of Sandhoff disease and to provide recurrence risk information to parents, the study team decided to evaluate HexB enzyme activity in the proband, which showed decreased activity consistent with Sandhoff disease. Subsequently, molecular and enzymatic results were returned to the family. Although initial clinical features were not specific for Sandhoff disease, subsequent consult with neurology identified early features such as seizures and difficulty tracking.

Discussion

Our study highlights the real-world experience of applying research ES across a range of presentations in 1 medical center. In this way, we have laid the foundation to establish genetically stratified medicine within 1 institution.

Recognizing that there are racial and ethnic disparities in genomics, our protocols were developed to be as accessible as possible. Our studies were structured in such a manner that barriers to participation could be reduced. This included remote enrollments to accommodate patients/parents who are unable to attend appointments in person, translating consent materials into Spanish, using CUIMC's approved short form of consent process for subjects who do not speak English or Spanish, and covering the cost of CLIA confirmations so that participants are not responsible for these costs when receiving a genetic diagnosis. Our enrollment accommodations combined with the diverse Washington Heights neighborhood of New York City were likely the driving factors behind our enrollment of a diverse study population with >50% of study subjects identifying from a group underrepresented in biomedical research. This indicates that study procedure design prioritizing enrollment strategies that address cost of testing, flexibility of enrollment location, and reduction of burdens from the consent process can facilitate enrollment of underrepresented groups in biomedical research and begin to address the long-standing disparities in genomic research.

Because many genetic conditions have multisystemic involvement, patients may be identified across a broad range of clinical departments. In our subject population, we experienced this with the gene *SCN2A,* which is typically associated with epileptic encephalopathy (OMIM 613721), episodic ataxia (OMIM 618924), and benign familial infantile seizures (OMIM 607745). As expected, we identified causative *SCN2A* variants in patients with epileptic encephalopathy and benign familial infantile seizures, including a parent who had a history of childhood seizures of unknown etiology. In addition, we identified an *SCN2A* causative variant in a fetus who was found to have arthrogryposis, micrognathia, bilateral ventriculomegaly, polymicrogyria, prominent cisterna magna, and polyhydramnios on prenatal ultrasound. Others have reported *SCN2A* as causing bilateral ventriculomegaly seen on prenatal ultrasound.[22] Recently, de novo variants in *SCN1A*, another epileptic encephalopathy gene, were also found in 3 patients with prenatally diagnosed arthrogryposis.[23] These examples highlight how a centralized infrastructure can assemble a cohort of patients from multiple different clinical areas and research groups, allowing an expansion in our knowledge of phenotypes and improved understanding of genomic diseases.

Similarly, we have identified causative *COL4A1* variants in our neurology, general genetics clinic, and fetal anomaly cohorts. *COL4A1* has been linked to small-vessel brain disease of variable severity and can include other findings such as eye defects, kidney involvement, and cardiac arrhythmias.[24] In the neurology cohort, we found causative variants in 1 pediatric and 1 adult subject. The pediatric subject had a history of an intrauterine stroke, congenital cataracts, and medically refractory epilepsy. The adult subject had a history of recurrent hemorrhagic strokes that started at age 39 years and had a family history of hemorrhagic strokes. An unrelated 3-year-old pediatric subject referred from the general genetics clinic had muscle weakness and magnetic resonance imaging findings suggestive of a leukodystrophy. The indication of the case identified through the fetal anomaly cohort was intracranial hemorrhage. These cases clearly show how variations in clinical expressivity of variants in certain genes can be found in patients of all ages and in multiple specialty clinics.

Table 2 Diagnostic yield of cby age at enrollment

	Total Cohort		Proband Age at Enrollment					
			Adult (>21 y)		Pediatric (21 y)		Prenatal[a]	
	N (%)	Diagnostic Yield[b] (%)	*n* (%)	Diagnostic Yield (%)	*n* (%)	Diagnostic Yield (%)	*n* (%)	Diagnostic Yield (%)
Total number of probands	4889 (100)	572 (11.7)	2664 (54.5)	201 (7.5)	1731 (35.4)	302 (17.4)	494 (10.1)	69 (14.0)
Analysis type								
Trio	1616 (33.1)	321 (19.9)	102 (6.3)	15 (14.7)	1020 (63.1)	237 (23.2)	494 (30.6)	69 (14.0)
Nontrio	3273 (66.9)	251 (7.7)	2562 (78.3)	186 (7.3)	711 (21.7)	65 (9.1)	-	-
Specialty cohort[c]								
CKD	2187 (44.7)	166 (7.6)	1909 (87.3)	147 (7.7)	278 (12.7)	19 (6.8)	-	-
Epilepsy	947 (19.4)	126 (13.3)	352 (37.2)	16 (4.5)	595 (62.8)	110 (18.5)	-	-
ID/DD/ASD	918 (18.8)	211 (23.0)	95 (10.3)	17 (17.9)	823 (89.7)	194 (23.6)	-	-
Fetal anomaly	494 (10.1)	69 (14.0)	-	-	-	-	494 (100.0)	69 (14.0)
Cerebral palsy	145 (3.0)	15 (10.3)	63 (43.4)	1 (1.6)	82 (56.6)	14 (17.1)	-	-
PICU	144 (2.9)	22 (15.3)	-	-	144 (100.0)	22 (15.3)	-	-
Immunode ciency	39 (0.8)	5 (12.8)	19 (48.7)	3 (15.8)	20 (51.3)	2 (10.0)	-	-
Ataxia	33 (0.7)	4 (12.1)	31 (93.9)	3 (9.7)	2 (6.1)	1 (50.0)	-	-
Parkinson disease	15 (0.3)	2 (13.3)	15 (100.0)	2 (13.3)	-	-	-	-

ASD, autism spectrum disorder; *CKD*, chronic kidney disease; *DD*, developmental delay; *ID*, intellectual disability; *PICU*, pediatric intensive care unit.
[a]Prenatal refers to a fetus enrolled before delivery.
[b]Diagnostic yield was calculated among probands only. Family members who may have also received a genetic result were not included in these calculations.
[c]Subjects may be included in multiple specialty cohorts. Additionally, 530 probands presented with other phenotypes and were not included in any specialty cohort.

Table 3 Most common genetic diagnoses

Diagnosis Gene	Proband Count	Inheritance of Variant(s)			Type of Variant(s)			Analysis Type		Specialty Cohort
		Inherited	De novo	Unknown	Protein-Truncating	Missense	Other	Trio	Nontrio	
COL4A5	35	0	0	35	11	23	1	0	35	CKD
COL4A3	17	0	0	19[a]	4	15[a]	0	0	17	CKD
COL4A4	14	0	0	18[b]	13[b]	5[b]	0	0	14	CKD
SCN1A	12	0	6[c]	6	6	5	1	6	6	Epilepsy, ID/DD/ASD, CP
PKD1	10	1	1	8	6	4	0	2	8	CKD
SCN2A	10	1	9[c,d]	0	0	10	0	10	0	Fetal anomaly, epilepsy, ID/DD/ASD, PICU, CP
NF1	9	4	2	3	6	2	1	6	3	CKD, epilepsy, ID/DD/ASD
TRPC6	8	0	0	8	0	8	0	0	8	CKD
NSD1	7	0	5	2	5	2	0	5	2	Epilepsy, ID/DD/ASD, PICU
UMOD	7	0	0	7	0	7	0	0	7	CKD
CACNA1A	6	0	3	3	2	4	0	3	3	Epilepsy, ID/DD/ASD, ataxia
EYA1	6	1	0	5	3	3	0	1	5	CKD, fetal anomaly
HNF1A	6	0	0	6	2	4	0	0	6	CKD, other[e]
NPHS2[f]	6	0	0	12	1	11	0	0	6	CKD
COL4A1	5	0	4	1	1	4	0	4	1	Fetal anomaly, epilepsy, CP, other[e]
CSNK2B	5	0	3	2	5	0	0	3	2	Epilepsy
PAX2	5	0	0	5	4	1	0	0	5	CKD
PTPN11	5	1	2	2	0	5	0	3	2	CKD, fetal anomaly, ID/DD/ASD

ASD, autism spectrum disorder; *CKD*, chronic kidney disease; *CP*, cerebral palsy; *DD*, developmental delay; *ID*, intellectual disability; *PICU*, pediatric intensive care unit.
[a]Two probands had an autosomal recessive genetic diagnosis caused by missense variants in *COL4A3*.
[b]Four probands had an autosomal recessive genetic diagnosis caused by protein-truncating or missense variants.
[c]One case had evidence of parental mosaicism.
[d]One case was mosaic for variant.
[e]Other specialty cohort or phenotype.
[f]Autosomal recessive condition.

Table 4 Reportable incidental ndings identi ed

Gene	Proband Count	Adult	Pediatric	Prenatal[a]
BRCA2	18	10	7	1
MYBPC3	6	4	2	0
PKP2	5	4	0	1
BRCA1	4	2	2	0
KCNQ1	4	3	0	1
PMS2	3	2	1	0
LDLR	3	2	1	0
MSH6	3	2	1	0
TNNT2	2	1	1	0
TP53	2	2	0	0
RYR1	2	1	1	0
DSP	2	1	1	0
MSH2	2	2	0	0
SCN5A	2	2	0	0
MYH7	2	2	0	0
PTEN	1	0	1	0
KCNH2	1	1	0	0
DSG2	1	0	1	0
HEXB[b]	1	0	0	1
SDHD	1	1	0	0
LMNA	1	1	0	0
F11[c]	1	0	0	1

[a]Prenatal refers to a fetus enrolled before delivery. Note that all incidental ndings were returned after delivery or end of pregnancy for this specialty cohort.

[b]Although this gene was not on the American College of Medical Genetics and Genomics incidental list, it was reported as a secondary nding because of the seriousness of the condition and to provide recurrence risk information to parents. This nding did not explain the fetal anomalies.

[c]Although this gene was not on the American College of Medical Genetics and Genomics incidental list, it was reported as a secondary nding because of potential medical management decisions for surgical correction of omphalocele. This nding did not explain the fetal anomalies.

Having broad ES protocols allows for the enrollment of subjects in clinics that may not have a large enough volume to support a stand-alone genetics research framework and allows for identification of phenotype expansions. For example, we have enrolled a cohort of patients with CP from a CP specialty clinic within the Department of Orthopedics. Many of the genetic diagnoses were associated with developmental disorders but not CP (*SCN1A*, *DOCK6*, *PPT1*, *ATM*, *SMARCB1*, and *ZSWIM6*).[25] In this way, we expanded the number of patients with a specific genetic diagnosis, and we were able to recognize genetic cohorts from across the medical center that could be involved in future genetically stratified clinical trials.

Our most common genetic diagnoses were in the *COL4A5*, *COL4A3*, and *COL4A4* genes reflecting our large CKD population. Variants in the *NF1* gene were also among the most common findings returned to subjects. Of note, all probands with this finding already had a clinical diagnosis of neurofibromatosis type 1 but had additional clinical findings that could not be explained by this diagnosis (eg, chorea, kidney issues, and contractures). One explanation is that these patients may have an atypical neurofibromatosis type 1 phenotype. Alternatively, these patients may have a secondary genetic diagnosis that is as of yet undiagnosed. In fact, a recent case report describes a patient with neurofibromatosis and West syndrome who was found to have a *KCNC2* variant in addition to an *NF1* variant.[26]

Importantly, these protocols also established a consistent process for clinical confirmation of diagnostic research findings, allowing these diagnoses to be returned to patients and their providers, and ultimately be incorporated into clinical care. Such results may be of clinical benefit by establishing or refining a diagnosis, informing recurrence risk and familial cascade testing, guiding additional referrals or evaluations, or informing treatment recommendations.[14] For instance, a 30-year-old male with a previous diagnosis of static encephalopathy of unknown origin was found to have compound heterozygous variants in *CYP27A1*, consistent with a diagnosis of cerebrotendinous xanthomatosis (OMIM 213700). The diagnosis was confirmed on biochemical testing, and the patient was started on chenodeoxycholic acid, which has been shown to stabilize and/or improve progression of the disease. In a second case, a 70-year-old male with extensive chronic white matter disease and cognitive impairment was found to have a pathogenic *NOTCH3* variant indicating a diagnosis of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL; OMIM 125310). Subsequently the subject was taken off of Aricept, as it was not expected to have any benefit to CADASIL-related cognitive impairment and was given guidance on minimization of stroke risks, which was not indicated before the CADASIL diagnosis. Finally, 2 subjects with uncontrolled epilepsy were started on new medications on the basis of the genetic diagnosis (pyridoxine for *PIGW* and everolimus for mosaic *AKT3*); unfortunately, in both cases, the treatment did not result in significant improvement. With respect to additional research referrals, genetic diagnoses from the study have enabled subjects to be potentially eligible for genetically based clinical research. Genetic diagnosis is particularly critical in determining clinical trial eligibility for conditions such as Alport syndrome (*COL4A3*, *COL4A4*, *COL4A5*), which may not be recognized on the basis of clinical symptoms alone. Genetic diagnoses also allow patients to participate in genetically focused natural history studies that establish end points for future clinical trials. For instance, a subject with a homozygous *NGLY1* variant went on to enroll in a natural history study at the National Institutes of Health.

To realize the full clinical benefit of this research, our institution ensured sufficient genetic counseling support for return of genetic results. On average, the research has supported 5 full-time genetic counselors; however, it is important to note that their responsibilities also included variant interpretation, patient consent, regulatory/IRB management, and sample coordination. In addition, many results were returned by the referring providers independent of the research genetic counselors. Having a team of centralized genetic counselors who were trained to work across all studies provided significant efficiencies. Furthermore, because this research shows the utility of genetic testing in

new clinical areas, clinical genetic counseling services have been expanded to support increased clinical testing.

Our combined infrastructure allows for a larger and more powerful data set for aggregate analyses because 94% of samples and data are available for other research use. This allows researchers to reanalyze unsolved cases with similar phenotypes. In addition, it allows for the ability to search cases at the gene/variant level if researchers are interested in a specific gene/variant. Diagnoses identified in these cohorts are displayed in IGMdx (igmdx.org), which is a publicly accessible data browser to promote and facilitate collaborations with other investigators.[27] In addition, we have created ATAVdb (http://atavdb.org/) that allows investigators to search for genes/variants in samples where permission for other research has been given.[28] By establishing mechanisms for centralized data storage and access, this system has fostered genetic training and research throughout the institution and has led to additional grant funding and research. Furthermore, by having a centralized record of genetic diagnoses made across different specialties, this model can allow for targeted precision medicine trials that are aimed at specific genetic alterations and subject phenotypes.

Protocols were written to allow for patient samples and genomic data to be coded and to incorporate phenotypic information from the EMR record. This allows for longitudinal data to be collected on patients and the ability to track disease progression and treatment responses. It also allows researchers to contact patients for enrollment in other research protocols such as clinical trials, natural history studies, or biomarker analyses if they meet certain criteria. Overall, this infrastructure allows for further research to be conducted and new protocols to be developed for stratified medicine clinical trials.

One of the strengths of our infrastructure in a single tertiary care center, as described in this paper, is that it allows for the customization of the research analysis on the basis of the cohort of interest (eg, the CKD analysis included a manually curated nephropathy-associated gene list). Given that not all cohorts underwent the same analysis, we cannot draw comparisons between diagnostic rates across cohorts. In addition, although the study population is large and diverse, there may have been other biases in the cohort. For instance, although we accepted a broad range of phenotypes, a health care provider was needed to refer a subject for enrollment. This may have skewed the cohort because providers may have been more likely to refer patients who are more severely affected. In addition, subjects may have been more motivated to enroll in research if they had complex medical issues.

In conclusion, we have shown that our study processes have enabled enrollment and ES/analysis of a phenotypically and demographically diverse cohort of patients within 1 tertiary care medical center. Because all data are stored centrally, new disease associations can be discovered across multiple clinical specialties, and we have the ability to recontact patients with updated genetic diagnoses. In addition, this infrastructure has allowed for the promotion of genetically stratified medicine within the CUIMC/NYPH health care system.

Data Availability

The data sets supporting this study are available at the following publicly accessible websites: igmdx.org and atavdb.org.

Acknowledgments

We would like to thank all the individuals who participated in the studies; all the providers who referred subjects for inclusion in the studies; and the laboratory, bioinformaticians, genetic counselors, and other team members who have supported this work.

This publication was supported in part by the National Center for Advancing Translational Sciences, National Institutes of Health, through grant number UL1TR001873 and UL1TR000040. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This work was supported in part by R01NS09459 (principal investigator: Heinzen).

Some study data were collected and managed using REDCap (Research Electronic Data Capture, Vanderbilt University; https://www.project-redcap.org/) tools hosted at Columbia University Irving Medical Center.

Author Information

Conceptualization: D.B.G., R.J.W., A.G.G., K.A.-Y.; Data Curation: N.L., N.V., A.R.-P., S.K., H.M., E.C., H.M.R., J.L.G., L.B.; Formal Analysis: N.L., N.V., A.R.-P, S.K., H.M., E.C., H.M.R., J.L.G., L.B., T.T.S.; Methodology: D.B.G., R.J.W., A.G.G., K.A.-Y.; Project Administration: L.B., H.M.R., J.L.G.; Writing-original draft: N.L., L.B.; Writing-review and editing: N.L., L.B., A.R.-P., H.M., S.K., N.V., J.L.G., H.M.R., E.C., T.T.S., R.J.W., K.A.-Y., A.G.G., D.B.G.

Ethics Declaration

The study was approved by the Columbia University Institutional Review Board (Institutional Review Board protocol numbers: AAAO6702, AAAO8410, AAAO8009, and AAAC7385). Written informed consent was obtained from all participants or their legal guardians/representatives.

Any individual-level data and/or clinical data have been de-identified.

Conflict of Interest

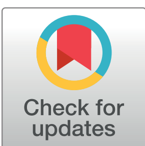
Dr. Gharavi has received research funding from the Renal Research Institute and Natera and has served as a consultant for Goldfinch Bio. All other authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2021.12.010) contains supplementary material, which is available to authorized users.

References

1. Tan TY, Lunke S, Chong B, et al. A head-to-head evaluation of the diagnostic efficacy and costs of trio versus singleton exome sequencing analysis. *Eur J Hum Genet*. 2019;27(12):1791–1799. http://doi.org/10.1038/s41431-019-0471-9.
2. Mapes BM, Foster CS, Kusnoor SV, et al. Diversity and inclusion for the All of Us research program: a scoping review. *PLoS One*. 2020;15(7):e0234962. http://doi.org/10.1371/journal.pone.0234962.
3. Petrovski S, Goldstein DB. Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol*. 2016;17(1):157. http://doi.org/10.1186/s13059-016-1016-y.
4. Bustamante CD, Burchard EG, De la Vega FM. Genomics for the world. *Nature*. 2011;475(7355):163–165. http://doi.org/10.1038/475163a.
5. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016;538(7624):161–164. http://doi.org/10.1038/538161a.
6. Koch L. Exploring human genomic diversity with gnomAD. *Nat Rev Genet*. 2020;21(8):448. http://doi.org/10.1038/s41576-020-0255-7.
7. Kim MS, Patel KP, Teng AK, Berens AJ, Lachance J. Genetic disease risks can be misestimated across global populations. *Genome Biol*. 2018;19(1):179. http://doi.org/10.1186/s13059-018-1561-7.
8. Smith HS, Franciskovich R, Lewis AM, et al. Outcomes of prior authorization requests for genetic testing in outpatient pediatric genetics clinics. *Genet Med*. 2021;23(5):950–955. http://doi.org/10.1038/s41436-020-01081-x.
9. Douglas MP, Parker SL, Trosman JR, Slavotinek AM, Phillips KA. Private payer coverage policies for exome sequencing (ES) in pediatric patients: trends over time and analysis of evidence cited. *Genet Med*. 2019;21(1):152–160. http://doi.org/10.1038/s41436-018-0043-3.
10. Trosman JR, Weldon CB, Slavotinek A, Norton ME, Douglas MP, Phillips KA. Perspectives of US private payers on insurance coverage for pediatric and prenatal exome sequencing: results of a study from the Program in prenatal and Pediatric Genomic Sequencing (P3EGS). *Genet Med*. 2020;22(2):283–291. http://doi.org/10.1038/s41436-019-0650-7.
11. Biswas S, Medne L, Devkota B, et al. A centralized approach for practicing genomic medicine. *Pediatrics*. 2020;145(3):e20190855. http://doi.org/10.1542/peds.2019-0855.
12. Rockowitz S, LeCompte N, Carmack M, et al. Children's rare disease cohorts: an integrative research and clinical genomics initiative. *NPJ Genom Med*. 2020;5:29. http://doi.org/10.1038/s41525-020-0137-0.
13. Mandl KD, Glauser T, Krantz ID, et al. The Genomics Research and Innovation Network: creating an interoperable, federated, genomics learning system. *Genet Med*. 2020;22(2):371–380. http://doi.org/10.1038/s41436-019-0646-3.
14. Groopman EE, Marasa M, Cameron-Christie S, et al. Diagnostic utility of exome sequencing for kidney disease. *N Engl J Med*. 2019;380(2):142–151. http://doi.org/10.1056/NEJMoa1806891.
15. Wolock CJ, Stong N, Ma CJ, et al. A case-control collapsing analysis identifies retinal dystrophy genes associated with ophthalmic disease in patients with no pathogenic ABCA4 variants. *Genet Med*. 2019;21(10):2336–2344. http://doi.org/10.1038/s41436-019-0495-0.
16. Lee K, Berg JS, Milko L, et al. High diagnostic yield of whole exome sequencing in participants with retinal dystrophies in a clinical ophthalmology setting. *Am J Ophthalmol*. 2015;160(2):354–363.e9. http://doi.org/10.1016/j.ajo.2015.04.026.
17. Zhu X, Petrovski S, Xie P, et al. Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med*. 2015;17(10):774–781. http://doi.org/10.1038/gim.2014.191.
18. Petrovski S, Aggarwal V, Giordano JL, et al. Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. *Lancet*. 2019;393(10173):758–767. http://doi.org/10.1016/S0140-6736(18)32042-7.
19. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565–574. http://doi.org/10.1038/gim.2013.73. Published correction appears in *Genet Med*. 2017;19(5):606.
20. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249–255. http://doi.org/10.1038/gim.2016.190.
21. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–424. http://doi.org/10.1038/gim.2015.30.
22. Bernardo S, Marchionni E, Prudente S, et al. Unusual association of SCN2A epileptic encephalopathy with severe cortical dysplasia detected by prenatal MRI. *Eur J Paediatr Neurol*. 2017;21(3):587–590. http://doi.org/10.1016/j.ejpn.2017.01.014.
23. Jaber D, Gitiaux C, Blesson S, et al. De novo mutations of *SCN1A* are responsible for arthrogryposis broadening the *SCN1A*-related phenotypes. *J Med Genet*. 2021;58(11):737–742. http://doi.org/10.1136/jmedgenet-2020-107166.
24. Plaisier E, Ronco P. COL4A1-related disorders. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews*. University of Washington; 2016.
25. May HJ, Fasheun JA, Bain JM, et al. Genetic testing in individuals with cerebral palsy. *Dev Med Child Neurol*. 2021;63(12):1448–1455. http://doi.org/10.1111/dmcn.14948.
26. Rademacher A, Schwarz N, Seiffert S, et al. Whole-exome sequencing in NF1-related West syndrome leads to the identification of KCNC2 as a novel candidate gene for epilepsy. *Neuropediatrics*. 2020;51(5):368–372. http://doi.org/10.1055/s-0040-1710524.
27. IGMdx. Institute for Genomic Medicine, Columbia University Irving Medical Center. Published February 5, 2020. http://igmdx.org/. Accessed December 14, 2021.
28. Ren Z, Povysil G, Hostyk JA, Cui H, Bhardwaj N, Goldstein DB. ATAV: a comprehensive platform for population-scale genomic analyses. *BMC Bioinformatics*. 2021;22(1):149. http://doi.org/10.1186/s12859-021-04071-1.



OPEN ACCESS

Citation: Manenti L, Maggiore U, Fiaccadori E, Meschi T, Antoni AD, Nouvenne A, et al. (2021) Reduced mortality in COVID-19 patients treated with colchicine: Results from a retrospective, observational study. PLoS ONE 16(3): e0248276. https://doi.org/10.1371/journal.pone.0248276

Editor: Antonio Cannatà, King's College London, UNITED KINGDOM

Received: October 12, 2020

Accepted: February 24, 2021

Published: March 24, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0248276

Copyright: © 2021 Manenti et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

RESEARCH ARTICLE

Reduced mortality in COVID-19 patients treated with colchicine: Results from a retrospective, observational study

Lucio Manenti[1☯], Umberto Maggiore[1☯]*, Enrico Fiaccadori[1], Tiziana Meschi[2], Anna Degli Antoni[3], Antonio Nouvenne[2], Andrea Ticinesi[2], Nicoletta Cerundolo[2], Beatrice Prati[2], Marco Delsante[1], Ilaria Gandoflini[1], Lorenzo Donghi[3], Micaela Gentile[1], Maria Teresa Farina[1], Vincenzo Oliva[1], Cristina Zambrano[1], Giuseppe Regolisti[4], Alessandra Palmisano[1], Caterina Caminiti[5], Enrico Cocchi[6], Carlo Ferrari[3], Leonardo V. Riella[7], Paolo Cravedi[8], Licia Peruzzi[6]

1 Dipartimento di Medicina e Chirurgia, Università di Parma e UO Nefrologia, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy, **2** Dipartimento Geriatrico-Riabilitativo, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy, **3** UO Malattie Infettive ed Epatologia, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy, **4** Dipartimento di Medicina e Chirurgia, Università di Parma e UO Clinica e Immunologia Medica, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy, **5** UO Ricerca e Innovazione, Azienda Ospedaliera-Universitaria di Parma, Parma, Italy, **6** Pediatric Nephrology Unit, Regina Margherita Children's Hospital, Città della Salute e della Scienza di Torino, Turin, Italy, **7** Division of Nephrology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States of America, **8** Renal Division, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America

☯ These authors contributed equally to this work.
* umberto.maggiore@unipr.it

Abstract

Objectives

Effective treatments for coronavirus disease 2019 (COVID-19) are urgently needed. We hypothesized that colchicine, by counteracting proinflammatory pathways implicated in the uncontrolled inflammatory response of COVID-19 patients, reduces pulmonary complications, and improves survival.

Methods

This retrospective study included 71 consecutive COVID-19 patients (hospitalized with pneumonia on CT scan or outpatients) who received colchicine and compared with 70 control patients who did not receive colchicine in two serial time periods at the same institution. We used inverse probability of treatment propensity-score weighting to examine differences in mortality, clinical improvement (using a 7-point ordinary scale), and inflammatory markers between the two groups.

Results

Amongst the 141 COVID-19 patients (118 [83.7%] hospitalized), 70 (50%) received colchicine. The 21-day crude cumulative mortality was 7.5% in the colchicine group and 28.5% in the control group (P = 0.006; adjusted hazard ratio: 0.24 [95%CI: 0.09 to 0.67]); 21-day

Funding: PC is supported by NIH NIAID grant 3U01AI063594-17S1. No external funding was received for this study.

Competing interests: The authors have declared that no competing interests exist.

clinical improvement occurred in 40.0% of the patients on colchicine and in 26.6% of control patients (adjusted relative improvement rate: 1.80 [95%CI: 1.00 to 3.22]). The strong association between the use of colchicine and reduced mortality was further supported by the diverging linear trends of percent daily change in lymphocyte count (P = 0.018), neutrophil-to-lymphocyte ratio (P = 0.003), and in C-reactive protein levels (P = 0.009). Colchicine was stopped because of transient side effects (diarrhea or skin rashes) in 7% of patients.

Conclusion

In this retrospective cohort study colchicine was associated with reduced mortality and accelerated recovery in COVID-19 patients. This support the rationale for current larger randomized controlled trials testing the safety/efficacy profile of colchicine in COVID-19 patients.

Introduction

Beginning in December 2019, a novel coronavirus, designated severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has caused an international outbreak of respiratory illness termed COVID-19 [1]. The full spectrum of COVID-19 ranges from mild, self-limiting respiratory tract illness to severe progressive pneumonia, multi-organ failure, and death. Cytokines and chemokines are thought to play an important role in the severity of complications during virus infections [2]. Patients with severe COVID-19 have higher serum levels of pro-inflammatory cytokines (TNF-α, IL-1, and IL-6) and chemokines (IL-8) compared to individuals with mild disease or healthy controls, and similar levels compared to patients with Severe Acute Respitatory Syndrome (SARS) or Middle East Respiratory Syndrome (MERS) [2]. The independent association between inflammatory markers and disease severity supports the concept that abnormal inflammatory response, rather than direct viral cytopathic effects, is the main cause of the life-threatening pulmonary complications in COVID-19 patients [3]. Various mechanisms have been postulated to explain the dysregulated immune response during SARS-CoV-2 infection. In particular, the viroporin envelope (E) protein, a minor virion structural component of SARS-CoV-2, has been shown to activate the NLR family pyrin domain containing 3 (NLRP3) inflammasome, eventually causing the release of cytokines and chemokines [4, 5].

Colchicine, an old drug that has been widely used in auto-immune and inflammatory disorders [6, 7], counteracts the assembly of the NLRP3 inflammasome [8], thereby reducing the release of IL-1b and an array of other interleukins, including IL-6, that are formed in response to danger signals [7–9]. Recently, colchicine has been successfully used in two cases of life-threatening post-transplant capillary leak syndrome [10]. These patients had required mechanically ventilation and hemodialysis for weeks before receiving colchicine, which quickly restored normal respiratory function and diuresis over 48 hrs [10].

Based on this background, we started prescribing colchicine as an off-label drug in health care outpatients, and shortly after in inpatients with COVID-19 and pneumonia on lung CT scan. Herein, we report the results of an observational retrospective study in which we used inverse probability of treatment weighting based on propensity score to undergo colchicine treatment, in order to assess the hypothesis that colchicine reduces mortality and time to clinical improvement in patients with COVID-19 pneumonia.

Patients and methods

Patients

This is an observational, retrospective study on COVID-19 patients followed from February 25th to April 8th, 2020 at the Parma University Hospital, a tertiary health-care Centre in Parma, Italy, which was designated as a COVID-19 hub by Italian health authorities. This retrospective study included COVID-19 patients (hospitalized with pneumonia on CT scan or outpatients). We included a series of consecutive patients who received colchicine for the treatment of COVID-19 from March 1sh to April 10th, 2020. The comparison group consisted of patients that were selected by random sampling amongst those admitted at the same hospital with a diagnosis of COVID-19 and pneumonia earlier in the pandemic (from March 1st to March 18th, 2020) and who could be matched 1:1 by age (± 10 years) and sex. Because a suitable age and sex match could only be found in 59 of the 71 patients, a 1:1 match of the same sex with the closest age was obtained in 22 cases. To reduce the risk of immortal time bias (i.e. patients on colchicine cannot die before taking colchicine) patients requiring intubation in the first 24 hours after admission were excluded. Data could not be eventually extracted in one patient in the colchicine group, leaving 70 patients in the colchicine group, and 71 in the control group. The study protocol was approved by the AVEN Ethics Committee on March 31st, 2020 (prot. n. 13306).

Criteria for COVID-19 diagnosis

We included all adult inpatients (aged $\geq$18 years) with a diagnosis of COVID-19 pneumonia based on: 1) CT typical findings (i.e. ground-glass opacities and/or patchy consolidation, and/or interstitial changes with a peripheral distribution), 2) positive nasopharyngeal swab test, and/or 3) serologic anti-SARS-CoV-2 antibody test. We also included health care personnel who received a diagnosis of COVID-19 based on typical symptoms and a positive nasopharyngeal swab test that were treated as outpatients and did not undergo a CT scan. As per institution protocol, all inpatients cases with clinical suspicion of COVID-19 underwent a lung CT-scan and a nasopharyngeal swab test at the time of admission. During the peak of the outbreak (i.e., at the time of patient enrollment), those that had typical clinical history and symptoms suggestive of COVID-19 along with CT findings indicative of viral pneumonia were diagnosed as COVID-19 regardless of the results of the swab test on admission. However, in those with a negative nasopharyngeal swab test at admission, diagnosis of infection was subsequently confirmed by a second nasopharyngeal swab test and/or positive test for antibodies against SARS--CoV-2.

Disease severity was quantified using a seven-point ordinal scale recommended by WHO and used by previous trials [11]. Clinical improvement was defined as a 2-point improvement on a 7-category ordinary scale which we measured every day until day 10, then at day 14, and 21) [11].

Study drug

Colchicine was administered orally 1 mg/day from day 1 up until clinical improvement or up to a maximum of 21 days, according to physicians' preferences. As per the drug information sheet and internal guidelines, the dose was adjusted for kidney function and drug to drug interaction. Colchicine was administered orally 1 mg/day from day 1 up until clinical improvement or up to a maximum of 21 days, according to physicians' preferences. As per the internal decision, the dose was adjusted for kidney function and drug to drug interaction. The dose had to be reduced to 0.5 mg/day if the patient developed severe diarrhea. In the case of acute or chronic kidney disease with eGFR $<$ 30mL/min not requiring dialysis, the dose was reduced to 0.5 mg/day; in

patients requiring dialysis, the dose was reduced to 0.5 mg every other day and was given after the end of each dialysis session. Patients with advanced liver impairment (up to Child-Pugh score B) could receive 0.5 mg every other day. Patients on antiretroviral drugs (including ritonavir or cobicistat) could receive one single dose of 1mg or 0.5mg according to the severity of the disease until one day from completion of antiviral treatment. Hydroxychloroquine and azithromycin did not require any dose adjustment. The decision to use anti-retroviral drugs, hydroxychloroquine, and azithromycin was up to the physician in charge of patient care.

Exclusion criteria were limited to advanced liver failure (Child-Pugh score C) and pregnancy. In particular, acute and chronic kidney failure of any degree did not represent a contraindication to the use of the drug.

Statistical analyses

Stata release 16 (2019, StataCorp, College Station, Tx, US) was used for all the analyses. A two-tailed P-value of less than 0.05 was regarded as statistically significant. In the main analysis, we used an inverse probability of treatment weighting that was based on propensity score to construct a weighted cohort of patients who differed with respect to the use of colchicine but were similar concerning other measured characteristics. Compared to the methods of matching, inverse probability of treatment weighting allows using information from every patient, without having to exclude patients that cannot find a suitable match, which is a desirable property in the presence of sparse data. To calculate the inverse probability of treatment weights, we estimated each patient's propensity to undergo colchicine treatment, using a logistic regression model that included predictor variables that had been selected based on their a priori possibility of confounding the relationship between colchicine use and mortality (age, sex, categorical variate indicating the severity of conditions at onset namely, non-hospitalized, hospitalized without oxygen, hospitalized and requiring supplemental oxygen, hospitalized requiring non-invasive ventilation, shortness of breath, cough, history of diabetes, history of hypertension, history of cancer, use of antibiotics, use of anti-retroviral drugs, use of hydroxychloroquine, use of i.v. steroids, use of tocilizumab). We assigned patients who received colchicine a weight of 1÷(propensity score) and those who did not receive colchicine a weight of 1÷(1–propensity score). To reduce the variability in the inverse probability of treatment–weighted models, we used stabilized weights [12]. Hazard-ratios (under the assumption of the absence of unmeasured confounding) that are estimated by propensity-score methods are more like the effects estimated in a randomized, controlled trial than those estimated by multivariable Cox regression [12]. We compared the distributions of categorical variables using the chi-square test in the unweighted cohort and the weighted logistic-regression models in the weighted cohort; of continuous variable by Mann-Whitney test. In the propensity-score–weighted cohort, we compared cumulative mortality between the colchicine group and the control group by plotting weighted survival functions and by estimating the hazard ratio for death associated with the use of colchicine with weighted Cox proportional-hazards models [12]. Change in lab parameters was estimated by fitting weighted random coefficients models via maximum likelihood which allowed for lack of consistency among subjects in the timing of the lab measurements. For the purpose of the analysis, C-reactive protein was log base 2 transformed (on unit decrease represent halving of C-reactive protein values).

Results

Patients

We included 141 consecutive patients with COVID 19. All patients were followed-up until death or 21 days after admission. In the 70 patients taking colchicine, the median time from

hospital admission to first colchicine administration was 4 days (interquartile range: 2 to 9), the median number of days on colchicine treatment was 6 (interquartile range: 2 to 13), with ten patients taking a single dose because of concomitant anti-protease treatment. Fiftynine patients (84%) started with 1 mg/day and eleven patients started with 0.5mg daily because of drug-to-drug interaction (eleven patients) and/or chronic renal failure (nine patients). **Table 1** and **S1 Table** summarize selected demographic and clinical characteristics of the study population before and after propensity-score weighting, respectively. Most baseline demographics and clinical characteristics were similar between the two groups. However, in the unweighted cohort, patients who received colchicine had, at admission, more often dyspnea, received more often supplemental oxygen and non- invasive mechanical ventilation (**Table 1**); a higher percentage of them were on antibiotics or antivirals at enrollment (**Table 1**); because of the higher disease severity in the colchicine group, more patients in the colchicine group received tocilizumab [10] (14.3%) in the colchicine vs. 4 (5.6%) in the control group; P = 0.086]. (**Table 1**). Those characteristics were balanced in the propensity-weighted cohort (**S1 Table**). In the propensity-weighted cohort, none of the characteristics statistically differed between the groups (**S1 Table**), and all the standardized differences were less than 20% (median +1.3%; range: -17.9 to +19.1%; interquartile range: -12.5 to +6.2%).

Outcomes

In the propensity-weighted cohort, the 21-day crude cumulative incidence of death was 7.5% In the colchicine group and 28.5% in the control group (P = 0.006; adjusted hazard ratio of death: 0.24 [95%CI: 0.09 to 0.67]) (**Fig 1**); 21-day clinical improvement occurred in 40.0% of the patients on colchicine and in 26.6% of control patients (P = 0.048); adjusted relative improvement rate: 1.80 (95%CI: 1.00 to 3.22). In exploratory subgroup analyses, the relative hazard reduction associated with colchicine was evident across different patient categories (**Fig 2**). However, in some categories, the hazard ratio could not be reliably estimated because of sparse data within some strata (i.e. patients with cancer, CKD, and receiving tocilizumab), and of no deaths (outpatients).

The strong association between the use of colchicine and reduced mortality was further supported in hospitalized patients by the diverging linear trends of log2 C-reactive protein levels and lymphocyte count since admission. In fact, log2 C-reactive protein had a sharper decrease in the colchicine group compared to the control group (P = 0.009, **Fig 3A**), a log reduction of 1 is equivalent to a halving of the concentration of C-reactive protein, in mg/L) while lymphocyte count showed a sharper increase in lymphocyte count in the colchicine group compared to the control group (P = 0.018) (**Fig 3B**). The improvement in lymphocyte count in the colchicine group compared to the control group was mirrored by a sharper improvement in the neutrophil-to-lymphocyte ratio (P = 0.003). Unlike patients in the control group, those who received colchicine had IL-6 serially measured because of the physicians' expectation that the drug would decrease IL-6 levels. Over the first week after admission, Log2 IL-6 levels decreased by almost two log units ($P<0.001$ for linear trend), implying a decrease to almost 25% of baseline levels (**Fig 4**). Because of the retrospective nature of the study, we did not have enough IL-6 measurements in the control group to make a reliable comparison.

Safety

Colchicine was well tolerated. Only four (7%) of the patients had to withdraw the drug because of side effects, two because of diarrhea, and two for skin rash.

Table 1. Demographic and clinical characteristics of the study population before propensity-score weighting.

	Control	Colchicine	P value*
	(N = 71)	(N = 70)	
Age—yrs	62.5 (14.5)	60.5 (13.4)	0.39
Male gender	49 (69.0)	51 (72.9)	0.62
Comorbidities			
Diabetes	13 (18.3)	11 (15.7)	0.68
Cancer	6 (8.4)	4 (5.7)	0.53
Hypertension	43 (60.6)	39 (55.7)	0.56
CKD	13 (18.3)	15 (21.4)	0.64
BMI[$]	25.0 (5.5)	25.9 (5.4)	0.54
Disease severity at diagnosis			
Non-hospitalized	13 (18.3)	10 (14.3)	0.004
Hosp. w/o O_2	24 (33.8)	8 (11.4)	
Hosp. with O_2	24 (33.8)	31 (40.3)	
Hosp. with NIV	10 (14.1)	21 (30.0)	
Clinical characteristics at diagnosis			
Fever	70 (98.6)	68 (97.1)	0.55
Dyspnea	26 (36.6)	50 (71.4)	<0.001
Cough	50 (70.4)	43 (61.4)	0.26
Arthro-myalgias	6 (8.4)	12 (17.1)	0.12
Diarrhea	5 (7.0)	9 (12.9)	0.25
Lab values at hospital admission (Inpatients)			
CRP- mg/dL	115.2 (4,>250)	116.6 (13,>250)	0.59
	n = 41	n = 54	
IL-6 –pg/mL	NA	127.6 (0.1, 860)	NA
	n = 0	n = 33	
PCT–ng/mL	0.17 (0,1.9)	0.17 (0,3)	0.99
	n = 26	n = 39	
Lymph.–N/mm^3	966 (532)	1072 (539)	0.33
	n = 45	n = 57	
Neut/Lymph ratio	5.4 (1.3,37)	5.2 (1.3,34)	0.62
	n = 22	n = 36	
D-dimer–ng/mL	869 (164,>9000)	1103(238,>9000)	0.79
	n = 25	n = 44	
Ferritin - μg/L	824(90,2594)	NA	NA
	n = 14		
sCreatinine–mg/dL	0.8 (0.5,7.2)	0.8 (0.4,21.0)	0.93
	n = 44	n = 54	
COVID-19 Treatment			
Use of antibiotics	39 (54.9)	60 (85.7)	<0.001
Antiviral treatment	45 (63.4)	40 (57.1)	0.45
Hydroxychloroquine	46 (64.8)	53 (75.7)	0.16
i.v. steroids	9 (12.7)	17 (24.3)	0.076
Tocilizumab	4 (5.6)	10 (14.3)	0.086

Categorical data are expressed as number (%), continuous data as average (SD), or median (range).

Lab values had missing data, therefore the number of available data is reported for each group.

*The P value for categorical variables was calculated by the chi-square test, for continuous variables by Mann-Whitney test.

[$]BMI was available in 70 of the 141 patients.

CKD, chronic kidney disease; NIV, non-invasive mechanical ventilation; Hosp., hospitalized; O_2, supplemental oxygen; CRP, C-reactive protein; PCT, procalcitonin; Lymph., Lymphocyte count; sCreatinine, serum creatinine, PTL, platelet. Neut/Lymph ratio, neutrophil to lymphocyte ratio, antiviral treatment, anti-retroviral drugs (lopinavir or lopinavir + ritonavir or cobicistat).

https://doi.org/10.1371/journal.pone.0248276.t001

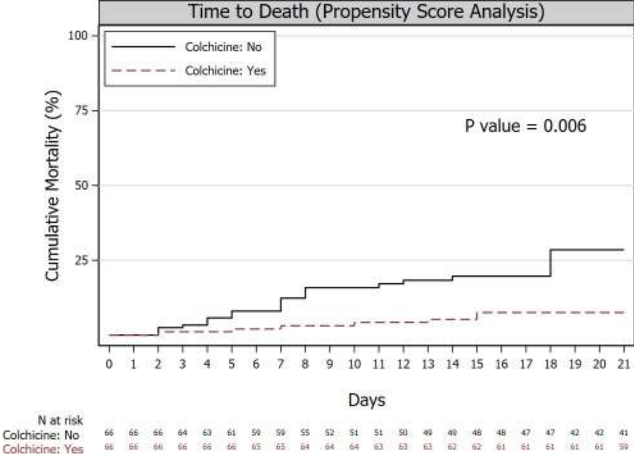


Fig 1. Time to death. Cumulative incidence of death since hospital admission (inpatients) or diagnosis (outpatients) in the two treatment groups. The 21-day crude cumulative incidence of death was 7.5% In the colchicine group and 28.5% in the control group (P = 0.006; adjusted hazard ratio of death: 0.24 [95%CI: 0.09 to 0.67]). The cumulative incidence curves and number at risk at the bottom of the figure refer to the cohort after inverse probability of treatment weighting (numbers at risk are rounded to the nearest integer).

https://doi.org/10.1371/journal.pone.0248276.g001

Discussion

In this observational retrospective study, using propensity score analysis, we found that colchicine administration was associated with a significant reduction in mortality and accelerated clinical improvement in patients with COVID-19.

The use of colchicine in COVID-19 has a sound biological rationale. The drug, which is well-known for its immunomodulatory properties in severe auto-inflammatory diseases [6], has been

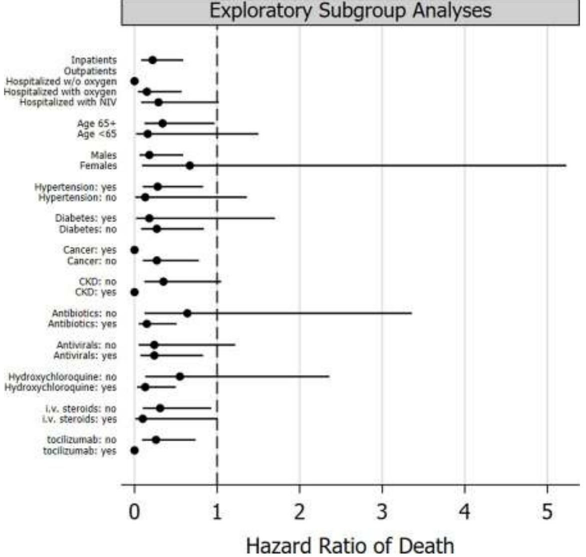


Fig 2. Exploratory subgroup analyses. Exploratory subgroup performed by Cox regression analyses adjusted by inverse probability of treatment weighting. The relative hazard reduction associated with colchicine was evident across different patient categories. However, in some categories the hazard ratio could not be reliably estimated because of sparse data (i.e. patients with cancer, with CKD, and receiving tocilizumab), and of no deaths in the stratum (outpatients). Solid circles represent the hazard ratio of death. Vertical bars represent 95 percent confidence intervals.

https://doi.org/10.1371/journal.pone.0248276.g002

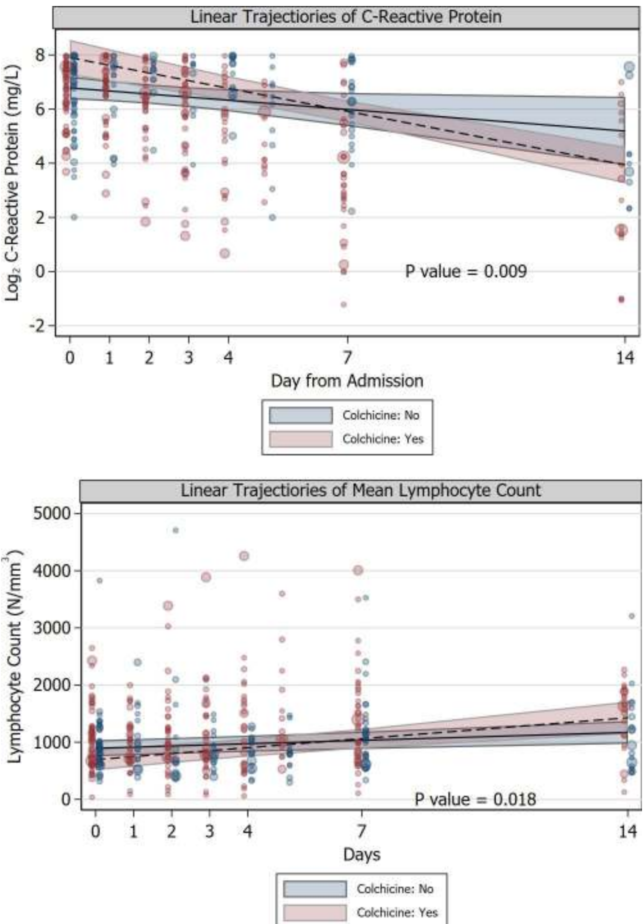


Fig 3. A-B. Linear trajectories since hospital admission of the mean of C-reactive protein and of lymphocyte count. Linear trajectories since hospital admission of the mean of C-reactive protein (Panel A), and of lymphocyte count (Panel B), in patients who received colchicine (red) or did not receive colchicine (blue). Trajectories express the average linear trends predicted from the propensity score-adjusted random coefficients model (see text). C-reactive protein is expressed as logarithm base 2, therefore1-unitdecrease implies halving of C-reactive protein levels. Log2 C-reactive protein had a sharper decrease in the colchicine group compared to the control group (P = 0.009); lymphocyte count showed a sharper increase in lymphocyte count in the colchicine group compared to the control group (P = 0.018). Bands represent 95% confidence intervals. Data values are indicated by circles the diameter of which is proportional to the inverse probability of treatment weight based on the propensity score that was used in the all the analyses.

https://doi.org/10.1371/journal.pone.0248276.g003

recently shown to modulate the release of circulating cytokines such as IL-6 [8, 9]. In addition, colchicine reduces lung injury in experimental acute respiratory distress syndrome (ARDS) [13]. Finally, colchicine may be an effective treatment for inflammation-induced thrombosis [14]. At variance with anti-IL6/IL-6 receptor antagonists, colchicine is widely available on the market, and inexpensive [6], and it is, therefore, suitable for massive use in countries where biologics are not easily available. We contend that colchicine use may be particularly attractive as a treatment to prevent the rapid progression of mild and moderate forms of COVID-19 to severe forms of respiratory failure and likewise it would be a useful drug also to treat a possible resurgence of SARS and MERS. Our study confirms the well-known safe toxicity profile of colchicine since the drug was withdrawn because of side effects (all of minor severity) in only 7% of treated patients.

Due to its retrospective nature, several baseline laboratory values and respiratory parameters were missing, therefore we could not fully adjust for all baseline differences between the

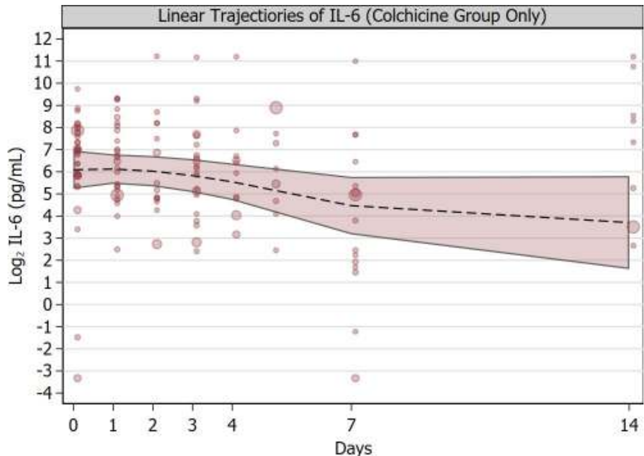


Fig 4. Linear trajectories of IL-6. Mean trajectory since hospital admission of log IL-6 serum levels (Panel A) in hospitalized patients who received colchicine. The dotted line is estimated based on random-coefficients mixed model with time included as polynomial variable.IL-6 serum level is expressed as logarithm base 2, therefore 1-unit decrease implie shalving of IL-6 serum levels. Bands represent 95% confidence intervals. Data values are indicated by circles the diameter of which is proportional to the inverse probability of treatment weight based on propensity score which was used in the analysis for comparison between treatment groups. The linear decreasing trend was statistically significant ($P<0.001$).

https://doi.org/10.1371/journal.pone.0248276.g004

two groups in the propensity score analyses. Nonetheless, patients receiving colchicine had, on average, more severe symptoms such as dyspnea compared to patients in the control group. In fact, because of the worse baseline conditions, patients on colchicine tended to receive more frequently tocilizumab and i.v. steroids. However, in our subgroup analyses (**Fig 2**) the effect of colchicine was also evident after excluding patients who received tocilizumab or i.v. steroids. During the study period, no other anti-inflammatory agent besides steroids, was used as for the standard of care at that time.

We cannot, however, exclude that differences in mortality between groups may be related to an underlying decreasing trend in mortality rates that might have occurred during the study period because of improved patient management. However, it is unlikely that an underlying trend accounts for all estimated effect of colchicine since the study period of the control and treatment group were very close to each other. Moreover, the timing of colchicine start and treatment duration was heterogeneous between patients in the colchicine group, the duration of treatment was often short because of the frequent use of protease-inhibitors-boosted anti-retroviral drugs, and colchicine was early withdrawn for an adverse event in 7% of patients. However, this study drawback should have biased our findings toward the null. Finally, although the selection of patients who could survive enough to receive colchicine may have caused an overestimation of the colchicine effect because of immortal time bias, the cumulative mortality curves diverged throughout the follow-up (**Fig 1**).

To the best of our knowledge, there is only one small randomized study [12], one observational study, which showed remarkably similar findings to our study [15], and two small case series of outpatients on colchicine in COVID-19 patients [16, 17], all showing clinical benefit. In the randomized study [12], which was performed in Greece, patient recruitment was terminated because of slow enrollment as a result of the rapid flattening of the curve of COVID-19. In that study, the clinical endpoint was defined as a 2-grade increase on an ordinal clinical scale that we used in our study, within a time frame of 21 days. The maintenance dosage of colchicine was the same as in our study. Of the 180 originally planned based on the clinical endpoint, only 105 patients were enrolled. Nonetheless, the clinical primary endpoint rate was 14.0% in the

control group (7 of 50 patients) and 1.8% in the colchicine group (1 of 55 patients) (odds ratio, 0.11; 95% CI, 0.01–0.96; P = 0.02), a finding which is fully consistent with the results of our retrospective observational study. The only large randomized study that so far showed a clear reduction in mortality among COVID-19 patients used dexamethasone [18]. Similar to colchicine, dexamethasone is another anti-inflammatory drug, which further supports the biological plausibility of our study finding [18]. In fact, early use of anti-inflammatory drugs may help to prevent complications such as pulmonary fibrosis, and thromboembolic disease.

In our study, the subgroup of health care personnel who did not require supplemental oxygen and were treated as outpatients had apparently the greatest benefit (**Fig 2**). These findings might imply that colchicine is most effective in the early stages of innate immune response. However, due to the small sample size, we cannot draw definite conclusions.

Our findings need to be confirmed by properly designed prospective trials. Indeed, we and others have started enrolling patients in multicenter randomized clinical trials on colchicine treatment for COVID-19. At the time of writing, twelve randomized studies on colchicine are registered on ClinicalTrial.gov. Some of these trials may eventually fail to provide results because, by the time the trials have started enrolling patients in a given country, the outbreak is declining and enrollment of patients may become difficult. Therefore, evidence coming from non-randomized studies may be fundamental to guide COVID-19 patient treatment.

In conclusion, pending the results from randomized control trials, our retrospective study provides evidence that colchicine may be a safe and effective drug for the treatment of COVID-19.

Supporting information

S1 Table. Demographic and clinical characteristics of the study population after propensity-score weighting.
(DOCX)

S1 Dataset.
(XLS)

S1 File. Stata code main analysis.
(TXT)

Author Contributions

Conceptualization: Lucio Manenti, Umberto Maggiore, Paolo Cravedi, Licia Peruzzi.

Data curation: Lucio Manenti.

Formal analysis: Lucio Manenti.

Writing – original draft: Lucio Manenti, Umberto Maggiore, Paolo Cravedi.

Writing – review & editing: Umberto Maggiore, Enrico Fiaccadori, Tiziana Meschi, Anna Degli Antoni, Antonio Nouvenne, Andrea Ticinesi, Nicoletta Cerundolo, Beatrice Prati, Marco Delsante, Ilaria Gandoflini, Lorenzo Donghi, Micaela Gentile, Maria Teresa Farina, Vincenzo Oliva, Cristina Zambrano, Giuseppe Regolisti, Alessandra Palmisano, Caterina Caminiti, Enrico Cocchi, Carlo Ferrari, Leonardo V. Riella, Paolo Cravedi, Licia Peruzzi.

References

1. Arshad Ali S, Baloch M, Ahmed N, Arshad Ali A, Iqbal A. The outbreak of Coronavirus Disease 2019 (COVID-19)-An emerging global health threat. J Infect Public Health. 2020; 13(4):644–646. https://doi.org/10.1016/j.jiph.2020.02.033 PMID: 32199792

2. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ et al. COVID-19: consider cytokine storm syndromes and immunosuppression. Lancet. 2020 Mar 28; 395(10229):1033–1034. Epub 2020 Mar 16. https://doi.org/10.1016/S0140-6736(20)30628-0 PMID: 32192578

3. Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y et al. Dysregulation of Immune Response in Patients With Coronavirus 2019 (COVID-19) in Wuhan, China. Clin Infect Dis. 2020 Jul 28; 71(15):762–768. https://doi.org/10.1093/cid/ciaa248 PMID: 32161940

4. Nieto-Torres JL, Verdiá-Báguena C, Jimenez-Guardeño JM, Regla-Nava JA, Castaño-Rodriguez C, Fernandez-Delgado R, et al. Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. Virology. 2015; 485:330–339. Epub 2015 Aug 29. https://doi.org/10.1016/j.virol.2015.08.010 PMID: 26331680

5. Castaño-Rodriguez C, Honrubia JM, Gutiérrez-Álvarez J, DeDiego ML, Nieto-Torres JL, Jimenez-Guardeño JM, et al. Role of Severe Acute Respiratory Syndrome Coronavirus Viroporins E, 3a, and 8a in Replication and Pathogenesis. mBio. 2018 May 22; 9(3):e02325–17. https://doi.org/10.1128/mBio.02325-17 PMID: 29789363

6. Karamanou M, Tsoucalas G, Pantos K, Androutsos G. Isolating Colchicine in 19th Century: An Old Drug Revisited. Curr Pharm Des. 2018; 24(6):654–658. https://doi.org/10.2174/1381612824666180115105850 PMID: 29336251

7. Imazio M, Brucato A, Cemin R, Ferrua S, Maggiolini S, Beqaraj F, et al. A randomized trial of colchicine for acute pericarditis. N Engl J Med. 2013 Oct 17; 369(16):1522–8. Epub 2013 Aug 31. https://doi.org/10.1056/NEJMoa1208536 PMID: 23992557

8. Nidorf SM, Thompson PL. Why Colchicine Should Be Considered for Secondary Prevention of Atherosclerosis: An Overview. Clin Ther. 2019; 41(1):41–48. https://doi.org/10.1016/j.clinthera.2018.11.016 PMID: 30591286

9. Martínez GJ, Robertson S, Barraclough J, Xia Q, Mallat Z, Bursill C, et al. Colchicine Acutely Suppresses Local Cardiac Production of Inflammatory Cytokines in Patients With an Acute Coronary Syndrome. J Am Heart Assoc. 2015 Aug 24; 4(8):e002128. https://doi.org/10.1161/JAHA.115.002128 PMID: 26304941

10. Cocchi E, Chiale F, Gianoglio B, Deorsola L, Pace Napoleone C, Fagioli F, et al. Colchicine: An Impressive Effect on Posttransplant Capillary Leak Syndrome and Renal Failure. Pediatrics. 2019 May; 143 (5):e20182820. https://doi.org/10.1542/peds.2018-2820 PMID: 31000685

11. Cao B, Wang Y, Wen D, Liu W, Wang J, Fan G, et al. A Trial of Lopinavir-Ritonavir in Adults Hospitalized with Severe Covid-19. N Engl J Med. 2020 May 7; 382(19):1787–1799. Epub 2020 Mar 18. https://doi.org/10.1056/NEJMoa2001282 PMID: 32187464

12. Deftereos SG, Giannopoulos G, Vrachatis DA, Siasos GD, Giotaki SG, Gargalianos P, et al. Effect of Colchicine vs Standard Care on Cardiac and Inflammatory Biomarkers and Clinical Outcomes in Patients Hospitalized With Coronavirus Disease 2019: The GRECCO-19 Randomized Clinical Trial. JAMA Netw Open. 2020 Jun 1; 3(6):e2013136. https://doi.org/10.1001/jamanetworkopen.2020.13136 PMID: 32579195

13. Dupuis J, Sirois MG, Rhéaume E, Nguyen QT, Clavet-Lanthier MÉ, Brand G, et al. Colchicine reduces lung injury in experimental acute respiratory distress syndrome. PLoS One. 2020 Dec 2; 15(12): e0242318. https://doi.org/10.1371/journal.pone.0242318 PMID: 33264297

14. Nonaka D, Takase H, Machii M, Ohno K. Colchicine therapy for deep vein thrombosis in a patient with vascular-type Behçet disease: A case report. Medicine (Baltimore). 2020 Apr; 99(16):e19814. https://doi.org/10.1097/MD.0000000000019814 PMID: 32312000

15. Scarsi M, Piantoni S, Colombo E, Airó P, Richini D, Miclini M, et al. Association between treatment with colchicine and improved survival in a single-centre cohort of adult hospitalised patients with COVID-19 pneumonia and acute respiratory distress syndrome. Ann Rheum Dis. 2020 Oct; 79(10):1286–1289. Epub 2020 Jul 30. https://doi.org/10.1136/annrheumdis-2020-217712 PMID: 32732245

16. Della-Torre E, Della-Torre F, Kusanovic M, Scotti R, Ramirez GA, Dagna L, et al. Treating COVID-19 with colchicine in community healthcare setting. Clin Immunol. 2020; 217:108490. https://doi.org/10.1016/j.clim.2020.108490 PMID: 32492478

17. Montealegre-Gomez G, Garavito E, Gomez-Lopez A, Rojas-Villarraga A, Parra-Medina R. [Colchicine: a potential therapeutic tool against COVID-19. Experience of 5 patients]. Reumatol Clin. 2020. Online ahead of print. https://doi.org/10.1016/j.reuma.2020.05.001 PMID: 32426001

18. Group RC, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in Hospitalized Patients with Covid-19—Preliminary Report. N Engl J Med. 2020. Online ahead of print. https://doi.org/10.1056/NEJMoa2021436 PMID: 32678530

Journal of Nephrology (2021) 34:1767–1781
https://doi.org/10.1007/s40620-020-00898-8

ORIGINAL ARTICLE

Clinical exome sequencing is a powerful tool in the diagnostic flow of monogenic kidney diseases: an Italian experience

Tiziana Vaisitti[1] · Monica Sorbini[1] · Martina Callegari[2] · Silvia Kalantari[2] · Valeria Bracciamà[2] · Francesca Arruga[1] · Silvia Bruna Vanzino[2] · Sabina Rendine[2] · Gabriele Togliatto[1] · Daniela Giachino[3,4] · Alessandra Pelle[3] · Enrico Cocchi[5] · Chiara Benvenuta[5] · Simone Baldovino[4,6] · Cristiana Rollino[6] · Roberta Fenoglio[6] · Savino Sciascia[6] · Michela Tamagnone[7] · Corrado Vitale[8] · Giovanni Calabrese[9] · Luigi Biancone[1,10] · Stefania Bussolino[11] · Silvana Savoldi[11] · Maurizio Borzumati[12] · Vincenzo Cantaluppi[13] · Fabio Chiappero[14] · Silvana Ungari[15] · Licia Peruzzi[5] · Dario Roccatello[4,6] · Antonio Amoroso[1,2] · Silvia Deaglio[1,2]

Received: 28 May 2020 / Accepted: 2 November 2020 / Published online: 23 November 2020
© The Author(s) 2020

Abstract

Background A considerable minority of patients on waiting lists for kidney transplantation either have no diagnosis (and fall into the subset of *undiagnosed* cases) because kidney biopsy was not performed or histological findings were non-specific, or do not fall into any well-defined clinical category. Some of these patients might be affected by a previously unrecognised monogenic disease.

Methods Through a multidisciplinary cooperative effort, we built an analytical pipeline to identify patients with chronic kidney disease (CKD) with a clinical suspicion of a monogenic condition or without a well-defined diagnosis. Following the stringent phenotypical and clinical characterization required by the flowchart, candidates meeting these criteria were further investigated by clinical exome sequencing followed by in silico analysis of 225 kidney-disease-related genes.

Results By using an ad hoc web-based platform, we enrolled 160 patients from 13 different Nephrology and Genetics Units located across the Piedmont region over 15 months. A preliminary "remote" evaluation based on well-defined inclusion criteria allowed us to define eligibility for NGS analysis. Among the 138 recruited patients, 52 (37.7%) were children and 86 (62.3%) were adults. Up to 48% of them had a positive family history for kidney disease. Overall, applying this workflow led to the identification of genetic variants potentially explaining the phenotype in 78 (56.5%) cases.

Conclusions These results underline the importance of clinical exome sequencing as a versatile and highly useful, non-invasive tool for genetic diagnosis of kidney diseases. Identifying patients who can benefit from targeted therapies, and improving the management of organ transplantation are further expected applications.

Keywords Next-generation sequencing · Chronic kidney failure · Transplantation · Renal monogenic disease

Introduction

The importance of genetic contributions in the development chronic kidney disease (CKD) is underlined by several observations: (1) inherited CKD (IKD) represents a high percentage of all CKDs [1–3], (2) the presence of a first-degree relative with end stage kidney disease (ESKD) confers a sevenfold increased risk of developing kidney failure [4], and (3) approximately 20–30% of patients report a positive family history of CKD in either a first- or second-degree relative [5, 6]. Thus, IKD represents one of the leading causes of CKD in children and adults, resulting in an increased risk of mortality, the need for organ transplantation, and high health care costs.

In the paediatric and young adult subset of patients, monogenic diseases represent up to 20% of patients who develop CKD before 25 years of age, with a variable diagnostic yield considering the different CKD categories [7–9].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40620-020-00898-8) contains supplementary material, which is available to authorized users.

✉ Antonio Amoroso
antonio.amoroso@unito.it

Extended author information available on the last page of the article

Changes in DNA sequence are usually single nucleotide variants (SNVs) and small insertions or deletions (indels), but larger deletions or insertions called copy number variants (CNVs) may also occur, particularly in syndromic children.

There are several monogenic inherited diseases that cause CKD, including developmental disorders, cystic and non-cystic ciliopathies, and glomerular and tubulo-interstitial diseases [10].

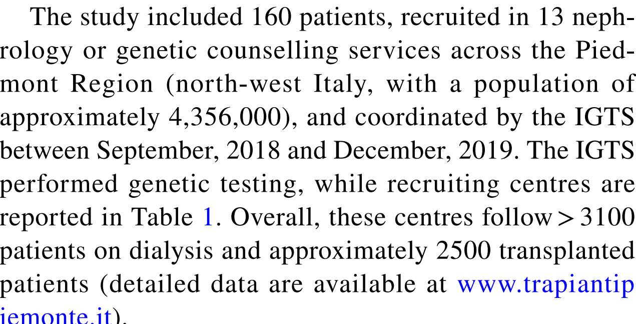
Establishing a genetic diagnosis strongly impacts patient management and prognosis [8, 11, 12], both by influencing treatment choices, as is the case for focal segmental glomerulosclerosis (FSGS), and by providing access to specific drugs, as is the case of vasopressin 2 antagonists for patients with autosomal dominant polycystic kidney disease (ADPKD). For these reasons, genetic testing is increasingly utilized in clinical nephrology due to accessibility of next generation sequencing (NGS) technologies [13, 14], which are non-invasive and cost-effective, and are becoming part of the diagnostic flow for several diseases, due to their decreasing costs, high throughput abilities and reduced sequencing times [15]. In this context, NGS technology enables us to simultaneously investigate hundreds of genes, thus opening up the possibility to rapidly identify genetic factors that underlie IKDs.

This study reports on the set up of an easy-to-use and accessible genetic testing platform which can be used to characterize undiagnosed cases of CKD eligible for NGS testing. Specifically, through this analytical pipeline, we aimed at (1) confirming diagnoses, particularly for patients in whom a monogenic condition was suspected, (2) finding the genetic cause of previously undiagnosed diseases, (3) identifying patients who could benefit from targeted therapies and (4) improving the management of organ transplantation, particularly in the living donor setting.

Materials and methods

Patients

This project relies on the multidisciplinary collaboration between the ImmunoGenetics and Transplant Biology Service (IGTS) of the Città della Salute e della Scienza University Hospital, the Centre of Research of Immunopathology and Rare Diseases-CMID, Centre of Coordination of the Interregional Network of Rare Diseases of Piedmont and Aosta Valley (San Giovanni Bosco Hub Hospital), the local ERK-net Member (Nephrology and Dialysis Unit, San Giovanni Bosco Hub Hospital and University of Turin), the Paediatric Nephrology Dialysis and Transplantation Unit (Regina Margherita Children's Hospital) and the Medical Genetics Unit at the San Luigi Gonzaga University Hospital.



The study included 160 patients, recruited in 13 nephrology or genetic counselling services across the Piedmont Region (north-west Italy, with a population of approximately 4,356,000), and coordinated by the IGTS between September, 2018 and December, 2019. The IGTS performed genetic testing, while recruiting centres are reported in Table 1. Overall, these centres follow > 3100 patients on dialysis and approximately 2500 transplanted patients (detailed data are available at www.trapiantipiemonte.it).

All patients included in the study provided written informed consent.

Set-up of the platform for genetic diagnosis of kidney diseases potentially leading to organ failure

We set up a web-based genetic service to provide initial genetic counselling to support regional nephrology centres in Piedmont that requested genetic evaluation (Fig. 1). Whenever possible, the referral centre provided IGTS with the patient's medical records including a detailed family history, clinical data from routine diagnostic procedures, parameters of kidney function, imaging data and biopsy results (https://www.cse.crtpiemonte.it/auth/CRT%20LoginGENnew.html).

The platform allowed remote multidisciplinary consultation in order to decide whether patients were eligible for this type of genetic test and to allocate patients with CKD into one of the following categories:

(a) Patients with a positive family history for CKD;
(b) Patients for whom genetic confirmation of the clinical diagnosis was required;
(c) Patients with CKD with no clinical diagnosis of a definite disease.

Genetic testing

DNA was extracted from blood samples, evaluated for integrity, and then processed for NGS analysis. Sequencing data were analysed by bio-informatics tools to identify, annotate and prioritize variants in order to generate a technical report. Variants were included in the genetic report that was then shared with the referring physician. Sanger sequencing on a second independent DNA extraction was performed to confirm NGS results. When possible, family segregation studies were performed.

The outcome of the genetic test was shared with the clinical team to plan the following steps (Fig. 1). Patients were referred to the closest genetic counselling centre.

Table 1 NGS cohort

Recruitment centre	N. of cases (n = 160)	Sex n (%)		Age at recruitment mean (min–max)		Eligibility	
		M	F	M	F	M	F
Regina Margherita Children's Hospital	52	24 (46.2)	28 (53.8)	9 (1–21)	8 (0–19)	21 26	
AOU San Luigi Gonzaga	19	12 (63.15)	7 (36.85)	41 (21–67)	34 (21–53)	12	6
San Giovanni Bosco Hospital	31	17 (54.8)	14 (45.2)	51 (22–77)	52 (19–67)	15	13
AOU Molinette Hospital	7	4 (57.2)	3 (42.8)	27 (18–35)	45 (32–55)	4	3
ASL CN1	15	10 (66.7)	5 (33.3)	58 (31–73)	53 (30–73)	8	3
ASL AL	6	4 (66.7)	2 (33.3)	46 (23–77)	57 (46–67)	3	2
CTO 2	1 (50)	1 (50)	30 (NA)	60 (NA)	1	0	
AO Ordine Mauriziano of Torino	4	1 (25)	3 (75)	37 (NA)	46 (27–57)	0	3
ASL TO3	6	1 (16.7)	5 (83.3)	51 (NA)	76 (72–83)	1	3
ASL TO4	9	7 (77.8)	2 (22.2)	54 (20–69)	48 (45–51)	5	1
SS of genetics Cuneo	3	2 (66.7)	1 (33.3)	56 (54–57)	45 (NA)	2	1
ASL VCO	4	4 (100)	0 (0)	45 (24–57)	NA	4	NA
ASL NO	2	2 (100)	0 (0)	33 (21–45)	NA	2 NA	

List of recruitment centres (Nephrology Units and Genetics Units) in Piedmont Region, and main features of the cohort included in the present study (n = 160). Number of cases, age at recruitment (mean age, min and max age) and eligibility for NGS are listed divided by gender (M: male; F: female). Recruiting centres are: San Giovanni Bosco Hospital; Regina Margherita Children's Hospital; AOU San Luigi Gonzaga: Azienda Ospedaliera Universitaria San Luigi Gonzaga; AOU Molinette Hospital: Azienda Ospedaliera Universitaria Molinette Hospital; ASL CN1: Azienda Sanitaria Locale—Cuneo, Mondovì and Savigliano; Struttura Semplice Genetics and Molecular Biology, ASL CN1 – Cuneo:ASL AL: Azienda Sanitaria Locale—Alessandria; CTO: Centro Traumatologico Ortopedico; ASL TO3: Azienda Sanitaria Locale—Collegno and Pinerolo; ASL TO4: Azienda Sanitaria Locale—Ciriè, Chivasso and Ivrea; ASL VCO: Azienda Sanitaria Locale del Verbano Cusio Ossola; AO Ordine Mauriziano di Torino: Azienda Ospedaliera Ordine Mauriziano di Torino; ASL NO: Azienda Sanitaria Locale di Novara

Diagnostic cohort

The validity of the platform and of the analytical pipeline was tested in a training cohort of 29 blindly tested patients for whom clinical and genetic diagnoses were already available. In each case previous genetic diagnosis was confirmed, suggesting that the adopted workflow is effective in the identification of monogenic kidney disease causative variants.

Of the 160 patients for whom genetic analysis was requested, 22 were excluded after a second re-evaluation (due to older age or confounding co-morbidities), while 138 were eligible for NGS analysis. Among them, 52 were children (< 18 years old, 37.7%), while 86 were adults (62.3%). Seventy-eight/138 (56%) were male [24 in the paediatric (46.2%) and 54 (62.8%) in the adult cohort, Table 2]. Sixty-seven out of 138 patients (48.5% in total; 34.6% in the paediatric and 57.0% in the adult subset) had a positive family history for kidney disease (Table 2).

Among the patients who were eligible for NGS, clinical suspicions were as follows: 32 (23.2%) patients who presented with clinical features compatible with ciliopathy, with or without liver involvement, most of whom (30/138; 21.7%) were diagnosed with polycystic kidneys; 21 patients (15.2%) who presented with a suspected glomerular disease, 11 (8%) with tubular diseases, 2 (1.4%) with nephrolithiasis/nephrocalcinosis, and 4 (2.9%) with haemolytic uraemic syndrome (HUS). Notably, a considerable percentage of these patients (60, 43.4%) presented with organ failure of unknown origin (Table 2). As expected, the great majority of paediatric patients (75%) were in stage I CKD, with a glomerular filtration rate ≥ 90 ml/min/1.73 m^2. Conversely, the majority of adult patients were in stages III-IV and presented important comorbidities, including diabetes (5.8%) and hypertension (69.8%) (Table 2).

Clinical exome sequencing and raw data processing

Libraries were prepared using the TruSight One Expanded Sequencing Kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Raw data were processed as reported in the Supplemental Methods (Online Resource). The choice of the clinical exome approach was dictated by the experimental need of performing a single sequencing, followed by flexible in silico analysis of organ-specific gene panels (e.g. kidney or liver), further tailored on the basis of the clinical suspicion, if available. The list of causative genes associated with kidney disease is updated twice a year, with eventual re-analysis of patients with negative genetic reports every 24 months, without the need for DNA re-sequencing.

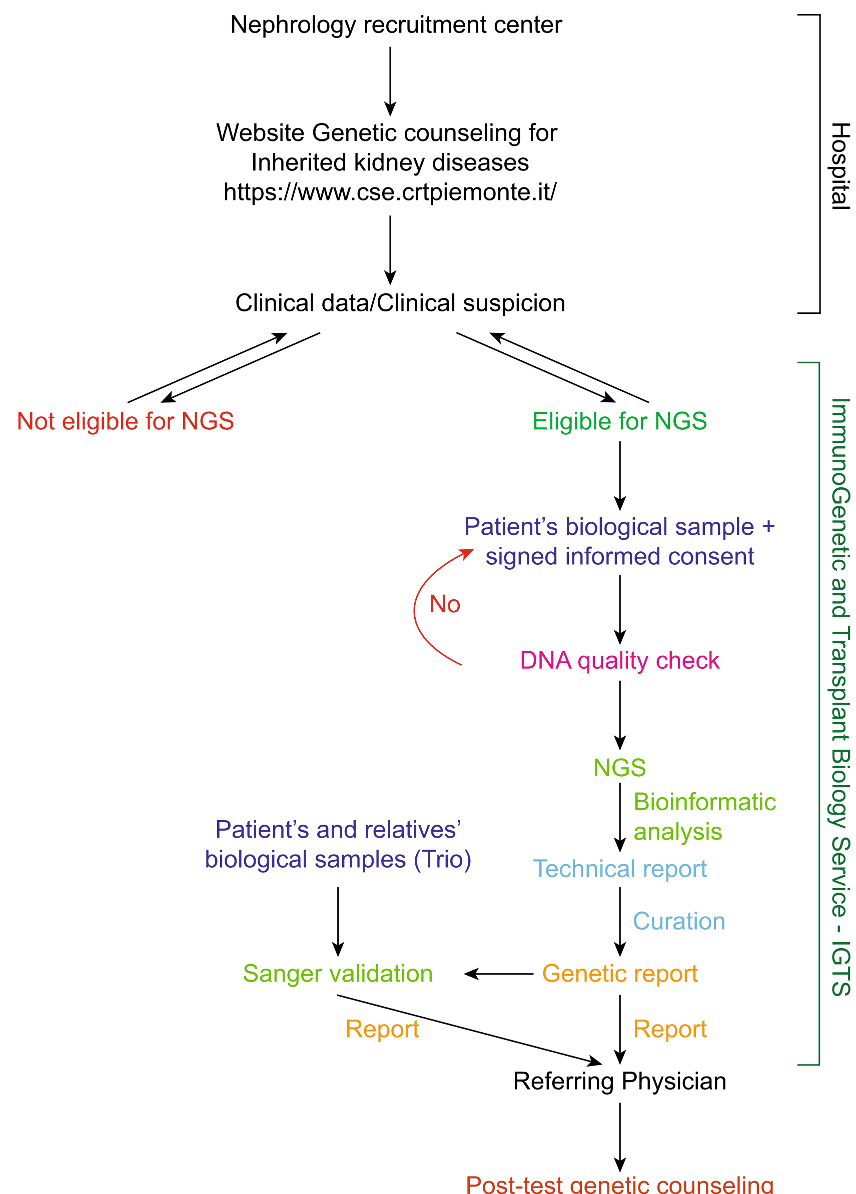


Fig. 1 Flowchart of the genetic counselling for inherited kidney diseases. Patients are recruited from the nephrology centres and clinical data are shared with the ImmunoGenetics and Transplant Biology Service (IGTS) through the website for genetic counselling for inherited kidney diseases. Eligibility is assessed based on familiarity, clinical suspicion, and available exams. For eligible patients, a biological sample is processed for NGS analysis. A genetic report is generated and then sent back to the referring physician. The last step provided by the Service is post-test genetic counselling

Design of an ad hoc pipeline of analysis to identify causative genes

To perform variant calling and identify causative variants, we designed an ad hoc pipeline of analysis based on sequential inclusion/exclusion steps. After reads were aligned to the GRCh37 as the reference genome using BWA, Isaac Aligner, GATK tools from Illumina, variants were processed using Variant Interpreter software, filtering-in mutations on the basis of a phenotype to genotype correlation. To this end, we generated a Clinical Phenotype to Genotype (CPTG) database by reviewing data from the literature and the main databases on inherited/orphan diseases (e.g., Online Mendelian Inheritance in Man-OMIM, Orphanet, PanelApp). CPTG associates clinical phenotypes to causative genes and is restricted to a panel of 225 genes related to kidney diseases (Supplementary Table 1, Online Resource). Variants were firstly filtered by using gene lists tailored on the basis of clinical suspicion, if available. If a pathogenic or a likely pathogenic variant was identified, analysis was stopped, otherwise, all 225 genes were tested. Synonymous variants were filtered out, while inclusion criteria for the remaining variants (non-synonymous, frame shift, splice site, non-sense) were (1) coverage $\geq 20\times$, (2) frequency ≥ 0.3 and (3) frequency in the overall population $\leq 1\%$ (to exclude polymorphisms which at the time of analysis were not known to be associated with clinical phenotype of kidney diseases), as reported in the 1000 Genomes Project (1 KG) and Exome Aggregation Consortium (ExAC) databases.

Table 2 Characteristics of patients eligible for NGS

Eligible cohort (n = 138)		
Features	Paediatric (n = 52)	Adults (n = 86)
Sex		
Female n (%)	28 (53.8)	32 (37.2)
Male n (%)	24 (46.2)	54 (62.8)
Positive family history n (%)	18 (34.6)	49 (57.0)
Age at onset mean (min–max)	3 (0–14)	37 (0–80)
Clinical suspicion		
CAKUT n (%)	3 (5.8)	0 (0)
Tubular disease n (%)	5 (9.6)	6 (7)
Ciliopathies n (%)	13 (25)	19 (22.1)
Nephrolithiasis/nephrocalcinosis n (%)	1 (1.9)	1 (1.2)
Glomerular disease n (%)	9 (17.3)	12 (13.9)
Haemolytic uraemic syndrome n (%)	1 (1.9)	3 (3.5)
Organ failure for unknown reasons n (%)	18 (34.6)	42 (48.8)
Others n (%)	2 (3.9)	3 (3.5)
Genetic diagnosis		
Cases with variants identified and in line with the clinical phenotype	30 (57.7)	48 (55.8)
Cases with no variants identified or incompatible with the clinical phenotype	22 (42.3)	38 (44.2)
CKD stage		
I	39 (75)	34 (39.6)
II	6 (11.5)	8 (9.3)
III	2 (3.9)	15 (17.5)
IV	1 (1.9)	11 (12.8)
V	0 (0)	10 (11.6)
Dialysis	0 (0)	4 (4.6)
Transplanted	4 (7.7)	4 (4.6)
Kidney biopsy performed	15 (28.8)	29 (33.7)
Imaging	41 (78.8)	69 (80.2)
Glomerular filtration rate (ml/min/1.73 m^2)		
≥ 90	39 (75)	34 (39.6)
60–89	6 (11.5)	7 (8.1)
30–59	2 (3.9)	16 (18.6)
15–29	1 (1.9)	12 (13.9)
< 15	4 (7.7)	17 (19.8)
Other characteristics		
Diabetes	0 (0)	5 (5.8)
Hypertension	8 (15.4)	60 (69.8)
Extra-renal features	21 (40.4)	28 (32.6)

Clinical details of the NGS-eligible study cohort (138 patients). Eligible patients are sub-divided on the basis of their gender, presence of a positive family history for kidney diseases, age at onset (mean, min and max age), clinical suspicion provided by clinicians at recruitment, results from genetic testing, chronic kidney disease (CKD) stage, availability of kidney biopsy or imaging data, glomerular filtration rate and other features. Number and percentage of cases are shown. CAKUT: congenital abnormalities of kidney and urinary tract

Inheritance mode was considered next. Specifically, if heterozygous mutations were found in genes associated with autosomal recessive (AR) diseases, they were carefully re-analysed to check for variants in genes known to be responsible for clinical phenotype in association with other genes (digenic diseases).

Filtered variants were then annotated (1) on the basis of the main public databases reporting associations between

gene variants and clinical phenotype, including OMIM, ClinVar, Human Genome Mutation Database (HGMD), the Genome Aggregation Database (GnomAD), database of single nucleotide polymorphism (dbSNP), Varsome, Exome Variant server (EVS) and (2) by considering the impact on protein structure or function by in silico prediction tools.

Variants classified as "pathogenic C5" and "likely pathogenic C4" were always included in the genetic report, as were "variants of unknown significance (VUS) C3" in genes associated with diseases with autosomal dominant (AD) or X-linked recessive (in males) mode of inheritance, while C3 variants in genes associated with diseases having AR mode of inheritance were reported only if they were in line with the clinical phenotype (Fig. 2). Confirmation by Sanger sequencing and family segregation studies were performed whenever possible.

Classification of the identified variants and their description in the genetic report were in line with The American College of Medical Genetics and Genomics (ACMG) policy statement on clinical sequencing (https

Fig. 2 Ad hoc pipeline of analysis. The pipeline is made up of several consecutive steps: phenotype-genotype correlation, filtering-in based on type of variant/frequency and disease list, inheritance model, variant annotation(s), manual curation and reporting of variants. For each step, specific actions and tools are indicated. *BWA* Burrows–Wheeler aligner, *GATK* genome analysis toolkit, *CPTG* clinical phenotype to genotype database, *Alt fr* altered allele frequency, *1 KG* 1000 Genomes Project, *ExAC* Exome Aggregation Consortium, *OMIM*: online mendelian inheritance in men, *HGMD* human genome mutation database, *GnomAD* the genome aggregation database, *dbSNP* database of single nucleotide polymorphism, *EVS* exome variant server

Phenotype-Genotype correlation
VARIANT CALLS (BWA -Isaac Alignment+GATK Identification)
CPTG database (Clinical Phenotype To Genotype database)
Filtering
CPTG database-Disease gene list
Clinical Exome
Digenic variants or compound heterozygosity
Synonymous Variants
Non-Synonymous Variants/ Nonsense/Splice site/Frameshift
Filtering-in Metrics (Coverage ≥20X; Alt fr ≥0.3; 1KG-Exac ≤1%)
Inheritance modeling
Heterozygous XL (♀)/AR
Homozygous XL (♂)
Hetero-/Homo-zygous AD Homozygous AR
Variant annotation(s)
Public Databases Annotation (OMIM/ClinVar/HGMD/GnomAD/dbSNP/Varsome/EVS) Internal Database Annotation (MoDiDi db)
In-silico prediction on protein: conservation/chemical properties (SIFT/PolyPhen-2/GERP/PhastCons/PhyloP/Grantham/Human Splicing Finder)
Manual Curation of Variants
In-depth review of top candidates (literature surf in PubMed/HGMD/ ClinVar/Google) Classification: **Pathogenic (C5)** **Likely pathogenic (C4)**
In-depth review of top candidates (literature surf in PubMed/HGMD/ ClinVar/Google) Classification: **VUS (C3)**
Reporting
Genetic report generation: variants included according to the ACMG guidelines

://www.acmg.net/) and with the Italian Society of Human Genetics (SIGU) [16].

Costs related to the NGS approach for the diagnosis of genetic kidney diseases

Overall, the cost of analysis per sample was differentiated on the basis of clinical suspicion: if a specific disease with < 3 causative genes was suspected the cost charged to the national health system was 1062 euros. For all other cases the cost charged to the national health system was 2262 euros.

Results

Overall genetic findings

Overall, by adopting the reported bio-informatics analysis pipeline, we detected 129 variants in 65 genes, with 28 patients carrying more than one variant. Interestingly, of all these variants, only 3 were recurrently present in more than one patient, while all the others were uniquely carried by individual patients.

Genetic variants were classified according to ACMG guidelines. In 78/138 (56.5%) patients, at least one variant was compatible with the clinical phenotype, as indicated in Table 3. In the remaining (60/138; 43.5%) patients, variants were either not present, or heterozygous in autosomal recessive genes or they were not in line with the clinical phenotype (not shown). Among patients for whom we identified variants compatible with the phenotype, 43 (55.1%) presented heterozygous variants in genes associated with autosomal dominant diseases, 16 (20.5%) were homozygous or compound heterozygous with variants in genes associated with autosomal recessive disease (among which 1 was a copy number loss) and 11 (14.1%) were characterized by variants in genes mapping on chromosome X (among which 2 were copy number losses). Lastly, 8 patients (10.3%) presented with variants in genes with both an autosomal dominant and autosomal recessive mode of inheritance (Table 3; Fig. 3a).

When considering all 129 variants in our cohort, 74 (57.4%) were missense, 10 (7.8%) nonsense, 19 (14.7%) frameshift, 7 (5.4%) indel, 15 (11.6%) variants affected the splicing regions and 4 (3.1%) were copy number variants (Fig. 3b).

Furthermore, when classifying all the variants identified by clinical exome sequencing according to ACMG guidelines to describe mutations in genes that cause Mendelian disorders, we found 27 variants defined as "pathogenic C5" (21.0%), 35 as "likely pathogenic C4" (27.1%) and 67 as "variants of unknown significance C3" (51.9%) (Table 3; Fig. 3c), considering that 28 patients were characterized by the presence of more than one variant with different classification.

Association between clinical and molecular diagnosis

The diagnosed cases, defining patients for whom genetic variants in line with the clinical phenotype were identified, were differentially distributed when considering the clinical suspicion categories (Table 3; Fig. 4). A high detection rate was obtained in glomerular diseases (14/21 cases; 66.7%), especially Alport disease and ciliopathies (22/32 cases; 68.8%), particularly ADPKD, while for tubular diseases and HUS, causative variants were identified in 4 out of 11, and 1 out of 4 cases, respectively. In the nephrolithiasis and nephrocalcinosis subset, one patient presented with a potentially causative variant in a relevant gene. With regard to the remaining categories, phenotype-related variants were detected in 50% of cases (4 out of 8). Moreover, our NGS approach identified the genetic culprit in a significant proportion of cases presenting with organ-failure of unknown origin (32/60 cases; 53.3%).

Among the cohort of patients with variants identified by NGS, all cases were validated by Sanger sequencing performed on a second independent aliquot of DNA. When possible, specifically in 23/52 paediatric patients, variant(s) were validated in the proband and in the trio. This analysis confirmed the segregation of variants in the family and helped clarify the clinical significance of "C3 VUS".

Discussion

In this study we describe an ad hoc-designed web-based platform built to connect regional Nephrology and Genetics centres to a centralized facility that provides genetic testing for patients with CKD. Herein we share our preliminary 15-month experience in applying this targeted sequencing to achieve a genetic diagnosis for undiagnosed patients with CKD in a well-selected cohort of patients from north-west Italy.

Our study presents two points that are worthy of interest. First, the feasibility of a centralized platform to support multidisciplinary consultation in patients with a high clinical suspicion of a monogenic condition. Second, an improvement in the diagnostic rate of patients with CKD and no previous definite diagnosis.

Our approach is based on a web-based platform as an extension of the existing regional transplant network. By using this platform, we attempted to optimize multidisciplinary consultations for patients for whom a monogenic condition was suspected. In order to explain the philosophy

Table 3 List of potentially diagnostic genetic variants

Patient ID	Disease category	Gene ID	MoI	All freq	Ref Seq	Coding seq/protein	Type of Variant	ACMG
#1	Glomerular disease	*COL4A5*	XL	Hem	NM_033380.2	c.1871G>A; p.(Gly624Asp)	missense	C5
		COL4A3	AR/AD	Het	NM_000091.4	c.3196C>T; p.(Pro1066Ser)	missense	C3
#2	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.8594G>A; p.(Arg2865Gln)	missense	C3
#3*	Unknown reasons	*DSTYK*	AD	Het	NM_015375.2	c.2215C>T; p.(Arg739Trp)	missense	C3
#4*	Unknown reasons	*SLC5A2*	AD	Het	NM_003041.3	c.1961A>G; p.(Asn654Ser)	missense	C5
#5*	Ciliopathies	*HNF1B*	AD	Het	NM_000458.2	c.826C>T; p.(Arg276*)	nonsense	C5
		PKD1	AD	Het	NM_001009944.2	c.7490-5C>G	splicing	C3
#6*	Glomerular disease	*COL4A5*	XL	Hem	NM_033380.2		CNV loss	C5
		COL4A6	XL	Hem	NM_001847.3		CNV loss	C5
#7*	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.8830G>T; p.(Val2944Phe)	missense	C3
		PKHD1	AR	Het	NM_138694.3	c.4075_4078dupCTGG; p.(Glu1360Alafs*19)	frameshift	C4
#8*	Glomerular disease	*WT1*	AD	Het	NM_024426.4	c.1373G>A; p.(Arg458Gln)	missense	C3
#9	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.763_775delCCTGCCCCCACCT; p.(Pro255Valfs*31)	frameshift	C4
#10*	Other	*LDLR*	AD	Het	NM_000527.4	c.313+1G>A	splicing	C5
#11	Other	*TTC21B*	AR	Hom	NM_024753.4	c.626C>T; p.(Pro209Leu)	missense	C5
		SLC4A1	AD/AR	Het	NM_000342.3	c.457C>A; p.(Leu153Met)	missense	C3
#12	Glomerular disease	*INF2*	AD	Het	NM_022489.3	c.212A>C; p.(Gln71Pro)	missense	C3
#13	Unknown reasons	*PKD2*	AD	Het	NM_000297.3	c.817_818 delCT; p.(Leu273Valfs*29)	frameshift	C5
#14*	Unknown reasons	*SLC12A3*	AR	Het	NM_000339.2	c.1964G>A; p.(Arg655His)	missense	C5
		SLC12A3	AR	Het	NM_000339.2	c.965C>T p.(Ala322Val)	missense	C3
		SLC12A3	AR	Het	NM_000339.2	c.1336-8C>A	splicing	C3
#15	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.1295C>T; p.(Ala432Val)	missense	C4
		SLC9A3R1	AD	Het	NM_004252.4	c.651delC; p.(Ala218Profs*27)	frameshift	C4
#16	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.9713-1G>A	splicing	C4
#17	Unknown reasons	*BBS2*	AR	Het	NM_031885.3	c.1864C>T; p.(Arg622*)	stop	C5
		BBS2	AR	Het	NM_031885.3	c.2067delA; p.(Lys689Asnfs*6)	frameshift	C4
#18	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.2659delT; p.(Trp887Glyfs*11)	frameshift	C4
#19	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c. 1145delG; p.(Gly382Valfs*83)	frameshift	C4
#20	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.11614G>A; p.(Glu3872Lys)	missense	C4
		PKD1	AD	Het	NM_001009944.2	c.5600A>G ;p.(Asn1867Ser)	missense	C3
#21	Unknown reasons	*COL4A4*	AR/AD	Het	NM_000092.4	c.755G>A; p.(Gly252Asp)	missense	C4
#22	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.7394_7396delTGTinsCGC; p.(LeuSer2465ProPro)	indel	C3
		PKHD1	AR	Het	NM_138694.3	c.10609A>C; p.(Asn3537His)	missense	C3
#23	Unknown reasons	*PKD2*	AD	Het	NM_000297.3	c.1096-6A>T	splicing	C3

Table 3 (continued)

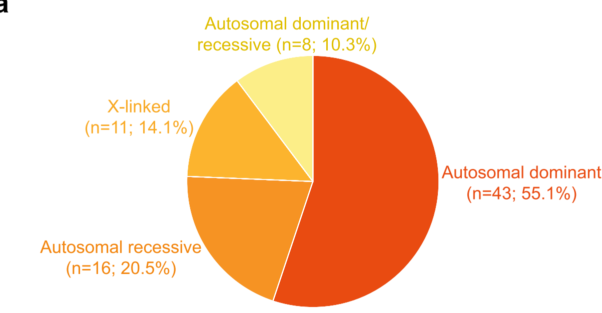
#24	Unknown reasons	*COL4A4*	AR/AD	Het	NM_000092.4	c.693+2T>C	splicing	C3
#25	Unknown reasons	*COL4A5*	XL	Hem	NM_033380.2	c.520G>C; p.(Gly174Arg)	missense	C5
#26*	Unknown reasons	*CLCN5*	XL	Hem	NM_001127898.3	c.1869delG; p.(Arg624Glyfs*32)	frameshift	C4
#27	Unknown reasons	*COL4A5*	XL	Hem	NM_033380.2	c.1075G>A; p.(Gly359Arg)	missense	C5
#28	Unknown reasons	*COL4A3*	AD/AR	Hom	NM_000091.4	c.3592G>A; p.(Gly1198Ser)	missense	C3
#29	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.12324G>A; p.(Trp4108*)	nonsense	C4
#30	Glomerular disease	*COL4A4*	AR/AD	Het	NM_000092.4	c.1460-6G>A	splicing	C3
		CFH	AR/AD	Het	NM_000186.3	c.3134-7T>C	splicing	C3
#31*	Unknown reasons	*NPHS1*	AR	Het	NM_004646.3	c.260G>T; p.(Gly87Val)	missense	C3
		NPHS1	AR	Het	NM_004646.3	c.349G>A; p.(Glu117Lys)	missense	C2
#32	Unknown reasons	*AGXT*	AR	Hom	NM_000030.2	c.33dupC; p.(Lys12Glnfs*156)	frameshift	C5
		HOGA1	AR	Het	NM_138413.3	c.700+5G>T	splicing	C5
#33	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.8990G>T; p.(Ser2997Ile)	missense	C4
#34	Unknown reasons	*ACTN4*	AD	Het	NM_004924.5	c.459C>G; p.(Phe153Leu)	missense	C3
#35	Unknown reasons	*COL4A1*	AD	Het	NM_001845.4	c.2566C>T; p.(Gln856*)	stop	C4
		SLC34A1	AR/AD	Het	NM_003052.4	c.840+1G>C	splicing	C4
		DSTYK	AD	Het	NM_015375.2	c.2215C>T; p.(Arg739Trp)	missense	C3
#36*	HUS	*CFHR1 - CFHR3*	AR	Hom			CNV loss	C3
#37*	Tubular disease	*ATP6V1B1*	AR	Het	NM_001692.3	c.1037C>G; p.(Pro346Arg)	missense	C4
		ATP6V1B1	AR	Het	NM_001692.3	c.33delG; p.(Leu12Serfs*8)	frameshift	C4
#38	Ciliopathies	*SLC5A2*	AD	Het	NM_003041.3	c.800G>A; p.(Arg267Gln)	missense	C3
#39	Ciliopathies	*TSC2*	AD	Het	NM_000548.4	c.2072G>A; p.(Arg691His)	missense	C3
#40	Glomerular disease	*COL4A5*	XL	Hem	NM_033380.2	c.3268T>C; p.(Tyr1090His)	missense	C3
#41	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.5648C>T; p.(Ala1883Val)	missense	C3
#42	Glomerular disease	*COL4A3*	AD/AR	Het	NM_000091.4	c.3437G>A p.(Gly1146Glu)	missense	C3
#43	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.8311G>Ap.(Glu2771Lys)	missense	C5
#44	Glomerular disease	*COL4A3*	AD/AR	Het	NM_000091.4	c.1006G>Tp.(Gly336Cys)	missense	C4
#45	Glomerular disease	*INF2*	AD	Het	NM_022489.3	c.1280_1285delCACCCC p.(Pro427_Pro428del)	indel	C3
#46	Ciliopathies	*PKD2*	AD	Het	NM_000297.3	c.1319G>C;p.(Arg440Thr)	missense	C3
#47	Ciliopathies	*TSC2*	AD	Het	NM_000548,4	c.1078A>G; p(Ile360Val)	missense	C3
		CYP24A1	AR	Het	NM_000782,4	c.1186C>T; p.(Arg396Trp)	missense	C5
#48	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.6583_6589delTGCCAGC; p.(Cys2195Glyfs*15)	frameshift	C5
#49	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.7119C>G; p.(Cys2373Trp)	missense	C3
#50	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.8293C>T; p.(Arg2765Cys)	missense	C3
#51	Unknown reasons	*COL4A5*	XL	Hem	NM_033380.2	c.2087G>A; p.(Gly696Asp)	missense	C3
#52	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.7288C>T; p.(Arg2430*)	nonsense	C5
		PKD1	AD	Het	NM_001009944.2	c.9454C>T; p.(Arg3152Trp)	missense	C3

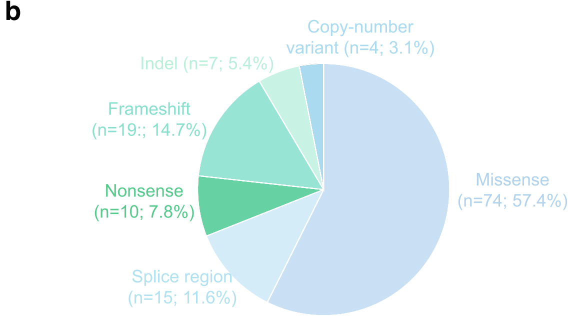
Table 3 (continued)

#53	Unknown reasons	*CLCN5*	XL	Hem	NM_001127898.2	c.1007T>C; p.(Leu336Pro)	missense	C4
		SLC4A1	AD	Het	NM_000342.3	c.118G>A; p.(Glu40Lys)	missense	C4
#54	Unknown reasons	*PKD1*	AD	Het	NM_001009944.2	c.6050C>T; p.(Ser2017Leu)	missense	C3
#55*	Unknown reasons	*SLC34A1*	AR	Hom	NM_003052.4	c.460_480dupATCCTGGTGACCGTGCTGGTG; p.(Ile154_Val160dup)	frameshift indel	C5
		SLC34A3	AR	Het	NM_001177316.1	c.756G>A p.(Gln252=)	splicing	C3
		SLC34A3	AR	Het	NM_001177316.1	c.561-8G>A	splicing	C3
#56	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.5963_5964insG p.(Glu1989*)	nonsense	C4
		PKD1	AD	Het	NM_001009944.2	c.7321G>A p.(Gly2441Ser)	missense	C3
#57*	Glomerular disease	*NPHS2*	AR	Hom	NM_014625.2	c.855_856delAA; p.(Arg286Thrfs*17)	frameshift	C4
#58	Ciliopathies	*PKD2*	AD	Het	NM_000297.3	c.1548+1G>A	splicing	C4
#59	Tubular disease	*PAX2*	AD	Het	NM_003990.3	c.821delC p.(Pro274Leufs*23)	frameshift	C4
#60	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.6424C>T; p.(Gln2142*)	nonsense	C4
#61*	Ciliopathies	*PKHD1*	AR	Het	NM_138694.3	c.131-1G>A	splicing	C4
		PKHD1	AR	Het	NM_138694.3	c.6731T>A p.(Leu2244His)	missense	C3
#62	Nephrolithiasis	*SLC7A9*	AR/AD	Het	NM_014270.4	c.313G>A; p.(Gly105Arg)	missense	C5
#63*	Unknown reasons	*SALL1*	AD	Het	NM_002968.2	c.475_477dupAGC; p.(Ser159dup)	indel	C3
		GRIP1	AD	Het	NM_021150.3	c.934T>C; p.(Cys312Arg)	missense	C3
#64*	Unknown reasons	*CD2AP*	AD	Het	NM_012120.2	c.730-4C>T	splicing	C3
#65	Unknown reasons	*COL4A3*	AR/AD	Het	NM_000091.4	c.2313_2330delACTCCC TGGACTTCCAGG; p.(Leu775_Gly780del)	Inframe deletion	C3
		PIGV	AR	Het	NM_001202554	c.439C>T; p.(Gln147*)	nonsense	C4
#66*	Unknown reasons	*BBS10*	AR	Hom	NM_024685.3	c.271dupT; p.(Cys91Leufs*5)	frameshift	C5
#67	Ciliopathies	*PKHD1*	AR	Het	NM_138694.3	c.4450G>A; p.(Ala1484Thr)	missense	C3
		PKHD1	AR	Het	NM_138694.3	c.12027C>G; p.(Tyr4009*)	nonsense	C3
#68	Glomerular disease	*COL4A5*	XL	Hem	NM_033380.2	c.4629C>A; p.(Tyr1543*)	stop	C4
#69	Tubular disease	*WNK4*	AD	Het	NM_032387.4	c.3554G>A; p.(Arg1185His)	missense	C3
#70	Glomerular disease	*COL4A6*	XL	Hem	NM_001847.3	c.3010C>G; p.(Pro1004Ala)	missense	C3
		COL4A3	AD/AR	Het	NM_000091.4	c.510A>T; p.(Lys170Asn)	missense	C3
		COL4A4	AD/AR	Het	NM_000092.4	c.594+5G>A	splicing	C3
#71	Unknown reasons	*SLC4A1*	AD	Het	NM_000342.3	c.1765C>T; p.(Arg589Cys)	missense	C5
#72*	Unknown reasons	*SLC34A1*	AR	Hom	NM_003052.4	c.272_292del TCCCCAAGCTGCGCCAGGCTG; p.(Val91_Ala97del)	Inframe deletion	C4
#73	Ciliopathies	*TSC1*	AD	Het	NM_000368.4	c.2626-3delC	splicing	C3
#74	Ciliopathies	*PKD1*	AD	Het	NM_001009944.2	c.6778_6780 delATT; p.(Ile2260del)	deletion	C4
#75*	Ciliopathies	*PKHD1*	AR	Het	NM_138694.3	c.4882C>G; p.(Pro1628Ala)	missense	C4
		PKHD1	AR	Het	NM_138694.3	c.9464A>G; p.(Tyr3155Cys)	missense	C3
#76	CAKUT	*SLC3A1*	AR/AD	Het	NM_000341.3	c.1400T>C; p.(Met467Thr)	missense	C5
#77*	Glomerular disease	*CUBN*	AR	Hom	NM_001081	c.164C>T; p.(Thr55Met)	missense	C3
		KANK4	NA	Het	NM_181712.4	c.2401T>C; p.(Tyr801His)	missense	C3
#78*	Tubular disease	*OCRL*	XL	Hem			CNV loss	C5

The table shows the list of 78 patients in whom a potentially diagnostic genetic variant may be present Asterisks indicate the family segregation studies that were carried out. When more than one variant is present, the ones potentially explaining the clinical phenotype are highlighted in blue

MoI mode of inheritance, *All. Freq* allele frequency, *Ref Seq* reference sequence, *AD* autosomal dominant, *AR* autosomal recessive, *XL* X-linked, *Het* heterozygous, *Hom* homozygous, *Hem* hemizygous, *CNV* copy number variation, *Indel* insertion/deletion, *HUS* haemolytic uraemic syndrome





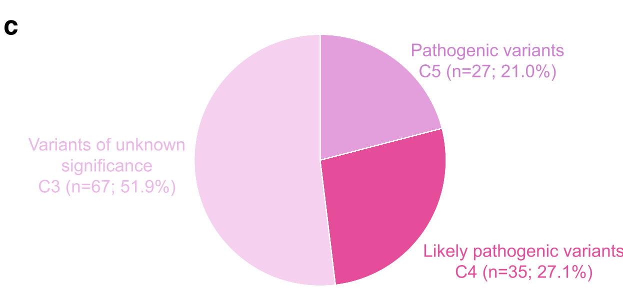


Fig. 3 Classification of the identified variants in the Piedmont cohort. **a** Number and percentage of patients having an autosomal dominant, autosomal recessive or X-linked disease on the basis of NGS-identified variants. **b** Classification of the identified variants as missense, nonsense, frameshift, insertion/deletion (indel) or affecting the splice site. Copy number variants (CNVs) are also represented. Number and percentage of variants belonging to the various categories is indicated in brackets. **c** Number and percentage of variants classified on the basis of the American College of Medical Genetics guidelines, considering pathogenic C5, likely pathogenic C4 and variants of unknown significance (VUS, C3)

and the practical issues of the platform two training courses for the nephrologists of the recruiting centres were organized in 2019. Moreover, in order to reduce waiting time, a rapid (mean waiting time of 3 days), web-based, pre-test genetic assessment was offered, with no need for patients to have in-person genetic counselling.

Patients' samples and informed consent were obtained through the Nephrology or Genetic Counselling Services, thus overcoming the need for the patient and his/her family to travel. The connection of the IGTS to the various nephrology Units throughout the Region was made possible by a capillary network of the Regional Centre for Transplantation. To speed up the connection between "the edge" and "the centre" of the hub, an IT platform was set-up allowing clinicians and geneticists to share clinical data and genetic reports. Recruited patients were initially evaluated by geneticists for their eligibility for NGS based on several criteria including family history and clinical data. Analyses of sequencing data and identification of the causative variants were performed based on an ad hoc pipeline.

Close to 300 patients were recruited between September, 2018 and March, 2020, and a final genetic report was available for 160 of them after a median time for genetic analysis of 6 months. The remaining 140 patients were in different steps of the diagnostic process at the time of this interim analysis.

In this study, we performed clinical exome sequencing followed by an in-silico analysis focused on selective genes in a cohort of 138/160 recruited patients affected by CKD. In 56.5% of cases NGS analysis was able to determine the molecular genetic cause of the disease, revealing 129 variants in 65 genes. These results are on average higher than those reported in the literature, likely due to patient preselection on the basis of positive family history and clinical suspicion [9].

With regard to the need to provide genetic confirmation of a previous clinical diagnosis, NGS analysis was able to confirm 68.8% of ciliopathies, a percentage that is in line with previous publications [17]. The detection rate was higher in glomerular diseases (66.7% vs. 14% reported in the literature) and nephrolithiasis (50% vs. 15–30%) [9]. This high percentage is due to selection of patients with a suspicion of Alport disease, at least based on biopsy results. In contrast, the percentage of solved cases presenting with Congenital Anomalies of the Kidneys and of the Urinary Tract (CAKUT) and haemolytic uraemic syndrome was quite low, with a considerable number of cases remaining undiagnosed. A reason for these results could be related to either the genetic heterogeneity of the disease, with many causative genes still to be identified, or to non-genetic causes [8].

In a considerable subset of the recruited cohort, patients were referred to genetic analysis because of a kidney disease of unknown origin. As expected, based on previous experience from other centres, this approach proved to be efficient in revealing causative variants: in a significant number of these cases, we were able to identify genetic variants that were in line with the clinical phenotype, thus helping clinicians in the management of these patients. Surprisingly, in our cohort, the percentage of patients for whom a genetic variant in line with the clinical phenotype was identified was not so different when considering paediatric (57.7%) and adult (55.8%) subgroups. One explanation is that our

Inner pie: Distribution of clinical suspicion
Outer pie: Distribution of variants compatible/incompatible with the clinical suspicion

CAKUT (3/138; 2.2%)
Haemolytic uraemic syndrome (4/138; 2.9%)
Tubular diseases (11/138; 8%)
Glomerular diseases (21/138; 15.2%)
Ciliopathies (32/138; 23.2%)
Nephrolithiasis Nephrocalcinosis (2/138; 1.4%)
Others (5/138; 3.6%)
Organ faillure for unknown reasons (60/138; 44.2%)

Patients with variants compatible with clinical phenotype 66.7%
Patients with no variants identified or incompatible with clinical phenotype 33.3%
Patients with variants compatible with clinical phenotype 25%
Patients with no variants identified or incompatible with clinical phenotype 75%
Patients with variants compatible with clinical phenotype 33.3%
Patients with no variants identified or incompatible with clinical phenotype 66.7%
Patients with variants compatible with clinical phenotype 36.4%
Patients with no variants identified or incompatible with clinical phenotype 63.6%
Patients with variants compatible with clinical phenotype 68.8%
Patients with no variants identified or incompatible with clinical phenotype 31.2%
Patients with variants compatible with clinical phenotype 50%
Patients with no variants identified or incompatible with clinical phenotype 50%
Patients with variants compatible with clinical phenotype 60%
Patients with no variants identified or incompatible with clinical phenotype 40%
Patients with variants identified compatible with clinical phenotype 53.3%
Patients with no variants identified or incompatible with clinical phenotype 46.7%

Fig. 4 Clinical and genetic diagnosis in the Piedmontese CKD cohort. Patient cohort is divided on the basis of the clinical suspicion (inner pie). Number and percentage of patients for each macro-category are indicated outside the outer pie, which instead represents the percentage of patients with identified causative variants (variants in line with the clinical phenotype) and patients with no causative variants identified or variants incompatible with the clinical phenotype for each disease category. Specific percentages of these cases are reported on the right with a colour-code legend

adult cohort was carefully selected for patients with a strong suspicion of an underlying genetic condition. In line with the selection of the cohort is the limited number of cases that were re-classified. Of note, 18 out of 60 patients lacking a definitive diagnosis were children. NGS application to this subgroup appeared to be a useful tool as it resulted in the detection of variants in an appreciable number of cases (10 out of 18; 55.5%), and provided a genetic explanation for their clinical condition.

Establishing a precise genetic diagnosis, especially for childhood-onset CKD, allows for pre-emptive screening for extra-renal manifestations. In some cases, the kidneys are not the only affected organs and variants in selective genes may cause syndromic diseases. In other cases, the phenotype is the result of hypomorphic mutations leading to variable expressivity and thus resulting in varying clinical manifestations. Moreover, it must be kept in mind that some disease-causing genes may manifest as de novo variants, with a non-inherited history. Finally, because of the high phenotype heterogeneity, several forms of IKDs may become evident only later in life, when patients reach ESKD. Establishing an early and accurate diagnosis will result in better patient management, improving quality of life, and avoiding useless treatments. Furthermore, it allows early screening of at-risk family members.

This technical approach has some known drawbacks. In exome sequencing, variants occurring in the intronic and promoter regions cannot be identified, and not all genomic regions are equally covered. Moreover, regions with high guanine-cytosine content, and high sequence homology with pseudogenes may be missed. Even detection of copy number variations or structural variants can be difficult and need to be further validated by alternative approaches. An additional limitation of this type of sequencing is represented by the detection of pathogenic variants in the *MUC-1* gene, represented by duplicated C or inserted A nucleotides within the

coding variable-number tandem repeats (VNTRs), which cannot be identified by exome or genome sequencing, but can only be identified by targeted analysis [18]. Finally, we have to underline that some genes known to be associated with specific CKD phenotypes are not included in this clinical exome panel, and thus variants occurring in these genes cannot be investigated. It is also worth pointing out that the list of genes involved in CKD is progressively expanding [9], therefore, applying the updated list of genes in the re-analysis of previously sequenced patients who received a non-conclusive or negative genetic diagnosis may result in the identification of causative genes. Likewise, variants of unknown significance identified by NGS can be re-classified over time, benefiting from periodic updates. These latter observations also justify the choice of the experimental approach adopted in this study based on clinical exome sequencing instead of limited and fixed targeted sequencing panels.

In conclusion, this study shows that clinical exome sequencing is a non-invasive, highly effective tool for genetic diagnosis if the program is supported by careful candidate selection. It can be useful in identifying patients who would benefit from targeted therapies, such as vasopressin 2 antagonists in the case of ADPKD. Furthermore, it may impact on therapy choices, particularly in the case of FSGS, and in the selection of the ideal family member as a kidney donor. This approach is especially applicable in geographic areas where the interaction between a robust nephrological network and genetic facilities is long-standing. Lastly, it can be cost-effective, especially if it is applied early in the diagnostic flow of the patient as it may (1) provide an early diagnosis and (2) avoid unnecessary treatment, while guiding the nephrologist towards the best management of the patient. For all these reasons, this approach could become, in well-characterized cases, an essential step of the diagnostic path.

Author contributions Conception and design: TV, DG, EC, SB, CR, RF, LP, DR, AA, SD. Material support: SBV, SR, GMT, AP, EC, CB, SB, CR, RF, SS, MT, CV, GC, LB, SB, SS, MB, VC, FC, SU, LP, DR. Acquisition, analysis and interpretation of data: TV, MS, MC, SK, VB, FA. Writing of the manuscript: TV, SS, AA, DR, SD.

Funding Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement. This work was supported by the Italian Ministry of Education-University and Research-MIUR, Progetto strategico di Eccellenza Dipartimentale #D15D18000410001 (to the Department of Medical Sciences, University of Turin; members: T. Vaisitti, L. Biancone, A. Amoroso and S. Deaglio); by the University of Turin Dept. of Medical Sciences “ex-60% Ricerca Locale 2018-grant”, RILO (to T. Vaisitti).

Compliance with ethical standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Consent to participate All patients in the present study provided written informed consent.

Ethics statement All participating patients signed an informed consent. Original signed forms are stored at the Immunogenetics and Transplant Biology Service.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Canadas-Garre M, Anderson K, Cappa R, Skelly R, Smyth LJ, McKnight AJ, Maxwell AP (2019) Genetic susceptibility to chronic kidney disease—some more pieces for the heritability puzzle. Front Genet 10:453. https://doi.org/10.3389/fgene.2019.00453
2. Chen TK, Knicely DH, Grams ME (2019) Chronic kidney disease diagnosis and management: a review. JAMA 322(13):1294–1304. https://doi.org/10.1001/jama.2019.14745
3. Satko SG, Freedman BI (2005) The familial clustering of renal disease and related phenotypes. Med Clin N Am 89(3):447–456. https://doi.org/10.1016/j.mcna.2004.11.011
4. Skrunes R, Svarstad E, Reisaeter AV, Vikse BE (2014) Familial clustering of ESRD in the Norwegian population. Clin J Am Soc Nephrol 9(10):1692–1700. https://doi.org/10.2215/CJN.01680214
5. Connaughton DM, Bukhari S, Conlon P, Cassidy E, O’Toole M, Mohamad M, Flanagan J, Butler T, O’Leary A, Wong L, O’Regan J, Moran S, O’Kelly P, Logan V, Griffin B, Griffin M, Lavin P, Little MA, Conlon P (2015) The Irish kidney gene project-prevalence of family history in patients with kidney disease in Ireland. Nephron 130(4):293–301. https://doi.org/10.1159/000436983
6. Freedman BI, Volkova NV, Satko SG, Krisher J, Jurkovitz C, Soucie JM, McClellan WM (2005) Population-based screening for family history of end-stage renal disease among incident dialysis patients. Am J Nephrol 25(6):529–535. https://doi.org/10.1159/000088491
7. Connaughton DM, Hildebrandt F (2019) Personalized medicine in chronic kidney disease by detection of monogenic mutations. Nephrol Dial Transplant. https://doi.org/10.1093/ndt/gfz0285365509

8. Groopman EE, Marasa M, Cameron-Christie S, Petrovski S, Aggarwal VS, Milo-Rasouly H, Li Y, Zhang J, Nestor J, Krithivasan P, Lam WY, Mitrotti A, Piva S, Kil BH, Chatterjee D, Reingold R, Bradbury D, DiVecchia M, Snyder H, Mu X, Mehl K, Balderes O, Fasel DA, Weng C, Radhakrishnan J, Canetta P, Appel GB, Bomback AS, Ahn W, Uy NS, Alam S, Cohen DJ, Crew RJ, Dube GK, Rao MK, Kamalakaran S, Copeland B, Ren Z, Bridgers J, Malone CD, Mebane CM, Dagaonkar N, Fellstrom BC, Haefliger C, Mohan S, Sanna-Cherchi S, Kiryluk K, Fleckner J, March R, Platt A, Goldstein DB, Gharavi AG (2019) Diagnostic utility of exome sequencing for kidney disease. N Engl J Med 380(2):142–151. https://doi.org/10.1056/NEJMoa1806891
9. Mann N, Braun DA, Amann K, Tan W, Shril S, Connaughton DM, Nakayama M, Schneider R, Kitzler TM, van der Ven AT, Chen J, Ityel H, Vivante A, Majmundar AJ, Daga A, Warejko JK, Lovric S, Ashraf S, Jobst-Schwan T, Widmeier E, Hugo H, Mane SM, Spaneas L, Somers MJG, Ferguson MA, Traum AZ, Stein DR, Baum MA, Daouk GH, Lifton RP, Manzi S, Vakili K, Kim HB, Rodig NM, Hildebrandt F (2019) Whole-exome sequencing enables a precision medicine approach for kidney transplant recipients. J Am Soc Nephrol 30(2):201–215. https://doi.org/10.1681/ASN.2018060575
10. Hildebrandt F (2010) Genetic kidney diseases. Lancet 375(9722):1287–1295. https://doi.org/10.1016/S0140-6736(10)60236-X
11. Ottlewski I, Munch J, Wagner T, Schonauer R, Bachmann A, Weimann A, Hentschel J, Lindner TH, Seehofer D, Bergmann C, Jamra RA, Halbritter J (2019) Value of renal gene panel diagnostics in adults waiting for kidney transplantation due to undetermined end-stage renal disease. Kidney Int 96(1):222–230. https://doi.org/10.1016/j.kint.2019.01.038
12. Mansilla MA, Sompallae RR, Nishimura CJ, Kwitek AE, Kimble MJ, Freese ME, Campbell CA, Smith RJ, Thomas CP (2019) Targeted broad-based genetic testing by next-generation sequencing informs diagnosis and facilitates management in patients with kidney diseases. Nephrol Dial Transplant. https://doi.org/10.1093/ndt/gfz173
13. van Nimwegen KJ, van Soest RA, Veltman JA, Nelen MR, van der Wilt GJ, Vissers LE, Grutters JP (2016) Is the $1000 genome as near as we think? A cost analysis of next-generation sequencing. Clin Chem 62(11):1458–1464. https://doi.org/10.1373/clinchem.2016.258632
14. Phillips KA, Deverka PA, Hooker GW, Douglas MP (2018) Genetic test availability and spending: where are we now? Where are we going? Health Aff (Millwood) 37(5):710–716. https://doi.org/10.1377/hlthaff.2017.1427
15. Stokman MF, Renkema KY, Giles RH, Schaefer F, Knoers NV, van Eerde AM (2016) The expanding phenotypic spectra of kidney diseases: insights from genetic studies. Nat Rev Nephrol 12(8):472–483. https://doi.org/10.1038/nrneph.2016.87nrneph.2016.87
16. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17(5):405–424. https://doi.org/10.1038/gim.2015.30
17. Bullich G, Domingo-Gallego A, Vargas I, Ruiz P, Lorente-Grandoso L, Furlano M, Fraga G, Madrid A, Ariceta G, Borregan M, Pinero-Fernandez JA, Rodriguez-Pena L, Ballesta-Martinez MJ, Llano-Rivas I, Menica MA, Ballarin J, Torrents D, Torra R, Ars E (2018) A kidney-disease gene panel allows a comprehensive genetic diagnosis of cystic and glomerular inherited kidney diseases. Kidney Int 94(2):363–371. https://doi.org/10.1016/j.kint.2018.02.027
18. Bleyer AJ, Kmoch S (1993) Autosomal dominant tubulointerstitial kidney disease, *MUC1*-related. In: Adam MP, Ardinger HH, Pagon RA et al (eds) GeneReviews®. University of Washington, Seattle

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Tiziana Vaisitti[1] · Monica Sorbini[1] · Martina Callegari[2] · Silvia Kalantari[2] · Valeria Bracciamà[2] · Francesca Arruga[1] · Silvia Bruna Vanzino[2] · Sabina Rendine[2] · Gabriele Togliatto[1] · Daniela Giachino[3,4] · Alessandra Pelle[3] · Enrico Cocchi[5] · Chiara Benvenuta[5] · Simone Baldovino[4,6] · Cristiana Rollino[6] · Roberta Fenoglio[6] · Savino Sciascia[6] · Michela Tamagnone[7] · Corrado Vitale[8] · Giovanni Calabrese[9] · Luigi Biancone[1,10] · Stefania Bussolino[11] · Silvana Savoldi[11] · Maurizio Borzumati[12] · Vincenzo Cantaluppi[13] · Fabio Chiappero[14] · Silvana Ungari[15] · Licia Peruzzi[5] · Dario Roccatello[4,6] · Antonio Amoroso[1,2] · Silvia Deaglio[1,2]

[1] Department of Medical Sciences, University of Turin, via Santena 19, 10126 Turin, Italy

[2] Immunogenetics and Transplant Biology Service, Città della Salute e della Scienza University Hospital, Turin, Italy

[3] Service of Genetic Counseling, San Luigi Gonzaga University Hospital, Orbassano, Turin, Italy

[4] Department of Clinical and Biological Sciences, University of Turin, Turin, Italy

[5] Pediatric Nephrology Dialysis and Transplantation Unit, Città della Salute e della Scienza University Hospital, Turin, Italy

[6] Nephrology and Dialysis Unit (ERKnet Member)-CMID, Center of Research of Immunopathology and Rare Diseases, San Giovanni Bosco Hospital, Turin, Italy

[7] Nephrology and Dialysis Unit ASL CN1, Cuneo, Italy

[8] Nephrology and Dialysis Unit, Ordine Mauriziano di Torino, Turin, Italy

[9] Nephrology and Dialysis Unit ASL AL, Alessandria, Italy

[10] Renal Transplantation Unit 'A. Vercellone,' Division of Nephrology Dialysis and Transplantation, Città della Salute e della Scienza University Hospital, Turin, Italy

[11] Nephrology and Dialysis Unit ASL TO4, Turin, Italy

[12] Nephrology and Dialysis Unit of Verbania ASL VCO, Verbano Cusio Ossola, Verbania, Italy

[13] Nephrology and Kidney Transplantation Unit, Maggiore Della Carità University Hospital, Novara, Italy

[14] Nephrology and Dialysis Unit ASL TO3, Turin, Italy

[15] Struttura Semplice Genetics and Molecular Biology, ASL CN1, Cuneo, Italy

Journal of the Neurological Sciences 418 (2020) 117111

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

journal homepage: www.elsevier.com/locate/jns

Clinical short communication

The potential impact of enhanced hygienic measures during the COVID-19 outbreak on hospital-acquired infections: A pragmatic study in neurological units

Emanuele Cerulli Irelli[a], Biagio Orlando[a], Enrico Cocchi[b], Alessandra Morano[a], Francesco Fattapposta[a], Vittorio Di Piero[a], Danilo Toni[a], Maria R. Ciardi[c], Anna T. Giallonardo[a], Giovanni Fabbrini[a,d], Alfredo Berardelli[a,d], Carlo Di Bonaventura[a,*]

[a] Department of Human Neurosciences, Policlinico Umberto I, Sapienza University of Rome, Italy
[b] Department of Public Health and Pediatrics, University of Turin, Turin, Italy
[c] Department of Public Health and Infectious Diseases, Sapienza University of Rome, Italy
[d] IRCCS Neuromed, Pozzilli, IS, Italy

ARTICLE INFO

Keywords:
COVID-19
Sars-CoV-2
Hospital infections
Stroke unit
Neurology unit

ABSTRACT

Objective: Hospital-acquired infections (HAIs) are frequent complications among acute patients hospitalized in neurological units, especially among those hospitalized for stroke. This study aimed to investigate if enhanced hygienic measures, including the systematic use of personal protective equipment (PPE), determined a decrease in HAI during the recent COVID-19 outbreak in "COVID-free" neurological units.

Methods: Patients hospitalized in neurology and stroke units of Policlinico Umberto I Hospital in Rome from March 8, 2020 and discharged prior to May 31, 2020 were included in the study and compared with patients hospitalized during the same period in 2019.

Results: A total of 319 patients were included in the study ($n = 103$ in 2020, $n = 216$ in 2019). Among patients hospitalized in 2019, the incidence of HAIs was 31.5% (95% confidence interval (CI): 0.25–0.38), compared with 23.3% (95% CI: 0.15–0.32) in 2020 ($p = 0.12$). Multivariable logistic regression showed that hospitalization during 2020 was independently associated with a lower risk of HAIs (odds ratio: 0.34, 95% CI:0.16–0.71, $p = 0.004$). Poisson regression models showed that hospitalization during 2020 was also independently associated with both a lower number of HAIs (relative risk [RR]: 0.56, 95% CI:0.38–0.81, $p = 0.01$) and a lower number of prescribed antibiotics per patient (RR: 0.66, 95% CI: 0.49–0.87, $p = 0.02$).

Conclusion: Our study design provides evidence regarding the impact of stricter hygienic measures, such as increased PPE use, on HAIs. Larger studies are needed to support the extension of preventive measures even after the COVID-19 outbreak in order to limit the occurrence of HAIs.

1. Introduction

The first autochthonous severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection-associated syndrome (COVID-19) in Italy was diagnosed in the second half of February and the infection rapidly spread across the country during the subsequent weeks [1,2]. During this COVID-19 pandemic, an extraordinary increase in the use of personal protective equipment (PPE) was implemented for both healthcare workers and patients [3,4]. Moreover, greater attention was given to personal hygienic precautions, such as hand washing and medical equipment disinfection [5]. Both PPE and hand hygiene are considered powerful preventive measures against hospital-acquired infections (HAIs) [6,7]. HAIs represent a frequent and serious complication among hospitalized patients, which can result in increased morbidity and mortality rates and costs [8]. Despite some contrasting results, HAIs represent a particularly frequent complication among hospitalized neurological patients, especially among those suffering from stroke or other conditions burdened by disability and dysphagia [9]. Recent European reports have depicted a challenging situation in Italy in terms of HAI-related deaths and antimicrobial resistance [10,11].

With this background, we aimed to analyze the incidence of HAIs

* Corresponding author.
E-mail address: c_dibonaventura@yahoo.it (C. Di Bonaventura).

https://doi.org/10.1016/j.jns.2020.117111
Received 7 July 2020; Received in revised form 12 August 2020; Accepted 26 August 2020
Available online 29 August 2020
0022-510X/ © 2020 Elsevier B.V. All rights reserved.

among patients hospitalized in neurological units during the COVID-19 outbreak in Italy as compared with the same period of the previous year.

2. Methods

The study was conducted according to STROBE guidelines as a retrospective case-control study and approved by the local Ethics committee. Data from patients hospitalized in the neurology and stroke units of Policlinico Umberto I Hospital in Rome from March 8, 2020 and discharged prior to May 31, 2020 were retrospectively reviewed (study group) and compared with hospitalized patients during the same period in 2019 (control group). The neurology and stroke divisions, which usually include a total of 42 beds, were considered "COVID-free" from the beginning of the pandemic and the number of beds was limited to 23. The following hygienic measures were adopted in our units during the COVID-19 outbreak: All patients wore a surgical mask during the hospital staying; healthcare workers systematically used protective face masks (FFP2 or surgical depending on whether the contact with patients was either direct or indirect), as well as disposable coats, gloves and safety goggles for eye protection; a distance of at least 1.5 m between hospital beds was also guaranteed. The study included patients who came to neurological units directly from the emergency room (ER). In order to ensure the homogeneity of the study setting and avoid possible confounding factors, patients transferred to our unit from wards other than the ER were excluded from the study. This study design gave us the opportunity to evaluate the impact of increased PPE use without the confounding effect of SARS-CoV-2 co-infection.

We collected demographic data, information on medical comorbidities, length and reason of hospitalization, number of invasive devices (urinary catheter, peripheral venous line, central venous line, nasogastric tubes, etc.), occurrence and number of HAIs, and the number of prescribed antibiotics. Antibiotic prophylaxis was never used in any patient during both years, and no antibiotic was prescribed before the occurrence of a HAI. HAIs were defined as infections occurring after at least 48 h of hospitalization [12]. We also included aspiration pneumonias occurring after this time point so as to encompass the whole spectrum of infections that may complicate neurological unit hospitalization.

The primary outcome of our study was to assess possible differences in HAIs and antibiotic prescriptions between study group patients hospitalized during the COVID-19 outbreak and control group patients hospitalized during the same period in 2019.

2.1. Statistical analysis

Descriptive statistic methods and data visualization were used to assess data distribution. Comparisons across means were performed with independent sample *t*-test in cases of normal distribution, whereas the Mann-Whitney *U* test was used in cases of non-normal distribution. Comparisons across proportions were performed with Fisher's exact test.

A multivariable logistic regression model (M0) was elaborated to evaluate the effect of hospitalization in 2020 on the presence of HAIs. The model was corrected for the number of invasive devices and medical comorbidities, age, sex, and diagnosis of ischemic/hemorrhagic stroke. The variable "number of comorbidities" was obtained by the sum of the medical comorbidities (i.e diabetes, arterial hypertension, history of previous cancer, obesity, history of previous stroke, chronic kidney failure, chronic obstructive pulmonary disease, peripheral arterial disease and neurodegenerative diseases) reported during patients' previous medical history. This choice allowed us to reduce the number of covariates included in the model and to obtain an event per variable (EPV) of 15.

Due to a Poisson distribution of the dependent variable, Poisson generalized linear models were elaborated to assess the effect of hospitalization in 2020 on the number of infections (M1) and antibiotics (M2) per patient. M1 and M2 were corrected for the same covariates as M0. Analyses were performed and figures generated using R 3.5.1 (R Project for Statistical Computing, Vienna, Austria).

Table 1
Demographic and clinical characteristics stratified by year of hospitalization.

	2019 *N* = 216 pts.	2020 *N* = 103 pts	*p* value
Age, n, mean (SD)	65.1 (18.8)	65.4 (17.7)	0.9
Female sex, n, %	110 (50.9)	40 (38.8)	0.04*
Arterial hypertension, n, %	133 (61.6)	62 (60.2)	0.8
Diabetes, n, %	35 (16.3)	24 (23.3)	0.13
Cardiopathy, n, %	61 (28.2)	32 (31.1)	0.6
COPD, n, %	16 (7.4)	8 (7.8)	0.9
History of cancer, n, %	26 (12)	15 (14.6)	0.5
Obesity, n, %	7 (3.2)	7 (6.8)	0.1
Neurodegenerative disease, n, %	23 (10.6)	16 (15.5)	0.2
Chronic kidney failure, n, %	9 (4.2)	6 (5.8)	0.5
Peripheral arterial disease, n, %	12 (5.6)	16 (15.5)	0.003*
History of previous stroke, n, %	22 (10.2)	12 (11.9)	0.6
History of MS or autoimmune disease, n, %	9 (4.2)	6 (5.9)	0.5
Number of comorbidities, n, mean (SD)	1.6 (1.3)	1.9 (1.5)	0.06
Length of hospitalization, n (SD)	14.2 (14.3)	11.4 (8.4)	0.2
Number of invasive devices, n, mean (SD)	1.5 (0.7)	1.7 (0.6)	0.002*
Central venous catheters, n, %	20 (9.3)	10 (9.7)	0.9
Number of antibiotics, n, mean (SD)	0.75 (1.5)	0.63 (1.3)	0.039*
Subjects admitted for stroke, n, %	114 (52.8)	48 (46.6)	0.3

Abbreviations: COPD = chronic obstructive pulmonary disease; MS = multiple sclerosis; SD = standard deviation. * p value less than 0.05.

3. Results

A total of 319 patients were included in the study according to study criteria. This figure included 216 control group patients hospitalized in 2019 and 103 study group patients hospitalized in 2020, representing a 52.3% decrease in the hospitalization rate in our units in 2020. Clinical and demographic data of the study and control groups are reported in Table 1. Ischemic stroke was the most frequent cause of hospitalization, accounting for 89 patients in 2019 and 41 in 2020, whereas hemorrhagic stroke was diagnosed in 25 and 7 subjects in 2019 and 2020, respectively. The causes of hospitalization in study and control patients are detailed in Fig. 1. Among patients with neurodegenerative diseases,

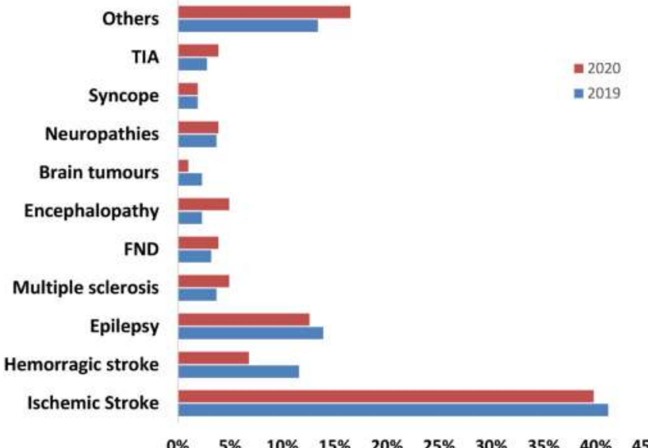


Fig. 1. Discharge diagnoses stratified by year of hospitalization.
Discharge diagnoses expressed as percentages. Patients hospitalized during 2019 are in blue while those hospitalized during 2020 are in red. Abbreviations: FND = functional neurological disorders; TIA = transient ischemic attack.

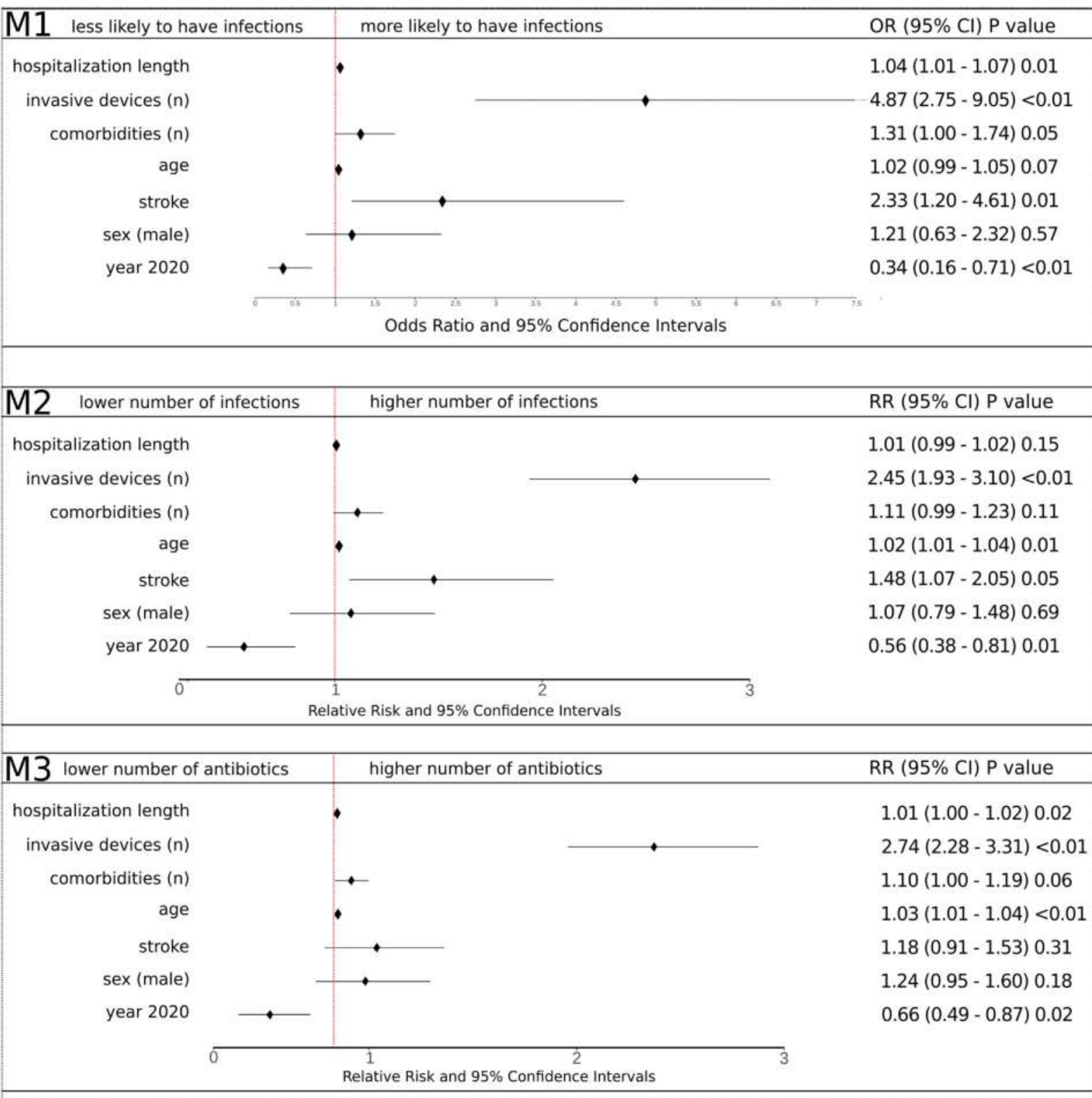


Fig. 2. Forest plot of regression analyses.
Forest plot of multivariable regression analyses. All independent variables in the models are represented. The upper panel shows the multivariable logistic regression model (M0) and relative factors associated with the absence/presence of hospital-acquired infections (HAIs). The X-axis represents the relative odds ratio (OR) and 95% confidence interval (CI) of each covariate. *P* values are reported on the right. The middle panel shows the multivariable Poisson regression model with the number of infections per patient as the outcome (M1). For each variable, the relative risk (RR) and 95% CI are reported and *P* values are shown on the right. The lower panel shows the multivariable Poisson regression model with the number of antibiotics administered per patient as the outcome (M2). For each variable, the RR and 95% CI are reported and P values are shown on the right.

the most frequent was Alzheimer's disease, observed in 9 subjects during 2019 and in 6 during 2020, followed by Parkinson's Disease, accounting for 8 and 3 patients in 2019 and 2020, respectively.

The overall number of HAIs considering both the years was 92 (28.8% of the patients). Among control patients, HAI incidence was 31.5% (95% confidence interval (CI): 0.25–0.38), as compared with 23.3% (95% CI: 0.15–0.32) in study patients (Fisher's exact test, $p = 0.12$). The number of HAIs per patient was 0,44 (95% CI: 0.34–0.55) among control patients, as compared with 0,26 (95% CI: 0.15–0.37) in study patients ($p = 0.02$).

According to M0, hospitalization in 2020 was independently associated with a decreased incidence of HAIs (odds ratio: 0.34, 95% CI: 0.16–0.71, $p = 0.004$) (Fig. 2).

Hospitalization in 2020 was also associated with a significant reduction in the number of HAIs per patient (M1: relative risk (RR): 0.56, 95% CI: 0.38–0.81, $p < 0.01$) (Fig. 2) and a significant reduction in the number of antibiotics per patient (M2: RR: 0.66, 95% CI: 0.49–0.87, $p = 0.02$) (Fig. 2).

4. Discussion

In our single-center case-control study, we found that the patients hospitalized between March 8–May 31, 2020 had a reduced likelihood of HAIs as compared with patients hospitalized during the same period in 2019. We also found a reduced number of antibiotic prescriptions per patient during the COVID-19 outbreak as compared with the same period in 2019. The documented reduced HAI risk could be related to increased hygienic measures and widespread use of PPE due to the

concurrent COVID-19 outbreak. Besides, the 1.5-m distance between hospital beds might have contributed as well. To the best of our knowledge, this is the first study reporting a possible association between reduced HAIs and stricter hygienic measures, including the use of PPE, among healthcare workers during the COVID-19 pandemic.

As regards the overall rate of HAIs in our cohort, we found a high incidence of infections as compared to literature [10]. This finding may be due to the high prevalence of stroke patients, who are known to be at high risk of HAIs [9,13], and to the frequent use of invasive devices (e.g. urinary and central catheters, nasogastric tubes, etc.) in these subjects, which is also associated with a higher HAI risk [11,14]. As expected, we also found that the use of invasive devices, a stroke diagnosis, and the length of hospitalization were significantly associated with an increased HAI risk, as previously reported by other authors [9,15].

In conclusion, our study found decreased HAI occurrence in patients hospitalized in our neurological unit during the COVID-19 outbreak with respect to the previous year, in spite of comparable clinical severity. Considering the possible association between reduced HAI risk and adherence to stricter hygienic measures, our data further highlights the great impact of precautionary measures, including the use of PPE, on hospitalized patients' care and management. These findings appear particularly relevant when considering the potential burden of HAIs on the Italian healthcare system in terms of both mortality and public spending.

5. Limitations

The main limitations of our study are the relatively small number of patients and the single-center retrospective design. Larger studies are needed to confirm our findings and to support the extension of preventive measures even after the SARS-CoV-2 outbreak in order to limit the occurrence of HAIs among hospitalized patients.

Funding/support

None.

Author contributions

Cerulli Irelli: Conceptualization, Data curation, Investigation, Formal analysis, Writing - original draft, review and editing. Orlando: Data curation, Methodology, Writing- review and editing. Cocchi: Conceptualization, Methodology, Software, Writing- review and editing. Morano: Data curation, Writing- review and editing. Fattapposta: Supervision, Writing- review and editing. Di Piero: Supervision, Writing-review and editing. Toni: Methodology, Supervision, Validation, Writing-review and editing. Ciardi: Validation, Writing-review and editing. Giallonardo: Data curation, Supervision, Writing- review and editing. Fabbrini: Methodology, Writing- review and editing. Berardelli: Methodology, Project administration, Supervision, Writing-review and editing. Di Bonaventura: Conceptualization, Data curation, Methodology, Supervision, Validation, Writing- original draft, review and editing.

Declaration of Competing Interest

None.

References

[1] A. Gagliano, P.G. Villani, F.M. Co', et al., COVID-19 epidemic in the Middle Province of northern Italy: impact, logistics, and strategy in the first line hospital [published online ahead of print, 2020 mar 24], Disaster Med. Public Health Prep. (2020) 1–5, https://doi.org/10.1017/dmp.2020.512.

[2] M. Gatto, E. Bertuzzo, L. Mari, et al., Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures, Proc. Natl. Acad. Sci. U. S. A. 117 (19) (2020) 10484–10491, https://doi.org/10.1073/pnas.2004978117.

[3] Centers for Disease Control and Prevention, Interim Infection Prevention and Control Recommendations for Patients with Suspected or Confirmed Coronavirus Disease 2019 (COVID-19) in Healthcare Settings, Available at, 2020. https://www.cdc.gov/coronavirus/2019-ncov/hcp/infection-control-recommendations.html.

[4] World Health Organization, Rational use of Personal Protective Equipment for Coronavirus Disease (COVID-19) and Considerations During Severe Shortages: Interim Guidance, Avalaible at, 2020. https://apps.who.int/iris/handle/10665/331695.

[5] N. Lotfinejad, A. Peters, D. Pittet, Hand hygiene and the novel coronavirus pandemic: The role of healthcare workers [published online ahead of print, 2020 Mar 19], J. Hosp. Infect. (2020), https://doi.org/10.1016/j.jhin.2020.03.017 S0195–6701(20)30116-X.

[6] World Health Organization. Guidelines on hand hygiene in health care, First Global Patient Safety Challenge Clean Care Is Safer Care, Available at: https://apps.who.int/iris/bitstream/handle/10665/44102/9789241597906_eng.pdf?sequence=1.

[7] B. Allegranzi, D. Pittet, Role of hand hygiene in healthcare-associated infection prevention, J. Hosp. Infect. 73 (4) (2009) 305–315, https://doi.org/10.1016/j.jhin.2009.04.019.

[8] M. Haque, M. Sartelli, J. McKimm, M. Abu Bakar, Health care-associated infections - an overview, Infect. Drug Resist. 11 (2018) 2321–2333 Published 2018 Nov 15 https://doi.org/10.2147/IDR.S177247.

[9] H.C. Emsley, S.J. Hopkins, Acute ischaemic stroke and infection: recent and emerging concepts, Lancet Neurol. 7 (4) (2008) 341–353, https://doi.org/10.1016/S1474-4422(08)70061-9.

[10] C. Suetens, K. Latour, T. Kärki, et al., Prevalence of healthcare-associated infections, estimated incidence and composite antimicrobial resistance index in acute care hospitals and long-term care facilities: results from two European point prevalence surveys, 2016 to 2017 [published correction appears in Euro Surveill. 2018 Nov;23(47):], Euro Surveill. 23 (46) (2018), https://doi.org/10.2807/1560-7917.ES.2018.23.46.1800516 1800516.

[11] European Centre for Disease Prevention and Control, Surveillance of Antimicrobial Resistance in Europe, Avalaible at, 2018. https://www.ecdc.europa.eu/en/publications-data/surveillance-antimicrobial-resistance-europe-2018.

[12] A. Revelas, Healthcare - associated infections: a public health problem, Niger Med. J. 53 (2) (2012) 59–64, https://doi.org/10.4103/0300-1652.103543.

[13] S. Aslanyan, C.J. Weir, H.C. Diener, M. Kaste, K.R. Lees, GAIN international steering committee and investigators. Pneumonia and urinary tract infection after acute ischaemic stroke: a tertiary analysis of the GAIN international trial, Eur. J. Neurol. 11 (1) (2004) 49–53, https://doi.org/10.1046/j.1468-1331.2003.00749.x.

[14] L.E. Nicolle, Catheter associated urinary tract infections, Antimicrob Resist Infect Control 3 (2014) 23 Published 2014 Jul 25 https://doi.org/10.1186/2047-2994-3-23.

[15] G. Pellizzer, P. Mantoan, L. Timillero, et al., Prevalence and risk factors for nosocomial infections in hospitals of the Veneto region, North-Eastern Italy, Infection 36 (2) (2008) 112–119, https://doi.org/10.1007/s15010-007-7092-x.

Clinical Genetic Screening in Adult Patients with Kidney Disease

Enrico Cocchi [1,2] *Jordan Gabriela Nestor* [1] *and Ali G. Gharavi* [1,3]

Abstract

Expanded accessibility of genetic sequencing technologies, such as chromosomal microarray and massively parallel sequencing approaches, is changing the management of hereditary kidney diseases. Genetic causes account for a substantial proportion of pediatric kidney disease cases, and with increased utilization of diagnostic genetic testing in nephrology, they are now also detected at appreciable frequencies in adult populations. Establishing a molecular diagnosis can have many potential benefits for patient care, such as guiding treatment, familial testing, and providing deeper insights on the molecular pathogenesis of kidney diseases. Today, with wider clinical use of genetic testing as part of the diagnostic evaluation, nephrologists have the challenging task of selecting the most suitable genetic test for each patient, and then applying the results into the appropriate clinical contexts. This review is intended to familiarize nephrologists with the various technical, logistical, and ethical considerations accompanying the increasing utilization of genetic testing in nephrology care.

CJASN 15: 1497–1510, 2020. doi: https://doi.org/10.2215/CJN.15141219

[1]Division of Nephrology and Center for Precision Medicine and Genomics, Department of Medicine, Columbia University, New York, New York
[2]Department of Pediatrics, Universita' degli Studi di Torino, Torino, Italy
[3]Insititute of Genomic Medicine, Columbia University, New York, New York

Correspondence: Dr. Ali Gharavi, Division of Nephrology and Center for Precision Medicine and Genomics, Columbia University College of Physicians and Surgeons, 1150 St. Nicholas Avenue, Russ Berrie Pavilion 413, New York, NY 10032. Email: ag2239@cumc.columbia.edu

Introduction

Kidney disease is associated with significant morbidity and mortality and affects over 20 million patients in the United States (1). A family history of nephropathy is reported in approximately 30% of cases, revealing the strong role of genetics in kidney disease (2–4). Genetic testing is increasingly used in clinical nephrology due to expanded utilization and accessibility of genetic sequencing technologies (5,6). As a result, Mendelian forms of kidney disease are increasingly detected in adult and pediatric patients. In fact, although historically more clinically apparent in pediatric populations, it is now clear that genetic forms of kidney disease are also highly prevalent in adults, with some studies reporting a Mendelian cause of kidney disease in up to 37% of adult cases (7,8). Establishing a genetic diagnosis has significant implications for nephrology care because it may inform prognosis (9–11) and selection of therapy (12,13), spare patients from undergoing invasive diagnostic procedures such as a kidney biopsy (14,15), and guide family planning (2).

Many genetic testing modalities are currently available (*e.g.*, targeted sequencing, microarrays, gene panels, genome-wide approaches, *etc.*), and selection of the most appropriate diagnostic sequencing approach is made on the basis of various factors. These include diagnostic yield of the different sequencing modalities, the patients' clinical picture, their preferences for the types of results that may emerge with broader sequencing approaches, out-of-pocket costs, and third-party payer coverage. This review is intended to familiarize clinical nephrologists with concepts relating to clinical genetic testing.

Mendelian Nephropathies, Genetic Testing Modalities, and Diagnostic Yields

The human genome is divided into protein-coding (approximately 1%, known as the "Exome," composed of approximately 408,659 exons) and noncoding sequences (approximately 99%) (16,17). Overall, it harbors approximately 20,000 genes, of which approximately 4100 are currently associated with Mendelian disorders (18). Human genomes differ greatly between individuals, and variations in genetic sequence are summarized in Figure 1 and include:

- Single nucleotide variants (SNVs): substitution of a single base.
- Small insertion-deletion (INDEL): insertion or deletion of approximately 2–1000 bases.
- Copy number variation: duplication/deletion that affects ≥1 kb in one or more loci. Copy number variants encompass 5%–10% of the human genome (19,20). The differentiation between small INDELs and copy number variants is specifically on the basis of the length of the affected DNA, reflecting a higher number of genes possibly involved in the latter.
- Chromosomal imbalance and rearrangements: deletions and duplications of entire chromosomes or segments of chromosomes. Inversions and translocations can also occur as a result of genome breakage followed by a rejoining of the broken ends in a different order than the original one.

These variations can potentially lead to a Mendelian disease. A recent survey identified 625 Mendelian disorders associated with kidney and urological traits (21), whereas the number of gene–disease associations continues to grow with the expanded use of massively parallel sequencing, a technological advancement that

 Copyright © 2020 by the American Society of Nephrology

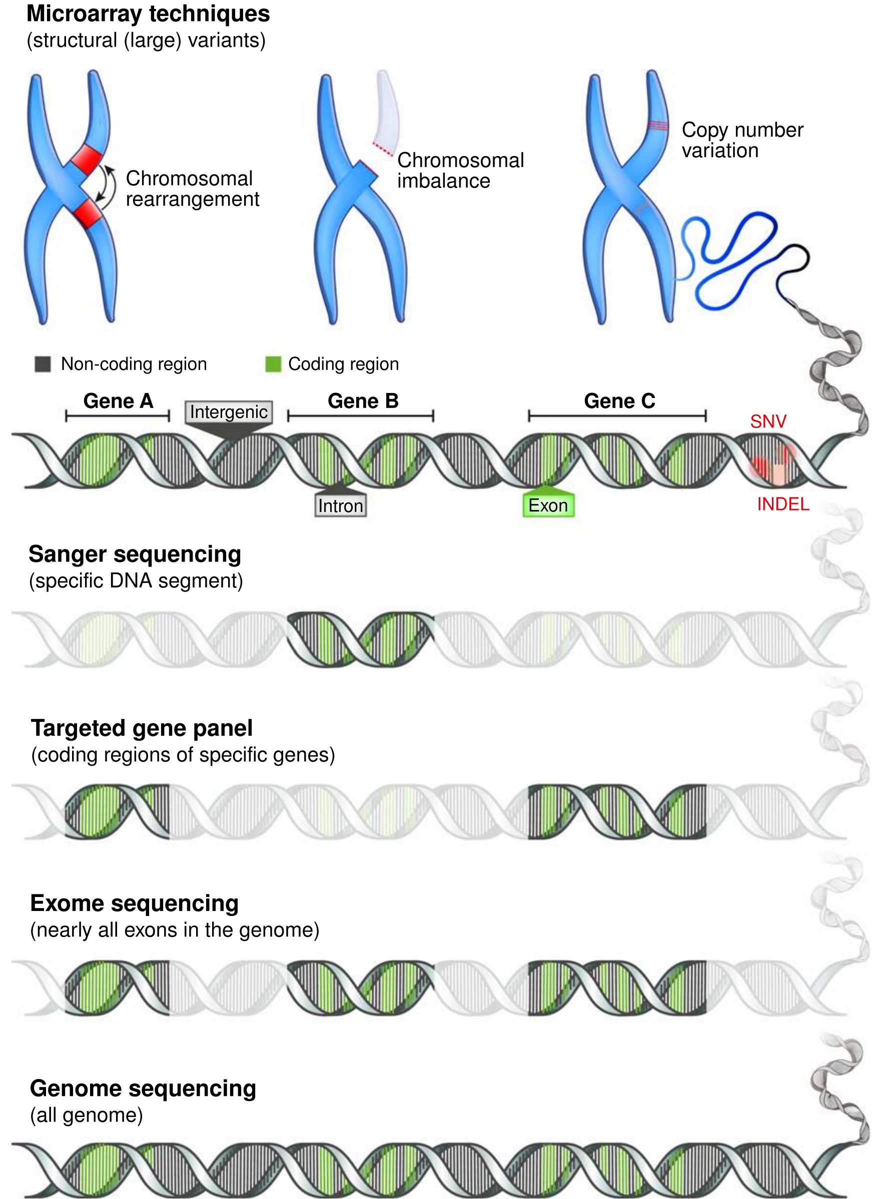


Figure 1. | **Genetic testing options.** Different classes of variants identifiable in human genome and how they are detected by the currently available genetic tests. The top part shows a "large-scale" section where an entire chromosome is represented (blue), as well as the relative large alterations (aka structural variants); chromosomal rearrangement (deletions, duplications, inversions, and translocations caused by a breakage in the DNA at two different locations, followed by a rejoining of the broken ends to produce a new chromosomal arrangement of genes); chromosomal imbalance (absence or duplication of a chromosomal portion) identifiable primarily through chromosomal array–based techniques; and copy number variation (a duplication or deletion that affects a large stretch of sequence, at least 1 kb). Small variants: single nucleotide variants (SNVs; substitution of a single base) and small insertion/deletions (INDELs; involving more than one base). We also see the coding (green), noncoding (dark gray), and regulatory (noncoding) portions of the genome that constitute each gene (gene A, gene B, and gene C) and how these are sequenced and analyzed through different techniques. Sanger sequencing is limited to a narrow portion of the genome, usually a single gene or small regions of a gene. Targeted gene panels: only coding portions of a specific set of genes are targeted. Exome sequencing captures nearly all of the coding sequences in an individual's genome. Genome sequencing covers nearly all regions of the genome (coding and noncoding regions) of an individual.

has increased the throughput of genomic sequencing (21). Importantly, diagnostic yield varies according to the categories of variants and choice of test. In Figure 1 and Table 1, we summarize technical and clinical aspects of different sequencing approaches, along with their respective benefits and drawbacks. Next, we discuss the major genomic diagnostic modalities.

Targeted Dideoxy Terminator (Sanger) Sequencing

This test aims to identify SNVs and INDELs in a specific gene or gene region (22). The execution is simple and results are highly reliable (error rate: 0.001%–1%) (23). Moreover, this method achieves a long reading length (approximately 800 bp) resulting in significant advantages for *de novo* variants confirmation and sequencing of

Table 1. Genetic testing options

Analysis Modality	Primary Scope	Advantages	Disadvantages	Uses
Sanger sequencing	• Identification of small variants (SNVs/INDELs) in a specific DNA region	• Simple technical execution • High analytical accuracy (error rate 0.001%–1%) • Fast and simple interpretation • No risk of secondary results • Long reading lengths (approximately 800 bp)	• Limited resolution (<1 kb) unsuitable for large structural variants • Time- and cost-inefficient for analysis of large DNA segments	• Confirmation of a specific suspected mutation in a gene • Confirmation of MPS-identified variants • Analysis of regions refractory to MPS analysis, such as repetitive regions
Chromosomal microarrays	• Identification of small chromosomal rearrangements/CNVs (≥200–400 kb)	• Higher resolution than standard karyotyping (50–100 kb) • Genome-wide analysis	• Cannot detect small mutations/CNVs • Limited ability to detect balanced chromosomal rearrangements, low-grade somatic mosaicism, and CNVs in pseudogenes and repetitive regions	• Patients with phenotype strongly suggestive of large rearrangements, such as multiple congenital anomalies and developmental diseases
Targeted MPS panels	• Identification of small variants (SNVs/INDELs) within genes of interest for the clinical phenotype	• Analysis of all genes possibly related to specific phenotypes • Diagnostic yield up to 50% (depending on the patient phenotype and genes selection method) • Restricted number of genes that minimizes risk of secondary findings and reduces analysis time	• Error rate approximately 0.5%–2% (MPS) • Short reading length (generally 50–300 bp) (MPS) • Limited reanalysis utility • Reliability depends on challenging panel design	• Patients with phenotypes pointing to specific disorders and with low genetic heterogeneity
Exome sequencing	• Identification of small variants (SNVs/INDELs) within coding regions of the genome	• Analysis of all coding regions in the genome • Unbiased approach increases diagnostic sensitivity • Cover almost all sites related to Mendelian diseases (approximately 85%)	• Coverage per base is generally lower than with targeted panels • Challenging and time-consuming interpretation (high number of candidate variants) • Potential for detection of secondary findings • Limited coverage in repetitive regions • Limited reliability for INDELs	• Patients undiagnosed with more specific methodologies • Screening of patients with undefined phenotype • Patients with heterogeneous/unspecific phenotype
Genome sequencing	• Identification of small variants (SNVs/INDELs) within coding and noncoding regions of the genome	• Identification of deep splicing and intronic variants unidentifiable with other techniques • Better analytic performance than exome sequencing • Efficient CNV identification • Extremely useful for reanalysis	• Maximizes results interpretation difficulty and time • Maximizes potential detection of secondary findings	• Patients undiagnosed with other methodologies • Screening of patients with undefined phenotype • Patients with heterogeneous/unspecific phenotype

SNV, single nucleotide variant; INDEL, insertion/deletion; MPS, massively parallel sequencing; CNV, copy number variation.

repetitive regions (24). *De novo* variants are those that are absent in parents and found only in the proband (*i.e.*, the individual undergoing genetic sequencing) as a result of a mutation in a parental germ cell. Repetitive regions are patterns of DNA fragments that occur in multiple copies, and can represent up to two thirds of the human genome (25). Sanger sequencing is generally used as a first-line test when there is strong clinical suspicion of a mutation in a specific gene, for screening at-risk family members for a known mutation, or as a confirmatory test for variants identified by massively parallel sequencing. It is also used for sequencing of specific genes or regions that are not attainable with massively parallel sequencing approaches (*e.g.*, guanine-cytosine-rich, highly repetitive segments, *etc.*).

Comparative Genomic Hybridization/Chromosomal Microarray Analysis

This represents an important improvement over classical karyotyping. Due to its high resolution (50–100 kb), this technique is able to identify smaller rearrangements than standard karyotyping (limited to a resolution of 1–2 Mb) in both coding and noncoding regions. Thus, this is the test of choice when a structural variant is suspected, such as in individuals with developmental disorders of the kidney. In fact, recent studies have detected a high frequency of copy number variation in children and young adults with congenital urinary tract malformations (26), implicating over 45 different genomic disorders, with six loci accounting for 65% of cases with pathogenic copy number variations (1q21, 4p16.1-p16.3, 16p11.2) (27,28). Overall diagnostic yield in children with congenital anomalies of kidney and urinary tract ranges between 10% and 17% (29,30) in nonsyndromic and syndromic cases (31,32). The drawbacks of chromosomal microarray include its limitations in detecting SNVs and INDELs, balanced chromosomal rearrangements, and deletion/duplications of <50,000 bp.

Massively Parallel Sequencing

This technique, which is sometimes called next-generation sequencing, was developed in the last decade. It allows for simultaneous sequencing of millions of DNA fragments, at a relatively low cost. It is best for identifying small variants, although large copy number variants can also be detected. Compared with Sanger sequencing, this is characterized by lower precision (error rate approximately 0.5%–2%) and shorter read length (generally 100–250 bp) (33). Possible applications differ in the size of genome sequenced, from specific regions to the entire genome. These include the approaches discussed below.

Targeted Gene Panels. If there is a limited number of genes that can cause a specific phenotype, a gene panel can be a cost-effective first-line test. This technique targets coding regions in a selected set of genes and is the best modality for conditions with a clearly distinct phenotype. Gene lists for different kidney phenotypes have been published (3,34,35). This approach leads to a diagnostic yield that is dependent on disease and patient selection procedures. Recent reports indicate diagnostic yields of 13% in children with congenital anomalies of kidney and urinary tract patients (3), approximately 20% in adult and children nephrotic syndrome cases (34), and up to 62% in glomerular disease and 78% in cystic diseases, in both adult and pediatric cases (35). Due to its targeted approach, gene panels may also simplify the interpretation of the results. In addition, panels are usually optimized for a specific set of genes, minimizing potential false negative findings due to incomplete sequence coverage. However, genes included in a particular panel vary between commercial services, and sometimes these panels fail to account for overlapping phenotypes. For example, nephrotic syndrome panels may not contain some congenital anomalies of kidney and urinary tract genes, such as *PAX2*, whose mutation may also manifest with proteinuria (36). The genes included in a specific panel must also be periodically updated as new genes are discovered.

Exome Sequencing. This sequencing approach examines nearly all coding regions of the genome (37). Exome sequencing proved its efficacy as a first-line test in both adult and pediatric patients, with a diagnostic yield extremely dependent on patient phenotype, age, and selection criteria. Recent studies demonstrate a diagnostic yield ranging from 11% in adult patients with FSGS (38), to 24% in primarily adult patients with congenital or cystic kidney disease (2). The yield is superior in pediatric cohorts: 32% in FSGS and 61% in cystic kidney disease (39). In general, the diagnostic yield outperforms targeted panels across different phenotypes in both adults and children (2,39,40). Importantly, exome sequencing allows for future reanalysis as new genes are discovered and analytic algorithms improve, without needing to resample the patient (41). As drawbacks, the coverage per base is generally lower than with targeted panels and can result in suboptimal coverage of some relevant genes (*e.g.*, *PKD1*, *GREB1L*) (42). Copy number variation detection with exome sequencing is possible but not always optimal, and the large amount of data generated can make interpretation challenging and time consuming (43). Several genomic regions relevant to nephrology are not optimally covered by exome sequencing and constitute blind spots that need to be recognized by clinicians ordering the test, such as duplicated region of *PKD1* or the *MUC1* variable number tandem repeat (a tandem repeat is a short repetitive nucleotide sequence that is generally difficult to sequence with short read MPS technology) (44). Importantly, a very promising application of exome sequencing is in the evaluation of nephropathies of unknown etiology, where, depending on the populations studied, it can have a diagnostic yield ranging from 17% (2) to 38% (8) in adult patients.

Genome Sequencing. This technique has the best analytic performance when compared with the aforementioned approaches, owing to its uniform coverage of the genome. Genome sequencing has superior coverage at a minimum depth of 10 reads compared to exome sequencing, and limits PCR artifacts, guanine-cytosine bias, and a high variance in allele fraction (42,45). It is also more efficient for copy number variant detection, even outperforming microarrays approaches (46). Genome sequencing can characterize noncoding regions, enabling detection of splicing or regulatory variants with large phenotypic effect, although this is currently limited by our incomplete understanding of the function of most noncoding regions (47,48). Thus, genome sequencing represents the most promising future of genetic

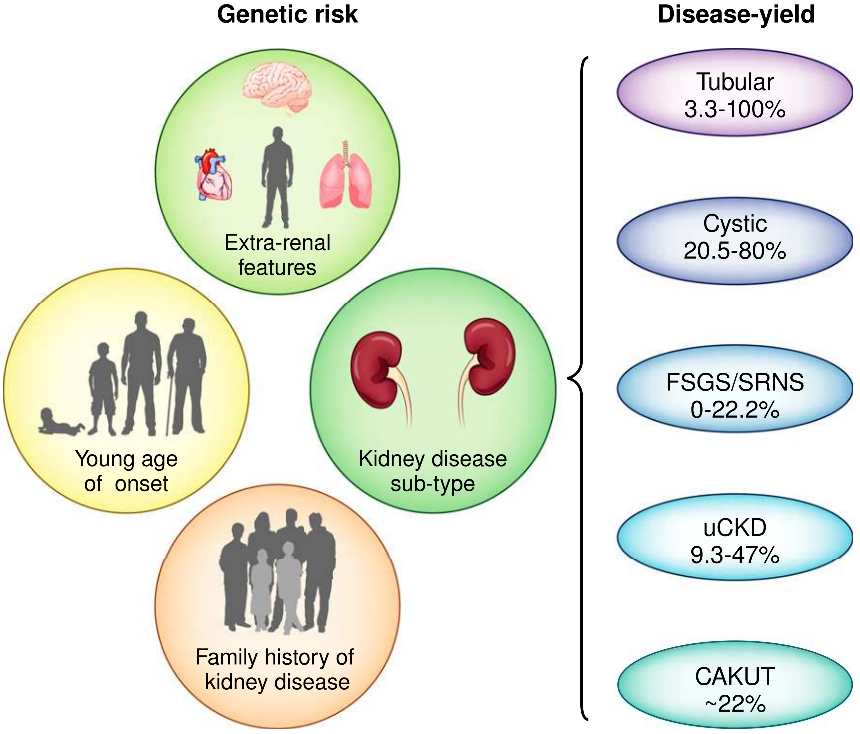


Figure 2. | **Clinical determinants of genetic risk and their influence on genetic test type.** The figure summarizes the clinical characteristics shaping the pretest risk of a genetic disease in a nephrology patient. Young age at onset, family history of kidney disease, and presence of extrarenal features are all predictive of genetic disease. Moreover, depending on the clinical diagnosis, the diagnostic yield of genetic testing varies. The yields of different modalities is shown in Figure 3. CAKUT, congenital anomalies of kidney and urinary tract; SRNS, steroid-resistant nephrotic syndrome; uCKD, CKD of unknown etiology.

testing because of its technical superiority, analytic performance, and mapping capacity, along with its progressively decreasing cost. Nonetheless, it still has blind spots shared with exome sequencing (*e.g.*, *PKD1* and *MUC1*). There is still a limited amount of research on the use of genome sequencing in the context of nephrology, but data support its superiority to traditional techniques in several other contexts (49,50). For example, genome sequencing had a significant improvement in diagnostic yield (+24%) when compared with targeted panels as a first-tier genetic test for pediatric disorders (51). Similar to exome sequencing, genome sequencing allows for future reanalysis and technologies application. Its potential drawbacks include the relatively high cost of the test (although this is progressively decreasing) and the possibility of detecting genetic findings not related to the primary indication for testing (*i.e.*, secondary findings).

Clinical Predictors of Diagnostic Yield

Several clinical factors influence the diagnostic potential of each test. To date, some studies support the use of exome sequencing as a first-line diagnostic approach for noncystic forms of genetic nephropathies. However, there is a paucity of data comparing the yield of targeted panel and exome/genome sequencing for different kidney phenotypes. In Figure 2, we summarize the clinical predictors of diagnostic yield and highlight optimal sequencing approaches on the basis of the broad kidney disease category. Then, in Figure 3, we compare the diagnostic yields achieved in genetic studies of diverse phenotypes using different sequencing techniques.

The main factors that contribute to variation in the diagnostic yield of the specific sequencing approach are discussed below.

Clinical Diagnosis

Cystic Kidney Disease. These disorders are predominantly attributable to mutations in *PKD1*/*PKD2* associated with adult-onset autosomal dominant polycystic kidney disease (the majority of cystic kidney diseases) and mutations in *PKHD1* in pediatric-onset autosomal recessive polycystic kidney disease. This clinical distinction is not absolute: 2% of patients with *PKD1/PKD2* mutations show a pediatric-onset severe phenotype clinically indistinguishable from autosomal recessive polycystic kidney disease (52), and other genes have been implicated as causal for cystic disease in both adults and pediatric patients (such as *GANAB* or *DNAJB11* in adults and *HNF1B* in both) (53,54). For cystic phenotypes, when there is a strong clinical suspicion of mutation in a specific gene, Sanger sequencing is the test of choice; otherwise, targeted panels are often an appropriate initial test because of the limited numbers of identified genes implicated to date. Furthermore, the assays can be optimized to cover the duplicated region of *PKD1* (45), which are otherwise not well covered by most other sequencing approaches. As such, the reported

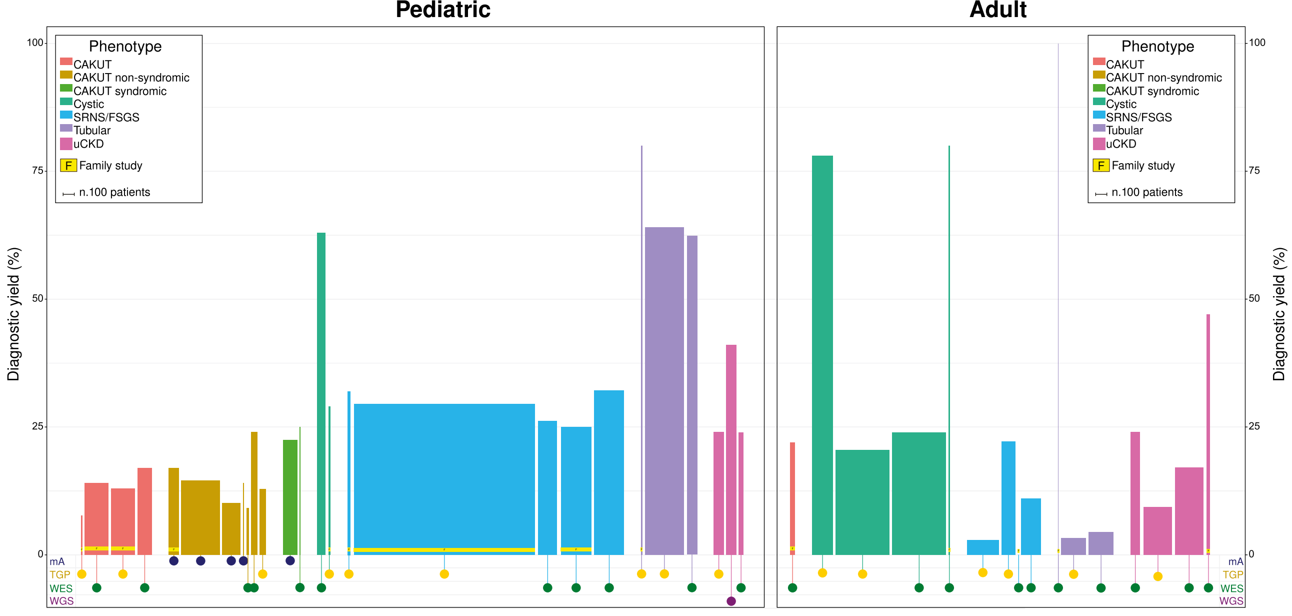


Figure 3. | **Diagnostic yield per phenotype and genetic test type.** The figure represents the diagnostic yield in different phenotype cohorts obtained through different genetic test type in pediatric and adult genetic studies on kidney disease. The *y* axis represents the percentage of diagnostic rate for the cohort (specified above each bar with the relative citation). The *x* axis represents the study under consideration; the width of the bar is dependent on the sample size for each study. The familial yellow flag indicates whether the study considered families (not individuals). The colored legend below the plot indicates the genetic test utilized in the study cohort. mA, chromosomal microarray techniques (especially in CAKUT); TGP, targeted gene panel; WES, whole exome sequencing; WGS, whole genome sequencing. The studies depicted here are presented in the Supplemental Material.

diagnostic yield of targeted panels for cystic kidney disease range in adult patients from 24% (2) to 88% (35), and 23%–80% with exome sequencing (2,8,55), based on the clinical characteristics of the cohort.

Congenital Anomalies of Kidney and Urinary Tract and Nephronophthisis. To date, these disorders are the most common cause of kidney disease in children, frequently involving congenital defects in other organ systems (56). With hundreds of genes implicated, they are highly heterogeneous in both phenotype and genotype. In addition, these phenotypes are often caused by structural genomic variants (57), detectable by chromosomal microarray (58), whose diagnostic yield is greatest in patients with kidney parenchymal malformations and those with extrarenal manifestations (27,31,59). The remaining cases are caused by SNVs in a single gene, identifiable with exome sequencing or targeted panels, which can provide a diagnostic yield of around 14% in recent reports (3) and is slightly superior in syndromic cases (25% versus 9% [32], 23% versus 15% [31,60]) in pediatric and young adult cohorts.

FSGS and Steroid-Resistant Nephrotic Syndrome. This is a genetically heterogeneous phenotype with over 30 known causal genes (61), affecting primarily children and young adults (62,63). The diagnostic yield can be as high as 32% (39) in recently analyzed cohorts with exome sequencing, and is higher in pediatric cases than in adult ones. Steroid-resistant nephrotic syndrome reached diagnostic yield of 25%–30% (39,64–66) in pediatric patients versus 12% in adults (67), whereas FSGS proved around 22% in pediatric patients (68) and 11% in adults (38). Importantly, a genetic diagnosis can have strong implications for patient management and should always be considered in early-onset disease, particularly in the presence of a clear family history. For example, most genetic forms of steroid-resistant nephrotic syndrome may not respond to immunosuppression therapy (69). Therefore, a positive finding may spare a patient from unwanted complications from this therapy (70). Similarly, a negative test may also guide management as it may influence clinical decision-making toward a longer and/or more aggressive immunosuppression course, and may provide prognostic insights too, as hereditary forms are associated with lower recurrence rate after transplantation (9). Finally, genetic testing can identify cases that may respond to therapy to glucocorticoids, such as cases owing to *PLCE1* (71), or cases that may benefit from other therapies that can be effectively treated with coenzyme Q10 (CoQ10) supplementation, such as steroid-resistant nephrotic syndromes owing to mutations in genes of the CoQ10 biosynthesis pathway (*ADCK4*, *COQ2*, *COQ6*, or *PDSS2*) (72,73). There are many genetic panels available as a first-line diagnostics, but exome sequencing can also be justifiable as a test of choice because of the rapid pace of discovery of new genes for nephrotic syndrome.

Tubulopathies (*e.g.*, Autosomal Dominant Tubulointerstitial Kidney Disease [ADTKD]). Recent studies show targeted panels to have the highest diagnostic yield in this disease category. In fact, diagnostic yield ranges from 62% (38,74) to 83% (75) in pediatric cohorts and 75% (75) to 100% (8) in adult cohorts. This is probably because of the highly characteristic clinical presentation and subsequent selection of cases. Importantly, recent findings suggest that trio analysis with exome sequencing, whereby both the

proband and their parents are sequenced for comparison, can significantly increase the targeted panels diagnostic performance because the availability of parental data enables segregation analysis and detection of *de novo* mutations (63). Several genes have been implicated in ADTKD to date, such as *UMOD, HNF1B, REN*, and *MUC1* in autosomal dominant tubulointerstitial kidney disease (76). Importantly, though commercial-targeted panels may include *MUC1*, the ADTKD is associated with variable-number tandem repeat mutations, which makes massively parallel sequencing approaches challenging because of the low coverage in these genomic regions (77).

Kidney Disease of Unknown Etiology. This is an important clinical category representing approximately 5%–15% of patients with kidney failure (1). If a thorough nephrology workup is nondiagnostic, genetic testing should be considered, particularly in the setting of early age of disease onset and a positive family history. The diagnostic yield in both adults and children ranges from 7% to 40% depending on patient characteristics and clinical history, with over two dozen different disorders described across developmental, glomerular, and tubulointerstitial categories (2,72,75). Given the heterogeneous disorders contributing to this disease category, exome sequencing had superior diagnostic performance as compared with targeted gene panels (2).

Family History and Extrarenal Manifestations

Studies indicate a three- to four-fold greater chance of a positive genetic finding in the setting of a positive family history of kidney disease (2,72). For example, a positive family history finding was much more likely in adults (53%) than in pediatric patients (12%) in a cohort of patients with monogenic nephrotic syndrome (34). It is worth noting that the clinical presentation can be different in different family members with the same mutation. For example, the same mutation in *PAX2*, responsible for papillorenal syndrome, may present as hearing or visual defect in some family members, but present as kidney diseases in other relatives (78). This variable expression motivates obtaining a thorough family history of disease in patients suspected of genetic disease.

In addition, multiorgan involvement is a compelling clinical finding suggestive of a genetic disorder. In a recent study of exome sequencing in adult nephrology patients, the presence of extrarenal features increased the diagnostic yield up to 69% (8). Although obtaining a detailed family history, or being able to decipher extrarenal features associated with a genetic diagnosis, may seem challenging for nongenetic experts, it is important to appreciate that these are key features suggestive of a genetic diagnosis and must be emphasized in medical education going forward. Moreover, collaborating with genetic counselors and/or clinical geneticists can ensure better collection of pertinent family and clinical data. Furthermore, the introduction of decision-support tools and electronic algorithms in medical health records, able to flag potential Mendelian disorders, will help clinicians in detecting genetic disorders and extrarenal features they might not have recognized (79).

Age of Onset

Mendelian disorders are more common in children (75), and age of onset has been shown to significantly influence the probability that the cause is an underlying genetic diagnosis for certain clinical subtypes. For example, genetic cystic kidney disease is almost three-fold more probable in pediatric cases than in adult ones (80). Hence, early onset of disease should increase suspicion for a genetic disease.

Genetic Data Analysis

Interpretation of genetic test results requires close collaboration between the clinician and the molecular pathologist, akin to the collaboration between the nephrologist and kidney pathologist in the evaluation of a biopsy. The interpretation of sequence data depends on the genomic tests used and requires gene- and variant-level analyses, phenotypic-level comparison, and determination of actionability. It also requires domain-specific expertise about the molecular and clinical features of disease. In an effort to standardize this process, the American College of Medical Genetics and Genomics (ACMG) has established guidelines for interpretation of sequence variants (81). Variants are classified as "pathogenic," "likely pathogenic," "variants of unknown significance (VUS)," "likely benign," or "benign." These guidelines recommend using terms such as "pathogenic" and "benign" for variants that are almost certainly disease causing or known to be benign, respectively; "likely pathogenic" and "likely benign" is used for variants with >90% certainty of either being disease causing or benign. VUS is used for variants for which there is insufficient evidence to classify in any of the other categories. As a general rule, clinicians should only act on variants classified as "pathogenic" and "likely pathogenic." These categories of variants are generally very rare in the population (typically with allele frequency <1:10,000), and have robust evidence of disease association on the basis of multiple prior reports, predicted effect on protein function, and potentially *in vitro* functional studies. Once a "pathogenic" or "likely pathogenic" variant has been identified, a second-level comparison is warranted to determine concordance of the patient's characteristics with the phenotype reported for the disease. Any discrepancy should be reviewed by both the molecular pathologist and clinician, and prior reports should be reviewed. Importantly, recent data suggest that many VUS will eventually be classified as benign (82–86) and can usually be resolved with subsequent clinical correlations, in discussions between the molecular diagnostician, clinician, and the patient.

Although a number of gene lists for nephropathy traits have been generated (2,3,21,34,35,87), nephrology lacks standardized, well-curated gene and variant lists. This will change as kidney disease is increasingly represented in initiatives such as ClinGen, which aims to define the clinical relevance of genes and variants for clinical and research use (88). The development of expertly-curated gene and variant lists is critical because of the many nuances in sequence interpretation. For example, although gene-disrupting mutations are often pathogenic, some disorders are almost exclusively caused by missense variants affecting specific domains of the protein

(*e.g.*, FSGS owing to mutations in the diaphanous inhibitory domains of *INF2* [89], or autosomal dominant tubulointerstitial kidney disease owing to mutations in exons 3 or 4 of *UMOD* [90], *etc*.). Therefore, loss-of-function (*e.g.*, nonsense, frameshift, splice, *etc*.) and missense variants outside of functional domains, even if exceedingly rare, are unlikely to be causal for these diseases (91). Conversely, a relatively common *NPHS2* variant (R229Q) can be disease causing if in *trans* with specific 3′ *NPHS2* mutations that exert a dominant-negative effect on the R229Q protein (92).

Currently, we still have an incomplete catalog of rare variants of the genome and limited understanding of their functional effect, hence many rare variants do not clearly fit into "pathogenic" or "benign" categories and are designated as VUS. A VUS classification frequently arises when a variant is novel and *in silico* algorithms predict it to be damaging, but there is insufficient data on its *in vivo* functional consequences and inheritance pattern. Classification as a VUS can also occur when a variant has been previously reported as disease causal, but there is contradictory data about its pathogenicity. This is the case of some population-specific polymorphisms or modestly rare variants that were erroneously thought to be extremely rare and reported as pathogenic (91). With the availability of larger and more diverse reference data sets, many previously reported pathogenic variants are now shown to have an allele frequency that exceeds the disease prevalence and can now be reclassified as benign (82,93). For example, a recent study showed that the top genes potentially contributing to most false positive results for kidney diagnostics were all identified before the availability of large publicly available databases (21). It is expected that the increasing number of genomes sequenced across populations of different ancestries will provide better estimates of variant allele frequencies, reducing the numbers of VUS and false positives. In addition, the availability of saturation mutagenesis of human genes (94) raises the possibility of obtaining functional information about every potential genetic variant in the future (in saturation mutagenesis, a single amino acid can be substituted to any of the other 19 possible substituents, which allows a comprehensive analysis of the function of amino acids at each position in the protein).

With the accelerating pace of new discoveries, an emerging challenge is the continuous review of data. Thus, a reanalysis of exomes may reveal a new diagnostic finding and revise a negative genetic test result. A recent report from the National Institutes of Health–sponsored Centers for Mendelian Genomics indicated that a new genetic discovery is made for every 28 exomes sequenced (95). This rapid pace of genetic discovery necessitates systematic updating of variant databases and gene panels, and provides clinicians and patients with the future prospect of achieving a diagnosis when the initial genetic findings are unrevealing. For example, a recent reevaluation of exome sequence data from two cohorts initially analyzed in 2012–2013 revealed that the diagnostic yield increased from 25% to 37% over a 5-year period, attributable mainly to newly discovered disease-causal genes (96). There are additional strategies for addressing unsolved cases, as shown by the Undiagnosed Disease Network of the National Institutes of Health (97): transcriptome analysis (also called RNA sequencing, which uses massive parallel sequencing approaches to reveal the sequence and quantity of RNA in a biologic sample) can reveal the possible effect of noncoding DNA sequence variants on splicing or allele-specific expression; or proteomic analysis can detect translation defects or point to metabolic defects. Although these approaches create opportunities for improving clinical care, they are resource intensive for the diagnostic laboratories and clinicians. Furthermore, sequence interpretation can be a burdensome process, particularly when each step is performed manually (98). Hence, semiautomated algorithms can rapidly characterize variant allele frequencies, identify previously reported pathogenic variants, incorporate results of *in silico* prediction tools, and expedite some of the more labor-intensive steps. There are also currently no guidelines and standards for reanalysis of genetic test results, and this challenge will need to be addressed in the next few years.

Return of Genetic Results and Clinical Implementation

The diagnosis of a Mendelian disorder, whether in the context of a diagnostic finding explicative of the patient's kidney disease or a medically actionable secondary finding, has the potential to meaningfully inform nephrology care (see Box 1; 99–101). The clinical value of actionable genetic findings underscores the importance of returning genetic findings to patients. Today, most experts agree that clinically valid, medically actionable genetic findings should be returned to adult patients who opted to receive them at the time of consent (102), including for individuals who underwent genetic sequencing through their participation in research (103). Importantly, however, genetic testing results can affect insurability and confidentiality, which many patients and providers may not fully realize. Although legislation exists through the 2009 Genetic Information Nondiscrimination Act to protect employers and health insurance carriers from discriminating against individuals on the basis of their genetic results (104,105), it is not fully comprehensive. Furthermore, clinicians should also be aware of the potential psychosocial implications of a genetic diagnosis for a patient, which may arise from multiple factors, including marginalization and stigmatization from members of their family and community. Therefore, it is essential that all patients thoroughly understand the implication of genetic testing and that genetic counseling services are available for them and their families.

On the clinical side, implementing genetic results into clinical care is an important and challenging task that can lead to effective targeted treatment; for example, patients with steroid-resistant nephrotic syndrome owing to mutations in genes of the CoQ10 biosynthesis pathway can be effectively treated with CoQ10 supplementation (72), and RNA interference therapies have been developed for primary hyperoxaluria type 1 caused by mutations in *AGXT* (105,106). In the same manner, secondary results affect patient health and need to be reported. The role of clinical nephrologists in this regard is generally limited to identifying and activating the right specialist for patient screening/follow-up. Moreover, ACMG guidelines

Case Study 1

A male patient with Focal Segmental Glomerulosclerosis on kidney biopsy is found to have a pathogenic large truncating variant in the *COL4A5* gene, diagnostic for Alport syndrome. In addition to establishing a molecular diagnosis for Alport syndrome, the type of variant may provide insights into the patient's clinical course, wherein Loss-of-Function variants, such as the one in this example, are associated with early onset of kidney failure, hearing loss and ocular abnormalities (116) and large deletions in the *COL4A5* gene are associated with increased risk for Anti-glomerular basement membrane disease post-transplantation (117). In addition, the genetic diagnosis may inform allograft donor selection among at-risk family members, as mothers of males with X-linked Alport syndrome are typically discouraged from serving as donors (118). Furthermore, conservative management of the patient's proteinuria with an agent that inhibits the Renin-Angiotensin-Aldosterone-system (RAAS) (e.g., ACE-Inhibitor or ARB) is recommended (3). Finally, a genetic diagnosis can also help identify disease-specific support groups, which can provide psychosocial support for affected individuals and their families, guide patients to clinical drug trials, registries, and other relevant resources.

Box 1. | Establishing a genetic diagnosis can support personalized nephrology care: A case study.

currently limit the list to 59 actionable genes in which mutations strongly affect patient prognosis and life expectancy (104), but larger sets are used in different laboratories (107). These challenging tasks require support to ensure the nephrologists' correct use of diagnostic genetic testing. This will also require collaborations with genetic counselors and clinical geneticists, development of clinical decision-support tools in electronic health records that address physician knowledge gaps in genomics, and the availability of remote consultation *via* telemedicine for providers with limited access to genetic services.

Genetic Testing and Clinical Trials

Genetics can affect the development and design of clinical trials in multiple ways. First, discoveries in human genetics can uncover new targets for therapy. In fact, recent analyses suggest that therapies targeting pathways discovered through human genetic studies have a greater chance of successful validation in clinical trials (108,109). A well known success story involves the identification of *PCSK9* mutations in rare forms of hyperlipidemia, which led to the development of treatment for common forms of this trait (110,111), and promising results are emerging in nephrology, such as chaperone-based therapies in Fabry disease (112). The incorporation of genetic analysis in clinical trials can also identify a subset of patients who may preferentially benefit from therapy, such as the discovery of clinical responsiveness to a tyrosine kinase inhibitor in patients with nonsmall cell lung cancer with specific somatic mutations in the *EGFR* gene (113). Similarly, *JAK1/JAK2* mutations in solid tumors can predict response to anti-PD1/PDL1 therapies (114). It is thus not surprising that different initiatives have been taken in this sense; for example, the Precision Medicine Initiative is an investment of $70 million to the National Cancer Institute to "scale up efforts to identify genomic drivers in cancer and apply that knowledge in the development of more effective approaches to cancer treatment" (115). In addition, several trials are now being developed to target specific molecular defects rather than the clinical category of disease, aiming to treat cancer according to underlying genomic alterations, regardless of clinical cancer type (116). Even if the trial is not targeting specific mutations, genetic testing may enhance the study power by identifying subsets of patients who would likely not respond to the investigation therapy. For example, exclusion of patients with mutations affecting kidney structure or function (such as *COL4A5* or *NPHS2*) may improve the power of studies targeting immunologic pathways for the treatment of steroid-resistant nephrotic syndrome and also minimize exposure to side effects in patients unlikely to benefit from such therapy. Moreover, identification of patients with genetic predisposition to outcomes such as cancer or cardiomyopathy may also help with adjudication of adverse events in clinical trials. Finally, pharmacogenomic analyses can help identify variants that affect drug absorption or metabolism, as exemplified by variants in *CYP2C9/VKORC1* for warfarin toxicity (117) or *TPMT* for azathioprine toxicity, enabling better assessment of dosage, safety, and side effects.

Conclusions

Achieving a precise diagnosis is a fundamental goal of medical practice. Although historically focused on pediatric populations, genetic testing has emerged as a powerful diagnostic tool in adult nephrology, with important implications for diagnosis and management. As genetic testing is increasingly incorporated into clinical practice, many challenges are yet to be addressed, such as optimizing the accuracy of variant interpretation, deploying genetic testing affordable on a large scale, and improving physician and patient education. More research is also needed to investigate the long-term effect of establishing a molecular diagnosis on health care utilization and outcomes, to facilitate third-party payer coverage for diagnostic testing. In this regard, there are emerging data on the cost benefit of genetic testing in fields such as diabetes or cancer (118–120), and similar research would benefit the nephrology field. With the rapid pace of discovery, we also anticipate that the nephrology field will increasingly deploy clinical trials stratified on specific molecular defects rather than the clinical category of disease, increasing the power and safety profile of studies.

Disclosures

A. Gharavi reports receiving grants from Renal Research Institute and participating on advisory boards for AstraZeneca Center for Genomics Research and Goldfinch Bio, outside the submitted work. All remaining authors have nothing to disclose.

Funding

The authors are supported by grants from the National Institutes of Health (DK080099, to A. Gharavi; and TL1TR001875, to J. Nestor), and the Columbia Precision Medicine Initiative (to A. Gharavi) from Columbia University.

Supplemental Material

This article contains the following supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.15141219/-/DCSupplemental.

Supplemental Material. References for Figure 3 studies.

References

1. Saran R, Robinson B, Abbott KC, Agodoa LYC, Bragg-Gresham J, Balkrishnan R, Bhave N, Dietrich X, Ding Z, Eggers PW, Gaipov A, Gillen D, Gipson D, Gu H, Guro P, Haggerty D, Han Y, He K, Herman W, Heung M, Hirth RA, HsiungJ-T, Hutton D, Inoue A, Jacobsen SJ, Jin Y, Kalantar-Zadeh K, Kapke A, KleineC-E, Kovesdy CP, Krueter W, Kurtz V, Li Y, Liu S, Marroquin MV, McCullough K, Molnar MZ, Modi Z, Montez-Rath M, Moradi H, Morgenstern H, Mukhopadhyay P, Nallamothu B, Nguyen DV, Norris KC, O'Hare AM, Obi Y, Park C, Pearson J, Pisoni R, Potukuchi PK, Repeck K, Rhee CM, Schaubel DE, Schrager J, Selewski DT, Shamraj R, Shaw SF, Shi JM, Shieu M, Sim JJ, Soohoo M, Steffick D, Streja E, Sumida K, Kurella Tamura M, Tilea A, Turf M, Wang D, Weng W, Woodside KJ, Wyncott A, Xiang J, Xin X, Yin M, You AS, Zhang X, Zhou H, Shahinian V: US renal data system 2018 annual data report: Epidemiology of kidney disease in the United States. *Am J Kidney Dis* 73[Suppl 1]: A7–A8, 2019
2. Groopman EE, Marasa M, Cameron-Christie S, Petrovski S, Aggarwal VS, Milo-Rasouly H, Li Y, Zhang J, Nestor J, Krithivasan P, Lam WY, Mitrotti A, Piva S, Kil BH, Chatterjee D, Reingold R, Bradbury D, DiVecchia M, Snyder H, Mu X, Mehl K, Balderes O, Fasel DA, Weng C, Radhakrishnan J, Canetta P, Appel GB, Bomback AS, Ahn W, Uy NS, Alam S, Cohen DJ, Crew RJ, Dube GK, Rao MK, Kamalakaran S, Copeland B, Ren Z, Bridgers J, Malone CD, Mebane CM, Dagaonkar N, Fellström BC, Haefliger C, Mohan S, Sanna-Cherchi S, Kiryluk K, Fleckner J, March R, Platt A, Goldstein DB, Gharavi AG: Diagnostic utility of exome sequencing for kidney disease. *N Engl J Med* 380: 142–151, 2019
3. van der Ven AT, Connaughton DM, Ityel H, Mann N, Nakayama M, Chen J, Vivante A, HwangD-Y, Schulz J, Braun DA, Schmidt JM, Schapiro D, Schneider R, Warejko JK, Daga A, Majmundar AJ, Tan W, Jobst-Schwan T, Hermle T, Widmeier E, Ashraf S, Amar A, Hoogstraaten CA, Hugo H, Kitzler TM, Kause F, Kolvenbach CM, Dai R, Spaneas L, Amann K, Stein DR, Baum MA, Somers MJG, Rodig NM, Ferguson MA, Traum AZ, Daouk GH, Bogdanović R, Stajić N, Soliman NA, Kari JA, El Desoky S, Fathy HM, Milosevic D, Al-Saffar M, Awad HS, Eid LA, Selvin A, Senguttuvan P, Sanna-Cherchi S, Rehm HL, MacArthur DG, Lek M, Laricchia KM, Wilson MW, Mane SM, Lifton RP, Lee RS, Bauer SB, Lu W, Reutter HM, Tasic V, Shril S, Hildebrandt F: Whole-exome sequencing identifies causative mutations in families with congenital anomalies of the kidney and urinary tract. *J Am Soc Nephrol* 29: 2348–2361, 2018
4. Connaughton DM, Bukhari S, Conlon P, Cassidy E, O'Toole M, Mohamad M, Flanagan J, Butler T, O'Leary A, Wong L, O'Regan J, Moran S, O'Kelly P, Logan V, Griffin B, Griffin M, Lavin P, Little MA, Conlon P: The Irish kidney gene project--prevalence of family history in patients with kidney disease in Ireland. *Nephron* 130: 293–301, 2015
5. van Nimwegen KJM, van Soest RA, Veltman JA, Nelen MR, van der Wilt GJ, Vissers LELM, Grutters JPC: Is the $1000 genome as near as we think? A cost analysis of next-generation sequencing. *Clin Chem* 62: 1458–1464, 2016
6. Phillips KA, Deverka PA, Hooker GW, Douglas MP: Genetic test availability and spending: Where are we now? Where are we going? *Health Aff (Millwood)* 37: 710–716, 2018
7. Lata S, Marasa M, Li Y, Fasel DA, Groopman E, Jobanputra V, Rasouly H, Mitrotti A, Westland R, Verbitsky M, Nestor J, Slater LM, D'Agati V, Zaniew M, Materna-Kiryluk A, Lugani F, Caridi G, Rampoldi L, Mattoo A, Newton CA, Rao MK, Radhakrishnan J, Ahn W, Canetta PA, Bomback AS, Appel GB, Antignac C, Markowitz GS, Garcia CK, Kiryluk K, Sanna-Cherchi S, Gharavi AG: Whole-exome sequencing in adults with chronic kidney disease: A pilot study. *Ann Intern Med* 168: 100–109, 2018
8. Connaughton DM, Kennedy C, Shril S, Mann N, Murray SL, Williams PA, Conlon E, Nakayama M, van der Ven AT, Ityel H, Kause F, Kolvenbach CM, Dai R, Vivante A, Braun DA, Schneider R, Kitzler TM, Moloney B, Moran CP, Smyth JS, Kennedy A, Benson K, Stapleton C, Denton M, Magee C, O'Seaghdha CM, Plant WD, Griffin MD, Awan A, Sweeney C, Mane SM, Lifton RP, Griffin B, Leavey S, Casserly L, de Freitas DG, Holian J, Dorman A, Doyle B, Lavin PJ, Little MA, Conlon PJ, Hildebrandt F: Monogenic causes of chronic kidney disease in adults. *Kidney Int* 95: 914–928, 2019
9. Morello W, Puvinathan S, Puccio G, Ghiggeri GM, Dello Strologo L, Peruzzi L, Murer L, Cioni M, Guzzo I, Cocchi E, Benetti E, Testa S, Ghio L, Caridi G, Cardillo M, Torelli R, Montini G: Post-transplant recurrence of steroid resistant nephrotic syndrome in children: The Italian experience [published online ahead of print Oct 15, 2019]. *J Nephrol* doi:10.1007/s40620-019-00660-9
10. Cornec-Le Gall E, AudrézetM-P, ChenJ-M, Hourmant M, MorinM-P, Perrichot R, Charasse C, Whebe B, Renaudineau E, Jousset P, GuillodoM-P, Grall-Jezequel A, Saliou P, Férec C, Le Meur Y: Type of PKD1 mutation influences renal outcome in ADPKD. *J Am Soc Nephrol* 24: 1006–1013, 2013
11. Nestor JG, Marasa M, Milo-Rasouly H, Groopman EE, Husain SA, Mohan S, Fernandez H, Aggarwal VS, Ahram DF, Vena N, Bogyo K, Bomback AS, Radhakrishnan J, Appel GB, Ahn W, Cohen DJ, Canetta PA, Dube GK, Rao MK, Morris HK, Crew RJ, Sanna-Cherchi S, Kiryluk K, Gharavi AG: Pilot study of return of genetic results to patients in adult nephrology. *J Am Soc Nephrol* 15: 651–664, 2020
12. Temme J, Peters F, Lange K, Pirson Y, Heidet L, Torra R, GrunfeldJ-P, Weber M, Licht C, MüllerG-A, Gross O: Incidence of renal failure and nephroprotection by RAAS inhibition in heterozygous carriers of X-chromosomal and autosomal recessive Alport mutations. *Kidney Int* 81: 779–783, 2012
13. MüllerR-U, Haas CS, Sayer JA: Practical approaches to the management of autosomal dominant polycystic kidney disease patients in the era of tolvaptan. *Clin Kidney J* 11: 62–69, 2018
14. Adam J, Connor TMF, Wood K, Lewis D, Naik R, Gale DP, Sayer JA: Genetic testing can resolve diagnostic confusion in Alport syndrome. *Clin Kidney J* 7: 197–200, 2014
15. Liapis H, Gaut JP: The renal biopsy in the genomic era. *Pediatr Nephrol* 28: 1207–1219, 2013
16. Marian AJ: Sequencing your genome: What does it mean? *Methodist DeBakey Cardiovasc J* 10: 3–6, 2014
17. Homo sapiens Annotation Report. Available at: https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Homo_sapiens/109/. Accessed February 27, 2020
18. Amberger JS, Bocchini CA, Scott AF, Hamosh A: OMIM.org: Leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res* 47: D1038–D1043, 2019
19. Malekpour SA, Pezeshk H, Sadeghi M: MSeq-CNV: Accurate detection of copy number variation from sequencing of multiple samples. *Sci Rep* 8: 4009, 2018
20. Zarrei M, MacDonald JR, Merico D, Scherer SW: A copy number variation map of the human genome. *Nat Rev Genet* 16: 172–183, 2015
21. Rasouly HM, Groopman EE, Heyman-Kantor R, Fasel DA, Mitrotti A, Westland R, Bier L, Weng C, Ren Z, Copeland B, Krithivasan P, Chung WK, Sanna-Cherchi S, Goldstein DB, Gharavi AG: The burden of candidate pathogenic variants for kidney and genitourinary disorders emerging from exome sequencing. *Ann Intern Med* 170: 11–21, 2019

22. PetersenB-S, Fredrich B, Hoeppner MP, Ellinghaus D, Franke A: Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet* 18: 14, 2017
23. Wang XV, Blades N, Ding J, Sultana R, Parmigiani G: Estimation of sequencing error rates in short reads. *BMC Bioinformatics* 13: 185, 2012
24. Morozova O, Marra MA: Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92: 255–264, 2008
25. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD: Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 7: e1002384, 2011
26. Snoek R, van Setten J, Keating BJ, Israni AK, Jacobson PA, Oetting WS, Matas AJ, Mannon RB, Zhang Z, Zhang W, Hao K, Murphy B, Reindl-Schwaighofer R, Heinzl A, Oberbauer R, Viklicky O, Conlon PJ, Stapleton CP, Bakker SJL, Snieder H, Peters EDJ, van der Zwaag B, Knoers NVAM, de Borst MH, van Eerde AM: *NPHP1* (Nephrocystin-1) gene deletions cause adult-onset ESRD. *J Am Soc Nephrol* 29: 1772–1779, 2018
27. Verbitsky M, Westland R, Perez A, Kiryluk K, Liu Q, Krithivasan P, Mitrotti A, Fasel DA, Batourina E, Sampson MG, Bodria M, Werth M, Kao C, Martino J, Capone VP, Vivante A, Shril S, Kil BH, Marasà M, Zhang JY, NaY-J, Lim TY, Ahram D, Weng PL, Heinzen EL, Carrea A, Piaggio G, Gesualdo L, Manca V, Masnata G, Gigante M, Cusi D, Izzi C, Scolari F, van Wijk JAE, Saraga M, Santoro D, Conti G, Zamboli P, White H, Drozdz D, Zachwieja K, Miklaszewska M, Tkaczyk M, Tomczyk D, Krakowska A, Sikora P, Jarmoliński T, Borszewska-Kornacka MK, Pawluch R, Szczepanska M, Adamczyk P, Mizerska-Wasiak M, Krzemien G, Szmigielska A, Zaniew M, Dobson MG, Darlow JM, Puri P, Barton DE, Furth SL, Warady BA, Gucev Z, Lozanovski VJ, Tasic V, Pisani I, Allegri L, Rodas LM, Campistol JM, Jeanpierre C, Alam S, Casale P, Wong CS, Lin F, Miranda DM, Oliveira EA, Simões-E-Silva AC, Barasch JM, Levy B, Wu N, Hildebrandt F, Ghiggeri GM, Latos-Bielenska A, Materna-Kiryluk A, Zhang F, Hakonarson H, Papaioannou VE, Mendelsohn CL, Gharavi AG, Sanna-Cherchi S: The copy number variation landscape of congenital anomalies of the kidney and urinary tract [published correction appears in *Nat Genet* 51: 764, 2019]. *Nat Genet* 51: 117–127, 2019
28. Siomou E, Mitsioni AG, Giapros V, Bouba I, Noutsopoulos D, Georgiou I: Copy-number variation analysis in familial nonsyndromic congenital anomalies of the kidney and urinary tract: Evidence for the causative role of a transposable element-associated genomic rearrangement. *Mol Med Rep* 15: 3631–3636, 2017
29. Caruana G, Wong MN, Walker A, Heloury Y, Webb N, Johnstone L, James PA, Burgess T, Bertram JF: Copy-number variation associated with congenital anomalies of the kidney and urinary tract. *Pediatr Nephrol* 30: 487–495, 2015
30. Sanna-Cherchi S, Kiryluk K, Burgess KE, Bodria M, Sampson MG, Hadley D, Nees SN, Verbitsky M, Perry BJ, Sterken R, Lozanovski VJ, Materna-Kiryluk A, Barlassina C, Kini A, Corbani V, Carrea A, Somenzi D, Murtas C, Ristoska-Bojkovska N, Izzi C, Bianco B, Zaniew M, Flogelova H, Weng PL, Kacak N, Giberti S, Gigante M, Arapovic A, Drnasin K, Caridi G, Curioni S, Allegri F, Ammenti A, Ferretti S, Goj V, Bernardo L, Jobanputra V, Chung WK, Lifton RP, Sanders S, State M, Clark LN, Saraga M, Padmanabhan S, Dominiczak AF, Foroud T, Gesualdo L, Gucev Z, Allegri L, Latos-Bielenska A, Cusi D, Scolari F, Tasic V, Hakonarson H, Ghiggeri GM, Gharavi AG: Copy-number disorders are a common cause of congenital kidney malformations. *Am J Hum Genet* 91: 987–997, 2012
31. Weber S, Moriniere V, Knüppel T, Charbit M, Dusek J, Ghiggeri GM, Jankauskiené A, Mir S, Montini G, Peco-Antic A, Wühl E, Zurowska AM, Mehls O, Antignac C, Schaefer F, Salomon R: Prevalence of mutations in renal developmental genes in children with renal hypodysplasia: Results of the ESCAPE study. *J Am Soc Nephrol* 17: 2864–2870, 2006
32. LeiT-Y, Fu F, Li R, Wang D, WangR-Y, JingX-Y, Deng Q, LiZ-Z, LiuZ-Q, Yang X, LiD-Z, Liao C: Whole-exome sequencing for prenatal diagnosis of fetuses with congenital anomalies of the kidney and urinary tract. *Nephrol Dial Transplant* 32: 1665–1675, 2017
33. Ashley EA: Towards precision medicine. *Nat Rev Genet* 17: 507–522, 2016
34. Sen ES, Dean P, Yarram-Smith L, Bierzynska A, Woodward G, Buxton C, Dennis G, Welsh GI, Williams M, Saleem MA: Clinical genetic testing using a custom-designed steroid-resistant nephrotic syndrome gene panel: Analysis and recommendations. *J Med Genet* 54: 795–804, 2017
35. Bullich G, Domingo-Gallego A, Vargas I, Ruiz P, Lorente-Grandoso L, Furlano M, Fraga G, Madrid Á, Ariceta G, Borregán M, Piñero-Fernández JA, Rodríguez-Peña L, Ballesta-Martínez MJ, Llano-Rivas I, Meñica MA, Ballarín J, Torrents D, Torra R, Ars E: A kidney-disease gene panel allows a comprehensive genetic diagnosis of cystic and glomerular inherited kidney diseases. *Kidney Int* 94: 363–371, 2018
36. Barua M, Stellacci E, Stella L, Weins A, Genovese G, Muto V, Caputo V, Toka HR, Charoonratana VT, Tartaglia M, Pollak MR: Mutations in PAX2 associate with adult-onset FSGS. *J Am Soc Nephrol* 25: 1942–1953, 2014
37. Patwardhan A, Harris J, Leng N, Bartha G, Church DM, Luo S, Haudenschild C, Pratt M, Zook J, Salit M, Tirch J, Morra M, Chervitz S, Li M, Clark M, Garcia S, Chandratillake G, Kirk S, Ashley E, Snyder M, Altman R, Bustamante C, Butte AJ, West J, Chen R: Achieving high-sensitivity for clinical applications using augmented exome sequencing. *Genome Med* 7: 71, 2015
38. Yao T, Udwan K, John R, Rana A, Haghighi A, Xu L, Hack S, Reich HN, Hladunewich MA, Cattran DC, Paterson AD, Pei Y, Barua M: Integration of genetic testing and pathology for the diagnosis of adults with FSGS. *Clin J Am Soc Nephrol* 14: 213–223, 2019
39. Rao J, Liu X, Mao J, Tang X, Shen Q, Li G, Sun L, Bi Y, Wang X, Qian Y, Wu B, Wang H, Zhou W, Ma D, Zheng B, Shen Y, Chen Z, Luan J, Wang X, Wang M, Dang X, Wang Y, Wu Y, Hou L, Sun S, Li Q, Liu X, Bai H, Yang Y, Shao X, Li Y, Zheng S, Han M, Liu C, Cao G, Zhao L, Qiu S, Dong Y, Zhu Y, Wang F, Zhang D, Li Y, Zhao L, Yang C, Luo X, Chen L, Jiang X, Zhang A, Xu H; for "Internet Plus" Nephrology Alliance of National Center for Children's Care: Genetic spectrum of renal disease for 1001 Chinese children based on a multicenter registration system. *Clin Genet* 96: 402–410, 2019
40. Mann N, Braun DA, Amann K, Tan W, Shril S, Connaughton DM, Nakayama M, Schneider R, Kitzler TM, van der Ven AT, Chen J, Ityel H, Vivante A, Majmundar AJ, Daga A, Warejko JK, Lovric S, Ashraf S, Jobst-Schwan T, Widmeier E, Hugo H, Mane SM, Spaneas L, Somers MJG, Ferguson MA, Traum AZ, Stein DR, Baum MA, Daouk GH, Lifton RP, Manzi S, Vakili K, Kim HB, Rodig NM, Hildebrandt F: Whole-exome sequencing enables a precision medicine approach for kidney transplant recipients. *J Am Soc Nephrol* 30: 201–215, 2019
41. Baker SW, Murrell JR, Nesbitt AI, Pechter KB, Balciuniene J, Zhao X, Yu Z, Denenberg EH, DeChene ET, Wilkens AB, Bhoj EJ, Guan Q, Dulik MC, Conlin LK, Abou Tayoun AN, Luo M, Wu C, Cao K, Sarmady M, Bedoukian EC, Tarpinian J, Medne L, Skraban CM, Deardorff MA, Krantz ID, Krock BL, Santani AB: Automated clinical exome reanalysis reveals novel diagnoses. *J Mol Diagn* 21: 38–48, 2019
42. Kingsmore SF, Cakici JA, Clark MM, Gaughran M, Feddock M, Batalov S, Bainbridge MN, Carroll J, Caylor SA, Clarke C, Ding Y, Ellsworth K, Farnaes L, Hildreth A, Hobbs C, James K, Kint CI, Lenberg J, Nahas S, Prince L, Reyes I, Salz L, Sanford E, Schols P, Sweeney N, Tokita M, Veeraraghavan N, Watkins K, Wigby K, Wong T, Chowdhury S, Wright MS, Dimmock D; RCIGM Investigators: A randomized, controlled trial of the analytic and diagnostic performance of singleton and trio, rapid genome and exome sequencing in ill infants. *Am J Hum Genet* 105: 719–733, 2019
43. Park JY, Clark P, Londin E, Sponziello M, Kricka LJ, Fortina P: Clinical exome performance for reporting secondary genetic findings. *Clin Chem* 61: 213–220, 2015
44. Ali H, Al-Mulla F, Hussain N, Naim M, Asbeutah AM, AlSahow A, Abu-Farha M, Abubaker J, Al Madhoun A, Ahmad S, Harris PC: PKD1 duplicated regions limit clinical utility of whole exome sequencing for genetic diagnosis of autosomal dominant polycystic kidney disease. *Sci Rep* 9: 4141, 2019
45. McCombie WR, Knowles JA, Mardis ER, McPherson JD: *Next-generation Sequencing in Medicine*, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 2019

46. Zhou B, Ho SS, Zhang X, Pattni R, Haraksingh RR, Urban AE: Whole-genome sequencing analysis of CNV using low-coverage and paired-end strategies is efficient and outperforms array-based CNV analysis. *J Med Genet* 55: 735–743, 2018
47. King K, Flinter FA, Nihalani V, Green PM: Unusual deep intronic mutations in the COL4A5 gene cause X linked Alport syndrome. *Hum Genet* 111: 548–554, 2002
48. LoY-F, Nozu K, Iijima K, Morishita T, HuangC-C, YangS-S, SytwuH-K, FangY-W, TsengM-H, LinS-H: Recurrent deep intronic mutations in the SLC12A3 gene responsible for Gitelman's syndrome. *Clin J Am Soc Nephrol* 6: 630–639, 2011
49. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BBA, Kleefstra T, Brunner HG, Vissers LELM, Veltman JA: Genome sequencing identifies major causes of severe intellectual disability. *Nature* 511: 344–347, 2014
50. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, Kingsmore SF: Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med* 3: 16, 2018
51. Lionel AC, Costain G, Monfared N, Walker S, Reuter MS, Hosseini SM, Thiruvahindrapuram B, Merico D, Jobling R, Nalpathamkalam T, Pellecchia G, Sung WWL, Wang Z, Bikangaga P, Boelman C, Carter MT, Cordeiro D, Cytrynbaum C, Dell SD, Dhir P, Dowling JJ, Heon E, Hewson S, Hiraki L, Inbar-Feigenberg M, Klatt R, Kronick J, Laxer RM, Licht C, MacDonald H, Mercimek-Andrews S, Mendoza-Londono R, Piscione T, Schneider R, Schulze A, Silverman E, Siriwardena K, Snead OC, Sondheimer N, Sutherland J, Vincent A, Wasserman JD, Weksberg R, Shuman C, Carew C, Szego MJ, Hayeems RZ, Basran R, Stavropoulos DJ, Ray PN, Bowdin S, Meyn MS, Cohn RD, Scherer SW, Marshall CR: Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med* 20: 435–443, 2018
52. Bergmann C: ARPKD and early manifestations of ADPKD: The original polycystic kidney disease and phenocopies. *Pediatr Nephrol* 30: 15–30, 2015
53. Porath B, Gainullin VG, Cornec-Le Gall E, Dillinger EK, Heyer CM, Hopp K, Edwards ME, Madsen CD, Mauritz SR, Banks CJ, Baheti S, Reddy B, Herrero JI, Bañales JM, Hogan MC, Tasic V, Watnick TJ, Chapman AB, Vigneau C, Lavainne F, AudrézetM-P, Ferec C, Le Meur Y, Torres VE, Harris PC; Genkyst Study Group, HALT Progression of Polycystic Kidney Disease Group; Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease: Mutations in GANAB, encoding the glucosidase IIα subunit, cause autosomal-dominant polycystic kidney and liver disease. *Am J Hum Genet* 98: 1193–1207, 2016
54. Verhave JC, Bech AP, Wetzels JFM, Nijenhuis T: Hepatocyte nuclear factor 1β-associated kidney disease: More than renal cysts and diabetes. *J Am Soc Nephrol* 27: 345–353, 2016
55. Braun DA, Schueler M, Halbritter J, Gee HY, Porath JD, Lawson JA, Airik R, Shril S, Allen SJ, Stein D, Al Kindy A, Beck BB, Cengiz N, Moorani KN, Ozaltin F, Hashmi S, Sayer JA, Bockenhauer D, Soliman NA, Otto EA, Lifton RP, Hildebrandt F: Whole exome sequencing identifies causative mutations in the majority of consanguineous or familial cases with childhood-onset increased renal echogenicity. *Kidney Int* 89: 468–475, 2016
56. Verbitsky M, Sanna-Cherchi S, Fasel DA, Levy B, Kiryluk K, Wuttke M, Abraham AG, Kaskel F, Köttgen A, Warady BA, Furth SL, Wong CS, Gharavi AG: Genomic imbalances in pediatric patients with chronic kidney disease. *J Clin Invest* 125: 2171–2178, 2015
57. Ma R, Deng L, Xia Y, Wei X, Cao Y, Guo R, Zhang R, Guo J, Liang D, Wu L: A clear bias in parental origin of de novo pathogenic CNVs related to intellectual disability, developmental delay and multiple congenital anomalies. *Sci Rep* 7: 44446, 2017
58. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH: Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86: 749–764, 2010
59. Li S, Han X, Wang Y, Chen S, Niu J, Qian Z, Li P, Jin L, Xu C: Chromosomal microarray analysis in fetuses with congenital anomalies of the kidney and urinary tract: A prospective cohort study and meta-analysis. *Prenat Diagn* 39: 165–174, 2019
60. Rasmussen M, Sunde L, Nielsen ML, Ramsing M, Petersen A, Hjortshøj TD, Olsen TE, Tabor A, Hertz JM, Johnsen I, Sperling L, Petersen OB, Jensen UB, Møller FG, Petersen MB, Lildballe DL: Targeted gene sequencing and whole-exome sequencing in autopsied fetuses with prenatally diagnosed kidney anomalies. *Clin Genet* 93: 860–869, 2018
61. Lovric S, Ashraf S, Tan W, Hildebrandt F: Genetic testing in steroid-resistant nephrotic syndrome: When and how? *Nephrol Dial Transplant* 31: 1802–1813, 2016
62. Rosenberg AZ, Kopp JB: Focal segmental glomerulosclerosis. *Clin J Am Soc Nephrol* 12: 502–517, 2017
63. Hogg R, Middleton J, Vehaskari VM: Focal segmental glomerulosclerosis--epidemiology aspects in children and adults. *Pediatr Nephrol* 22: 183–186, 2007
64. Warejko JK, Tan W, Daga A, Schapiro D, Lawson JA, Shril S, Lovric S, Ashraf S, Rao J, Hermle T, Jobst-Schwan T, Widmeier E, Majmundar AJ, Schneider R, Gee HY, Schmidt JM, Vivante A, van der Ven AT, Ityel H, Chen J, Sadowski CE, Kohl S, Pabst WL, Nakayama M, Somers MJG, Rodig NM, Daouk G, Baum M, Stein DR, Ferguson MA, Traum AZ, Soliman NA, Kari JA, El Desoky S, Fathy H, Zenker M, Bakkaloglu SA, Müller D, Noyan A, Ozaltin F, Cadnapaphornchai MA, Hashmi S, Hopcian J, Kopp JB, Benador N, Bockenhauer D, Bogdanovic R, Stajić N, Chernin G, Ettenger R, Fehrenbach H, Kemper M, Munarriz RL, Podracka L, Büscher R, Serdaroglu E, Tasic V, Mane S, Lifton RP, Braun DA, Hildebrandt F: Whole exome sequencing of patients with steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol* 13: 53–62, 2018
65. Sadowski CE, Lovric S, Ashraf S, Pabst WL, Gee HY, Kohl S, Engelmann S, Vega-Warner V, Fang H, Halbritter J, Somers MJ, Tan W, Shril S, Fessi I, Lifton RP, Bockenhauer D, El-Desoky S, Kari JA, Zenker M, Kemper MJ, Mueller D, Fathy HM, Soliman NA, Hildebrandt F; SRNS Study Group: A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. *J Am Soc Nephrol* 26: 1279–1289, 2015
66. Bierzynska A, McCarthy HJ, Soderquest K, Sen ES, Colby E, Ding WY, Nabhan MM, Kerecuk L, Hegde S, Hughes D, Marks S, Feather S, Jones C, Webb NJA, Ognjanovic M, Christian M, Gilbert RD, Sinha MD, Lord GM, Simpson M, Koziell AB, Welsh GI, Saleem MA: Genomic and clinical profiling of a national nephrotic syndrome cohort advocates a precision medicine approach to disease management. *Kidney Int* 91: 937–947, 2017
67. Gribouval O, Boyer O, Hummel A, Dantal J, Martinez F, Sberro-Soussan R, Etienne I, Chauveau D, Delahousse M, Lionet A, Allard J, Pouteil Noble C, TêteM-J, Heidet L, Antignac C, Servais A: Identification of genetic causes for sporadic steroid-resistant nephrotic syndrome in adults. *Kidney Int* 94: 1013–1022, 2018
68. Trautmann A, Lipska-Ziętkiewicz BS, Schaefer F: Exploring the clinical and genetic spectrum of steroid resistant nephrotic syndrome: The PodoNet registry. *Front Pediatr* 6: 200, 2018
69. Mason AE, Sen ES, Bierzynska A, Colby E, Afzal M, Dorval G, Koziell AB, Williams M, Boyer O, Welsh GI, Saleem MA: UK RaDaR/NephroS Study: Response to first course of intensified immunosuppression in genetically stratified steroid resistant nephrotic syndrome [published online ahead of print April 21, 2020]. *Clin J Am Soc Nephrol* doi: 10.2215/CJN.13371019
70. Sampson MG, Gillies CE, Robertson CC, Crawford B, Vega-Warner V, Otto EA, Kretzler M, Kang HM: Using population genetics to interrogate the monogenic nephrotic syndrome diagnosis in a case cohort. *J Am Soc Nephrol* 27: 1970–1983, 2016
71. Hinkes B, Wiggins RC, Gbadegesin R, Vlangos CN, Seelow D, Nürnberg G, Garg P, Verma R, Chaib H, Hoskins BE, Ashraf S, Becker C, Hennies HC, Goyal M, Wharram BL, Schachter AD, Mudumana S, Drummond I, Kerjaschki D, Waldherr R, Dietrich A, Ozaltin F, Bakkaloglu A, Cleper R, Basel-Vanagaite L, Pohl M, Griebel M, Tsygin AN, Soylu A, Müller D, Sorli CS, Bunney TD, Katan M, Liu J, Attanasio M, O'toole JF, Hasselbacher K, Mucha

B, Otto EA, Airik R, Kispert A, Kelley GG, Smrcka AV, Gudermann T, Holzman LB, Nürnberg P, Hildebrandt F: Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. *Nat Genet* 38: 1397–1405, 2006

72. Ashraf S, Gee HY, Woerner S, Xie LX, Vega-Warner V, Lovric S, Fang H, Song X, Cattran DC, Avila-Casado C, Paterson AD, Nitschké P, Bole-Feysot C, Cochat P, Esteve-Rudd J, Haberberger B, Allen SJ, Zhou W, Airik R, Otto EA, Barua M, Al-Hamed MH, Kari JA, Evans J, Bierzynska A, Saleem MA, Böckenhauer D, Kleta R, El Desoky S, Hacihamdioglu DO, Gok F, Washburn J, Wiggins RC, Choi M, Lifton RP, Levy S, Han Z, Salviati L, Prokisch H, Williams DS, Pollak M, Clarke CF, Pei Y, Antignac C, Hildebrandt F: ADCK4 mutations promote steroid-resistant nephrotic syndrome through CoQ10 biosynthesis disruption. *J Clin Invest* 123: 5179–5189, 2013
73. Diomedi-Camassei F, Di Giandomenico S, Santorelli FM, Caridi G, Piemonte F, Montini G, Ghiggeri GM, Murer L, Barisoni L, Pastore A, Muda AO, Valente ML, Bertini E, Emma F: COQ2 nephropathy: A newly described inherited mitochondriopathy with primary renal involvement. *J Am Soc Nephrol* 18: 2773–2780, 2007
74. Ashton EJ, Legrand A, Benoit V, Roncelin I, Venisse A, ZennaroM-C, Jeunemaitre X, Iancu D, Van't Hoff WG, Walsh SB, Godefroid N, Rotthier A, Del Favero J, Devuyst O, Schaefer F, Jenkins LA, Kleta R, Dahan K, Vargas-Poussou R, Bockenhauer D: Simultaneous sequencing of 37 genes identified causative mutations in the majority of children with renal tubulopathies. *Kidney Int* 93: 961–967, 2018
75. Mallett AJ, McCarthy HJ, Ho G, Holman K, Farnsworth E, Patel C, Fletcher JT, Mallawaarachchi A, Quinlan C, Bennetts B, Alexander SI: Massively parallel sequencing and targeted exomes in familial kidney disease can diagnose underlying genetic disorders. *Kidney Int* 92: 1493–1506, 2017
76. EckardtK-U, Alper SL, Antignac C, Bleyer AJ, Chauveau D, Dahan K, Deltas C, Hosking A, Kmoch S, Rampoldi L, Wiesener M, Wolf MT, Devuyst O; Kidney Disease: Improving Global Outcomes: Autosomal dominant tubulointerstitial kidney disease: Diagnosis, classification, and management--A KDIGO consensus report. *Kidney Int* 88: 676–683, 2015
77. Kirby A, Gnirke A, Jaffe DB, Barešová V, Pochet N, Blumenstiel B, Ye C, Aird D, Stevens C, Robinson JT, Cabili MN, Gat-Viks I, Kelliher E, Daza R, DeFelice M, Hůlková H, Sovová J, Vylet'al P, Antignac C, Guttman M, Handsaker RE, Perrin D, Steelman S, Sigurdsson S, Scheinman SJ, Sougnez C, Cibulskis K, Parkin M, Green T, Rossin E, Zody MC, Xavier RJ, Pollak MR, Alper SL, Lindblad-Toh K, Gabriel S, Hart PS, Regev A, Nusbaum C, Kmoch S, Bleyer AJ, Lander ES, Daly MJ: Mutations causing medullary cystic kidney disease type 1 lie in a large VNTR in MUC1 missed by massively parallel sequencing. *Nat Genet* 45: 299–303, 2013
78. Adam J, Browning AC, Vaideanu D, Heidet L, Goodship JA, Sayer JA: A wide spectrum of phenotypes in a family with renal coloboma syndrome caused by a PAX2 mutation. *Clin Kidney J* 6: 410–413, 2013
79. Bastarache L, Hughey JJ, Hebbring S, Marlo J, Zhao W, Ho WT, Van Driest SL, McGregor TL, Mosley JD, Wells QS, Temple M, Ramirez AH, Carroll R, Osterman T, Edwards T, Ruderfer D, Velez Edwards DR, Hamid R, Cogan J, Glazer A, WeiW-Q, Feng Q, Brilliant M, Zhao ZJ, Cox NJ, Roden DM, Denny JC: Phenotype risk scores identify patients with unrecognized Mendelian disease patterns. *Science* 359: 1233–1239, 2018
80. Connaughton DM, Hildebrandt F: Personalized medicine in chronic kidney disease by detection of monogenic mutations. *Nephrol Dial Transplant* 35: 390–397, 2020
81. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee: Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the association for molecular pathology. *Genet Med* 17: 405–424, 2015
82. Shah N, HouY-CC, YuH-C, Sainger R, Caskey CT, Venter JC, Telenti A: Identification of misclassified ClinVar variants via disease population prevalence. *Am J Hum Genet* 102: 609–619, 2018
83. Mighton C, Charames GS, Wang M, ZakoorK-R, Wong A, Shickh S, Watkins N, Lebo MS, Bombard Y, Lerner-Ellis J: Variant classification changes over time in BRCA1 and BRCA2. *Genet Med* 21: 2248–2254, 2019
84. SoRelle JA, Thodeson DM, Arnold S, Gotway G, Park JY: Clinical utility of reinterpreting previously reported genomic epilepsy test results for pediatric patients. *JAMA Pediatr* 173: e182302, 2019
85. Macklin S, Durand N, Atwal P, Hines S: Observed frequency and challenges of variant reclassification in a hereditary cancer clinic. *Genet Med* 20: 346–350, 2018
86. Deignan JL, Chung WK, Kearney HM, Monaghan KG, Rehder CW, Chao EC; ACMG Laboratory Quality Assurance Committee: Points to consider in the reevaluation and reanalysis of genomic test results: A statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 21: 1267–1270, 2019
87. Parsa A, Fuchsberger C, Köttgen A, O'Seaghdha CM, Pattaro C, de Andrade M, Chasman DI, Teumer A, Endlich K, Olden M, ChenM-H, Tin A, Kim YJ, Taliun D, Li M, Feitosa M, Gorski M, Yang Q, Hundertmark C, Foster MC, Glazer N, Isaacs A, Rao M, Smith AV, O'Connell JR, Struchalin M, Tanaka T, Li G, HwangS-J, Atkinson EJ, Lohman K, Cornelis MC, Johansson A, Tönjes A, Dehghan A, Couraki V, Holliday EG, Sorice R, Kutalik Z, Lehtimäki T, Esko T, Deshmukh H, Ulivi S, Chu AY, Murgia F, Trompet S, Imboden M, Kollerits B, Pistis G, Harris TB, Launer LJ, Aspelund T, Eiriksdottir G, Mitchell BD, Boerwinkle E, Schmidt H, Hofer E, Hu F, Demirkan A, Oostra BA, Turner ST, Ding J, Andrews JS, Freedman BI, Giulianini F, Koenig W, Illig T, Döring A, WichmannH-E, Zgaga L, Zemunik T, Boban M, Minelli C, Wheeler HE, Igl W, Zaboli G, Wild SH, Wright AF, Campbell H, Ellinghaus D, Nöthlings U, Jacobs G, Biffar R, Ernst F, Homuth G, Kroemer HK, Nauck M, Stracke S, Völker U, Völzke H, Kovacs P, Stumvoll M, Mägi R, Hofman A, Uitterlinden AG, Rivadeneira F, Aulchenko YS, Polasek O, Hastie N, Vitart V, Helmer C, Wang JJ, Stengel B, Ruggiero D, Bergmann S, Kähönen M, Viikari J, Nikopensius T, Province M, Colhoun H, Doney A, Robino A, Krämer BK, Portas L, Ford I, Buckley BM, Adam M, ThunG-A, Paulweber B, Haun M, Sala C, Mitchell P, Ciullo M, Vollenweider P, Raitakari O, Metspalu A, Palmer C, Gasparini P, Pirastu M, Jukema JW, Probst-Hensch NM, Kronenberg F, Toniolo D, Gudnason V, Shuldiner AR, Coresh J, Schmidt R, Ferrucci L, van Duijn CM, Borecki I, Kardia SLR, Liu Y, Curhan GC, Rudan I, Gyllensten U, Wilson JF, Franke A, Pramstaller PP, Rettig R, Prokopenko I, Witteman J, Hayward C, Ridker PM, Bochud M, Heid IM, Siscovick DS, Fox CS, Kao WL, Böger CA: Common variants in Mendelian kidney disease genes and their association with renal function. *J Am Soc Nephrol* 24: 2105–2117, 2013
88. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS; ClinGen: ClinGen--the clinical genome resource. *N Engl J Med* 372: 2235–2242, 2015
89. Brown EJ, Schlöndorff JS, Becker DJ, Tsukaguchi H, Tonna SJ, Uscinski AL, Higgs HN, Henderson JM, Pollak MR: Mutations in the formin gene INF2 cause focal segmental glomerulosclerosis [published correction appears in *Nat Genet* 42: 361, 2010]. *Nat Genet* 42: 72–76, 2010
90. Devuyst O, Olinger E, Rampoldi L: Uromodulin: From physiology to rare and complex kidney disorders. *Nat Rev Nephrol* 13: 525–544, 2017
91. Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, Margulies DM, Loscalzo J, Kohane IS: Genetic misdiagnoses and the potential for health disparities. *N Engl J Med* 375: 655–665, 2016
92. Tory K, Menyhárd DK, Woerner S, Nevo F, Gribouval O, Kerti A, Stráner P, Arrondel C, Huynh Cong E, Tulassay T, Mollet G, Perczel A, Antignac C: Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. *Nat Genet* 46: 299–304, 2014
93. Whiffin N, Roberts AM, Minikel E, Zappala Z, Walsh R, O'Donnell-Luria AH, Karczewski KJ, Harrison SM, Thomson KL,

Sage H, Ing AY, Barton PJR, Funke B, Cook SA, MacArthur DG, Ware JS: Using high-resolution variant frequencies empowers clinical genome interpretation and enables investigation of genetic architecture. *Am J Hum Genet* 104: 187–190, 2019
94. Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, Janizek JD, Huang X, Starita LM, Shendure J: Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 562: 217–222, 2018
95. Posey JE, O'Donnell-Luria AH, Chong JX, Harel T, Jhangiani SN, Coban Akdemir ZH, Buyske S, Pehlivan D, Carvalho CMB, Baxter S, Sobreira N, Liu P, Wu N, Rosenfeld JA, Kumar S, Avramopoulos D, White JJ, Doheny KF, Witmer PD, Boehm C, Sutton VR, Muzny DM, Boerwinkle E, Günel M, Nickerson DA, Mane S, MacArthur DG, Gibbs RA, Hamosh A, Lifton RP, Matise TC, Rehm HL, Gerstein M, Bamshad MJ, Valle D, Lupski JR; Centers for Mendelian Genomics: Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet Med* 21: 798–812, 2019
96. Liu P, Meng L, Normand EA, Xia F, Song X, Ghazi A, Rosenfeld J, Magoulas PL, Braxton A, Ward P, Dai H, Yuan B, Bi W, Xiao R, Wang X, Chiang T, Vetrini F, He W, Cheng H, Dong J, Gijavanekar C, Benke PJ, Bernstein JA, Eble T, Eroglu Y, Erwin D, Escobar L, Gibson JB, Gripp K, Kleppe S, Koenig MK, Lewis AM, Natowicz M, Mancias P, Minor L, Scaglia F, Schaaf CP, Streff H, Vernon H, Uhles CL, Zackai EH, Wu N, Sutton VR, Beaudet AL, Muzny D, Gibbs RA, Posey JE, Lalani S, Shaw C, Eng CM, Lupski JR, Yang Y: Reanalysis of clinical exome sequencing data. *N Engl J Med* 380: 2478–2480, 2019
97. Gahl WA, Mulvihill JJ, Toro C, Markello TC, Wise AL, Ramoni RB, Adams DR, Tifft CJ; UDN: The NIH undiagnosed diseases program and Network: Applications to modern medicine. *Mol Genet Metab* 117: 393–400, 2016
98. Jalali Sefid Dashti M, Gamieldien J: A practical guide to filtering and prioritizing genetic variants. *Biotechniques* 62: 18–30, 2017
99. Tan R, Colville D, Wang YY, Rigby L, Savige J: Alport retinopathy results from "severe" COL4A5 mutations and predicts early renal failure. *Clin J Am Soc Nephrol* 5: 34–38, 2010
100. Savige J, Ariani F, Mari F, Bruttini M, Renieri A, Gross O, Deltas C, Flinter F, Ding J, Gale DP, Nagel M, Yau M, Shagam L, Torra R, Ars E, Hoefele J, Garosi G, Storey H: Expert consensus guidelines for the genetic diagnosis of Alport syndrome. *Pediatr Nephrol* 34: 1175–1189, 2019
101. Savige J, Gregory M, Gross O, Kashtan C, Ding J, Flinter F: Expert guidelines for the management of Alport syndrome and thin basement membrane nephropathy. *J Am Soc Nephrol* 24: 364–375, 2013
102. Jarvik GP, Amendola LM, Berg JS, Brothers K, Clayton EW, Chung W, Evans BJ, Evans JP, Fullerton SM, Gallego CJ, Garrison NA, Gray SW, Holm IA, Kullo IJ, Lehmann LS, McCarty C, Prows CA, Rehm HL, Sharp RR, Salama J, Sanderson S, Van Driest SL, Williams MS, Wolf SM, Wolf WA, Burke W; eMERGE Act-ROR Committee and CERC Committee; CSER Act-ROR Working Group: Return of genomic results to research participants: The floor, the ceiling, and the choices in between. *Am J Hum Genet* 94: 818–826, 2014
103. Bombard Y, Brothers KB, Fitzgerald-Butt S, Garrison NA, Jamal L, James CA, Jarvik GP, McCormick JB, Nelson TN, Ormond KE, Rehm HL, Richer J, Souzeau E, Vassy JL, Wagner JK, Levy HP: The responsibility to recontact research participants after reinterpretation of genetic and genomic research results. *Am J Hum Genet* 104: 578–595, 2019
104. Zawati MH, Parry D, Thorogood A, Nguyen MT, Boycott KM, Rosenblatt D, Knoppers BM: Reporting results from whole-genome and whole-exome sequencing in clinical practice: A proposal for Canada? *J Med Genet* 51: 68–70, 2014
105. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, McKelvey KD, Ormond KE, Richards CS, Vlangos CN, Watson M, Martin CL, Miller DT: Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): A policy statement of the American College of Medical Genetics and Genomics [published correction appears in *Genet Med* 19: 484, 2017]. *Genet Med* 19: 249–255, 2017
106. Liebow A, Li X, Racie T, Hettinger J, Bettencourt BR, Najafian N, Haslett P, Fitzgerald K, Holmes RP, Erbe D, Querbes W, Knight J: An investigational RNAi therapeutic targeting glycolate oxidase reduces oxalate production in models of primary Hyperoxaluria. *J Am Soc Nephrol* 28: 494–503, 2017
107. O'Daniel JM, McLaughlin HM, Amendola LM, Bale SJ, Berg JS, Bick D, Bowling KM, Chao EC, Chung WK, Conlin LK, Cooper GM, Das S, Deignan JL, Dorschner MO, Evans JP, Ghazani AA, Goddard KA, Gornick M, Farwell Hagman KD, Hambuch T, Hegde M, Hindorff LA, Holm IA, Jarvik GP, Knight Johnson A, Mighion L, Morra M, Plon SE, Punj S, Richards CS, Santani A, Shirts BH, Spinner NB, Tang S, Weck KE, Wolf SM, Yang Y, Rehm HL: A survey of current practices for genomic sequencing test interpretation and reporting processes in US laboratories. *Genet Med* 19: 575–582, 2017
108. Morgan P, Brown DG, Lennard S, Anderton MJ, Barrett JC, Eriksson U, Fidock M, Hamrén B, Johnson A, March RE, Matcham J, Mettetal J, Nicholls DJ, Platz S, Rees S, Snowden MA, Pangalos MN: Impact of a five-dimensional framework on R&D productivity at AstraZeneca. *Nat Rev Drug Discov* 17: 167–181, 2018
109. Dugger SA, Platt A, Goldstein DB: Drug development in the era of precision medicine. *Nat Rev Drug Discov* 17: 183–196, 2018
110. Abifadel M, Varret M, RabèsJ-P, Allard D, Ouguerram K, Devillers M, Cruaud C, Benjannet S, Wickham L, Erlich D, Derré A, Villéger L, Farnier M, Beucler I, Bruckert E, Chambaz J, Chanu B, LecerfJ-M, Luc G, Moulin P, Weissenbach J, Prat A, Krempf M, Junien C, Seidah NG, Boileau C: Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 34: 154–156, 2003
111. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, Edelberg JM, Goodman SG, Hanotin C, Harrington RA, Jukema JW, Lecorps G, Mahaffey KW, Moryusef A, Pordy R, Quintero K, Roe MT, Sasiela WJ, TambyJ-F, Tricoci P, White HD, Zeiher AM; ODYSSEY OUTCOMES Committees and Investigators: Alirocumab and cardiovascular outcomes after acute coronary syndrome. *N Engl J Med* 379: 2097–2107, 2018
112. Germain DP, Nicholls K, Giugliani R, Bichet DG, Hughes DA, Barisoni LM, Colvin RB, Jennette JC, Skuban N, Castelli JP, Benjamin E, Barth JA, Viereck C: Efficacy of the pharmacologic chaperone migalastat in a subset of male patients with the classic phenotype of Fabry disease and migalastat-amenable variants: Data from the phase 3 randomized, multicenter, double-blind clinical trial and extension study. *Genet Med* 21: 1987–1997, 2019
113. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350: 2129–2139, 2004
114. Nowicki TS, Hu-Lieskovan S, Ribas A: Mechanisms of resistance to PD-1 and PD-L1 blockade. *Cancer J* 24: 47–53, 2018
115. Goossens N, Nakagawa S, Sun X, Hoshida Y: Cancer biomarker discovery and validation. *Transl Cancer Res* 4: 256–269, 2015
116. Abrams J, Conley B, Mooney M, Zwiebel J, Chen A, Welch JJ, Takebe N, Malik S, McShane L, Korn E, Williams M, Staudt L, Doroshow J: National Cancer Institute's precision medicine initiatives for the new National Clinical Trials Network. *Am Soc Clin Oncol Educ Book* 34: 71–76, 2014
117. Madian AG, Wheeler HE, Jones RB, Dolan ME: Relating human genetic variation to variation in drug responses. *Trends Genet* 28: 487–495, 2012
118. GoodSmith MS, Skandari MR, Huang ES, Naylor RN: The impact of biomarker screening and cascade genetic testing on the cost-effectiveness of MODY genetic testing. *Diabetes Care* 42: 2247–2255, 2019
119. Sun L, Brentnall A, Patel S, Buist DSM, Bowles EJA, Evans DGR, Eccles D, Hopper J, Li S, Southey M, Duffy S, Cuzick J, Dos Santos Silva I, Miners A, Sadique Z, Yang L, Legood R, Manchanda R: A cost-effectiveness analysis of multigene testing for all patients with breast cancer [published online ahead of print Oct 3, 2019]. *JAMA Oncol* doi:10.1001/jamaoncol.2019.3323
120. Petelin L, Hossack L, Shanahan M, Mitchell G, Liew D, James PA, Trainer AH: Cost-effectiveness of long-term clinical management of BRCA pathogenic variant carriers. *Genet Med* 22: 831–839, 2020

Published online ahead of print. Publication date available at www.cjasn.org.

Clinical Genetic Screening in Adult Patients with Kidney Disease

Enrico Cocchi,[1,2] *Jordan Gabriela Nestor*,[1] *and Ali G. Gharavi*[1,3]

Abstract

Expanded accessibility of genetic sequencing technologies, such as chromosomal microarray and massively parallel sequencing approaches, is changing the management of hereditary kidney diseases. Genetic causes account for a substantial proportion of pediatric kidney disease cases, and with increased utilization of diagnostic genetic testing in nephrology, they are now also detected at appreciable frequencies in adult populations. Establishing a molecular diagnosis can have many potential benefits for patient care, such as guiding treatment, familial testing, and providing deeper insights on the molecular pathogenesis of kidney diseases. Today, with wider clinical use of genetic testing as part of the diagnostic evaluation, nephrologists have the challenging task of selecting the most suitable genetic test for each patient, and then applying the results into the appropriate clinical contexts. This review is intended to familiarize nephrologists with the various technical, logistical, and ethical considerations accompanying the increasing utilization of genetic testing in nephrology care.

CJASN 15: 1497–1510, 2020. doi: https://doi.org/10.2215/CJN.15141219

[1]Division of Nephrology and Center for Precision Medicine and Genomics, Department of Medicine, Columbia University, New York, New York
[2]Department of Pediatrics, Universita' degli Studi di Torino, Torino, Italy
[3]Insititute of Genomic Medicine, Columbia University, New York, New York

Correspondence: Dr. Ali Gharavi, Division of Nephrology and Center for Precision Medicine and Genomics, Columbia University College of Physicians and Surgeons, 1150 St. Nicholas Avenue, Russ Berrie Pavilion 413, New York, NY 10032. Email: ag2239@cumc.columbia.edu

Introduction

Kidney disease is associated with significant morbidity and mortality and affects over 20 million patients in the United States (1). A family history of nephropathy is reported in approximately 30% of cases, revealing the strong role of genetics in kidney disease (2–4). Genetic testing is increasingly used in clinical nephrology due to expanded utilization and accessibility of genetic sequencing technologies (5,6). As a result, Mendelian forms of kidney disease are increasingly detected in adult and pediatric patients. In fact, although historically more clinically apparent in pediatric populations, it is now clear that genetic forms of kidney disease are also highly prevalent in adults, with some studies reporting a Mendelian cause of kidney disease in up to 37% of adult cases (7,8). Establishing a genetic diagnosis has significant implications for nephrology care because it may inform prognosis (9–11) and selection of therapy (12,13), spare patients from undergoing invasive diagnostic procedures such as a kidney biopsy (14,15), and guide family planning (2).

Many genetic testing modalities are currently available (*e.g.*, targeted sequencing, microarrays, gene panels, genome-wide approaches, *etc.*), and selection of the most appropriate diagnostic sequencing approach is made on the basis of various factors. These include diagnostic yield of the different sequencing modalities, the patients' clinical picture, their preferences for the types of results that may emerge with broader sequencing approaches, out-of-pocket costs, and third-party payer coverage. This review is intended to familiarize clinical nephrologists with concepts relating to clinical genetic testing.

Mendelian Nephropathies, Genetic Testing Modalities, and Diagnostic Yields

The human genome is divided into protein-coding (approximately 1%, known as the "Exome," composed of approximately 408,659 exons) and noncoding sequences (approximately 99%) (16,17). Overall, it harbors approximately 20,000 genes, of which approximately 4100 are currently associated with Mendelian disorders (18). Human genomes differ greatly between individuals, and variations in genetic sequence are summarized in Figure 1 and include:

- Single nucleotide variants (SNVs): substitution of a single base.
- Small insertion-deletion (INDEL): insertion or deletion of approximately 2–1000 bases.
- Copy number variation: duplication/deletion that affects ≥1 kb in one or more loci. Copy number variants encompass 5%–10% of the human genome (19,20). The differentiation between small INDELs and copy number variants is specifically on the basis of the length of the affected DNA, reflecting a higher number of genes possibly involved in the latter.
- Chromosomal imbalance and rearrangements: deletions and duplications of entire chromosomes or segments of chromosomes. Inversions and translocations can also occur as a result of genome breakage followed by a rejoining of the broken ends in a different order than the original one.

These variations can potentially lead to a Mendelian disease. A recent survey identified 625 Mendelian disorders associated with kidney and urological traits (21), whereas the number of gene–disease associations continues to grow with the expanded use of massively parallel sequencing, a technological advancement that

 Copyright © 2020 by the American Society of Nephrology

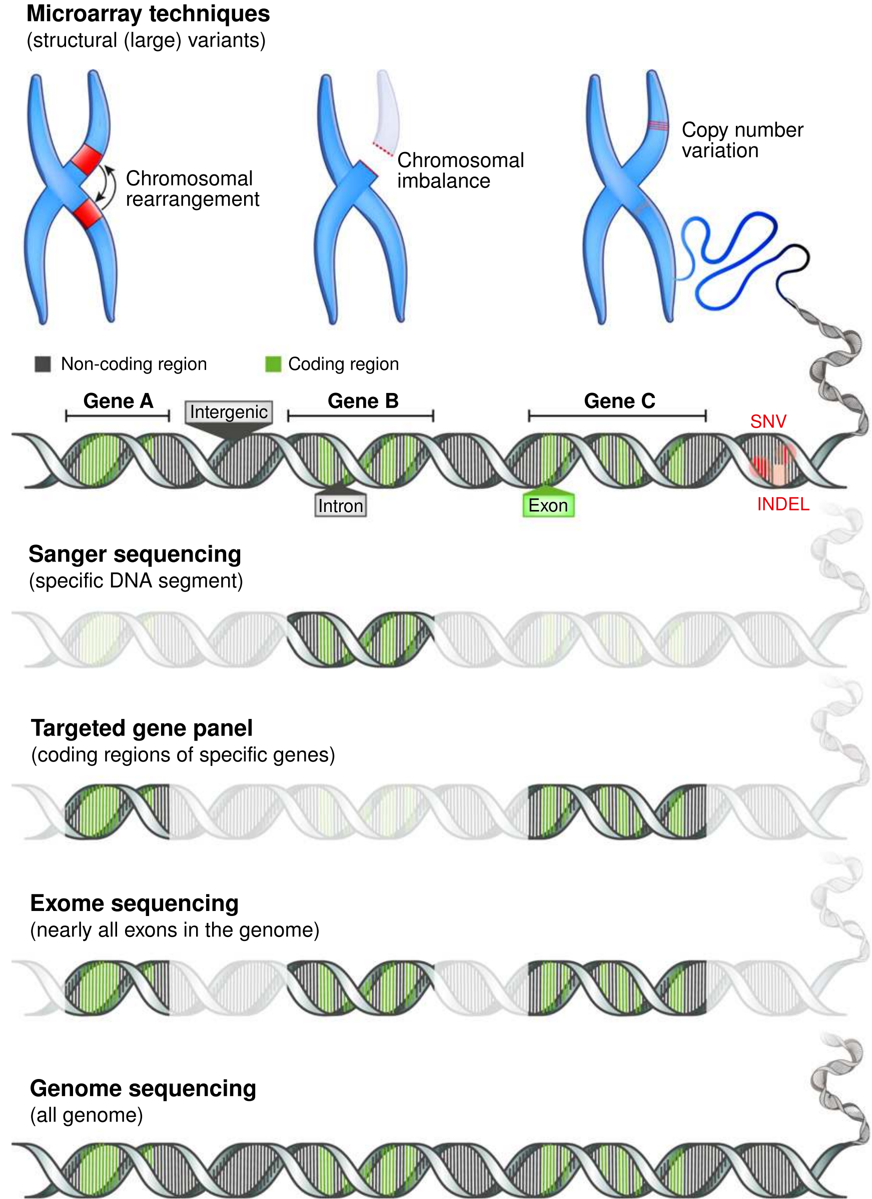


Figure 1. | **Genetic testing options.** Different classes of variants identifiable in human genome and how they are detected by the currently available genetic tests. The top part shows a "large-scale" section where an entire chromosome is represented (blue), as well as the relative large alterations (aka structural variants); chromosomal rearrangement (deletions, duplications, inversions, and translocations caused by a breakage in the DNA at two different locations, followed by a rejoining of the broken ends to produce a new chromosomal arrangement of genes); chromosomal imbalance (absence or duplication of a chromosomal portion) identifiable primarily through chromosomal array–based techniques; and copy number variation (a duplication or deletion that affects a large stretch of sequence, at least 1 kb). Small variants: single nucleotide variants (SNVs; substitution of a single base) and small insertion/deletions (INDELs; involving more than one base). We also see the coding (green), noncoding (dark gray), and regulatory (noncoding) portions of the genome that constitute each gene (gene A, gene B, and gene C) and how these are sequenced and analyzed through different techniques. Sanger sequencing is limited to a narrow portion of the genome, usually a single gene or small regions of a gene. Targeted gene panels: only coding portions of a specific set of genes are targeted. Exome sequencing captures nearly all of the coding sequences in an individual's genome. Genome sequencing covers nearly all regions of the genome (coding and noncoding regions) of an individual.

has increased the throughput of genomic sequencing (21). Importantly, diagnostic yield varies according to the categories of variants and choice of test. In Figure 1 and Table 1, we summarize technical and clinical aspects of different sequencing approaches, along with their respective benefits and drawbacks. Next, we discuss the major genomic diagnostic modalities.

Targeted Dideoxy Terminator (Sanger) Sequencing

This test aims to identify SNVs and INDELs in a specific gene or gene region (22). The execution is simple and results are highly reliable (error rate: 0.001%–1%) (23). Moreover, this method achieves a long reading length (approximately 800 bp) resulting in significant advantages for *de novo* variants confirmation and sequencing of

Table 1. Genetic testing options

Analysis Modality	Primary Scope	Advantages	Disadvantages	Uses
Sanger sequencing	• Identification of small variants (SNVs/INDELs) in a specific DNA region	• Simple technical execution • High analytical accuracy (error rate 0.001%–1%) • Fast and simple interpretation • No risk of secondary results • Long reading lengths (approximately 800 bp)	• Limited resolution (<1 kb) unsuitable for large structural variants • Time- and cost-inefficient for analysis of large DNA segments	• Confirmation of a specific suspected mutation in a gene • Confirmation of MPS-identified variants • Analysis of regions refractory to MPS analysis, such as repetitive regions
Chromosomal microarrays	• Identification of small chromosomal rearrangements/CNVs (≥200–400 kb)	• Higher resolution than standard karyotyping (50–100 kb) • Genome-wide analysis	• Cannot detect small mutations/CNVs • Limited ability to detect balanced chromosomal rearrangements, low-grade somatic mosaicism, and CNVs in pseudogenes and repetitive regions	• Patients with phenotype strongly suggestive of large rearrangements, such as multiple congenital anomalies and developmental diseases
Targeted MPS panels	• Identification of small variants (SNVs/INDELs) within genes of interest for the clinical phenotype	• Analysis of all genes possibly related to specific phenotypes • Diagnostic yield up to 50% (depending on the patient phenotype and genes selection method) • Restricted number of genes that minimizes risk of secondary findings and reduces analysis time	• Error rate approximately 0.5%–2% (MPS) • Short reading length (generally 50–300 bp) (MPS) • Limited reanalysis utility • Reliability depends on challenging panel design	• Patients with phenotypes pointing to specific disorders and with low genetic heterogeneity
Exome sequencing	• Identification of small variants (SNVs/INDELs) within coding regions of the genome	• Analysis of all coding regions in the genome • Unbiased approach increases diagnostic sensitivity • Cover almost all sites related to Mendelian diseases (approximately 85%)	• Coverage per base is generally lower than with targeted panels • Challenging and time-consuming interpretation (high number of candidate variants) • Potential for detection of secondary findings • Limited coverage in repetitive regions • Limited reliability for INDELs	• Patients undiagnosed with more specific methodologies • Screening of patients with undefined phenotype • Patients with heterogeneous/unspecific phenotype
Genome sequencing	• Identification of small variants (SNVs/INDELs) within coding and noncoding regions of the genome	• Identification of deep splicing and intronic variants unidentifiable with other techniques • Better analytic performance than exome sequencing • Efficient CNV identification • Extremely useful for reanalysis	• Maximizes results interpretation difficulty and time • Maximizes potential detection of secondary findings	• Patients undiagnosed with other methodologies • Screening of patients with undefined phenotype • Patients with heterogeneous/unspecific phenotype

SNV, single nucleotide variant; INDEL, insertion/deletion; MPS, massively parallel sequencing; CNV, copy number variation.

repetitive regions (24). *De novo* variants are those that are absent in parents and found only in the proband (*i.e.*, the individual undergoing genetic sequencing) as a result of a mutation in a parental germ cell. Repetitive regions are patterns of DNA fragments that occur in multiple copies, and can represent up to two thirds of the human genome (25). Sanger sequencing is generally used as a first-line test when there is strong clinical suspicion of a mutation in a specific gene, for screening at-risk family members for a known mutation, or as a confirmatory test for variants identified by massively parallel sequencing. It is also used for sequencing of specific genes or regions that are not attainable with massively parallel sequencing approaches (*e.g.*, guanine-cytosine-rich, highly repetitive segments, *etc.*).

Comparative Genomic Hybridization/Chromosomal Microarray Analysis

This represents an important improvement over classical karyotyping. Due to its high resolution (50–100 kb), this technique is able to identify smaller rearrangements than standard karyotyping (limited to a resolution of 1–2 Mb) in both coding and noncoding regions. Thus, this is the test of choice when a structural variant is suspected, such as in individuals with developmental disorders of the kidney. In fact, recent studies have detected a high frequency of copy number variation in children and young adults with congenital urinary tract malformations (26), implicating over 45 different genomic disorders, with six loci accounting for 65% of cases with pathogenic copy number variations (1q21, 4p16.1-p16.3, 16p11.2) (27,28). Overall diagnostic yield in children with congenital anomalies of kidney and urinary tract ranges between 10% and 17% (29,30) in nonsyndromic and syndromic cases (31,32). The drawbacks of chromosomal microarray include its limitations in detecting SNVs and INDELs, balanced chromosomal rearrangements, and deletion/duplications of <50,000 bp.

Massively Parallel Sequencing

This technique, which is sometimes called next-generation sequencing, was developed in the last decade. It allows for simultaneous sequencing of millions of DNA fragments, at a relatively low cost. It is best for identifying small variants, although large copy number variants can also be detected. Compared with Sanger sequencing, this is characterized by lower precision (error rate approximately 0.5%–2%) and shorter read length (generally 100–250 bp) (33). Possible applications differ in the size of genome sequenced, from specific regions to the entire genome. These include the approaches discussed below.

Targeted Gene Panels. If there is a limited number of genes that can cause a specific phenotype, a gene panel can be a cost-effective first-line test. This technique targets coding regions in a selected set of genes and is the best modality for conditions with a clearly distinct phenotype. Gene lists for different kidney phenotypes have been published (3,34,35). This approach leads to a diagnostic yield that is dependent on disease and patient selection procedures. Recent reports indicate diagnostic yields of 13% in children with congenital anomalies of kidney and urinary tract patients (3), approximately 20% in adult and children nephrotic syndrome cases (34), and up to 62% in glomerular disease and 78% in cystic diseases, in both adult and pediatric cases (35). Due to its targeted approach, gene panels may also simplify the interpretation of the results. In addition, panels are usually optimized for a specific set of genes, minimizing potential false negative findings due to incomplete sequence coverage. However, genes included in a particular panel vary between commercial services, and sometimes these panels fail to account for overlapping phenotypes. For example, nephrotic syndrome panels may not contain some congenital anomalies of kidney and urinary tract genes, such as *PAX2*, whose mutation may also manifest with proteinuria (36). The genes included in a specific panel must also be periodically updated as new genes are discovered.

Exome Sequencing. This sequencing approach examines nearly all coding regions of the genome (37). Exome sequencing proved its efficacy as a first-line test in both adult and pediatric patients, with a diagnostic yield extremely dependent on patient phenotype, age, and selection criteria. Recent studies demonstrate a diagnostic yield ranging from 11% in adult patients with FSGS (38), to 24% in primarily adult patients with congenital or cystic kidney disease (2). The yield is superior in pediatric cohorts: 32% in FSGS and 61% in cystic kidney disease (39). In general, the diagnostic yield outperforms targeted panels across different phenotypes in both adults and children (2,39,40). Importantly, exome sequencing allows for future reanalysis as new genes are discovered and analytic algorithms improve, without needing to resample the patient (41). As drawbacks, the coverage per base is generally lower than with targeted panels and can result in suboptimal coverage of some relevant genes (*e.g.*, *PKD1*, *GREB1L*) (42). Copy number variation detection with exome sequencing is possible but not always optimal, and the large amount of data generated can make interpretation challenging and time consuming (43). Several genomic regions relevant to nephrology are not optimally covered by exome sequencing and constitute blind spots that need to be recognized by clinicians ordering the test, such as duplicated region of *PKD1* or the *MUC1* variable number tandem repeat (a tandem repeat is a short repetitive nucleotide sequence that is generally difficult to sequence with short read MPS technology) (44). Importantly, a very promising application of exome sequencing is in the evaluation of nephropathies of unknown etiology, where, depending on the populations studied, it can have a diagnostic yield ranging from 17% (2) to 38% (8) in adult patients.

Genome Sequencing. This technique has the best analytic performance when compared with the aforementioned approaches, owing to its uniform coverage of the genome. Genome sequencing has superior coverage at a minimum depth of 10 reads compared to exome sequencing, and limits PCR artifacts, guanine-cytosine bias, and a high variance in allele fraction (42,45). It is also more efficient for copy number variant detection, even outperforming microarrays approaches (46). Genome sequencing can characterize noncoding regions, enabling detection of splicing or regulatory variants with large phenotypic effect, although this is currently limited by our incomplete understanding of the function of most noncoding regions (47,48). Thus, genome sequencing represents the most promising future of genetic

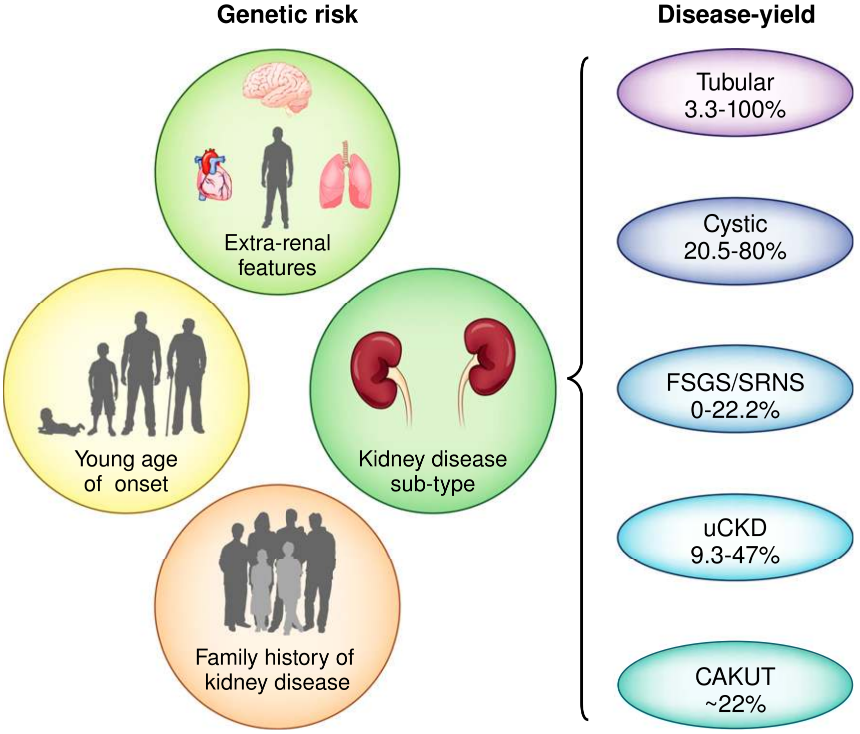


Figure 2. | **Clinical determinants of genetic risk and their influence on genetic test type.** The figure summarizes the clinical characteristics shaping the pretest risk of a genetic disease in a nephrology patient. Young age at onset, family history of kidney disease, and presence of extrarenal features are all predictive of genetic disease. Moreover, depending on the clinical diagnosis, the diagnostic yield of genetic testing varies. The yields of different modalities is shown in Figure 3. CAKUT, congenital anomalies of kidney and urinary tract; SRNS, steroid-resistant nephrotic syndrome; uCKD, CKD of unknown etiology.

testing because of its technical superiority, analytic performance, and mapping capacity, along with its progressively decreasing cost. Nonetheless, it still has blind spots shared with exome sequencing (*e.g.*, *PKD1* and *MUC1*). There is still a limited amount of research on the use of genome sequencing in the context of nephrology, but data support its superiority to traditional techniques in several other contexts (49,50). For example, genome sequencing had a significant improvement in diagnostic yield (+24%) when compared with targeted panels as a first-tier genetic test for pediatric disorders (51). Similar to exome sequencing, genome sequencing allows for future reanalysis and technologies application. Its potential drawbacks include the relatively high cost of the test (although this is progressively decreasing) and the possibility of detecting genetic findings not related to the primary indication for testing (*i.e.*, secondary findings).

Clinical Predictors of Diagnostic Yield

Several clinical factors influence the diagnostic potential of each test. To date, some studies support the use of exome sequencing as a first-line diagnostic approach for noncystic forms of genetic nephropathies. However, there is a paucity of data comparing the yield of targeted panel and exome/genome sequencing for different kidney phenotypes. In Figure 2, we summarize the clinical predictors of diagnostic yield and highlight optimal sequencing approaches on the basis of the broad kidney disease category. Then, in Figure 3, we compare the diagnostic yields achieved in genetic studies of diverse phenotypes using different sequencing techniques.

The main factors that contribute to variation in the diagnostic yield of the specific sequencing approach are discussed below.

Clinical Diagnosis

Cystic Kidney Disease. These disorders are predominantly attributable to mutations in *PKD1*/*PKD2* associated with adult-onset autosomal dominant polycystic kidney disease (the majority of cystic kidney diseases) and mutations in *PKHD1* in pediatric-onset autosomal recessive polycystic kidney disease. This clinical distinction is not absolute: 2% of patients with *PKD1/PKD2* mutations show a pediatric-onset severe phenotype clinically indistinguishable from autosomal recessive polycystic kidney disease (52), and other genes have been implicated as causal for cystic disease in both adults and pediatric patients (such as *GANAB* or *DNAJB11* in adults and *HNF1B* in both) (53,54). For cystic phenotypes, when there is a strong clinical suspicion of mutation in a specific gene, Sanger sequencing is the test of choice; otherwise, targeted panels are often an appropriate initial test because of the limited numbers of identified genes implicated to date. Furthermore, the assays can be optimized to cover the duplicated region of *PKD1* (45), which are otherwise not well covered by most other sequencing approaches. As such, the reported

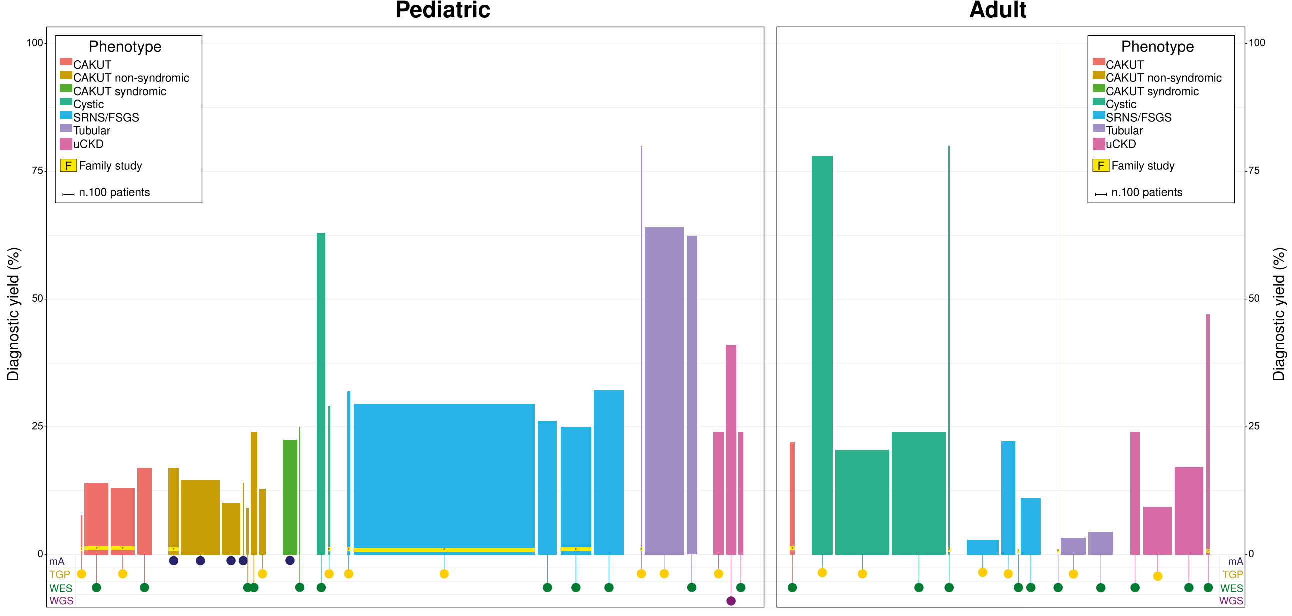


Figure 3. | **Diagnostic yield per phenotype and genetic test type.** The figure represents the diagnostic yield in different phenotype cohorts obtained through different genetic test type in pediatric and adult genetic studies on kidney disease. The *y* axis represents the percentage of diagnostic rate for the cohort (specified above each bar with the relative citation). The *x* axis represents the study under consideration; the width of the bar is dependent on the sample size for each study. The familial yellow flag indicates whether the study considered families (not individuals). The colored legend below the plot indicates the genetic test utilized in the study cohort. mA, chromosomal microarray techniques (especially in CAKUT); TGP, targeted gene panel; WES, whole exome sequencing; WGS, whole genome sequencing. The studies depicted here are presented in the Supplemental Material.

diagnostic yield of targeted panels for cystic kidney disease range in adult patients from 24% (2) to 88% (35), and 23%–80% with exome sequencing (2,8,55), based on the clinical characteristics of the cohort.

Congenital Anomalies of Kidney and Urinary Tract and Nephronophthisis. To date, these disorders are the most common cause of kidney disease in children, frequently involving congenital defects in other organ systems (56). With hundreds of genes implicated, they are highly heterogeneous in both phenotype and genotype. In addition, these phenotypes are often caused by structural genomic variants (57), detectable by chromosomal microarray (58), whose diagnostic yield is greatest in patients with kidney parenchymal malformations and those with extrarenal manifestations (27,31,59). The remaining cases are caused by SNVs in a single gene, identifiable with exome sequencing or targeted panels, which can provide a diagnostic yield of around 14% in recent reports (3) and is slightly superior in syndromic cases (25% versus 9% [32], 23% versus 15% [31,60]) in pediatric and young adult cohorts.

FSGS and Steroid-Resistant Nephrotic Syndrome. This is a genetically heterogeneous phenotype with over 30 known causal genes (61), affecting primarily children and young adults (62,63). The diagnostic yield can be as high as 32% (39) in recently analyzed cohorts with exome sequencing, and is higher in pediatric cases than in adult ones. Steroid-resistant nephrotic syndrome reached diagnostic yield of 25%–30% (39,64–66) in pediatric patients versus 12% in adults (67), whereas FSGS proved around 22% in pediatric patients (68) and 11% in adults (38). Importantly, a genetic diagnosis can have strong implications for patient management and should always be considered in early-onset disease, particularly in the presence of a clear family history. For example, most genetic forms of steroid-resistant nephrotic syndrome may not respond to immunosuppression therapy (69). Therefore, a positive finding may spare a patient from unwanted complications from this therapy (70). Similarly, a negative test may also guide management as it may influence clinical decision-making toward a longer and/or more aggressive immunosuppression course, and may provide prognostic insights too, as hereditary forms are associated with lower recurrence rate after transplantation (9). Finally, genetic testing can identify cases that may respond to therapy to glucocorticoids, such as cases owing to *PLCE1* (71), or cases that may benefit from other therapies that can be effectively treated with coenzyme Q10 (CoQ10) supplementation, such as steroid-resistant nephrotic syndromes owing to mutations in genes of the CoQ10 biosynthesis pathway (*ADCK4*, *COQ2*, *COQ6*, or *PDSS2*) (72,73). There are many genetic panels available as a first-line diagnostics, but exome sequencing can also be justifiable as a test of choice because of the rapid pace of discovery of new genes for nephrotic syndrome.

Tubulopathies (*e.g.*, Autosomal Dominant Tubulointerstitial Kidney Disease [ADTKD]). Recent studies show targeted panels to have the highest diagnostic yield in this disease category. In fact, diagnostic yield ranges from 62% (38,74) to 83% (75) in pediatric cohorts and 75% (75) to 100% (8) in adult cohorts. This is probably because of the highly characteristic clinical presentation and subsequent selection of cases. Importantly, recent findings suggest that trio analysis with exome sequencing, whereby both the

proband and their parents are sequenced for comparison, can significantly increase the targeted panels diagnostic performance because the availability of parental data enables segregation analysis and detection of *de novo* mutations (63). Several genes have been implicated in ADTKD to date, such as *UMOD, HNF1B, REN*, and *MUC1* in autosomal dominant tubulointerstitial kidney disease (76). Importantly, though commercial-targeted panels may include *MUC1*, the ADTKD is associated with variable-number tandem repeat mutations, which makes massively parallel sequencing approaches challenging because of the low coverage in these genomic regions (77).

Kidney Disease of Unknown Etiology. This is an important clinical category representing approximately 5%–15% of patients with kidney failure (1). If a thorough nephrology workup is nondiagnostic, genetic testing should be considered, particularly in the setting of early age of disease onset and a positive family history. The diagnostic yield in both adults and children ranges from 7% to 40% depending on patient characteristics and clinical history, with over two dozen different disorders described across developmental, glomerular, and tubulointerstitial categories (2,72,75). Given the heterogeneous disorders contributing to this disease category, exome sequencing had superior diagnostic performance as compared with targeted gene panels (2).

Family History and Extrarenal Manifestations

Studies indicate a three- to four-fold greater chance of a positive genetic finding in the setting of a positive family history of kidney disease (2,72). For example, a positive family history finding was much more likely in adults (53%) than in pediatric patients (12%) in a cohort of patients with monogenic nephrotic syndrome (34). It is worth noting that the clinical presentation can be different in different family members with the same mutation. For example, the same mutation in *PAX2*, responsible for papillorenal syndrome, may present as hearing or visual defect in some family members, but present as kidney diseases in other relatives (78). This variable expression motivates obtaining a thorough family history of disease in patients suspected of genetic disease.

In addition, multiorgan involvement is a compelling clinical finding suggestive of a genetic disorder. In a recent study of exome sequencing in adult nephrology patients, the presence of extrarenal features increased the diagnostic yield up to 69% (8). Although obtaining a detailed family history, or being able to decipher extrarenal features associated with a genetic diagnosis, may seem challenging for nongenetic experts, it is important to appreciate that these are key features suggestive of a genetic diagnosis and must be emphasized in medical education going forward. Moreover, collaborating with genetic counselors and/or clinical geneticists can ensure better collection of pertinent family and clinical data. Furthermore, the introduction of decision-support tools and electronic algorithms in medical health records, able to flag potential Mendelian disorders, will help clinicians in detecting genetic disorders and extrarenal features they might not have recognized (79).

Age of Onset

Mendelian disorders are more common in children (75), and age of onset has been shown to significantly influence the probability that the cause is an underlying genetic diagnosis for certain clinical subtypes. For example, genetic cystic kidney disease is almost three-fold more probable in pediatric cases than in adult ones (80). Hence, early onset of disease should increase suspicion for a genetic disease.

Genetic Data Analysis

Interpretation of genetic test results requires close collaboration between the clinician and the molecular pathologist, akin to the collaboration between the nephrologist and kidney pathologist in the evaluation of a biopsy. The interpretation of sequence data depends on the genomic tests used and requires gene- and variant-level analyses, phenotypic-level comparison, and determination of actionability. It also requires domain-specific expertise about the molecular and clinical features of disease. In an effort to standardize this process, the American College of Medical Genetics and Genomics (ACMG) has established guidelines for interpretation of sequence variants (81). Variants are classified as "pathogenic," "likely pathogenic," "variants of unknown significance (VUS)," "likely benign," or "benign." These guidelines recommend using terms such as "pathogenic" and "benign" for variants that are almost certainly disease causing or known to be benign, respectively; "likely pathogenic" and "likely benign" is used for variants with >90% certainty of either being disease causing or benign. VUS is used for variants for which there is insufficient evidence to classify in any of the other categories. As a general rule, clinicians should only act on variants classified as "pathogenic" and "likely pathogenic." These categories of variants are generally very rare in the population (typically with allele frequency <1:10,000), and have robust evidence of disease association on the basis of multiple prior reports, predicted effect on protein function, and potentially *in vitro* functional studies. Once a "pathogenic" or "likely pathogenic" variant has been identified, a second-level comparison is warranted to determine concordance of the patient's characteristics with the phenotype reported for the disease. Any discrepancy should be reviewed by both the molecular pathologist and clinician, and prior reports should be reviewed. Importantly, recent data suggest that many VUS will eventually be classified as benign (82–86) and can usually be resolved with subsequent clinical correlations, in discussions between the molecular diagnostician, clinician, and the patient.

Although a number of gene lists for nephropathy traits have been generated (2,3,21,34,35,87), nephrology lacks standardized, well-curated gene and variant lists. This will change as kidney disease is increasingly represented in initiatives such as ClinGen, which aims to define the clinical relevance of genes and variants for clinical and research use (88). The development of expertly-curated gene and variant lists is critical because of the many nuances in sequence interpretation. For example, although gene-disrupting mutations are often pathogenic, some disorders are almost exclusively caused by missense variants affecting specific domains of the protein

(*e.g.*, FSGS owing to mutations in the diaphanous inhibitory domains of *INF2* [89], or autosomal dominant tubulointerstitial kidney disease owing to mutations in exons 3 or 4 of *UMOD* [90], *etc*.). Therefore, loss-of-function (*e.g.*, nonsense, frameshift, splice, *etc*.) and missense variants outside of functional domains, even if exceedingly rare, are unlikely to be causal for these diseases (91). Conversely, a relatively common *NPHS2* variant (R229Q) can be disease causing if in *trans* with specific 3′ *NPHS2* mutations that exert a dominant-negative effect on the R229Q protein (92).

Currently, we still have an incomplete catalog of rare variants of the genome and limited understanding of their functional effect, hence many rare variants do not clearly fit into "pathogenic" or "benign" categories and are designated as VUS. A VUS classification frequently arises when a variant is novel and *in silico* algorithms predict it to be damaging, but there is insufficient data on its *in vivo* functional consequences and inheritance pattern. Classification as a VUS can also occur when a variant has been previously reported as disease causal, but there is contradictory data about its pathogenicity. This is the case of some population-specific polymorphisms or modestly rare variants that were erroneously thought to be extremely rare and reported as pathogenic (91). With the availability of larger and more diverse reference data sets, many previously reported pathogenic variants are now shown to have an allele frequency that exceeds the disease prevalence and can now be reclassified as benign (82,93). For example, a recent study showed that the top genes potentially contributing to most false positive results for kidney diagnostics were all identified before the availability of large publicly available databases (21). It is expected that the increasing number of genomes sequenced across populations of different ancestries will provide better estimates of variant allele frequencies, reducing the numbers of VUS and false positives. In addition, the availability of saturation mutagenesis of human genes (94) raises the possibility of obtaining functional information about every potential genetic variant in the future (in saturation mutagenesis, a single amino acid can be substituted to any of the other 19 possible substituents, which allows a comprehensive analysis of the function of amino acids at each position in the protein).

With the accelerating pace of new discoveries, an emerging challenge is the continuous review of data. Thus, a reanalysis of exomes may reveal a new diagnostic finding and revise a negative genetic test result. A recent report from the National Institutes of Health–sponsored Centers for Mendelian Genomics indicated that a new genetic discovery is made for every 28 exomes sequenced (95). This rapid pace of genetic discovery necessitates systematic updating of variant databases and gene panels, and provides clinicians and patients with the future prospect of achieving a diagnosis when the initial genetic findings are unrevealing. For example, a recent reevaluation of exome sequence data from two cohorts initially analyzed in 2012–2013 revealed that the diagnostic yield increased from 25% to 37% over a 5-year period, attributable mainly to newly discovered disease-causal genes (96). There are additional strategies for addressing unsolved cases, as shown by the Undiagnosed Disease Network of the National Institutes of Health (97): transcriptome analysis (also called RNA sequencing, which uses massive parallel sequencing approaches to reveal the sequence and quantity of RNA in a biologic sample) can reveal the possible effect of noncoding DNA sequence variants on splicing or allele-specific expression; or proteomic analysis can detect translation defects or point to metabolic defects. Although these approaches create opportunities for improving clinical care, they are resource intensive for the diagnostic laboratories and clinicians. Furthermore, sequence interpretation can be a burdensome process, particularly when each step is performed manually (98). Hence, semiautomated algorithms can rapidly characterize variant allele frequencies, identify previously reported pathogenic variants, incorporate results of *in silico* prediction tools, and expedite some of the more labor-intensive steps. There are also currently no guidelines and standards for reanalysis of genetic test results, and this challenge will need to be addressed in the next few years.

Return of Genetic Results and Clinical Implementation

The diagnosis of a Mendelian disorder, whether in the context of a diagnostic finding explicative of the patient's kidney disease or a medically actionable secondary finding, has the potential to meaningfully inform nephrology care (see Box 1; 99–101). The clinical value of actionable genetic findings underscores the importance of returning genetic findings to patients. Today, most experts agree that clinically valid, medically actionable genetic findings should be returned to adult patients who opted to receive them at the time of consent (102), including for individuals who underwent genetic sequencing through their participation in research (103). Importantly, however, genetic testing results can affect insurability and confidentiality, which many patients and providers may not fully realize. Although legislation exists through the 2009 Genetic Information Nondiscrimination Act to protect employers and health insurance carriers from discriminating against individuals on the basis of their genetic results (104,105), it is not fully comprehensive. Furthermore, clinicians should also be aware of the potential psychosocial implications of a genetic diagnosis for a patient, which may arise from multiple factors, including marginalization and stigmatization from members of their family and community. Therefore, it is essential that all patients thoroughly understand the implication of genetic testing and that genetic counseling services are available for them and their families.

On the clinical side, implementing genetic results into clinical care is an important and challenging task that can lead to effective targeted treatment; for example, patients with steroid-resistant nephrotic syndrome owing to mutations in genes of the CoQ10 biosynthesis pathway can be effectively treated with CoQ10 supplementation (72), and RNA interference therapies have been developed for primary hyperoxaluria type 1 caused by mutations in *AGXT* (105,106). In the same manner, secondary results affect patient health and need to be reported. The role of clinical nephrologists in this regard is generally limited to identifying and activating the right specialist for patient screening/follow-up. Moreover, ACMG guidelines

Case Study 1

A male patient with Focal Segmental Glomerulosclerosis on kidney biopsy is found to have a pathogenic large truncating variant in the *COL4A5* gene, diagnostic for Alport syndrome. In addition to establishing a molecular diagnosis for Alport syndrome, the type of variant may provide insights into the patient's clinical course, wherein Loss-of-Function variants, such as the one in this example, are associated with early onset of kidney failure, hearing loss and ocular abnormalities (116) and large deletions in the *COL4A5* gene are associated with increased risk for Anti-glomerular basement membrane disease post-transplantation (117). In addition, the genetic diagnosis may inform allograft donor selection among at-risk family members, as mothers of males with X-linked Alport syndrome are typically discouraged from serving as donors (118). Furthermore, conservative management of the patient's proteinuria with an agent that inhibits the Renin-Angiotensin-Aldosterone-system (RAAS) (e.g., ACE-Inhibitor or ARB) is recommend ed (3). Finally, a genetic diagnosis can also help identify disease-specific support groups, which can provide psychosocial support for affected individuals and their families, guide patients to clinical drug trials, registries, and other relevant resources.

Box 1. | Establishing a genetic diagnosis can support personalized nephrology care: A case study.

currently limit the list to 59 actionable genes in which mutations strongly affect patient prognosis and life expectancy (104), but larger sets are used in different laboratories (107). These challenging tasks require support to ensure the nephrologists' correct use of diagnostic genetic testing. This will also require collaborations with genetic counselors and clinical geneticists, development of clinical decision-support tools in electronic health records that address physician knowledge gaps in genomics, and the availability of remote consultation *via* telemedicine for providers with limited access to genetic services.

Genetic Testing and Clinical Trials

Genetics can affect the development and design of clinical trials in multiple ways. First, discoveries in human genetics can uncover new targets for therapy. In fact, recent analyses suggest that therapies targeting pathways discovered through human genetic studies have a greater chance of successful validation in clinical trials (108,109). A well known success story involves the identification of *PCSK9* mutations in rare forms of hyperlipidemia, which led to the development of treatment for common forms of this trait (110,111), and promising results are emerging in nephrology, such as chaperone-based therapies in Fabry disease (112). The incorporation of genetic analysis in clinical trials can also identify a subset of patients who may preferentially benefit from therapy, such as the discovery of clinical responsiveness to a tyrosine kinase inhibitor in patients with nonsmall cell lung cancer with specific somatic mutations in the *EGFR* gene (113). Similarly, *JAK1/JAK2* mutations in solid tumors can predict response to anti-PD1/PDL1 therapies (114). It is thus not surprising that different initiatives have been taken in this sense; for example, the Precision Medicine Initiative is an investment of $70 million to the National Cancer Institute to "scale up efforts to identify genomic drivers in cancer and apply that knowledge in the development of more effective approaches to cancer treatment" (115). In addition, several trials are now being developed to target specific molecular defects rather than the clinical category of disease, aiming to treat cancer according to underlying genomic alterations, regardless of clinical cancer type (116). Even if the trial is not targeting specific mutations, genetic testing may enhance the study power by identifying subsets of patients who would likely not respond to the investigation therapy. For example, exclusion of patients with mutations affecting kidney structure or function (such as *COL4A5* or *NPHS2*) may improve the power of studies targeting immunologic pathways for the treatment of steroid-resistant nephrotic syndrome and also minimize exposure to side effects in patients unlikely to benefit from such therapy. Moreover, identification of patients with genetic predisposition to outcomes such as cancer or cardiomyopathy may also help with adjudication of adverse events in clinical trials. Finally, pharmacogenomic analyses can help identify variants that affect drug absorption or metabolism, as exemplified by variants in *CYP2C9/VKORC1* for warfarin toxicity (117) or *TPMT* for azathioprine toxicity, enabling better assessment of dosage, safety, and side effects.

Conclusions

Achieving a precise diagnosis is a fundamental goal of medical practice. Although historically focused on pediatric populations, genetic testing has emerged as a powerful diagnostic tool in adult nephrology, with important implications for diagnosis and management. As genetic testing is increasingly incorporated into clinical practice, many challenges are yet to be addressed, such as optimizing the accuracy of variant interpretation, deploying genetic testing affordable on a large scale, and improving physician and patient education. More research is also needed to investigate the long-term effect of establishing a molecular diagnosis on health care utilization and outcomes, to facilitate third-party payer coverage for diagnostic testing. In this regard, there are emerging data on the cost benefit of genetic testing in fields such as diabetes or cancer (118–120), and similar research would benefit the nephrology field. With the rapid pace of discovery, we also anticipate that the nephrology field will increasingly deploy clinical trials stratified on specific molecular defects rather than the clinical category of disease, increasing the power and safety profile of studies.

Disclosures

A. Gharavi reports receiving grants from Renal Research Institute and participating on advisory boards for AstraZeneca Center for Genomics Research and Goldfinch Bio, outside the submitted work. All remaining authors have nothing to disclose.

Funding

The authors are supported by grants from the National Institutes of Health (DK080099, to A. Gharavi; and TL1TR001875, to J. Nestor), and the Columbia Precision Medicine Initiative (to A. Gharavi) from Columbia University.

Supplemental Material

This article contains the following supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.15141219/-/DCSupplemental.

Supplemental Material. References for Figure 3 studies.

References

1. Saran R, Robinson B, Abbott KC, Agodoa LYC, Bragg-Gresham J, Balkrishnan R, Bhave N, Dietrich X, Ding Z, Eggers PW, Gaipov A, Gillen D, Gipson D, Gu H, Guro P, Haggerty D, Han Y, He K, Herman W, Heung M, Hirth RA, HsiungJ-T, Hutton D, Inoue A, Jacobsen SJ, Jin Y, Kalantar-Zadeh K, Kapke A, KleineC-E, Kovesdy CP, Krueter W, Kurtz V, Li Y, Liu S, Marroquin MV, McCullough K, Molnar MZ, Modi Z, Montez-Rath M, Moradi H, Morgenstern H, Mukhopadhyay P, Nallamothu B, Nguyen DV, Norris KC, O'Hare AM, Obi Y, Park C, Pearson J, Pisoni R, Potukuchi PK, Repeck K, Rhee CM, Schaubel DE, Schrager J, Selewski DT, Shamraj R, Shaw SF, Shi JM, Shieu M, Sim JJ, Soohoo M, Steffick D, Streja E, Sumida K, Kurella Tamura M, Tilea A, Turf M, Wang D, Weng W, Woodside KJ, Wyncott A, Xiang J, Xin X, Yin M, You AS, Zhang X, Zhou H, Shahinian V: US renal data system 2018 annual data report: Epidemiology of kidney disease in the United States. *Am J Kidney Dis* 73[Suppl 1]: A7–A8, 2019
2. Groopman EE, Marasa M, Cameron-Christie S, Petrovski S, Aggarwal VS, Milo-Rasouly H, Li Y, Zhang J, Nestor J, Krithivasan P, Lam WY, Mitrotti A, Piva S, Kil BH, Chatterjee D, Reingold R, Bradbury D, DiVecchia M, Snyder H, Mu X, Mehl K, Balderes O, Fasel DA, Weng C, Radhakrishnan J, Canetta P, Appel GB, Bomback AS, Ahn W, Uy NS, Alam S, Cohen DJ, Crew RJ, Dube GK, Rao MK, Kamalakaran S, Copeland B, Ren Z, Bridgers J, Malone CD, Mebane CM, Dagaonkar N, Fellström BC, Haefliger C, Mohan S, Sanna-Cherchi S, Kiryluk K, Fleckner J, March R, Platt A, Goldstein DB, Gharavi AG: Diagnostic utility of exome sequencing for kidney disease. *N Engl J Med* 380: 142–151, 2019
3. van der Ven AT, Connaughton DM, Ityel H, Mann N, Nakayama M, Chen J, Vivante A, HwangD-Y, Schulz J, Braun DA, Schmidt JM, Schapiro D, Schneider R, Warejko JK, Daga A, Majmundar AJ, Tan W, Jobst-Schwan T, Hermle T, Widmeier E, Ashraf S, Amar A, Hoogstraaten CA, Hugo H, Kitzler TM, Kause F, Kolvenbach CM, Dai R, Spaneas L, Amann K, Stein DR, Baum MA, Somers MJG, Rodig NM, Ferguson MA, Traum AZ, Daouk GH, Bogdanović R, Stajić N, Soliman NA, Kari JA, El Desoky S, Fathy HM, Milosevic D, Al-Saffar M, Awad HS, Eid LA, Selvin A, Senguttuvan P, Sanna-Cherchi S, Rehm HL, MacArthur DG, Lek M, Laricchia KM, Wilson MW, Mane SM, Lifton RP, Lee RS, Bauer SB, Lu W, Reutter HM, Tasic V, Shril S, Hildebrandt F: Whole-exome sequencing identifies causative mutations in families with congenital anomalies of the kidney and urinary tract. *J Am Soc Nephrol* 29: 2348–2361, 2018
4. Connaughton DM, Bukhari S, Conlon P, Cassidy E, O'Toole M, Mohamad M, Flanagan J, Butler T, O'Leary A, Wong L, O'Regan J, Moran S, O'Kelly P, Logan V, Griffin B, Griffin M, Lavin P, Little MA, Conlon P: The Irish kidney gene project--prevalence of family history in patients with kidney disease in Ireland. *Nephron* 130: 293–301, 2015
5. van Nimwegen KJM, van Soest RA, Veltman JA, Nelen MR, van der Wilt GJ, Vissers LELM, Grutters JPC: Is the $1000 genome as near as we think? A cost analysis of next-generation sequencing. *Clin Chem* 62: 1458–1464, 2016
6. Phillips KA, Deverka PA, Hooker GW, Douglas MP: Genetic test availability and spending: Where are we now? Where are we going? *Health Aff (Millwood)* 37: 710–716, 2018
7. Lata S, Marasa M, Li Y, Fasel DA, Groopman E, Jobanputra V, Rasouly H, Mitrotti A, Westland R, Verbitsky M, Nestor J, Slater LM, D'Agati V, Zaniew M, Materna-Kiryluk A, Lugani F, Caridi G, Rampoldi L, Mattoo A, Newton CA, Rao MK, Radhakrishnan J, Ahn W, Canetta PA, Bomback AS, Appel GB, Antignac C, Markowitz GS, Garcia CK, Kiryluk K, Sanna-Cherchi S, Gharavi AG: Whole-exome sequencing in adults with chronic kidney disease: A pilot study. *Ann Intern Med* 168: 100–109, 2018
8. Connaughton DM, Kennedy C, Shril S, Mann N, Murray SL, Williams PA, Conlon E, Nakayama M, van der Ven AT, Ityel H, Kause F, Kolvenbach CM, Dai R, Vivante A, Braun DA, Schneider R, Kitzler TM, Moloney B, Moran CP, Smyth JS, Kennedy A, Benson K, Stapleton C, Denton M, Magee C, O'Seaghdha CM, Plant WD, Griffin MD, Awan A, Sweeney C, Mane SM, Lifton RP, Griffin B, Leavey S, Casserly L, de Freitas DG, Holian J, Dorman A, Doyle B, Lavin PJ, Little MA, Conlon PJ, Hildebrandt F: Monogenic causes of chronic kidney disease in adults. *Kidney Int* 95: 914–928, 2019
9. Morello W, Puvinathan S, Puccio G, Ghiggeri GM, Dello Strologo L, Peruzzi L, Murer L, Cioni M, Guzzo I, Cocchi E, Benetti E, Testa S, Ghio L, Caridi G, Cardillo M, Torelli R, Montini G: Post-transplant recurrence of steroid resistant nephrotic syndrome in children: The Italian experience [published online ahead of print Oct 15, 2019]. *J Nephrol* doi:10.1007/s40620-019-00660-9
10. Cornec-Le Gall E, AudrézetM-P, ChenJ-M, Hourmant M, MorinM-P, Perrichot R, Charasse C, Whebe B, Renaudineau E, Jousset P, GuillodoM-P, Grall-Jezequel A, Saliou P, Férec C, Le Meur Y: Type of PKD1 mutation influences renal outcome in ADPKD. *J Am Soc Nephrol* 24: 1006–1013, 2013
11. Nestor JG, Marasa M, Milo-Rasouly H, Groopman EE, Husain SA, Mohan S, Fernandez H, Aggarwal VS, Ahram DF, Vena N, Bogyo K, Bomback AS, Radhakrishnan J, Appel GB, Ahn W, Cohen DJ, Canetta PA, Dube GK, Rao MK, Morris HK, Crew RJ, Sanna-Cherchi S, Kiryluk K, Gharavi AG: Pilot study of return of genetic results to patients in adult nephrology. *J Am Soc Nephrol* 15: 651–664, 2020
12. Temme J, Peters F, Lange K, Pirson Y, Heidet L, Torra R, GrunfeldJ-P, Weber M, Licht C, MüllerG-A, Gross O: Incidence of renal failure and nephroprotection by RAAS inhibition in heterozygous carriers of X-chromosomal and autosomal recessive Alport mutations. *Kidney Int* 81: 779–783, 2012
13. MüllerR-U, Haas CS, Sayer JA: Practical approaches to the management of autosomal dominant polycystic kidney disease patients in the era of tolvaptan. *Clin Kidney J* 11: 62–69, 2018
14. Adam J, Connor TMF, Wood K, Lewis D, Naik R, Gale DP, Sayer JA: Genetic testing can resolve diagnostic confusion in Alport syndrome. *Clin Kidney J* 7: 197–200, 2014
15. Liapis H, Gaut JP: The renal biopsy in the genomic era. *Pediatr Nephrol* 28: 1207–1219, 2013
16. Marian AJ: Sequencing your genome: What does it mean? *Methodist DeBakey Cardiovasc J* 10: 3–6, 2014
17. Homo sapiens Annotation Report. Available at: https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Homo_sapiens/109/. Accessed February 27, 2020
18. Amberger JS, Bocchini CA, Scott AF, Hamosh A: OMIM.org: Leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res* 47: D1038–D1043, 2019
19. Malekpour SA, Pezeshk H, Sadeghi M: MSeq-CNV: Accurate detection of copy number variation from sequencing of multiple samples. *Sci Rep* 8: 4009, 2018
20. Zarrei M, MacDonald JR, Merico D, Scherer SW: A copy number variation map of the human genome. *Nat Rev Genet* 16: 172–183, 2015
21. Rasouly HM, Groopman EE, Heyman-Kantor R, Fasel DA, Mitrotti A, Westland R, Bier L, Weng C, Ren Z, Copeland B, Krithivasan P, Chung WK, Sanna-Cherchi S, Goldstein DB, Gharavi AG: The burden of candidate pathogenic variants for kidney and genitourinary disorders emerging from exome sequencing. *Ann Intern Med* 170: 11–21, 2019

22. PetersenB-S, Fredrich B, Hoeppner MP, Ellinghaus D, Franke A: Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet* 18: 14, 2017
23. Wang XV, Blades N, Ding J, Sultana R, Parmigiani G: Estimation of sequencing error rates in short reads. *BMC Bioinformatics* 13: 185, 2012
24. Morozova O, Marra MA: Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92: 255–264, 2008
25. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD: Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 7: e1002384, 2011
26. Snoek R, van Setten J, Keating BJ, Israni AK, Jacobson PA, Oetting WS, Matas AJ, Mannon RB, Zhang Z, Zhang W, Hao K, Murphy B, Reindl-Schwaighofer R, Heinzl A, Oberbauer R, Viklicky O, Conlon PJ, Stapleton CP, Bakker SJL, Snieder H, Peters EDJ, van der Zwaag B, Knoers NVAM, de Borst MH, van Eerde AM: *NPHP1* (Nephrocystin-1) gene deletions cause adult-onset ESRD. *J Am Soc Nephrol* 29: 1772–1779, 2018
27. Verbitsky M, Westland R, Perez A, Kiryluk K, Liu Q, Krithivasan P, Mitrotti A, Fasel DA, Batourina E, Sampson MG, Bodria M, Werth M, Kao C, Martino J, Capone VP, Vivante A, Shril S, Kil BH, Marasà M, Zhang JY, NaY-J, Lim TY, Ahram D, Weng PL, Heinzen EL, Carrea A, Piaggio G, Gesualdo L, Manca V, Masnata G, Gigante M, Cusi D, Izzi C, Scolari F, van Wijk JAE, Saraga M, Santoro D, Conti G, Zamboli P, White H, Drozdz D, Zachwieja K, Miklaszewska M, Tkaczyk M, Tomczyk D, Krakowska A, Sikora P, Jarmoliński T, Borszewska-Kornacka MK, Pawluch R, Szczepanska M, Adamczyk P, Mizerska-Wasiak M, Krzemien G, Szmigielska A, Zaniew M, Dobson MG, Darlow JM, Puri P, Barton DE, Furth SL, Warady BA, Gucev Z, Lozanovski VJ, Tasic V, Pisani I, Allegri L, Rodas LM, Campistol JM, Jeanpierre C, Alam S, Casale P, Wong CS, Lin F, Miranda DM, Oliveira EA, Simões-E-Silva AC, Barasch JM, Levy B, Wu N, Hildebrandt F, Ghiggeri GM, Latos-Bielenska A, Materna-Kiryluk A, Zhang F, Hakonarson H, Papaioannou VE, Mendelsohn CL, Gharavi AG, Sanna-Cherchi S: The copy number variation landscape of congenital anomalies of the kidney and urinary tract [published correction appears in *Nat Genet* 51: 764, 2019]. *Nat Genet* 51: 117–127, 2019
28. Siomou E, Mitsioni AG, Giapros V, Bouba I, Noutsopoulos D, Georgiou I: Copy-number variation analysis in familial nonsyndromic congenital anomalies of the kidney and urinary tract: Evidence for the causative role of a transposable element-associated genomic rearrangement. *Mol Med Rep* 15: 3631–3636, 2017
29. Caruana G, Wong MN, Walker A, Heloury Y, Webb N, Johnstone L, James PA, Burgess T, Bertram JF: Copy-number variation associated with congenital anomalies of the kidney and urinary tract. *Pediatr Nephrol* 30: 487–495, 2015
30. Sanna-Cherchi S, Kiryluk K, Burgess KE, Bodria M, Sampson MG, Hadley D, Nees SN, Verbitsky M, Perry BJ, Sterken R, Lozanovski VJ, Materna-Kiryluk A, Barlassina C, Kini A, Corbani V, Carrea A, Somenzi D, Murtas C, Ristoska-Bojkovska N, Izzi C, Bianco B, Zaniew M, Flogelova H, Weng PL, Kacak N, Giberti S, Gigante M, Arapovic A, Drnasin K, Caridi G, Curioni S, Allegri F, Ammenti A, Ferretti S, Goj V, Bernardo L, Jobanputra V, Chung WK, Lifton RP, Sanders S, State M, Clark LN, Saraga M, Padmanabhan S, Dominiczak AF, Foroud T, Gesualdo L, Gucev Z, Allegri L, Latos-Bielenska A, Cusi D, Scolari F, Tasic V, Hakonarson H, Ghiggeri GM, Gharavi AG: Copy-number disorders are a common cause of congenital kidney malformations. *Am J Hum Genet* 91: 987–997, 2012
31. Weber S, Moriniere V, Knüppel T, Charbit M, Dusek J, Ghiggeri GM, Jankauskiené A, Mir S, Montini G, Peco-Antic A, Wühl E, Zurowska AM, Mehls O, Antignac C, Schaefer F, Salomon R: Prevalence of mutations in renal developmental genes in children with renal hypodysplasia: Results of the ESCAPE study. *J Am Soc Nephrol* 17: 2864–2870, 2006
32. LeiT-Y, Fu F, Li R, Wang D, WangR-Y, JingX-Y, Deng Q, LiZ-Z, LiuZ-Q, Yang X, LiD-Z, Liao C: Whole-exome sequencing for prenatal diagnosis of fetuses with congenital anomalies of the kidney and urinary tract. *Nephrol Dial Transplant* 32: 1665–1675, 2017
33. Ashley EA: Towards precision medicine. *Nat Rev Genet* 17: 507–522, 2016
34. Sen ES, Dean P, Yarram-Smith L, Bierzynska A, Woodward G, Buxton C, Dennis G, Welsh GI, Williams M, Saleem MA: Clinical genetic testing using a custom-designed steroid-resistant nephrotic syndrome gene panel: Analysis and recommendations. *J Med Genet* 54: 795–804, 2017
35. Bullich G, Domingo-Gallego A, Vargas I, Ruiz P, Lorente-Grandoso L, Furlano M, Fraga G, Madrid Á, Ariceta G, Borregán M, Piñero-Fernández JA, Rodríguez-Peña L, Ballesta-Martínez MJ, Llano-Rivas I, Meñica MA, Ballarín J, Torrents D, Torra R, Ars E: A kidney-disease gene panel allows a comprehensive genetic diagnosis of cystic and glomerular inherited kidney diseases. *Kidney Int* 94: 363–371, 2018
36. Barua M, Stellacci E, Stella L, Weins A, Genovese G, Muto V, Caputo V, Toka HR, Charoonratana VT, Tartaglia M, Pollak MR: Mutations in PAX2 associate with adult-onset FSGS. *J Am Soc Nephrol* 25: 1942–1953, 2014
37. Patwardhan A, Harris J, Leng N, Bartha G, Church DM, Luo S, Haudenschild C, Pratt M, Zook J, Salit M, Tirch J, Morra M, Chervitz S, Li M, Clark M, Garcia S, Chandratillake G, Kirk S, Ashley E, Snyder M, Altman R, Bustamante C, Butte AJ, West J, Chen R: Achieving high-sensitivity for clinical applications using augmented exome sequencing. *Genome Med* 7: 71, 2015
38. Yao T, Udwan K, John R, Rana A, Haghighi A, Xu L, Hack S, Reich HN, Hladunewich MA, Cattran DC, Paterson AD, Pei Y, Barua M: Integration of genetic testing and pathology for the diagnosis of adults with FSGS. *Clin J Am Soc Nephrol* 14: 213–223, 2019
39. Rao J, Liu X, Mao J, Tang X, Shen Q, Li G, Sun L, Bi Y, Wang X, Qian Y, Wu B, Wang H, Zhou W, Ma D, Zheng B, Shen Y, Chen Z, Luan J, Wang X, Wang M, Dang X, Wang Y, Wu Y, Hou L, Sun S, Li Q, Liu X, Bai H, Yang Y, Shao X, Li Y, Zheng S, Han M, Liu C, Cao G, Zhao L, Qiu S, Dong Y, Zhu Y, Wang F, Zhang D, Li Y, Zhao L, Yang C, Luo X, Chen L, Jiang X, Zhang A, Xu H; for "Internet Plus" Nephrology Alliance of National Center for Children's Care: Genetic spectrum of renal disease for 1001 Chinese children based on a multicenter registration system. *Clin Genet* 96: 402–410, 2019
40. Mann N, Braun DA, Amann K, Tan W, Shril S, Connaughton DM, Nakayama M, Schneider R, Kitzler TM, van der Ven AT, Chen J, Ityel H, Vivante A, Majmundar AJ, Daga A, Warejko JK, Lovric S, Ashraf S, Jobst-Schwan T, Widmeier E, Hugo H, Mane SM, Spaneas L, Somers MJG, Ferguson MA, Traum AZ, Stein DR, Baum MA, Daouk GH, Lifton RP, Manzi S, Vakili K, Kim HB, Rodig NM, Hildebrandt F: Whole-exome sequencing enables a precision medicine approach for kidney transplant recipients. *J Am Soc Nephrol* 30: 201–215, 2019
41. Baker SW, Murrell JR, Nesbitt AI, Pechter KB, Balciuniene J, Zhao X, Yu Z, Denenberg EH, DeChene ET, Wilkens AB, Bhoj EJ, Guan Q, Dulik MC, Conlin LK, Abou Tayoun AN, Luo M, Wu C, Cao K, Sarmady M, Bedoukian EC, Tarpinian J, Medne L, Skraban CM, Deardorff MA, Krantz ID, Krock BL, Santani AB: Automated clinical exome reanalysis reveals novel diagnoses. *J Mol Diagn* 21: 38–48, 2019
42. Kingsmore SF, Cakici JA, Clark MM, Gaughran M, Feddock M, Batalov S, Bainbridge MN, Carroll J, Caylor SA, Clarke C, Ding Y, Ellsworth K, Farnaes L, Hildreth A, Hobbs C, James K, Kint CI, Lenberg J, Nahas S, Prince L, Reyes I, Salz L, Sanford E, Schols P, Sweeney N, Tokita M, Veeraraghavan N, Watkins K, Wigby K, Wong T, Chowdhury S, Wright MS, Dimmock D; RCIGM Investigators: A randomized, controlled trial of the analytic and diagnostic performance of singleton and trio, rapid genome and exome sequencing in ill infants. *Am J Hum Genet* 105: 719–733, 2019
43. Park JY, Clark P, Londin E, Sponziello M, Kricka LJ, Fortina P: Clinical exome performance for reporting secondary genetic findings. *Clin Chem* 61: 213–220, 2015
44. Ali H, Al-Mulla F, Hussain N, Naim M, Asbeutah AM, AlSahow A, Abu-Farha M, Abubaker J, Al Madhoun A, Ahmad S, Harris PC: PKD1 duplicated regions limit clinical utility of whole exome sequencing for genetic diagnosis of autosomal dominant polycystic kidney disease. *Sci Rep* 9: 4141, 2019
45. McCombie WR, Knowles JA, Mardis ER, McPherson JD: *Next-generation Sequencing in Medicine*, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 2019

46. Zhou B, Ho SS, Zhang X, Pattni R, Haraksingh RR, Urban AE: Whole-genome sequencing analysis of CNV using low-coverage and paired-end strategies is efficient and outperforms array-based CNV analysis. *J Med Genet* 55: 735–743, 2018
47. King K, Flinter FA, Nihalani V, Green PM: Unusual deep intronic mutations in the COL4A5 gene cause X linked Alport syndrome. *Hum Genet* 111: 548–554, 2002
48. LoY-F, Nozu K, Iijima K, Morishita T, HuangC-C, YangS-S, SytwuH-K, FangY-W, TsengM-H, LinS-H: Recurrent deep intronic mutations in the SLC12A3 gene responsible for Gitelman's syndrome. *Clin J Am Soc Nephrol* 6: 630–639, 2011
49. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BBA, Kleefstra T, Brunner HG, Vissers LELM, Veltman JA: Genome sequencing identifies major causes of severe intellectual disability. *Nature* 511: 344–347, 2014
50. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, Kingsmore SF: Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med* 3: 16, 2018
51. Lionel AC, Costain G, Monfared N, Walker S, Reuter MS, Hosseini SM, Thiruvahindrapuram B, Merico D, Jobling R, Nalpathamkalam T, Pellecchia G, Sung WWL, Wang Z, Bikangaga P, Boelman C, Carter MT, Cordeiro D, Cytrynbaum C, Dell SD, Dhir P, Dowling JJ, Heon E, Hewson S, Hiraki L, Inbar-Feigenberg M, Klatt R, Kronick J, Laxer RM, Licht C, MacDonald H, Mercimek-Andrews S, Mendoza-Londono R, Piscione T, Schneider R, Schulze A, Silverman E, Siriwardena K, Snead OC, Sondheimer N, Sutherland J, Vincent A, Wasserman JD, Weksberg R, Shuman C, Carew C, Szego MJ, Hayeems RZ, Basran R, Stavropoulos DJ, Ray PN, Bowdin S, Meyn MS, Cohn RD, Scherer SW, Marshall CR: Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med* 20: 435–443, 2018
52. Bergmann C: ARPKD and early manifestations of ADPKD: The original polycystic kidney disease and phenocopies. *Pediatr Nephrol* 30: 15–30, 2015
53. Porath B, Gainullin VG, Cornec-Le Gall E, Dillinger EK, Heyer CM, Hopp K, Edwards ME, Madsen CD, Mauritz SR, Banks CJ, Baheti S, Reddy B, Herrero JI, Bañales JM, Hogan MC, Tasic V, Watnick TJ, Chapman AB, Vigneau C, Lavainne F, AudrézetM-P, Ferec C, Le Meur Y, Torres VE, Harris PC; Genkyst Study Group, HALT Progression of Polycystic Kidney Disease Group; Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease: Mutations in GANAB, encoding the glucosidase IIα subunit, cause autosomal-dominant polycystic kidney and liver disease. *Am J Hum Genet* 98: 1193–1207, 2016
54. Verhave JC, Bech AP, Wetzels JFM, Nijenhuis T: Hepatocyte nuclear factor 1β-associated kidney disease: More than renal cysts and diabetes. *J Am Soc Nephrol* 27: 345–353, 2016
55. Braun DA, Schueler M, Halbritter J, Gee HY, Porath JD, Lawson JA, Airik R, Shril S, Allen SJ, Stein D, Al Kindy A, Beck BB, Cengiz N, Moorani KN, Ozaltin F, Hashmi S, Sayer JA, Bockenhauer D, Soliman NA, Otto EA, Lifton RP, Hildebrandt F: Whole exome sequencing identifies causative mutations in the majority of consanguineous or familial cases with childhood-onset increased renal echogenicity. *Kidney Int* 89: 468–475, 2016
56. Verbitsky M, Sanna-Cherchi S, Fasel DA, Levy B, Kiryluk K, Wuttke M, Abraham AG, Kaskel F, Köttgen A, Warady BA, Furth SL, Wong CS, Gharavi AG: Genomic imbalances in pediatric patients with chronic kidney disease. *J Clin Invest* 125: 2171–2178, 2015
57. Ma R, Deng L, Xia Y, Wei X, Cao Y, Guo R, Zhang R, Guo J, Liang D, Wu L: A clear bias in parental origin of de novo pathogenic CNVs related to intellectual disability, developmental delay and multiple congenital anomalies. *Sci Rep* 7: 44446, 2017
58. Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH: Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86: 749–764, 2010
59. Li S, Han X, Wang Y, Chen S, Niu J, Qian Z, Li P, Jin L, Xu C: Chromosomal microarray analysis in fetuses with congenital anomalies of the kidney and urinary tract: A prospective cohort study and meta-analysis. *Prenat Diagn* 39: 165–174, 2019
60. Rasmussen M, Sunde L, Nielsen ML, Ramsing M, Petersen A, Hjortshøj TD, Olsen TE, Tabor A, Hertz JM, Johnsen I, Sperling L, Petersen OB, Jensen UB, Møller FG, Petersen MB, Lildballe DL: Targeted gene sequencing and whole-exome sequencing in autopsied fetuses with prenatally diagnosed kidney anomalies. *Clin Genet* 93: 860–869, 2018
61. Lovric S, Ashraf S, Tan W, Hildebrandt F: Genetic testing in steroid-resistant nephrotic syndrome: When and how? *Nephrol Dial Transplant* 31: 1802–1813, 2016
62. Rosenberg AZ, Kopp JB: Focal segmental glomerulosclerosis. *Clin J Am Soc Nephrol* 12: 502–517, 2017
63. Hogg R, Middleton J, Vehaskari VM: Focal segmental glomerulosclerosis--epidemiology aspects in children and adults. *Pediatr Nephrol* 22: 183–186, 2007
64. Warejko JK, Tan W, Daga A, Schapiro D, Lawson JA, Shril S, Lovric S, Ashraf S, Rao J, Hermle T, Jobst-Schwan T, Widmeier E, Majmundar AJ, Schneider R, Gee HY, Schmidt JM, Vivante A, van der Ven AT, Ityel H, Chen J, Sadowski CE, Kohl S, Pabst WL, Nakayama M, Somers MJG, Rodig NM, Daouk G, Baum M, Stein DR, Ferguson MA, Traum AZ, Soliman NA, Kari JA, El Desoky S, Fathy H, Zenker M, Bakkaloglu SA, Müller D, Noyan A, Ozaltin F, Cadnapaphornchai MA, Hashmi S, Hopcian J, Kopp JB, Benador N, Bockenhauer D, Bogdanovic R, Stajić N, Chernin G, Ettenger R, Fehrenbach H, Kemper M, Munarriz RL, Podracka L, Büscher R, Serdaroglu E, Tasic V, Mane S, Lifton RP, Braun DA, Hildebrandt F: Whole exome sequencing of patients with steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol* 13: 53–62, 2018
65. Sadowski CE, Lovric S, Ashraf S, Pabst WL, Gee HY, Kohl S, Engelmann S, Vega-Warner V, Fang H, Halbritter J, Somers MJ, Tan W, Shril S, Fessi I, Lifton RP, Bockenhauer D, El-Desoky S, Kari JA, Zenker M, Kemper MJ, Mueller D, Fathy HM, Soliman NA, Hildebrandt F; SRNS Study Group: A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. *J Am Soc Nephrol* 26: 1279–1289, 2015
66. Bierzynska A, McCarthy HJ, Soderquest K, Sen ES, Colby E, Ding WY, Nabhan MM, Kerecuk L, Hegde S, Hughes D, Marks S, Feather S, Jones C, Webb NJA, Ognjanovic M, Christian M, Gilbert RD, Sinha MD, Lord GM, Simpson M, Koziell AB, Welsh GI, Saleem MA: Genomic and clinical profiling of a national nephrotic syndrome cohort advocates a precision medicine approach to disease management. *Kidney Int* 91: 937–947, 2017
67. Gribouval O, Boyer O, Hummel A, Dantal J, Martinez F, Sberro-Soussan R, Etienne I, Chauveau D, Delahousse M, Lionet A, Allard J, Pouteil Noble C, TêteM-J, Heidet L, Antignac C, Servais A: Identification of genetic causes for sporadic steroid-resistant nephrotic syndrome in adults. *Kidney Int* 94: 1013–1022, 2018
68. Trautmann A, Lipska-Ziętkiewicz BS, Schaefer F: Exploring the clinical and genetic spectrum of steroid resistant nephrotic syndrome: The PodoNet registry. *Front Pediatr* 6: 200, 2018
69. Mason AE, Sen ES, Bierzynska A, Colby E, Afzal M, Dorval G, Koziell AB, Williams M, Boyer O, Welsh GI, Saleem MA: UK RaDaR/NephroS Study: Response to first course of intensified immunosuppression in genetically stratified steroid resistant nephrotic syndrome [published online ahead of print April 21, 2020]. *Clin J Am Soc Nephrol* doi: 10.2215/CJN.13371019
70. Sampson MG, Gillies CE, Robertson CC, Crawford B, Vega-Warner V, Otto EA, Kretzler M, Kang HM: Using population genetics to interrogate the monogenic nephrotic syndrome diagnosis in a case cohort. *J Am Soc Nephrol* 27: 1970–1983, 2016
71. Hinkes B, Wiggins RC, Gbadegesin R, Vlangos CN, Seelow D, Nürnberg G, Garg P, Verma R, Chaib H, Hoskins BE, Ashraf S, Becker C, Hennies HC, Goyal M, Wharram BL, Schachter AD, Mudumana S, Drummond I, Kerjaschki D, Waldherr R, Dietrich A, Ozaltin F, Bakkaloglu A, Cleper R, Basel-Vanagaite L, Pohl M, Griebel M, Tsygin AN, Soylu A, Müller D, Sorli CS, Bunney TD, Katan M, Liu J, Attanasio M, O'toole JF, Hasselbacher K, Mucha

B, Otto EA, Airik R, Kispert A, Kelley GG, Smrcka AV, Gudermann T, Holzman LB, Nürnberg P, Hildebrandt F: Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. *Nat Genet* 38: 1397–1405, 2006
72. Ashraf S, Gee HY, Woerner S, Xie LX, Vega-Warner V, Lovric S, Fang H, Song X, Cattran DC, Avila-Casado C, Paterson AD, Nitschké P, Bole-Feysot C, Cochat P, Esteve-Rudd J, Haberberger B, Allen SJ, Zhou W, Airik R, Otto EA, Barua M, Al-Hamed MH, Kari JA, Evans J, Bierzynska A, Saleem MA, Böckenhauer D, Kleta R, El Desoky S, Hacihamdioglu DO, Gok F, Washburn J, Wiggins RC, Choi M, Lifton RP, Levy S, Han Z, Salviati L, Prokisch H, Williams DS, Pollak M, Clarke CF, Pei Y, Antignac C, Hildebrandt F: ADCK4 mutations promote steroid-resistant nephrotic syndrome through CoQ10 biosynthesis disruption. *J Clin Invest* 123: 5179–5189, 2013
73. Diomedi-Camassei F, Di Giandomenico S, Santorelli FM, Caridi G, Piemonte F, Montini G, Ghiggeri GM, Murer L, Barisoni L, Pastore A, Muda AO, Valente ML, Bertini E, Emma F: COQ2 nephropathy: A newly described inherited mitochondriopathy with primary renal involvement. *J Am Soc Nephrol* 18: 2773–2780, 2007
74. Ashton EJ, Legrand A, Benoit V, Roncelin I, Venisse A, ZennaroM-C, Jeunemaitre X, Iancu D, Van't Hoff WG, Walsh SB, Godefroid N, Rotthier A, Del Favero J, Devuyst O, Schaefer F, Jenkins LA, Kleta R, Dahan K, Vargas-Poussou R, Bockenhauer D: Simultaneous sequencing of 37 genes identified causative mutations in the majority of children with renal tubulopathies. *Kidney Int* 93: 961–967, 2018
75. Mallett AJ, McCarthy HJ, Ho G, Holman K, Farnsworth E, Patel C, Fletcher JT, Mallawaarachchi A, Quinlan C, Bennetts B, Alexander SI: Massively parallel sequencing and targeted exomes in familial kidney disease can diagnose underlying genetic disorders. *Kidney Int* 92: 1493–1506, 2017
76. EckardtK-U, Alper SL, Antignac C, Bleyer AJ, Chauveau D, Dahan K, Deltas C, Hosking A, Kmoch S, Rampoldi L, Wiesener M, Wolf MT, Devuyst O; Kidney Disease: Improving Global Outcomes: Autosomal dominant tubulointerstitial kidney disease: Diagnosis, classification, and management--A KDIGO consensus report. *Kidney Int* 88: 676–683, 2015
77. Kirby A, Gnirke A, Jaffe DB, Barešová V, Pochet N, Blumenstiel B, Ye C, Aird D, Stevens C, Robinson JT, Cabili MN, Gat-Viks I, Kelliher E, Daza R, DeFelice M, Hůlková H, Sovová J, Vylet'al P, Antignac C, Guttman M, Handsaker RE, Perrin D, Steelman S, Sigurdsson S, Scheinman SJ, Sougnez C, Cibulskis K, Parkin M, Green T, Rossin E, Zody MC, Xavier RJ, Pollak MR, Alper SL, Lindblad-Toh K, Gabriel S, Hart PS, Regev A, Nusbaum C, Kmoch S, Bleyer AJ, Lander ES, Daly MJ: Mutations causing medullary cystic kidney disease type 1 lie in a large VNTR in MUC1 missed by massively parallel sequencing. *Nat Genet* 45: 299–303, 2013
78. Adam J, Browning AC, Vaideanu D, Heidet L, Goodship JA, Sayer JA: A wide spectrum of phenotypes in a family with renal coloboma syndrome caused by a PAX2 mutation. *Clin Kidney J* 6: 410–413, 2013
79. Bastarache L, Hughey JJ, Hebbring S, Marlo J, Zhao W, Ho WT, Van Driest SL, McGregor TL, Mosley JD, Wells QS, Temple M, Ramirez AH, Carroll R, Osterman T, Edwards T, Ruderfer D, Velez Edwards DR, Hamid R, Cogan J, Glazer A, WeiW-Q, Feng Q, Brilliant M, Zhao ZJ, Cox NJ, Roden DM, Denny JC: Phenotype risk scores identify patients with unrecognized Mendelian disease patterns. *Science* 359: 1233–1239, 2018
80. Connaughton DM, Hildebrandt F: Personalized medicine in chronic kidney disease by detection of monogenic mutations. *Nephrol Dial Transplant* 35: 390–397, 2020
81. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee: Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the association for molecular pathology. *Genet Med* 17: 405–424, 2015
82. Shah N, HouY-CC, YuH-C, Sainger R, Caskey CT, Venter JC, Telenti A: Identification of misclassified ClinVar variants via disease population prevalence. *Am J Hum Genet* 102: 609–619, 2018
83. Mighton C, Charames GS, Wang M, ZakoorK-R, Wong A, Shickh S, Watkins N, Lebo MS, Bombard Y, Lerner-Ellis J: Variant classification changes over time in BRCA1 and BRCA2. *Genet Med* 21: 2248–2254, 2019
84. SoRelle JA, Thodeson DM, Arnold S, Gotway G, Park JY: Clinical utility of reinterpreting previously reported genomic epilepsy test results for pediatric patients. *JAMA Pediatr* 173: e182302, 2019
85. Macklin S, Durand N, Atwal P, Hines S: Observed frequency and challenges of variant reclassification in a hereditary cancer clinic. *Genet Med* 20: 346–350, 2018
86. Deignan JL, Chung WK, Kearney HM, Monaghan KG, Rehder CW, Chao EC; ACMG Laboratory Quality Assurance Committee: Points to consider in the reevaluation and reanalysis of genomic test results: A statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 21: 1267–1270, 2019
87. Parsa A, Fuchsberger C, Köttgen A, O'Seaghdha CM, Pattaro C, de Andrade M, Chasman DI, Teumer A, Endlich K, Olden M, ChenM-H, Tin A, Kim YJ, Taliun D, Li M, Feitosa M, Gorski M, Yang Q, Hundertmark C, Foster MC, Glazer N, Isaacs A, Rao M, Smith AV, O'Connell JR, Struchalin M, Tanaka T, Li G, HwangS-J, Atkinson EJ, Lohman K, Cornelis MC, Johansson A, Tönjes A, Dehghan A, Couraki V, Holliday EG, Sorice R, Kutalik Z, Lehtimäki T, Esko T, Deshmukh H, Ulivi S, Chu AY, Murgia F, Trompet S, Imboden M, Kollerits B, Pistis G, Harris TB, Launer LJ, Aspelund T, Eiriksdottir G, Mitchell BD, Boerwinkle E, Schmidt H, Hofer E, Hu F, Demirkan A, Oostra BA, Turner ST, Ding J, Andrews JS, Freedman BI, Giulianini F, Koenig W, Illig T, Döring A, WichmannH-E, Zgaga L, Zemunik T, Boban M, Minelli C, Wheeler HE, Igl W, Zaboli G, Wild SH, Wright AF, Campbell H, Ellinghaus D, Nöthlings U, Jacobs G, Biffar R, Ernst F, Homuth G, Kroemer HK, Nauck M, Stracke S, Völker U, Völzke H, Kovacs P, Stumvoll M, Mägi R, Hofman A, Uitterlinden AG, Rivadeneira F, Aulchenko YS, Polasek O, Hastie N, Vitart V, Helmer C, Wang JJ, Stengel B, Ruggiero D, Bergmann S, Kähönen M, Viikari J, Nikopensius T, Province M, Colhoun H, Doney A, Robino A, Krämer BK, Portas L, Ford I, Buckley BM, Adam M, ThunG-A, Paulweber B, Haun M, Sala C, Mitchell P, Ciullo M, Vollenweider P, Raitakari O, Metspalu A, Palmer C, Gasparini P, Pirastu M, Jukema JW, Probst-Hensch NM, Kronenberg F, Toniolo D, Gudnason V, Shuldiner AR, Coresh J, Schmidt R, Ferrucci L, van Duijn CM, Borecki I, Kardia SLR, Liu Y, Curhan GC, Rudan I, Gyllensten U, Wilson JF, Franke A, Pramstaller PP, Rettig R, Prokopenko I, Witteman J, Hayward C, Ridker PM, Bochud M, Heid IM, Siscovick DS, Fox CS, Kao WL, Böger CA: Common variants in Mendelian kidney disease genes and their association with renal function. *J Am Soc Nephrol* 24: 2105–2117, 2013
88. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS; ClinGen: ClinGen--the clinical genome resource. *N Engl J Med* 372: 2235–2242, 2015
89. Brown EJ, Schlöndorff JS, Becker DJ, Tsukaguchi H, Tonna SJ, Uscinski AL, Higgs HN, Henderson JM, Pollak MR: Mutations in the formin gene INF2 cause focal segmental glomerulosclerosis [published correction appears in *Nat Genet* 42: 361, 2010]. *Nat Genet* 42: 72–76, 2010
90. Devuyst O, Olinger E, Rampoldi L: Uromodulin: From physiology to rare and complex kidney disorders. *Nat Rev Nephrol* 13: 525–544, 2017
91. Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, Margulies DM, Loscalzo J, Kohane IS: Genetic misdiagnoses and the potential for health disparities. *N Engl J Med* 375: 655–665, 2016
92. Tory K, Menyhárd DK, Woerner S, Nevo F, Gribouval O, Kerti A, Stráner P, Arrondel C, Huynh Cong E, Tulassay T, Mollet G, Perczel A, Antignac C: Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. *Nat Genet* 46: 299–304, 2014
93. Whiffin N, Roberts AM, Minikel E, Zappala Z, Walsh R, O'Donnell-Luria AH, Karczewski KJ, Harrison SM, Thomson KL,

Sage H, Ing AY, Barton PJR, Funke B, Cook SA, MacArthur DG, Ware JS: Using high-resolution variant frequencies empowers clinical genome interpretation and enables investigation of genetic architecture. *Am J Hum Genet* 104: 187–190, 2019
94. Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, Janizek JD, Huang X, Starita LM, Shendure J: Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 562: 217–222, 2018
95. Posey JE, O'Donnell-Luria AH, Chong JX, Harel T, Jhangiani SN, Coban Akdemir ZH, Buyske S, Pehlivan D, Carvalho CMB, Baxter S, Sobreira N, Liu P, Wu N, Rosenfeld JA, Kumar S, Avramopoulos D, White JJ, Doheny KF, Witmer PD, Boehm C, Sutton VR, Muzny DM, Boerwinkle E, Günel M, Nickerson DA, Mane S, MacArthur DG, Gibbs RA, Hamosh A, Lifton RP, Matise TC, Rehm HL, Gerstein M, Bamshad MJ, Valle D, Lupski JR; Centers for Mendelian Genomics: Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet Med* 21: 798–812, 2019
96. Liu P, Meng L, Normand EA, Xia F, Song X, Ghazi A, Rosenfeld J, Magoulas PL, Braxton A, Ward P, Dai H, Yuan B, Bi W, Xiao R, Wang X, Chiang T, Vetrini F, He W, Cheng H, Dong J, Gijavanekar C, Benke PJ, Bernstein JA, Eble T, Eroglu Y, Erwin D, Escobar L, Gibson JB, Gripp K, Kleppe S, Koenig MK, Lewis AM, Natowicz M, Mancias P, Minor L, Scaglia F, Schaaf CP, Streff H, Vernon H, Uhles CL, Zackai EH, Wu N, Sutton VR, Beaudet AL, Muzny D, Gibbs RA, Posey JE, Lalani S, Shaw C, Eng CM, Lupski JR, Yang Y: Reanalysis of clinical exome sequencing data. *N Engl J Med* 380: 2478–2480, 2019
97. Gahl WA, Mulvihill JJ, Toro C, Markello TC, Wise AL, Ramoni RB, Adams DR, Tifft CJ; UDN: The NIH undiagnosed diseases program and Network: Applications to modern medicine. *Mol Genet Metab* 117: 393–400, 2016
98. Jalali Sefid Dashti M, Gamieldien J: A practical guide to filtering and prioritizing genetic variants. *Biotechniques* 62: 18–30, 2017
99. Tan R, Colville D, Wang YY, Rigby L, Savige J: Alport retinopathy results from "severe" COL4A5 mutations and predicts early renal failure. *Clin J Am Soc Nephrol* 5: 34–38, 2010
100. Savige J, Ariani F, Mari F, Bruttini M, Renieri A, Gross O, Deltas C, Flinter F, Ding J, Gale DP, Nagel M, Yau M, Shagam L, Torra R, Ars E, Hoefele J, Garosi G, Storey H: Expert consensus guidelines for the genetic diagnosis of Alport syndrome. *Pediatr Nephrol* 34: 1175–1189, 2019
101. Savige J, Gregory M, Gross O, Kashtan C, Ding J, Flinter F: Expert guidelines for the management of Alport syndrome and thin basement membrane nephropathy. *J Am Soc Nephrol* 24: 364–375, 2013
102. Jarvik GP, Amendola LM, Berg JS, Brothers K, Clayton EW, Chung W, Evans BJ, Evans JP, Fullerton SM, Gallego CJ, Garrison NA, Gray SW, Holm IA, Kullo IJ, Lehmann LS, McCarty C, Prows CA, Rehm HL, Sharp RR, Salama J, Sanderson S, Van Driest SL, Williams MS, Wolf SM, Wolf WA, Burke W; eMERGE Act-ROR Committee and CERC Committee; CSER Act-ROR Working Group: Return of genomic results to research participants: The floor, the ceiling, and the choices in between. *Am J Hum Genet* 94: 818–826, 2014
103. Bombard Y, Brothers KB, Fitzgerald-Butt S, Garrison NA, Jamal L, James CA, Jarvik GP, McCormick JB, Nelson TN, Ormond KE, Rehm HL, Richer J, Souzeau E, Vassy JL, Wagner JK, Levy HP: The responsibility to recontact research participants after reinterpretation of genetic and genomic research results. *Am J Hum Genet* 104: 578–595, 2019
104. Zawati MH, Parry D, Thorogood A, Nguyen MT, Boycott KM, Rosenblatt D, Knoppers BM: Reporting results from whole-genome and whole-exome sequencing in clinical practice: A proposal for Canada? *J Med Genet* 51: 68–70, 2014
105. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, McKelvey KD, Ormond KE, Richards CS, Vlangos CN, Watson M, Martin CL, Miller DT: Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): A policy statement of the American College of Medical Genetics and Genomics [published correction appears in *Genet Med* 19: 484, 2017]. *Genet Med* 19: 249–255, 2017
106. Liebow A, Li X, Racie T, Hettinger J, Bettencourt BR, Najafian N, Haslett P, Fitzgerald K, Holmes RP, Erbe D, Querbes W, Knight J: An investigational RNAi therapeutic targeting glycolate oxidase reduces oxalate production in models of primary Hyperoxaluria. *J Am Soc Nephrol* 28: 494–503, 2017
107. O'Daniel JM, McLaughlin HM, Amendola LM, Bale SJ, Berg JS, Bick D, Bowling KM, Chao EC, Chung WK, Conlin LK, Cooper GM, Das S, Deignan JL, Dorschner MO, Evans JP, Ghazani AA, Goddard KA, Gornick M, Farwell Hagman KD, Hambuch T, Hegde M, Hindorff LA, Holm IA, Jarvik GP, Knight Johnson A, Mighion L, Morra M, Plon SE, Punj S, Richards CS, Santani A, Shirts BH, Spinner NB, Tang S, Weck KE, Wolf SM, Yang Y, Rehm HL: A survey of current practices for genomic sequencing test interpretation and reporting processes in US laboratories. *Genet Med* 19: 575–582, 2017
108. Morgan P, Brown DG, Lennard S, Anderton MJ, Barrett JC, Eriksson U, Fidock M, Hamrén B, Johnson A, March RE, Matcham J, Mettetal J, Nicholls DJ, Platz S, Rees S, Snowden MA, Pangalos MN: Impact of a five-dimensional framework on R&D productivity at AstraZeneca. *Nat Rev Drug Discov* 17: 167–181, 2018
109. Dugger SA, Platt A, Goldstein DB: Drug development in the era of precision medicine. *Nat Rev Drug Discov* 17: 183–196, 2018
110. Abifadel M, Varret M, RabèsJ-P, Allard D, Ouguerram K, Devillers M, Cruaud C, Benjannet S, Wickham L, Erlich D, Derré A, Villéger L, Farnier M, Beucler I, Bruckert E, Chambaz J, Chanu B, LecerfJ-M, Luc G, Moulin P, Weissenbach J, Prat A, Krempf M, Junien C, Seidah NG, Boileau C: Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 34: 154–156, 2003
111. Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, Edelberg JM, Goodman SG, Hanotin C, Harrington RA, Jukema JW, Lecorps G, Mahaffey KW, Moryusef A, Pordy R, Quintero K, Roe MT, Sasiela WJ, TambyJ-F, Tricoci P, White HD, Zeiher AM; ODYSSEY OUTCOMES Committees and Investigators: Alirocumab and cardiovascular outcomes after acute coronary syndrome. *N Engl J Med* 379: 2097–2107, 2018
112. Germain DP, Nicholls K, Giugliani R, Bichet DG, Hughes DA, Barisoni LM, Colvin RB, Jennette JC, Skuban N, Castelli JP, Benjamin E, Barth JA, Viereck C: Efficacy of the pharmacologic chaperone migalastat in a subset of male patients with the classic phenotype of Fabry disease and migalastat-amenable variants: Data from the phase 3 randomized, multicenter, double-blind clinical trial and extension study. *Genet Med* 21: 1987–1997, 2019
113. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350: 2129–2139, 2004
114. Nowicki TS, Hu-Lieskovan S, Ribas A: Mechanisms of resistance to PD-1 and PD-L1 blockade. *Cancer J* 24: 47–53, 2018
115. Goossens N, Nakagawa S, Sun X, Hoshida Y: Cancer biomarker discovery and validation. *Transl Cancer Res* 4: 256–269, 2015
116. Abrams J, Conley B, Mooney M, Zwiebel J, Chen A, Welch JJ, Takebe N, Malik S, McShane L, Korn E, Williams M, Staudt L, Doroshow J: National Cancer Institute's precision medicine initiatives for the new National Clinical Trials Network. *Am Soc Clin Oncol Educ Book* 34: 71–76, 2014
117. Madian AG, Wheeler HE, Jones RB, Dolan ME: Relating human genetic variation to variation in drug responses. *Trends Genet* 28: 487–495, 2012
118. GoodSmith MS, Skandari MR, Huang ES, Naylor RN: The impact of biomarker screening and cascade genetic testing on the cost-effectiveness of MODY genetic testing. *Diabetes Care* 42: 2247–2255, 2019
119. Sun L, Brentnall A, Patel S, Buist DSM, Bowles EJA, Evans DGR, Eccles D, Hopper J, Li S, Southey M, Duffy S, Cuzick J, Dos Santos Silva I, Miners A, Sadique Z, Yang L, Legood R, Manchanda R: A cost-effectiveness analysis of multigene testing for all patients with breast cancer [published online ahead of print Oct 3, 2019]. *JAMA Oncol* doi:10.1001/jamaoncol.2019.3323
120. Petelin L, Hossack L, Shanahan M, Mitchell G, Liew D, James PA, Trainer AH: Cost-effectiveness of long-term clinical management of BRCA pathogenic variant carriers. *Genet Med* 22: 831–839, 2020

Published online ahead of print. Publication date available at www.cjasn.org.

Nephrol Dial Transplant (2021) 36: 1389–1398
doi: 10.1093/ndt/gfaa092
Advance Access publication 24 June 2020

ORIGINAL ARTICLE

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

The switch from proteasome to immunoproteasome is increased in circulating cells of patients with fast progressive immunoglobulin A nephropathy and associated with defective CD46 expression

Licia Peruzzi[1,2], Rosanna Coppo[1], Enrico Cocchi[2], Elisa Loiacono[1], Massimilano Bergallo[3], Monica Bodria[4], Luca Vergano[2], Alexandra Krutova[5], Maria Luisa Russo[1], Alessandro Amore[2], Sigrid Lundberg[6], Dita Maixerova[7], Vladimir Tesar[7], Agnieszka Perkowska-Ptasińska[8], Magdalena Durlik[8], Dimitris Goumenos[9], Marios Papasotiriou[9], Kresimir Galesic[10], Luka Toric[10], Aikaterini Papagianni[11], Maria Stangou[11], Malgorzata Mizerska-Wasiak[12], Loreto Gesualdo[13], Eustacchio Montemurno[13], Luisa Benozzi[14], Stefano Cusinato[14], Tomasz Hryszko[15], Marian Klinger[16], Dorota Kamińska[17] and Magdalena Krajewska[17]; on behalf of the VALIGA study group of the ERA-EDTA Immunonephrology Working Group

[1]Fondazione Ricerca Molinette, Regina Margherita Hospital, Turin, Italy, [2]Department of Nephrology, Dialysis and Transplantation, Regina Margherita Hospital, Turin, Italy, [3]Department of Nephrology, Dialysis and Transplantation, Public Health and Pediatric Sciences, University of Turin, Turin, Italy, [4]IRCCS Istituto Giannina Gaslini, Genova, Italy, [5]Regional Children's Clinical Hospital, Vladivostok, Russia, [6]Department of Clinical Sciences, Nephrology, Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden, [7]Department of Nephrology, First Faculty of Medicine and General University Hospital, Prague, Czech Republic, [8]Department of Transplantation Medicine and Nephrology, Warsaw Medical University, Warsaw, Poland, [9]Department of Nephrology, University Hospital of Patras, Patras, Greece, [10]Department of Nephrology, Dubrava University Hospital, Zagreb, Croatia, [11]Department of Nephrology, Hippokration General Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece, [12]Department of Pediatrics and Nephrology, Medical University of Warsaw, Warsaw, Poland, [13]Department of Nephrology, Emergency and Transplantation, University of Bari, Bari, Italy, [14]Department of Nephrology, Borgomanero Hospital, Borgomanero, Italy, [15]Department of Nephrology, Transplantation and Dialysis, Medical University of Bialystok, Bialystok, Poland, [16]Department of Internal Medicine, Opole University, Poland and [17]Department of Nephrology and Transplantation Medicine, Wroclaw Medical University, Wroclaw, Poland

Correspondence to: Rosanna Coppo; E-mail: rosanna.coppo@unito.it; Twitter handle: @ERAEDTA

ABSTRACT

The proteasome to immunoproteasome (iPS) switch consists of β1, β2 and β5 subunit replacement by low molecular weight protein 2 (LMP2), LMP7 and multicatalytic endopeptidase-like complex-1 (MECL1) subunits, resulting in a more efficient peptide preparation for major histocompatibility complex 1 (MHC-I) presentation. It is activated by toll-like receptor (TLR) agonists and interferons and may also be influenced by genetic variation. In a previous study we found an iPS upregulation in peripheral cells of patients with immunoglobulin A nephropathy (IgAN). We aimed to investigate in 157 IgAN patients enrolled through the multinational Validation Study of the Oxford Classification of IgAN (VALIGA) study the relationships between iPS switch and estimated glomerular filtration rate (eGFR) modifications from renal biopsy to sampling. Patients had a previous long follow-up (6.4 years in median) that allowed an accurate calculation of their slope of renal function decline. We also evaluated the effects of the *PSMB8/PSMB9* locus (rs9357155) associated with IgAN in genome-wide association studies and the expression of messenger RNAs (mRNAs) encoding for TLRs and CD46, a C3 convertase inhibitor, acting also on T-regulatory cell promotion, found to have reduced expression in progressive IgAN. We detected an upregulation of LMP7/β5 and LMP2/β1 switches. We observed no genetic effect of rs9357155. TLR4 and TLR2 mRNAs were found to be significantly associated with iPS switches, particularly TLR4 and LMP7/β5 ($P < 0.0001$). The LMP7/β5 switch was

© The Author(s) 2020. Published by Oxford University Press on behalf of ERA-EDTA. All rights reserved.

KEY LEARNING POINTS

What is already known about this subject?

- The mechanisms underlying the variable course of immunoglobulin A nephropathy (IgAN) remain unknown. New markers for disease progression are needed to improve risk stratification.
- Innate and adaptive immunity is involved in the development and progression of IgAN. The modification of the proteasome (PS) in immunoproteasome (iPS) is a key factor for optimal lymphocyte activation. We previously reported an upregulation of iPS in patients with IgAN, but the role of these changes in the progression of IgAN has never been investigated.
- The complement regulator CD46 (or membrane cofactor protein) modulates at a cellular level the C3 convertase activity of the lectin and alternative complement pathways. Both these pathways are involved in the pathogenesis and progression of IgAN. CD46 is not only involved in the regulation of complement activation, but is also expressed on T-cell membrane, where it promotes a switch in T helper type 1 to type 1 regulatory T cell differentiation. We recently reported a low expression of CD46 mRNA in peripheral white blood cells (WBCs) of patients with progressive IgAN.

What this study adds?

- We found that patients with IgAN presenting with rapid estimated glomerular filtration rate loss were characterized by a combination of high PS to iPS LMP7/β5 switch and low CD46 expression in WBCs. This is the first report, to our knowledge, of the interrelation ship between these two actors in innate and adaptive immunity control in patients with IgAN.

What impact this may have on practice or policy?

- A multivariate logistic regression model categorizing patients by different levels of kidney disease progression showed a high prediction value for the combination of high LMP7/β5 and low CD46 expression in peripheral cells. This may represent a biomarker for identifying patients at risk of progression to be tested in larger cohorts in prospective studies.

significantly associated with the rate of eGFR loss ($P = 0.026$), but not with eGFR at biopsy. Fast progressors (defined as the loss of eGFR >75th centile, i.e. -1.91 mL/min/1.73 m^2/year) were characterized by significantly elevated LMP7/β5 mRNA ($P = 0.04$) and low CD46 mRNA expression ($P < 0.01$). A multivariate logistic regression model, categorizing patients by different levels of kidney disease progression, showed a high prediction value for the combination of high LMP7/β5 and low CD46 expression.

Keywords: biomarkers, CD46, complement, IgA nephropathy, immune proteasome, progression, risk factors

ADDITIONAL CONTENT

An author video to accompany this article is available at: https://academic.oup.com/ndt/pages/author_videos.

INTRODUCTION

Immunoglobulin A nephropathy (IgAN) is a disease with variable course, and patients who progress to end-stage renal disease (ESRD) have an extremely variable rate of glomerular filtration rate (GFR) decline. Over the last decade, the pathogenic events leading to the development of IgAN have been elucidated and several factors have been identified [1, 2]. The production of excessive amounts of poorly galactosylated IgA1 molecules (Gd-IgA1) along with the synthesis of IgG autoantibodies, leads to immune complex formation and deposition in the mesangium, with local complement activation and subsequent glomerular damage [3, 4]. However, the mechanisms underlying different kidney disease progression rates remain unknown. Hence there is great interest in new markers for disease progression, with the hope of improving risk stratification [5].

The production of Gd-IgA1 is initiated in mucosal sites by surface receptors, such as toll-like receptors (TLRs), specifically recognizing pathogen-associated molecular patterns expressed by multiple pathogens [6–8]. The recognition of invading microbes by TLRs on dendritic cells (DCs) in the mucosae induces cytokine production and enhances antigen presentation to naive T cells, resulting in antigen-specific adaptive immune response activation. After ligand binding, TLRs trigger nuclear factor-κB (NF-κB) and interferon (IFN) regulatory factor transcription in DCs [9, 10]. IFN-γ, tumour necrosis factor (TNF)-α and TLR agonists then induce the replacement of the proteasome (PS) catalytic units β1, β2 and β5 (encoded by the genes *PSMB5*, *PSMB6* and *PSMB7*), with low molecular weight proteins (LMP2 and LMP7) and a multicatalytic endopeptidase-like complex (MECL-1) (encoded by the genes *PSMB8, PSMB9 and PSMB10*, respectively), transforming the PS into immunoproteasome (iPS) [10–12]. These modifications result in an enhanced catalytic property and a 'professionalization' of PS in presenting peptides to major histocompatibility complex class I, leading to optimal lymphocyte activation. We previously reported an upregulation of iPS in patients with IgAN [13], but the role of these changes in the progression of IgAN has never been investigated.

Among the factors playing a role in inflammation and renal damage after macromolecular IgA mesangial deposition, complement activation is of great interest given the increasing number of complement-targeting drugs [14]. Several studies have tried to elucidate the possible role of defective complement control in IgAN, with conflicting results, but unanimously concluding on a possible defective regulation of an alternative complement pathway factor (CFH-H) via CFH-related protein [15–17]. We focused our interest on the complement regulator Cluster of differentiation-46 (CD46) (or membrane cofactor protein), which modulates at cellular level the C3 convertase

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

Table 1. Demographic and clinical data at renal biopsy and at sampling in the cohort of 157 patients with IgAN

Study cohort	No. of patients: 157		Gender: females/males (%) 51/106 (32.5/67.5)		
Clinical data at renal biopsy	Age (years)	eGFR (mL/min/1.73 m^2)	Proteinuria (g/day/1.73 m^2)	MAP (mmHg)	Biopsy features (%)
	36.8 (23.1–49.4)	70.92 (48.48–98.72)	1.1 (0.4–2.04)	100 (86.7–106.7)	• M1 54.7 • E1 21.6 • S1 57.3 • T1–T2 29.3 • C1 12.7
Time from biopsy to sampling (years)	6.4 (2.8–10.7)				
Clinical data at sampling	Age (years)	eGFR (mL/min/1.73 m^2)	Proteinuria (g/day/1.73 m^2)	MAP (mmHg)	Therapy (%)
	44.4 (31.9–56.4)	73.3 (45.9–89.8)	0.4 (0.17–0.8)	93.3 (84.3–97.8)	• RASB 83.4 • Cs/Is 34.4 • (Cs/Is at s. 17.8)
Follow-up data	Duration of follow-up (years)		Time-averaged proteinuria (g/day/1.73 m^2)		
	6.4 (2.8–10.7)		0.74 (0.32–1.31)		
Clinical outcomes	Rate of eGFR loss (mL/min/1.73 m^2/year)		50% loss of eGFR from baseline		
	−0.41 (−1.91–0.87)		5.7%		

eGFR calculated by modified Schwartz or MDRD formula (see Methods section). MEST from 124 patients. M1, mesangial hypercellularity (>50 of glomeruli with mesangial hypercellularity); E1, presence of endocapillary hypercellularity; S1, presence of segmental glomerular sclerosis; T1–2, tubular atrophy/interstitial fibrosis in ≥25% of renal biopsy tissues; C1, presence of any crescents; Cs, corticosteroids; Is, Immunosuppressive drugs (exposure at any previous time); Cs/Is at s, exposure at sampling. Values are expressed as median (IQR) or percentage. Partially modified from Coppo *et al.* [19].

activity of the lectin and alternative complement pathways. We investigated the transcriptional expression in peripheral white blood cells (WBCs) of CD46 in 157 patients enrolled by the Validation Study of the Oxford Classification of IgAN (VALIGA) [18] and found a significantly lower expression of CD46 messenger RNA (mRNA) in patients with progressive IgAN [19]. CD46 is not only involved in the regulation of complement activation, but is also expressed on T-cell membrane, where it promotes a switch in T helper type 1 (Th1) to type 1 regulatory T cell (Tr1) differentiation [20, 21]. Hence CD46 emerges as a key sensor of immune activation and a core modulator of adaptive immunity [22].

We explored the correlations between iPS switch and estimated glomerualr filtration rate (eGFR) decline in 157 patients with IgAN enrolled in the international VALIGA study [18], where the renal biopsies from multiple centres were centrally scored and the long follow-up available allowed an accurate calculation of GFR slope before sampling. We also evaluated the single nucleotide polymorphism (SNP) rs9357155 at the *PSMB8*/*PSMB9* locus that encodes for LMP2/LMP7, previously found to be involved in the genetic susceptibility to IgAN [23, 24], and the expression of mRNAs encoding for TLRs and CD46.

We found that patients with IgAN presenting with rapid eGFR loss were characterized by a combination of high PS to iPS LMP7/β5 switch and low CD46 expression. This is the first report, to our knowledge, of an interrelationship between these two actors in innate and adaptive immunity control in patients with IgAN.

RESULTS

Categorization of progressive and fast progressive patients

The 157 subjects with IgAN from the VALIGA cohort had at the time of sampling a previous observation period of 6.4 years [interquartile range (IQR) 2.8–10.7] after renal biopsy, which allowed for a precise calculation of GRF slope (Table 1 reports data at renal biopsy and at sampling). Based on the median value of the eGFR slope of −0.41 mL/min/1.73 m^2/year, patients were categorized into 79 'progressors' with a median eGFR loss of −1.91 mL/min/1.73 m^2/year and 78 'non-progressors', who had a median improvement in eGFR of +0.89 (Supplementary data, Table S1). Based on the 75th centile of high rate of eGFR loss (−1.91 mL/min/1.73 m^2/year), patients were categorized into 40 'fast progressors', with a median eGFR decline of −3.8 (IQR −8.2 to −2.9) mL/min/1.73 m^2/year, and 117 'non-fast progressors', with a mild eGFR loss or some improvement during the follow-up [median −0.2 (IQR −0.8–2.2) mL/min/1.73 m^2/year] (Table 2 and Figure 1).

No difference in clinical data at renal biopsy or at sampling was found between progressors and non-progressors, with the exception of significantly higher T1–2 scores (Supplementary data, Table S1). A clinical feature of fast progressors (Table 2) was greater degree of proteinuria at renal biopsy and at sampling at the end of follow-up. The endpoint of 50% decline in eGFR or ESRD (reached in five patients with eGFR < 15 mL/min at sampling) was significantly more frequent in fast progressors; $P < 0.0001$. Treatment with corticosteroid/immunosuppressive (Cs/IS) drugs and renin–angiotensin blockers was adopted with similar frequency in the two groups (Table 2).

Switch from PS to iPS in circulating blood cells of patients with IgAN

In the 157 IgAN patients investigated, a significant switch from PS to iPS was detected for LMP7/β5 [median 1.66 (IQR 1.31–2.14) versus 1.16 (0.95–1.41) in healthy controls, $P < 0.001$] and for LMP2/β1 [median 1.22 (IQR 0.94–2.06) versus 0.65 (0.35–1.00), $P < 0.001$]. No significant modification was found for the third switch, i.e. MECL-1/β2. LMP7/β5 and LMP2/β1 were significantly correlated with proteinuria at sampling ($P = 0.03$ and $P = 0.004$, respectively), but no correlation was found with eGFR at sampling (Supplementary data, Table S2

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

Table 2. Clinical data in patients with IgAN with different velocity of progression, defined as fast progressors and non-fast progressors

	Non-fast progressors $n=117$ (74.5%)	Fast progressors $n=40$ (25.5%)	P-value
Clinical data at renal biopsy			
Gender (female), n (%)	42 (35.9)	9 (22.5)	0.17
Age (years)	36.3 (21.8–50.1)	37.2 (24.3–46.7)	0.99
eGFR (mL/min/1.73 m^2)	70.5 (50.9–92.9)	73.1 (38.9–101.9)	0.78
Proteinuria (g/day/1.73 m^2)	1.0 (0.4–1.9)	1.8 (0.7–2.5)	0.04
MAP (mmHg)	100 (87.2–107)	100 (87–103.3)	0.91
Biopsy features, %			
M1	52.6	62.1	0.40
E1	26.3	0.07	0.04
S1	61.1	44.8	0.14
T1–2	23.2	44.4	0.01
C1	11.6	17.2	0.53
Clinical data at sampling		–	
Age (years)	43.9 (31–56.4)	44.6 (34.5–55.3)	0.72
Duration of follow-up (years)	6.7 (3.2–11.2)	5.6 (2.2–11.3)	0.48
MAP (mmHg)	120.0 (108.3–123.3)	118.3 (107.9–129.3)	0.62
Proteinuria (g/day/1.73 m^2)	0.3 (0.2–0.8)	0.6 (0.2–1.7)	0.02
Time-averaged proteinuria (g/day/1.73 m^2)	0.7 (0.3–1.2)	0.9 (0.5–1.7)	0.11
RASB treatment, %	86.5	73.3	0.10
Cs/Is treatment, %	33.3	36.7	0.83
Clinical outcomes	–	–	–
Rate of eGFR loss (mL/min/1.73 m^2/year)	0.2 (−0.8–2.2)	−3.8 (−8.2–−2.9)	P < 0.0001

The 75th percentile of the high rate of eGFR loss detected in the cohort of 157 patients was used to categorize patients into fast progressors (eGFR decrease ≥−1.91 mL/min/1.73 m^2/year) and non-fast progressors (eGFR loss <−1.91/mL/min/m^2/year or no loss) during follow-up. Values are expressed as median (IQR) unless stated otherwise.

reports these correlations). No correlation was found with therapy at sampling and the investigated parameters. The two switches were similar in progressors versus non-progressors but significantly higher in fast progressors compared with non-fast progressors (Table 3) [LMP7/β5 1.87 (IQR 1.38–2.94) versus 1.58 (1.28–2.09), P = 0.04 (Figure 2B) and LMP2/β1 1.6 (IQR 1–5.55) versus 1.18 (0.94–1.73), P = 0.04]. However, only the LMP7/β5 switch was significantly correlated with the eGFR slope (P = 0.027) (Figure 2A).

Correlations between iPS switch, TLR and CD46 expressions

In the entire cohort of IgAN patients, a strong correlation was found between the switch from PS to iPS and TLR mRNAs (Table 4), particularly between LMP7/β5 and TLR4 mRNA expression (P < 0.0001) (Figure 3).

A significant difference was detected in CD46 mRNA expression in peripheral blood cells of patients with fast progressing IgAN versus non-fast progressors (P = 0.007) (Table 3 and Figure 4A). LMP7/β5 switch and CD46 mRNA expressions showed inverse trends of correlation with respect to the fast progression status (Figure 4B). The fast progressor group showed a trend for a negative correlation while in the non-fast progressor group there was a trend for a positive correlation. In a logistic regression on fast progression status in the cohort of IgAN investigated that considered clinical data at renal biopsy and categorized MEST-C (mesangial hypercellularity, endocapillary hypercellularity, segmental glomerulosclerosis, tubular atrophy–interstitial fibrosis and crescent formations) scores, we observed that the association between LMP7/β5 switch and CD46mRNA expression had a strong predictor effect (Supplementary data, Table S3). In fact, patients with both low expression of CD46 and high expression of LMP7/β5 switch in peripheral blood were at high risk for progression of IgAN {odds ratio [OR] 15.8 [95% confidence interval (CI) 2.16–158.34] (Figure 5). This distribution of risk according to the LMP7/β5-CD46 group of pertinence indicates the value of such categorization in predicting the fast progression of IgAN in our cohort sample (model area under the curve (AUC) 0.78, P = 0.004].

Risk allele

The frequency of the risk allele in the SNP rs9357155-C was 0.87 in patients compared with 0.83 healthy Utah residents with European ancestry controls from HapMap-III, and the SNP was not associated with progressive IgAN (Table 3). No association was found with PS/iPS switches.

MATERIALS AND METHODS

Clinical data set and definitions

Definitions of clinical variables and outcomes followed the original VALIGA study [18]. Briefly, eGFR was estimated using the four-variable Modification of Diet in Renal Disease formula and proteinuria expressed in g/day. Mean arterial pressure (MAP) was calculated as one-third of the pulse pressure. At renal biopsy, 41 subjects were children (<18 years old): in these subjects, eGFR was estimated using the Schwartz formula (constant $K = 0.55$) with a maximum eGFR set at 120 mL/min/1.73 m^2. Proteinuria was expressed in g/day/1.73 m^2; MAP was adjusted for gender and age, as in the original report. At sampling, three subjects were <18 years old. ESRD was defined as eGFR <15 mL/min/1.73 m^2 in all patients. Time-averaged proteinuria was determined for each year of observation.

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

Immunosuppressive treatment was considered as intent to treat regardless of the type or duration. Renal biopsies were scored according to the Oxford classification: mesangial hypercellularity, M0/M1 (less than or equal to >50% of glomeruli with >4 mesangial cells/area); endocapillary hypercellularity, E0/E1 (present/absent); segmental glomerulosclerosis, S0/S1 (present/absent); and tubular atrophy/interstitial fibrosis, T0/T1/T2 (<25, >25 <50, >50%) [25]. Moreover, the presence of crescents (C1–C2) was considered [26].

Quantification of mRNA expressions

Blood samples were collected in PAXgene tubes (QIAGEN, Hilden, Germany) in patients with IgAN and healthy controls as previously described [19]. RNA was extracted from WBCs with the PAXgene RNA System Kit (QIAGEN). Quantitative real-time polymerase chain reaction was performed using TaqMan Universal PCR MasterMix (Life Technologies, Carlsbad, CA, USA). Probes for the detection of mRNAs encoding for PS subunits (β1, β2 and β5), iPS subunits (LMP2, LMP7 and MECL-1), TLR2, TLR4 and CD46 were used as previously reported [19, 27, 28]. Relative quantification of target genes expression was performed with the $\Delta\Delta C_t$ method and the relative fold changes were determined as previously detailed, hence results are expressed in corresponding arbitrary units [13].

Genotype data

The genotypes for rs2071543 were generated as part of the previously published studies [8, 24]. The quality control analyses were identical to the ones described in these studies.

Statistical methods

The rate of renal function decline (eGFR slope) was determined by fitting a straight line through available eGFR data from renal biopsy to sampling, as previously described [18]. Disease progression and velocity were defined according to the median and higher 75th centile value of eGFR slope in the study cohort, as described above. The functional form of all continuous variables was assessed. Data were tested for normal distribution using the Shapiro–Wilk test. Descriptive variables were presented as median (IQR, owing to non-normal distributions) or frequency (count) and compared across relevant groups using the Mann–Whitney U or Wilcoxon signed-rank test for continuous variables and Fisher exact test for categorical variables. The Oxford MEST-C scores were categorized as reported. Therapy was the intent to treat regardless of the type or duration, as in Coppo *et al.* [19]. Exposure to Cs/Is at sampling was also considered. Spearman rank correlation coefficient was used to investigate correlations, due to the non-linearity of assessed data. Group tests were two-sided, with P-values <0.05 considered statistically significant. Univariate and multivariate regression analyses were performed to calculate associations between iPS expression and other clinical variables, including disease progression. A reciever operating characteristics (ROC) curve served to identify the optimal threshold point of CD46 mRNA and LMP7/β5 expression based on disease progression velocity status. Multivariate logistic regression analyses were chosen to

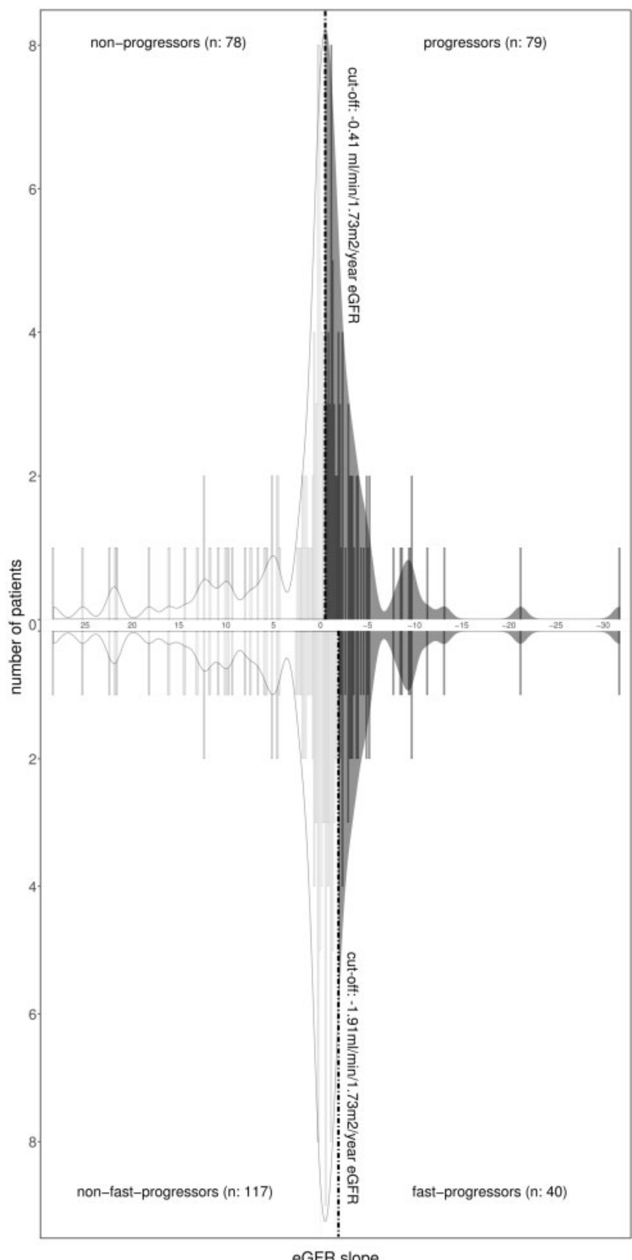


FIGURE 1: Identification of progressive and fast progressive subjects in the cohort of IgAN patients. Histogram of frequency distribution of eGFR slopes in progressors and non-progressors and fast progressors and non-fast progressors. The median value of the eGFR slope in the 157 patients investigated (−0.41 mL/min/1.73 m^2/year) allowed a categorization into progressors and non-progressors. The 75th centile of eGFR decrease (−1.91/mL/min/m^2/year) allowed a categorization of the patients into 40 'fast progressors', with median eGFR loss of −3.8 (IQR −8.2 to −2.9) mL/min/1.73 m^2/year, and 117 'non-fast progressors' with a median eGFR loss of −0.2 (IQR −0.8–2.2) mL/min/1.73 m^2/year. The change of threshold from the 50th to 25th centile identifies a more severe su-population of IgAN patients with faster progression.

investigate the accuracy of disease progression prediction models. The best predictive model including clinical data at renal biopsy (sex, age, eGFR, proteinuria and MAP) and MEST-C scores categorized into 0 (no positive score) versus 1 (at least one score > 0) was implemented with a categorical variable

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

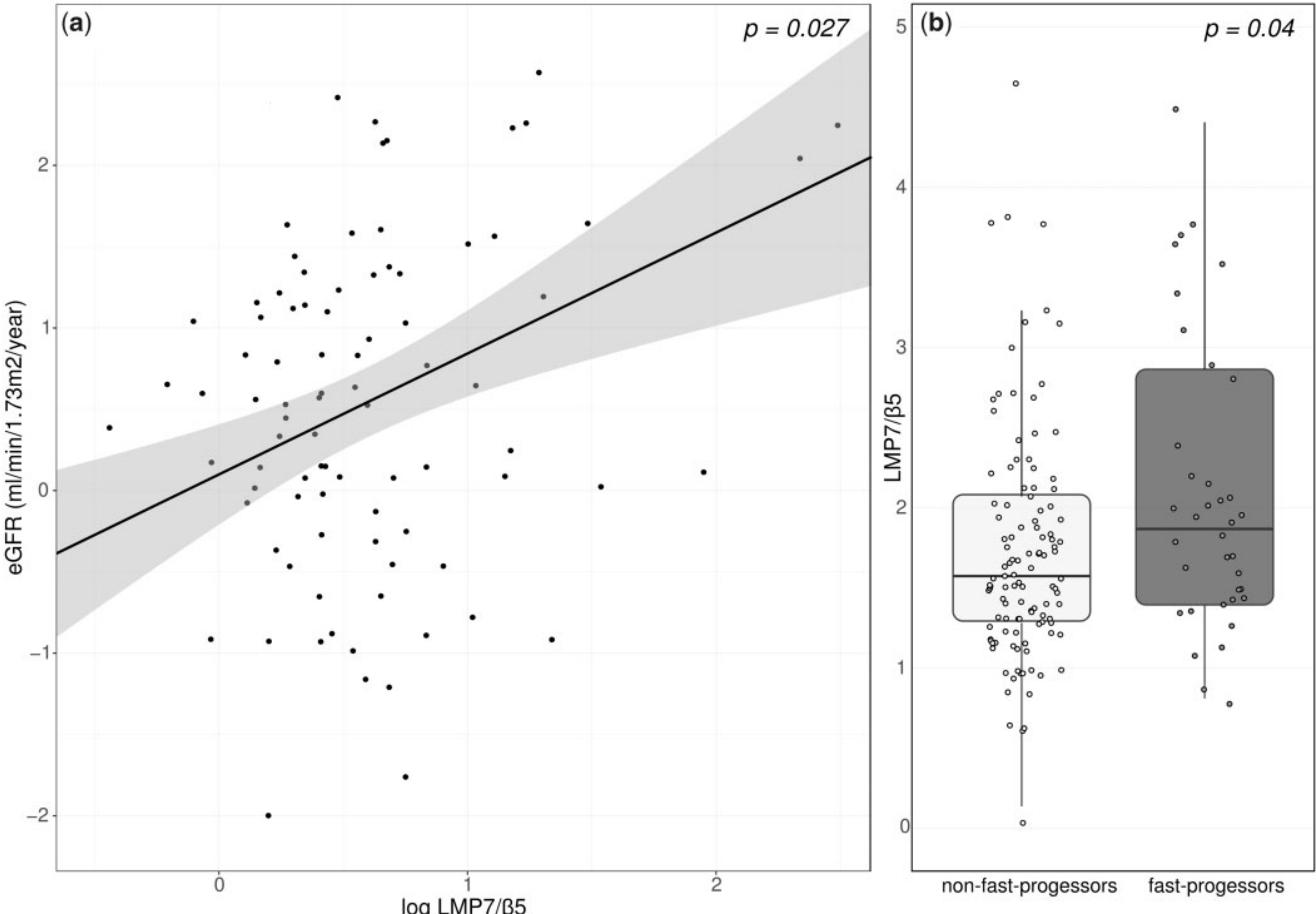


FIGURE 2: Correlation between LMP7/β5 switch in peripheral WBCs and velocity of progression of IgAN. (**A**) Correlation between LMP7/β5 switch and eGFR slope (P = 0.027). (**B**) LMP7/β5 switch in fast progressors and non-fast progressors (P = 0.04).

Table 3. Switches from PS to iPS (LMP7/β5, LMP2/β1, MECL-1/β2), TLR and CD46 mRNA expressions in peripheral WBCs and *PSMB8*/*PSMB9* SNP rs9357155 genotype alleles

Variable	Non-fast progressors (*n* = 117)	Fast progressors (*n* = 40)	P-value
LMP7/β5	1.58 (1.28–2.09)	1.87 (1.38–2.94)	0.04
LMP2/β1	1.18 (0.94–1.73)	1.6 (1–5.55)	0.04
MECL-1/β2	0.85 (0.64–1.17)	0.79 (0.6–1.34)	0.83
TLR2mRNA	1.19 (0.8–1.89)	1.56 (0.86–3.53)	0.12
TLR4mRNA	1.34 (0.85–1.95)	1.3 (0.89–3.01)	0.45
CD46 mRNAlog	−0.13 (−0.58 to 0.29)	−0.55 (−1.07 to 0.01)	0.007
PSMB8 SNP rs9357155	TT: 1.5%	TT: 0.0%	0.74
	CT: 20.0%	CT: 16.0%	
	CC: 78.5%	CC: 84.0%	

LMP7/β5, LMP2/β1, MECL-1/β2 are expressed as a ratio of mRNA expressions and TLR2 and TLR mRNA expressions as arbitrary units. CD46 mRNA expression is reported in log U (details in the Methods section). rs9357155 SNP genotypes at the *PSMB8*/*PSMB9* locus are expressed as the frequency of the alleles.

defined by LMP7/β5 and CD46 mRNA combined expression thresholds, identified through the ROC of univariate models on disease progression velocity. Patients were categorized into four risk groups, from lower (CD46 overexpressed and LMP7/β5 underexpressed) to higher predicted risk (CD46 underexpressed and LMP7β5 overexpressed). Overall model fit was assessed using the Akaike's information criterion, AUC and log-likelihood test. Analyses were performed and figures generated using R 3.4.4 (R Foundation, Vienna, Austria).

DISCUSSION

We previously reported an upregulation of the iPS in peripheral blood mononuclear cells of patients with IgAN in a single-centre exploratory study [13]. The increased expression of iPS was particularly significant for the LMP7/β5 switch that correlated with persistent proteinuria.

This study investigated a larger group of 157 patients with IgAN from multiple European centres contributing to the VALIGA cohort with centrally reviewed renal pathology scores

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

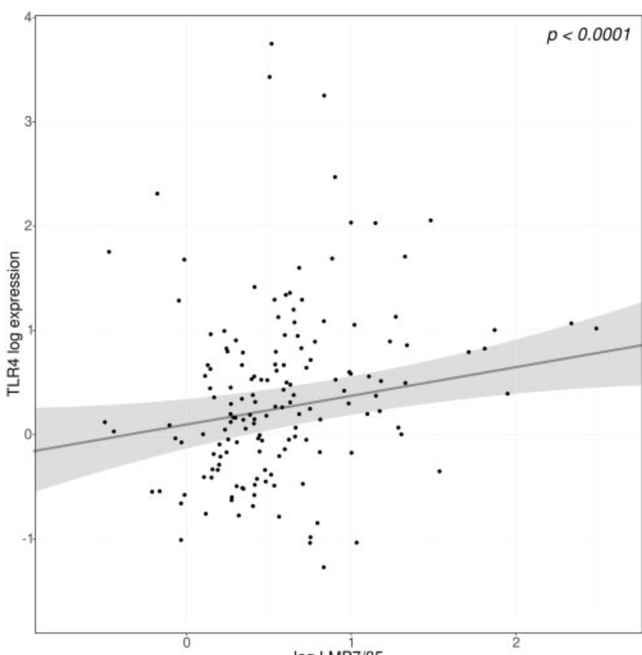


FIGURE 3: Correlation between LMP7/β5 switch and TLR4 mRNA expression in peripheral WBCs. Spearman's ρ = 0.31, P = 9.02e-05 (<0.0001).

and complete long-term follow-up after renal biopsy [19]. The switch from PS to iPS was assessed as a ratio of subunit mRNA expression, as in the previous study of our group [19], as well as in recent literature [29]. The individual PS and iPS proteins are present in cells in various combinations and do not change after cell maturation [30]. In this selected cohort, we confirmed the upregulation of iPS, with increased switch of LMP7/β5 and LMP2/β1, comparied with healthy controls. No evidence of genetic conditioning was detectable for the risk allele in the SNP rs9357155-C of the *PSMB8/9* gene encoding for LMP7/β5 and LMP2/β1.

The novelty of the present report is the finding of a significant association between the rate of eGFR decline and the increased expression of the LMP7/β5 in peripheral cells in patients with IgAN. Patients with fast progression (arbitrarily defined by an eGFR slope >75th centile in the whole group) had significantly higher PS-iPS switch, independent from the level of renal function at sampling or at renal biopsy. LMP7/β5 and LMP2/β1 switch correlated with the velocity of yearly loss of eGFR and proteinuria at sampling. The remarkable finding is that the same subcohort of fast progressing patients had a concomitant significant reduction in CD46 mRNA expression in peripheral cells. The combination of increased PS–iPS switch and reduced CD46 expression was strongly predictive of a fast progressor status.

Recent data have suggested a dysregulation in gut-associated lymphoid tissue (GALT) in patients with IgAN. In these patients, genetic determinants, gut dysbiosis and reactions to diet components may play a combined role [31]. iPS is highly represented in the GALT, acting in defence against pathogens [12]. In inflammatory bowel disease, an increase in LMP7/β5 switch was reported at a local intestinal mucosa level [32, 33]. LMP7 and LMP2 are encoded by the *PSMB8* and *PSMB9* genes

Table 4. Correlation between the PS to IPS switches (LMP7/β5, LMP2/β1, MECL-1/β2) and TLR2 or TLR4 mRNA expressions

	TLR2		TLR4	
	ρ	P-value	ρ	P-value
LMP7/β5	0.122	0.12	0.307	<0.0001
LMP2/β1	0.241	0.002	0.243	0.002
MECL-1/β2	0.216	0.006	0.274	0.0005

ρ, Spearman correlation test.

and are involved in NF-κB activation, maintenance of the intestinal epithelial barrier and local control of inflammatory responses to infection [24]. In a previous genome-wide association study, rs9357155 at the *PSMB8/9* locus was found to be associated with increased risk of IgAN [23]. However, this SNP was not associated with increased iPS switch in our study.

We observed a highly significant association between LMP7/β5 switch and TLR4 mRNA expression in circulating cells. TLR4 is highly expressed at the intestinal mucosa surface, as a receptor for Gram-positive bacterial lipopolysaccharide [34, 35]. We previously detected a significant increased expression of TLR4 mRNA in peripheral mononuclear cells of patients with IgAN [27]. TLR4 activation induces a release of NF-κB and IFN-β [34] aimed at defending against pathogens while maintaining tolerance to commensal bacteria. The concomitant increased expression of LMP7/β5 and TLR4 mRNAs suggests the hypothesis that an altered GALT response to intestinal components may favour an increased iPS switch in immune-competent cells that later reach the systemic circulation.

iPS is involved also in autoimmunity, and an up-regulation of iPS subunits has been detected in human autoimmune disorders [32, 36, 37]. In IgAN, an autoimmune reaction against Gd-IgA1 leads to the production of autoantibodies correlated with the progression risk [38]. Cells of the adaptive immune system are regulated by iPS, mainly by LMP7/β5 switch, inducing a reduction in regulatory T (Treg) cell differentiation and feedback. This event may lead to uncontrolled immune response and persistent T-cell activation [39]. A reduced circulation of Treg cells in patients with IgAN was reported [40] and we previously reported a lower expression of the transcriptional factor forkhead box protein P3 in peripheral mononuclear cells of children with primary IgAN and IgA vasculitis with nephritis, with increased iPS switch [28]. Treg cells play a powerful anti-inflammatory role, maintaining tolerance through the direct contact or the release of anti-inflammatory cytokines and balancing the pro-inflammatory effects of Th17. Patients with progressive IgAN have an increased renal presence of Th17-positive cells [40, 41]. These data collectively suggest that the increased expression of iPS may favour autoimmune reactions, defective regulatory T-cell control and persistent inflammatory conditions favouring disease progression in IgAN.

The coexistance of iPS high switch and CD46 mRNA low expression in patients with progressive IgAN is unlikely to be a chance finding, since the effect estimates for these associations are large. CD46 modulates at the cellular level C3 convertase activity of the lectin and alternative complement pathways; its genetic defect can lead to uncontrolled complement activation at

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

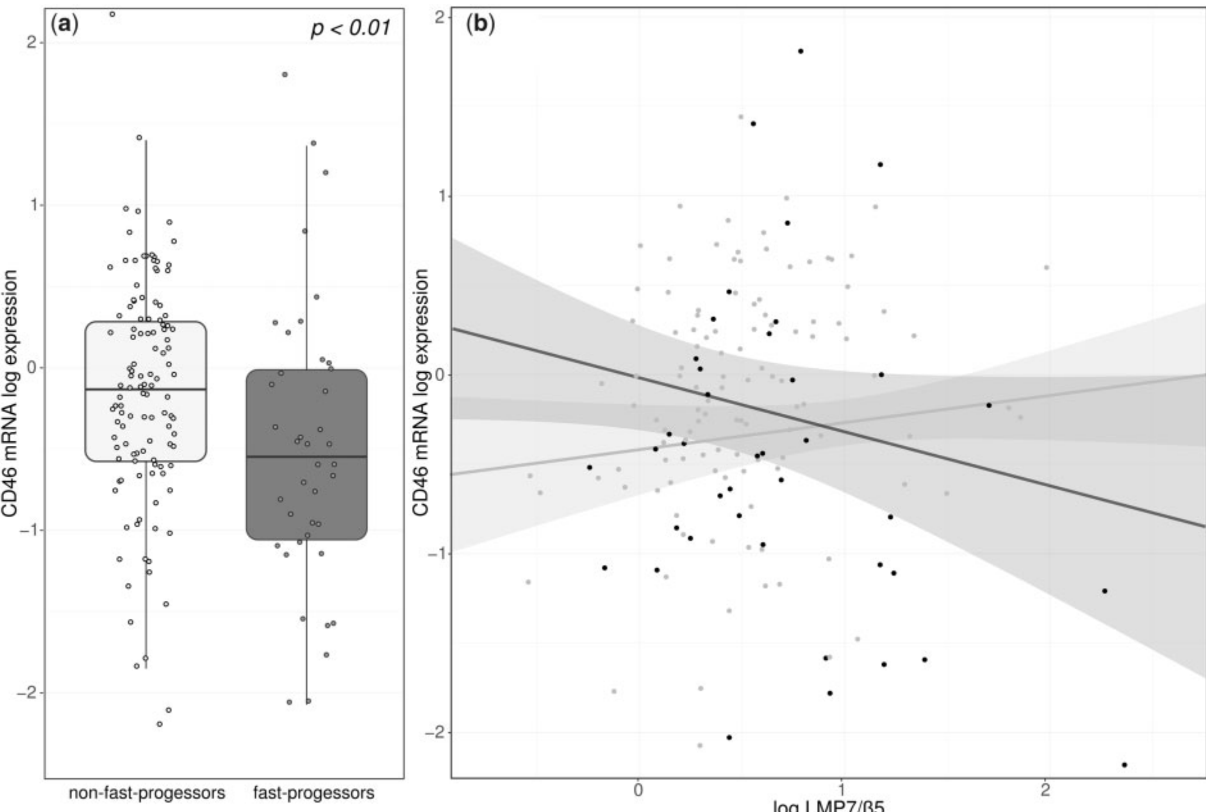


FIGURE 4: CD46 mRNA expression in peripheral WBCs in non-fast progressors and fast progressors and correlation with LMP7/β5 switch. (**A**) CD46 expression in WBCs in fast progressors and non-fast progressors (Wilcoxon–Mann–Whitney test P-value = 0.007). (**B**) LMP7/β5 switch and CD46 mRNA expression are oppositely correlated depending on fast progression status. The graph shows the opposite trends of correlation of CD46 and LMP7/β5 expressions in patients grouped by fast progression status. The fast progressor group shows a negative trend (Spearman ρ = −0.21), with overexpression of LMP7/β5 coupled with low CD46 expression and vice versa. The non-fast progressor group shows a positive correlation trend (Spearman ρ = 0.16). Both Spearman correlation tests result were non-significant. The expression values are log-transformed to avoid the graphical crushing effect of non-normally distributed values. The Spearman correlation test was operated on crude levels of expression to avoid a biased correlation index.

the endothelial level and atypical haemolytic uraemic syndrome [22]. Apart from this function, CD46 has been recently identified as a costimulatory molecule for T-cell activation; CD46-costimulated human T cells induce a Tr1 phenotype with high production of the anti-inflammatory interleukin (IL)-10 [41]. As reported above, Tregs are reduced in patients with IgAN. The CD46-mediated differentiation pathway is defective in several chronic inflammatory diseases, underlying the importance of CD46 in controlling T-cell function. CD46-mediated Tr1 differentiation is altered in patients with multiple sclerosis [42], presenting with impaired IL-10 secretion upon CD46 costimulation [41, 43] in patients with lupus [44]. Hence CD46 is a multifunctional molecule, which can be down-regulated by genetic conditioning and by acquired factors, playing a role in immune activation and as a modulator of adaptive immunity. We observed that the decreased expression of CD46 mRNA in the peripheral cells of patients with IgAN significantly improved prediction of the progression rate when added to clinical and pathology features [19].

Here we report that the association of reduced CD46 mRNA expression and enhanced iPS switch in peripheral cells increases the risk of IgAN progression.

SUPPLEMENTARY DATA

Supplementary data are available at ndt online.

ACKNOWLEDGEMENTS

We thank Drs. Krzysztof Kiryluk and Ali Gharavi from Columbia University for providing genotype data for the *PSMB8/9* locus and for helpful feedback on the study design.

FUNDING

The study was supported by the Immunopathology Working Group of the ERA-EDTA. The VALIGA study was granted by the first research call of the ERA-EDTA in 2009. Preliminary and partial data were presented as an oral communication at the ERA-EDTA 2016 Congress.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. We declare that the results presented in this article have not been published previously in whole or in part. The study had approval from

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

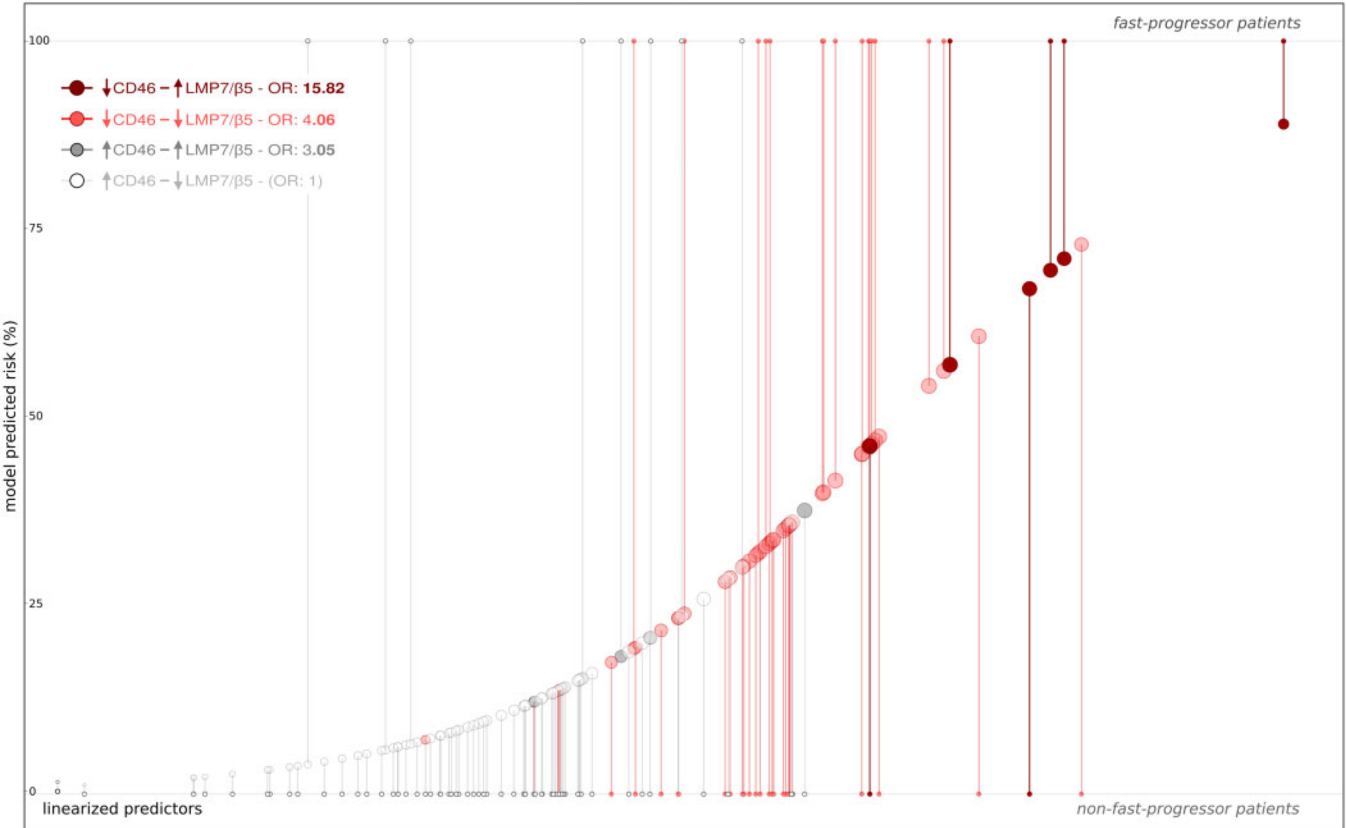


FIGURE 5: LMP7/β5 and CD46 expression are strongly predictive for rapid progression. The graph shows the patients as grouped by LMP7/β5 switches and CD46 expression thresholds and its strong effect on rapid progression prediction. This is a graphical representation of the logistic model utilized for the prediction of rapid progression in our patients (Supplementary data, Table S3): the dots lying on the horizontal line at level 1 of the y-axis are the fast progressor patients, the ones lying on level 0 of the y-axis are the non-fast progressor patients. The S-shaped positioned dots between 0 and 1 (on the vertical axis) are the estimated risk of each patient based on the predictor variable values, transformed by the link logit function that characterizes binomial distribution. Each combination of predictor variables defines an estimated relative risk for each patient that, thanks to a logit transformation, lies between 0 (no risk) and 1 (full risk), assuming an asymptotic distribution. The vertical lines on the graph connect each patient's real status with the respective estimated risk, representing the residual of each combination. Dots are coloured based on LMP7/β5 and CD46 expression threshold group of pertinence. Thresholds of over/underexpressions are based on univariate ROC bootstrapped best thresholds for disease progression velocity. Patients with both underexpression of CD46 mRNA and overexpression of LMP7/β5 are estimated to be closer to 1 on the y-axis (higher risk, as shown by an OR of 15.82 by the model), with respect to patients with both overexpression of CD46mRNA and underexpression of LMP7/β5, that are closer to 0 on the y-axis (lower risk) by the model. Groups with intermediate estimated risk tend to have only underexpression of CD46 mRNA (OR 4.06) or overexpression of LMP7/β5 (OR 3.03). This distribution of risk according to the LMP7/β5-CD46mRNA group of pertinence testify to the great value of such categorization in predicting the rapid progression of IgAN patients in our sample.

local ethics committee and was performed in accordance with the Declaration of Helsinki.

REFERENCES

1. Suzuki H, Kiryluk K, Novak J *et al.* The pathophysiology of IgA nephropathy. *J Am Soc Nephrol* 2011; 22: 1795–1803
2. Yeo SC, Cheung CK, Barratt J. New insights into the pathogenesis of IgA nephropathy. *Pediatr Nephrol* 2018; 33: 763–777
3. Maillard N, Wyatt RJ, Julian BA *et al.* Current understanding of the role of complement in IgA nephropathy. *J Am Soc Nephrol* 2015; 26: 1503–1512
4. Floege J, Daha MR. IgA nephropathy: new insights into the role of complement. *Kidney Int* 2018; 94: 16–18
5. Coppo R. Biomarkers and targeted new therapies for IgA nephropathy. *Pediatr Nephrol* 2017; 32: 725–731
6. Kang W, Kudsk KA. Is there evidence that the gut contributes to mucosal immunity in humans? *J Parenter Enteral Nutr* 2007; 31: 246–258
7. Lu Y, Li X, Liu S *et al.* Toll-like receptors and inflammatory bowel disease. *Front Immunol* 2018; 9: 72
8. Kiryluk K, Novak J. The genetics and immunobiology of IgA nephropathy. *J Clin Invest* 2014; 124: 2325–2332
9. Toubi E, Shoenfeld Y. Toll-like receptors and their role in the development of autoimmune diseases. *Autoimmunity* 2004; 37: 183–188
10. Kammerl IE, Meiners S. Proteasome function shapes innate and adaptive immune responses. *Am J Physiol Lung Cell Mol Physiol* 2016; 311: L328–L336
11. Meyer-Schwesinger C. The ubiquitin–proteasome system in kidney physiology and disease. *Nat Rev Nephrol.* 2019; 15: 393–411
12. Ferrington DA, Gregerson DS. Immunoproteasomes: structure, function, and antigen presentation. *Prog Mol Biol Transl Sci* 2012; 109: 75–112
13. Coppo R, Camilla R, Alfarano A *et al.* Upregulation of the immunoproteasome in peripheral blood mononuclear cells of patients with IgA nephropathy. *Kidney Int* 2009; 75: 536–541
14. Rizk DV, Maillard N, Julian BA *et al.* The emerging role of complement proteins as a target for therapy of IgA nephropathy. *Front Immunol* 2019; 10: 1–14

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

15. Jullien P, Laurent B, Claisse G *et al.* Deletion variants of CFHR1 and CFHR3 associate with mesangial immune deposits but not with progression of IgA nephropathy. *J Am Soc Nephrol* 2018; 29: 661–669
16. Zhai YL, Meng SJ, Zhu L *et al.* Rare variants in the complement factor H-related protein 5 gene contribute to genetic susceptibility to IgA nephropathy. *J Am Soc Nephrol* 2016; 27: 2894–2905
17. Xie J, Kiryluk K, Li Y *et al.* Fine mapping implicates a deletion of CFHR1 and CFHR3 in protection from IgA nephropathy in Han Chinese. *J Am Soc Nephrol* 2016; 27: 3187–3194
18. Coppo R, Troyanov S, Bellur S *et al.* Validation of the Oxford classification of IgA nephropathy in cohorts with different presentations and treatments. *Kidney Int* 2014; 86: 828–836
19. Coppo R, Peruzzi L, Loiacono E *et al.* Defective gene expression of the membrane complement inhibitor CD46 in patients with progressive immunoglobulin A nephropathy. *Nephrol Dial Transplant* 2019; 34: 587–596
20. Cardone J, Le Friec G, Kemper C. CD46 in innate and adaptive immunity: an update. *Clin Exp Immunol* 2011; 164: 301–311
21. Fuchs A, Atkinson JP, Fremeaux-Bacchi V *et al.* CD46-induced human Treg enhance B-cell responses. *Eur J Immunol* 2009; 39: 3097–3109
22. Yamamoto H, Fara AF, Dasgupta P *et al.* CD46: the "multitasker" of complement proteins. *Int J Biochem Cell Biol* 2013; 45: 2808–2820
23. Gharavi AG, Kiryluk K, Choi M *et al.* Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nat Genet* 2011; 43: 321–329
24. Kiryluk K, Li Y, Sanna-Cherchi S *et al.* Geographic differences in genetic susceptibility to IgA nephropathy: GWAS replication study and geospatial risk analysis. *PLoS Genet* 2012; 8: e1002765
25. Cattran DC, Coppo R, Cook HT *et al.* The Oxford classification of IgA nephropathy: rationale, clinicopathological correlations, and classification. *Kidney Int* 2009; 76: 534–545
26. Trimarchi H, Barratt J, Cattran DC *et al.* Oxford classification of IgA nephropathy 2016: an update from the IgA nephropathy classification working group. *Kidney Int* 2017; 91: 1014–1021
27. Coppo R, Camilla R, Amore A *et al.* Toll-like receptor 4 expression is increased in circulating mononuclear cells of patients with immunoglobulin A nephropathy. *Clin Exp Immunol* 2010; 159: 73–81
28. Donadio ME, Loiacono E, Peruzzi L *et al.* Toll-like receptors, immunoproteasome and regulatory T cells in children with Henoch-Schönlein purpura and primary IgA nephropathy. *Pediatr Nephrol* 2014; 29: 1545–1551
29. Chandrasekaran A, Adkins LJ, Seltzer HM *et al.* Age-dependent effects of immunoproteasome deficiency on mouse adenovirus type 1 pathogenesis. *J Virol* 2019; 93: e00569
30. Guillaume B, Chapiro J, Stroobant V *et al.* Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proc Natl Acad Sci USA* 2010; 107: 18599–18604
31. Coppo R. The gut–kidney axis in IgA nephropathy: role of microbiota and diet on genetic predisposition. *Pediatr Nephrol* 2018; 33: 53–61
32. Fitzpatrick LR, Small JS, Poritz LS *et al.* Enhanced intestinal expression of the proteasome subunit low molecular mass polypeptide 2 in patients with inflammatory bowel disease. *Dis Colon Rectum* 2007; 50: 337–350
33. Schmidt N, Gonzalez E, Visekruna A *et al.* Targeting the proteasome: partial inhibition of the proteasome by bortezomib or deletion of the immunosubunit LMP7 attenuates experimental colitis. *Gut* 2010; 59: 896–906
34. Neal MD, Leaphart C, Levy R *et al.* Enterocyte TLR4 mediates phagocytosis and translocation of bacteria across the intestinal barrier. *J Immunol* 2006; 176: 3070–3079
35. Fukata M, Hernandez Y, Conduah D *et al.* Innate immune signaling by toll-like receptor-4 (TLR4) shapes the inflammatory microenvironment in colitis-associated tumors. *Inflamm Bowel Dis* 2009; 15: 997–1006
36. Eleftheriadis T, Pissas G, Antoniadi G *et al.* CD8+ T-cell auto-reactivity is dependent on the expression of the immunoproteasome subunit LMP7 in exposed to lipopolysaccharide antigen presenting cells and epithelial target cells. *Autoimmunity* 2013; 46: 439–445
37. Ghannam K, Martinez-Gamboa L, Spengler L *et al.* Upregulation of immunoproteasome subunits in myositis indicates active inflammation with involvement of antigen presenting cells, CD8 T-cells and IFNγ. *PLoS One* 2014; 9: e104048
38. Berthoux F, Suzuki H, Thibaudin L *et al.* Autoantibodies targeting galactose-deficient IgA1 associate with progression of IgA nephropathy. *J Am Soc Nephrol* 2012; 23: 1579–1587
39. Kalim KW, Basler M, Kirk CJ *et al.* Immunoproteasome subunit LMP7 deficiency and inhibition suppresses Th1 and Th17 but enhances regulatory T cell differentiation. *J Immunol* 2012; 189: 4182–4193
40. Lin FJ, Jiang GR, Shan JP *et al.* Imbalance of regulatory T cells to Th17 cells in IgA nephropathy. *Scand J Clin Lab Invest* 2012; 72: 221–229
41. van Es LA, de Heer E, Vleming LJ *et al.* GMP-17-positive T-lymphocytes in renal tubules predict progression in early stages of IgA nephropathy. *Kidney Int* 2008; 73: 1426–1433
42. Ni Choileain S, Astier AL. CD46 plasticity and its inflammatory bias in multiple sclerosis. *Arch Immunol Ther Exp (Warsz)* 2011; 59: 49–59
43. Schramm EC, Roumenina LT, Rybkine T *et al.* Mapping interactions between complement C3 and regulators using mutations in atypical hemolytic uremic syndrome. *Blood* 2015; 125: 2359–2369
44. Ellinghaus U, Cortini A, Pinder CL *et al.* Dysregulated CD46 shedding interferes with Th1-contraction in systemic lupus erythematosus. *Eur J Immunol* 2017; 47: 1200–1210

Received: 20.11.2019; Editorial decision: 5.3.2020

Downloaded from https://academic.oup.com/ndt/article/36/8/1389/5862352 by guest on 24 February 2023

ORIGINAL RESEARCH
published: 05 June 2020
doi: 10.3389/fped.2020.00234

Pre-discharge Cardiorespiratory Monitoring in Preterm Infants. the CORE Study

Francesco Cresi*, Enrico Cocchi, Elena Maggiora, Alice Pirra, Federica Logrippo, Maria Chiara Ariotti, Chiara Peila, Enrico Bertino and Alessandra Coscia

Neonatal Intensive Care Unit, City of Health and Science—University of Turin, Turin, Italy

Objective: Ensuring cardiorespiratory (CR) stability is essential for a safe discharge. The aim of this study was to assess the impact of a new pre-discharge protocol named CORE on the risk of hospital readmission (RHR).

Methods: Preterm infants admitted in our NICU between 2015 and 2018 were randomly assigned to CORE (exposed) or to standard (not-exposed) discharge protocol. CORE included 24 h-clinical observation, followed by 24 h-instrumental CR monitoring only for high-risk infants. RHR 12 months after discharge and length of stay represent the primary and secondary outcomes, respectively.

Results: Three hundred and twenty three preterm infants were enrolled. Exposed infants had a lower RHR (log-rank $p < 0.05$). The difference was especially marked 3 months after discharge (9.09 vs. 21.6%; $p = 0.004$). The hospital length of stay in exposed and not-exposed infants was 39(26–58) and 43(26–68) days, respectively ($p = 0.16$).

Conclusions: The CORE protocol could help neonatologists to define the best timing for discharge reducing RHR without lengthening hospital stay.

Keywords: cardiorespiratory stability, safe discharge, hospital readmission, length of stay, NICU discharge

OPEN ACCESS

Edited by:
Eugene Dempsey,
University College Cork, Ireland

Reviewed by:
Kazumichi Fujioka,
Kobe University, Japan
Hercília Guimarães,
University of Porto, Portugal

****Correspondence:***
Francesco Cresi
francesco.cresi@unito.it

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: *12 March 2020*
Accepted: *17 April 2020*
Published: *05 June 2020*

Citation:
Cresi F, Cocchi E, Maggiora E, Pirra A, Logrippo F, Ariotti MC, Peila C, Bertino E and Coscia A (2020) Pre-discharge Cardiorespiratory Monitoring in Preterm Infants. the CORE Study. Front. Pediatr. 8:234. doi: 10.3389/fped.2020.00234

INTRODUCTION

Preterm birth is the most important determinant of adverse infant outcomes in terms of survival and short and long-term health complications affecting quality of life (1). Preterm infants, especially those born with a very low birth weight (VLBW), are more likely to suffer from major morbidity such as respiratory distress and subsequent bronchopulmonary dysplasia (BPD), necrotizing enterocolitis, intraventricular hemorrhage, retinopathy of prematurity, persistent patent ductus arteriosus, and sepsis and are at higher risk of mortality than infants born at term (2–8).

Since preterm birth predisposes infants to higher health risks and an increased rate of re-hospitalization (9, 10), hospital discharge represents a delicate process and determining an appropriate criteria is challenging. Thus, the discharge of a high-risk infant should be approached as a multidisciplinary process with an overall view on the infant's health course in the Neonatal Intensive Care Unit (NICU) aiming at providing families with the instruments and knowledge necessary for a safe return home and improving post-discharge outcomes (11).

Currently, discharge decision making varies widely among neonatologists. The decision to discharge is primarily based on the infant's medical status (demonstration of functional maturation including physiological competencies of thermoregulation, control of breathing, respiratory stability, feeding skills, and weight gain) but its success, including avoiding early admission to

emergency department, is complicated by several factors, above all by the adequate competence and suitability of families, the availability of support services and the pressure to contain hospital costs (11–13).

Moreover, data (9, 10, 14) shows that during the first year of life, hospital readmission rate increases with decreasing gestational age at birth, ranging from 13% in infants born at 35 weeks gestation to 31% in infants born at ≤ 25 weeks gestation (9). Therefore, preterm and VLBW infants are significantly more likely to be readmitted than infants born full term with appropriate weight (9, 10) These data are important especially in view of what the American Academy of Pediatrics "Policy Statement-Hospital Stay for Healthy Term Newborns" (14) which states *"the risk of hospital readmission for infants discharged from the NICU can be seen as an indicator of an inadequate assessment by health care providers of the newborn's readiness for discharge, a lack of resources and/or an inability of a parent to provide early newborn care, or inappropriate and/or untimely availability of, or access to, outpatient care"* and focus the attention on the central relevance they have in the cost-benefit analysis of discharge strategies, which should not only consider the birth hospital admission but also the median and long term health resources utilization.

In the NICU of Sant'Anna Hospital in Turin, we developed a pre-discharge protocol, named CORE (Cardio Observation and Respiratory Evaluation), to guarantee a safe discharge home for preterm infants.

The aim of the study was to assess the effectiveness of the new pre-discharge CORE protocol to evaluate the CR stability and its effect on the risk of hospital readmission.

MATERIALS AND METHODS

Study Design, Setting, and Population

In Sant'Anna Hospital NICU the pre-discharge CORE protocol to evaluate cardiorespiratory (CR) stability was established in 2015, as described below. At admission, preterm infants with 25^{+0}-33^{+6} weeks gestational age (GA) admitted to Sant'Anna Hospital NICU from November 2015 to January 2018 were randomly assigned to one of the three medical-nursing teams of the NICU. Each team followed the same diagnostic and therapeutic protocols except for the pre-discharge evaluation. Thus, infants assigned to one team were managed with CORE protocol (exposed) and infants assigned to the other two teams were managed following our unit discharge standard protocol (not-exposed). According to them, infants were considered ready for discharge if they achieved ≥ 1,600 g weight with stable weight gain, stable thermoregulation, spontaneous breathing, full oral feeding by breast or bottle and normal vital signs for at least 48 h before discharge. Infants with congenital abnormalities, major cardiac disorders and intraventricular hemorrhage of grade 2 or higher were excluded from the study.

According to the given risk factors, we divided infants into two subgroups: low-risk infants and high-risk infants (GA <28 weeks at birth and/or post-menstrual age ≤34 weeks at first clinical observation and/or history of mechanical ventilation >24 h, and/or need for supplemental oxygen and/or evidence of extreme CR events in the last 2 weeks).

We considered all patients enrolled who were readmitted within 12 months after discharge. In the analysis, we considered the whole group and the subgroups of infants with respiratory disorders and infants with apnea and/or apparent life-threatening events.

Risk of hospital readmission (RHR) during the first year after discharge and length of stay (LOS) represent the primary and secondary outcome, respectively.

Standard Discharge Protocol

Preterm infants were considered ready to discharge according to the achievement of the following characteristics and competencies: GA >34 weeks, weight >1,600 g, stable weight gain and thermoregulation, spontaneous breathing, full oral feeding by breast or bottle without CR compromise. The absence of apnea and desaturation episodes requiring any type of intervention during the last week of clinical observation assessed by nurses using a not-recording pulse-oximetry.

CORE Protocol

The CORE protocol was a three-step process to evaluate CR stability. The CORE protocol flow-chart is outlined in **Figure 1**.

The first step was a clinical evaluation to screen infants ready for discharge according to the achievement of some specific characteristics and competencies: GA >34 weeks, weight >1,600 g, stable weight gain and thermoregulation, spontaneous breathing, full oral feeding by breast or bottle without CR compromise. The second step was a 24-h clinical observation (CO) administered by nurses in monitored patients by using a 24 h structured diary. Low risk infants with normal CO were discharged. High risk infants with normal CO underwent 24-h instrumental monitoring (IM), as a third and final step. High-risk infants were discharged after a normal IM. Infants with pathological CO or IM underwent a clinical evaluation to identify and treat any medical problem (infections, BPD, nasal obstruction, cardiopathies, neurological issues, etc.) and repeated CO or IM after a week.

Infants with three pathological IM due to CR events were considered at risk of sudden infant death syndrome (SIDS) and referred to our SIDS center which provides patients with a home cardiorespiratory monitor capable of transmitting data and organizes follow-up visits.

Infants with pathological IM because of transcutaneous blood oxygen saturation <90% (CT90) exceeding 3% of the monitoring time repeated IM with oxygen supplementation. Infants with three consecutive pathological IM and/or need for supplemental oxygen were discharged with home oxygen equipment and a home monitor. These infants were referred to Pneumology Department which scheduled subsequent

Abbreviations: BPD, Bronchopulmonary Dysplasia; CO, 24 h Clinical Observation; CORE, Cardio Observation and Respiratory Evaluation; CR, Cardiorespiratory; CT90%, percentage of the monitoring time with $SatO_2$ < 90%; GA, Gestational Age; HR, Heart Rate; RHR, Risk of Hospital Readmission; IM, 24 h Instrumental Monitoring; LOS, Length of Stay; NICU, Neonatal Intensive Care Unit; PMA, Post Menstrual Age; SIDS, Sudden Infant Death Syndrome; $SatO_2$, transcutaneous blood Oxygen Saturation; VLBW, Very Low Birth Weight.

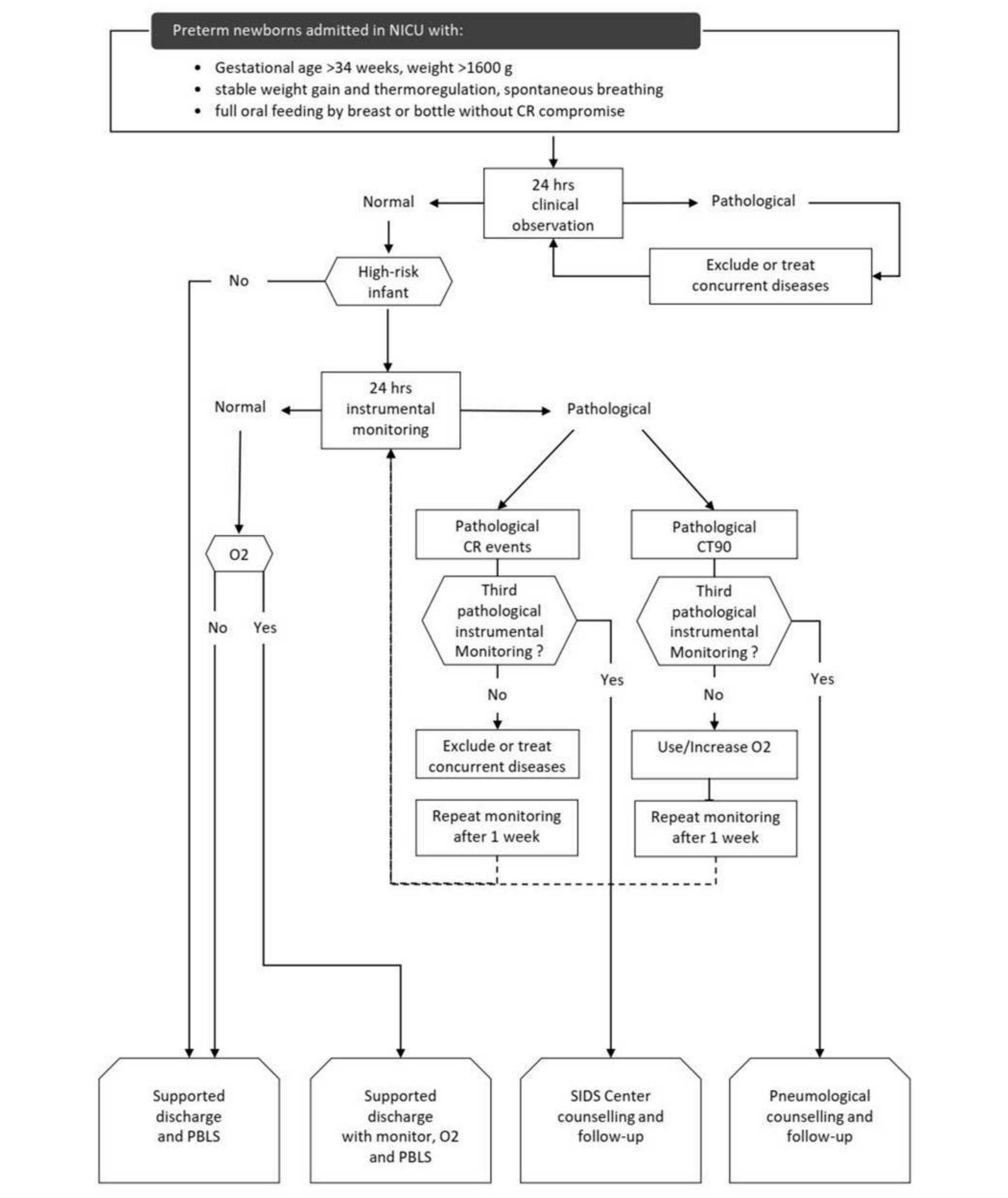


FIGURE 1 | CORE protocol flow-chart. high-risk infants: GA < 28 weeks at birth and/or post-menstrual age ≤ 34 weeks at first clinical observation and/or history of mechanical ventilation >24 h, and/or need for supplemental oxygen and/or evidence of extreme CR events in the last 2 weeks. CR: cardiorespiratory. O2:

(Continued)

FIGURE 1 | supplemental oxygen given by nasal cannula. CR events: desaturation (SpO2 <80% for at least 4 s); apnea (absence of thoracic movements for at least 20 s or at least 5 s if associated with SatO2 <80% or HR <80 bpm); bradycardia (HR <80 bpm for 15 s or <60 bpm for 5 s). Extreme CR events were defined as apnea lasting more than 30 s and/or HR <60 bpm for 10 s and/or <50 bpm for 15 s), HR <80 bpm and/or SatO2 <80% for 3 or more min, CR events associated with clinical signs and need for resuscitation. CT90: time with transcutaneous blood oxygen saturation <90%, in percent 24 h. clinical observation: nurse-administered clinical observation. It was considered pathological if 3 or more CR events and/or 1 or more extreme CR events occurred. 24 h. instrumental monitoring: cardiorespiratory recording. It was considered pathological if 12 or more CR events (events index >0.5 CR events/h) and/or 1 or more extreme CR events occurred and/or if the CT90 was >3%. PBLS, pediatric basic life support; SIDS, sudden infant death syndrome.

24 h-respiratory monitoring after discharge and follow-up visits. All parents were provided with pre-discharge education on SIDS prevention, safe sleep practices and basic infant cardiopulmonary resuscitation.

Techniques

CO consists of a 24-h nurse-administered clinical observation. During CO, heart rate (HR) and transcutaneous blood oxygen saturation ($SatO_2$) were continuously measured by a pulse-oximetry sensor placed on the right wrist or foot using Masimo Radical-7 pulse-oximetry (Masimo Corp., Irvine, CA, USA). Monitor setup included alarms for HR <80 bpm and $SatO_2$ <80%. A structured diary was used to mark any CR event (apnea, bradycardia, desaturation), pulse-oximetry parameters during these events (lowest HR, lowest $SatO_2$, duration of the event) and the kind of action subsequently undertaken (tactile stimulation, repositioning, aspiration, ventilation, etc).

IM consists of a 24-h CR recording using Getemed Vitaguard 3,100 system (Getemed Medizin und Information Stechnik AG, Teltow, Germany), equipped with Signal Extraction Technology (Masimo Corp., Irvine, CA, USA) (15). During IM a diary was completed by the caregiver to note the time and duration of meals, sleeping periods and any other situation that could potentially influence the results of the monitoring.

HR, $SatO_2$ and respiratory rate (RR) were measured during IM by a pulse-oximetry sensor placed on the wrist or foot and three cardiac electrodes placed on the chest. The recorded data were analyzed by a trained operator, using VitaWin3® evaluation software, version 3.3.

CR events were defined as:

- desaturation (SpO_2 <80% for at least 4 s);
- apnea (absence of thoracic movements for at least 20 s or at least 5 s if associated with $SatO_2$ <80% or HR <80 bpm);
- bradycardia (HR <80 bpm for 15 s or <60 bpm for 5 s);
- combined event (a combination of two of the events above);
- complete event (the combination of the three events above);

extreme CR events were defined as:

- apnea lasting more than 30 s and/or HR <60 bpm for 10 s and/or <50 bpm for 15 s).
- HR <80 bpm and/or $SatO_2$ <80% for 3 or more min
- CR events associated with clinical signs (changes in skin color, muscle tone or state of consciousness) and need for resuscitation (tactile stimulation, ventilation, etc.)

CO was considered pathological if 3 or more CR events and/or 1 or more extreme CR events occurred.

IM was considered pathological if 12 or more CR events (events index >0.5 CR events/h) and/or 1 or more extreme CR events occurred and/or if the CT90 was >3%.

Statistical Analysis

Data were tested for normal distribution using the Shapiro-Wilk Test and the Kolmogorov-Smirnov test. Descriptive continuous variables were presented as median and interquartile (if non-normally distributed) or mean and standard deviation (if normally distributed). Categorical variables were presented as frequency and counts.

The primary outcome was evaluated by survival analysis with non-parametric distribution. Secondary outcomes were evaluated by Mann-Whitney-U or Wilcoxon signed-rank test for continuous variables (if non-normally distributed) or T Student test (if normally distributed). Fisher exact test and Chi-squared test served for categorical variables. Group tests were two-sided with $p < 0.05$ considered significant. All analyses were performed, and all figures generated using Python 3.5.3 ("G. van Rossum, Python tutorial, Technical Report CS-R9526, Centrum voor Wiskundeen Informatica (CWI), Amsterdam, May 1995.") and R 3.4.1 (R Core Team 2017, Vienna, Austria).

Ethics

Written informed consent was obtained by parents of all included patients. The study was approved by the Ethics Committee of the Sant'Anna-Regina Margherita Children's Hospital (Protocol no. 0000064, 02/01/2019).

RESULTS

On a total of 323 enrolled infants, 110 (34.1%) infants were managed according to CORE protocol. Therefore, the control group included a total of 213 (65.9%) infants that were managed following our unit discharge standards.

High-risk infants were 117 (36.2%), 47/110 (42.7%) exposed vs. 70/213 (32, 86%) not exposed to CORE protocol ($p = 0.081$). The main characteristics of the two study populations are summarized in **Table 1**.

During the first year after discharge 89/323 (27.55%) infants were readmitted in our hospital, 21/110 (19.09%) exposed vs. 68/213 (31.92%) not-exposed ($p = 0.018$) with a relative risk reduction of 42.61%. Overall, the main cause of readmission were respiratory disorders (47.19%) followed by gastrointestinal disorders (13.48%) and non-respiratory infections (7.87%).

Out of 117 high-risk infants enrolled, 33/117 (37.08%) of them were readmitted, 9/47 (19.15%) exposed vs. 24/68 (35.29%) not-exposed ($p = 0.059$) with a relative risk reduction of 54.26%.

TABLE 1 | Demographic and anthropometric characteristics of the study population. Data presented as median (IQR) or n (%).

All infants	**Exposed (110)**	**Not-exposed (213)**	***p*-value**
Gestational Age (weeks)	30.64 (28.43–32.29)	30.86 (28.71–32.29)	0.373
Birth weight (g)	1250 (967.5–1560)	1300 (995–1610)	0.370
Apgar at 1 min	6 (5–8)	7 (5–8)	0.732
Apgar at 5 min	8 (7–8)	8 (7–9)	0.187
Male	59 (53.6%)	91 (42.7%)	0.613
Cesarean section	79 (71.8%)	163 (76.5%)	0.324
High-risk infants	**Exposed (47)**	**Not-exposed (70)**	***p*-value**
Gestational Age (weeks)	28.0 (26.7–29.7)	27.6 (26.0–30.0)	0.612
Birth weight (g)	990 (790–1190)	990 (800–1200)	0.461
Apgar at 1 min	6 (5–8)	6 (4–7)	0.447
Apgar at 5 min	8 (7–8)	8 (7–8)	0.965
Male	26 (55.3%)	28 (40.0%)	0.113
Cesarean section	29 (61.7%)	112 (72.9%)	0.211
Low-risk infants	**Exposed (63)**	**Not-exposed (143)**	***p*-value**
Gestational Age (weeks)	31.9 (30.3–32.6)	31.7 (30.0–32.7)	0.378
Birth weight (g)	1490 (1180–1700)	1400 (1160–1700)	0.623
Apgar at 1 min	6 (5–8)	7 (5–8)	0.486
Apgar at 5 min	8 (7–8)	8 (7–9)	0.198
Male	33 (52.4%)	63 (44.1%)	0.290
Cesarean section	51 (81.0%)	112(78.3%)	0.627

42/323 (13.00%) infants were readmitted for respiratory disorders; 11/110 (10.00%) exposed vs. 31/213 (14.55%) not-exposed ($p = 0.24$) with a relative risk reduction of 31.29%. 11/323 (3.41%) were readmitted for apnea and/or apparent life-threatening events; 3/110 (2.73%) exposed vs. 8/213 (3.76%) not-exposed ($p = 0.63$) with a relative risk reduction of 27.39%.

Considering the primary outcome of the study (RHR in the first 12 months after discharge), we detected a significant reduction of RHR in the infants exposed to CORE protocol (log-rank $p < 0.05$), as it is clearly visible in the Kaplan-Meier hospital readmission free survival curve (**Figure 2**). This reduction was especially pronounced during the first 3 months after discharge during which 10/110 exposed (9.09%) vs. 46/213 not-exposed infants (21.6%) were readmitted to hospital ($p = 0.004$). The distribution of RHR, according to gestational age and time after discharge, is represented in **Figure 3A** and it shows a reduction of risk in the exposed group, especially in infants with GA < 28 weeks during the first 3 months after discharge. In this time period, we observed a reduction of hospital readmissions in the infants exposed to CORE (1.71%) vs. not-exposed (15.38%), especially in high-risk infants ($p = 0.002$) (**Figure 3B**).

Finally, LOS did not differ between exposed and not-exposed infants with a LOS of 43 (26–68) days vs. 39 (26–58) days, respectively ($p = 0.09$).

In exposed and not-exposed high-risk infants LOS was, respectively, 67 (51–83) and 58 (47–74) days ($p = 0.867$).

According to the CORE protocol, 47 (42.7%) of exposed infants were included in the high-risk group and underwent IM. At their first IM they had a median age of 61 (43–66) days, PMA of 36.6 (35.4–38.2) weeks and weight of 1,910 (1,700–2,110) grams. Thirteen infants (27.7%) needed a second IM because of an events-index higher than 0.5 in 5 (38.5%) cases, a CT90% higher than 3% in 2 (15.4%) cases and a combination of both in the remaining cases. All these 13 infants showed better results at the second IM. However, 6 (46.15%) infants had a second pathological IM and repeated a third IM that was normal in 2 (33.33%) infants. Four out of 47 infants were referred to the Pneumology Department and discharged home with CR monitor and/or oxygen. Cardiorespiratory data from 24 h instrumental IM recorded at the first IM and at discharge are reported in **Table 2**.

DISCUSSION

The American Academy of Pediatrics, in 2008 (12), and the Canadian Pediatric Society, in 2014 (13), outlined the discharge criteria most followed by the NICUs worldwide. One of the most important requirements identified by both the institutions is the detection of physiological stability from a cardiorespiratory perspective, however no literature defining cardiorespiratory stability in the newborn is available so far. Documenting an "apnea free period" (ranging from 3 to 10 days) (16–20) is one of the most common methods to assess CR stability before discharge. The AAP recommends that all preterm infants should undergo screening for CR events while in car seats before discharge (21, 22), as premature infants are at risk for desaturation and apnea when placed in upright car seats. Moreover, some studies focused on objective methodologies such as 24-h pulse oximetry (23–28) or polysomnography assessment (24, 26, 28). However, neither AAP nor CPS specify which should be considered the best technique to objectively assess the CR function and stability before discharging a preterm infant (12, 13).

In this context, we developed the pre-discharge CORE protocol to assess the readiness of infants to a safe discharged. Our results revealed that infants exposed to the pre-discharge CORE protocol had a significantly lower RHR, with the main gain in hospital readmission free-survival observed in the first 3 months after discharge. Furthermore, we observed a reduction of about 1/3 of hospitalizations for respiratory problems in infants exposed to the CORE protocol.

To our knowledge, few studies (23–29) evaluated the impact of a NICU pre-discharge CR stability monitoring on RHR. Recently, Chandrasekharan et al. (30) evaluated the impact of the implementation of a standardized protocol for the duration of observation in VLBW infants with apnea/bradycardia events before discharge over different epochs.

Although our study contains methodological differences with that of Chandrasekharan et al. (30), the main results are comparable, underlying how the application of a pre-discharge protocol based on the evaluation and management of CR events is effective in reducing RHR without extending LOS.

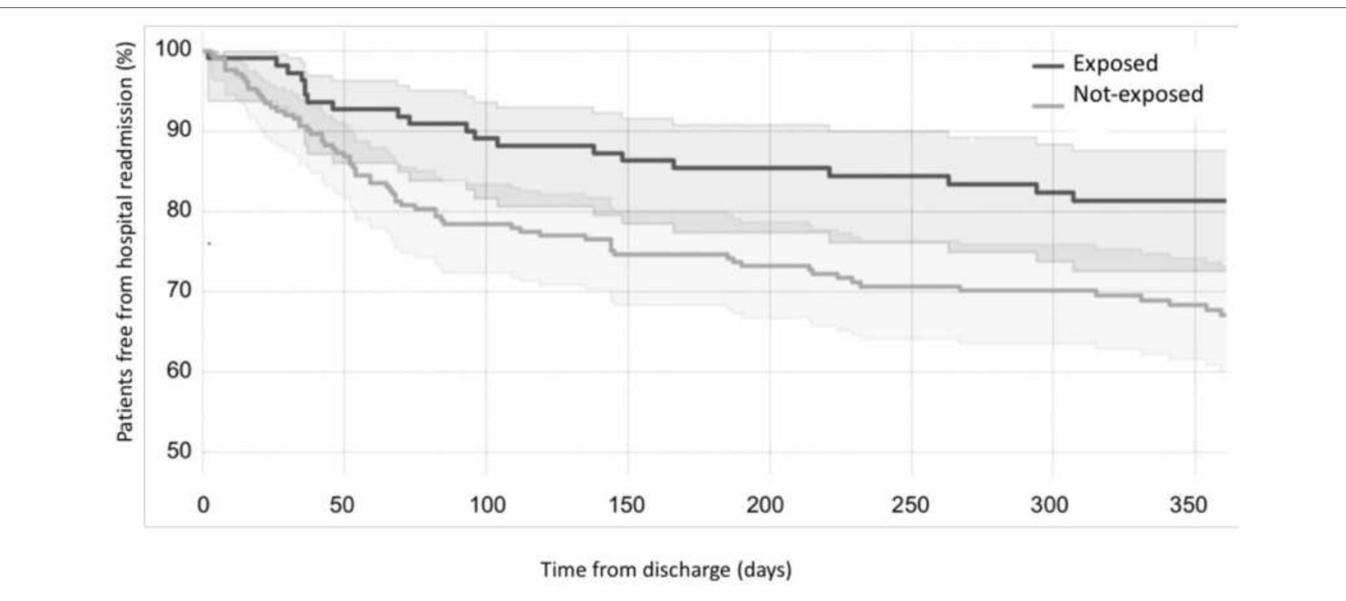


FIGURE 2 | Effects of CORE protocol on hospital readmission. Kaplan-Meier curves showing the proportion of individuals free from hospital readmission over time (black curve = exposed to CORE group; gray curve = not-exposed to CORE group).

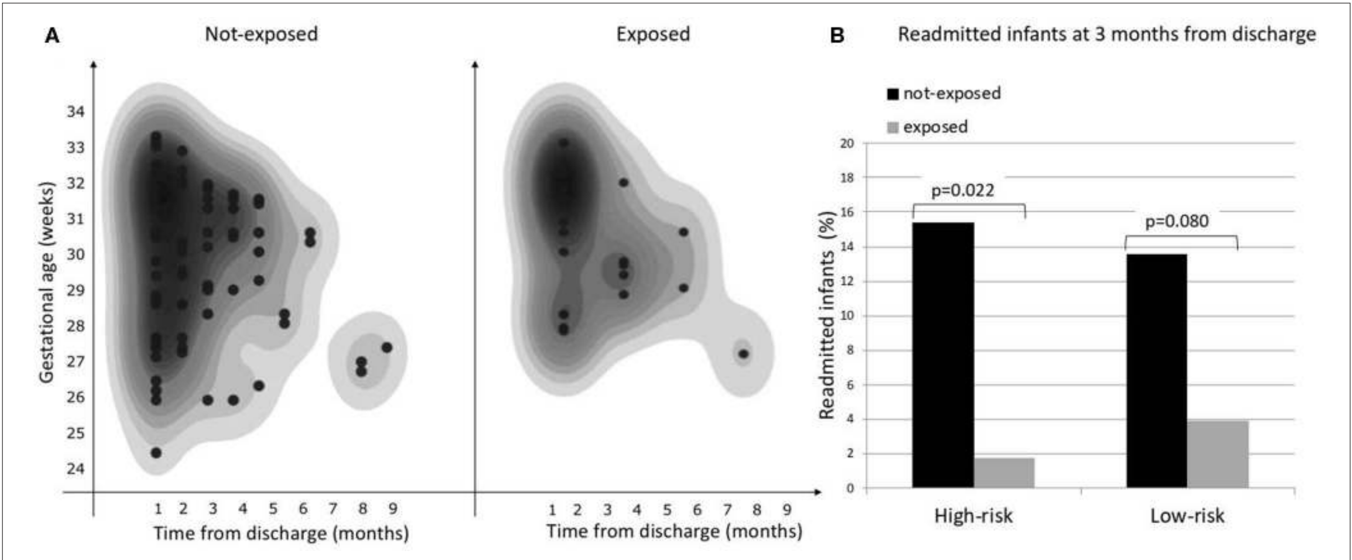


FIGURE 3 | (A) Overall risk of hospital readmission by gestational age and time from discharge, **(B)** Proportion of hospital readmissions during the first 3 months from discharge in exposed and not-exposed infants in high- and low-risk groups.

The main differences between our study and that of Chandrasekharan et al. (30) are due to the structure of the protocol, to the characteristics of the population studied, to the study design (comparison between different epochs), and to the technique used to detect CR events. The protocol proposed by Chandrasekharan et al. (30) is based on determining the duration of the "spell-free" period, by clinical observation, in a group of VLBW infants otherwise ready for discharge. Our protocol is a three-step process based on a clinical evaluation to screen infants ready for discharge, 24-h clinical observation (CO) to asses CR stability and 24-h instrumental monitoring (IM), performed after a normal CO, in high-risk infants only. The choice to perform IM after a normal CO in high-risk infants was based on the hypothesis that CO might not be sufficiently accurate to detect all CR events present in high-risk infants (19, 20, 23, 28). On the other hand, submitting these patients to IM only after obtaining a normal CO limited the number of repetitions of IM, which requires long reporting times by highly trained medical personnel.

TABLE 2 | Cardiorespiratory data from 24 h instrumental IM recorded at the first IM and at discharge in high-risk infants.

	CR data at the first IM (47 patients)		CR data at discharge (47 patients)	
	mean(SD)	median(IQR)	mean(SD)	median(IQR)
Age (days)	59.91 (26.86)	61.00 (44.00–76.00)	62.55 (28.53)	61.00 (45.00–76.50)
PMA (weeks)	36.73 (2.41)	36.60 (35.40–38.20)	37.10 (2.66)	36.60 (35.40–38.20)
Weight (grams)	1925.0 (289.3)	1910.0 (1703.8–2093.8)	1999.6 (328.7)	1977.5 (1757.5–2242.5)
All CR events (n.)	14.60 (24.92)	6.00 (3.50–13.00)	9.06 (11.97)	5.00 (2.00–9.50)
Apneas > 20 s (n.)	0.80 (2.40)	0.00 (0.00–0.00)	0.64 (1.98)	0.00 (0.00–0.00)
Desaturations (n.)	14.08 (29.15)	4.68 (0.53–10.31)	8.95 (15.54)	3.69 (0.00–8.14)
Bradycardia (n.)	1.17 (2.60)	0.00 (0.00–1.00)	0.96 (2.38)	0.00 (0.00–1.00)
Apneas and desaturations (n.)	0.58 (1.73)	0.00 (0.00–0.00)	0.29 (1.07)	0.00 (0.00–0.00)
Apneas and bradycardia (n.)	0.17 (0.76)	0.00 (0.00–0.00)	0.17 (0.59)	0.00 (0.00–0.00)
Desaturations and bradycardia (n.)	0.63 (1.65)	0.00 (0.00–0.00)	0.48 (1.55)	0.00 (0.00–0.00)
Complete events (n.)	1.01 (3.45)	0.00 (0.00–0.00)	0.25 (0.66)	0.00 (0.00–0.00)
Extreme events (n.)	0.00 (0.00)	0.00 (0.00–0.00)	0.00 (0.00)	0.00 (0.00–0.00)
CT90 (%)	2.32 (2.85)	1.15 (0.58–2.83)	1.55 (1.48)	0.95 (0.50–2.28)
Events index	0.77 (1.31)	0.34 (0.16–0.68)	0.49 (0.74)	0.23 (0.11–0.47)

High-risk infants consisted of a high number of severely preterm newborns most affected by apnea of prematurity and at risk of frequent short episodes of blood oxygen desaturation invisible during CO. The IM provides accurate information on blood oxygen saturation levels to calculate the CT90 which is important to establish whether supplemental oxygen is required. As a matter of fact, almost 28% of high-risk infants had pathological IM after a normal CO. This data suggest that CO alone should not be considered enough for a safe discharge home of high-risk infants (19, 20, 23, 28).

Despite this limitation, CO alone could contribute to evaluating CR stability as we found a reduction of RHR also in low-risk infants that did not undergo IM.

The main methodological issues of our study were the definition of CR events and the thresholds to define IM as pathological. We defined CR events according to Ramanathan et al. (31), but it is important to note that different authors (19, 20, 27, 32–35) used different criteria lacking a unanimous consensus.

We found that the most frequent event was blood oxygen desaturations while apnea, bradycardia and extreme CR event were much less present, and this is consistent with previously published data (32). However, the frequency of CR events that can be considered physiological in a preterm infant near discharge is not known. Arbitrarily, we considered pathological an event index > 0.5 events/h and a CT90 > 3%. We found this threshold effective in recognizing patients with cardiorespiratory instability reducing their RHR by delaying their discharge without significantly increase the LOS of the whole study population. It should be noted that the choice of lower threshold values might reduce or delay hospital readmission, but it could probably increase the number of patients with pathological IM dramatically increasing LOS of the whole study population.

Currently, there is no consensus to define a normal CR monitoring in near-discharge preterm infants (19, 24, 25, 28). However, IM performed in our study provide data from a group of near-discharge preterm infants with homogeneous post-menstrual age that appeared to be stable at the clinical judgment (**Table 2**) and we think that these data can represent first step toward reference values and that they can be useful to clinicians in reporting IM recordings.

The limitations of our study include the sample size that does not allow separate analysis of the causes of hospital readmission while it could be interesting to evaluate the protective effect of the exposition to CORE protocol to specific causes of hospital readmission, however, our data demonstrated a reduction of almost 1/3 of readmissions due to respiratory problems in the subjects exposed to the CORE protocol. Another limit of the study is that RHR could be underestimated in not-exposed infants. While readmission data of the exposed group were obtained from the emergency department registry of our hospital and matched with those obtained during the follow-up visits provided by the CORE protocol, readmission data of not-exposed infants were not. It is possible that a small part of the readmissions of not-exposed infants was lost because they were admitted to other hospitals. However, this bias would not have reduced the significance of our results, but rather increased the differences in RHR observed between the two groups.

CONCLUSION

This report described how the application of a new pre-discharge CR monitoring protocol could help clinicians in defining the best timing for NICU discharge. The main outcome achieved is a marked reduction in RHR, especially in the first months after discharge, for preterm infants to whom this protocol was applied without lengthening hospital stay. This study also underlies that CR maturation and stability should be assessed through instrumental monitoring and not only by simple observation strategies. However, further large multicenter and prospective trials evaluating this protocol are required before widespread distribution and application in NICUs.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comitato etico interaziendale della Città della Salute e della Scienza di Torino. The study was approved by the Ethics Committee of the Sant'Anna-Regina Margherita Children's Hospital (Protocol no. 0000064, 02/01/2019). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FC conceptualized and designed the study, drafted the initial manuscript, reviewed, and revised the manuscript. EC conceptualized and designed the study, performed acquisition and analysis of data, reviewed, and revised the manuscript. FL, EM, and MA performed cardiorespiratory monitoring analysis, drafted the initial manuscript, and reviewed and revised the manuscript. AP performed cardiorespiratory monitoring analysis, administered basic infant cardiopulmonary resuscitation, drafted the initial manuscript, and reviewed and revised the manuscript. AC, CP, and EB coordinated and supervised data collection and reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

The following investigators participated in this study: Cecilia Giusto, Martina Cavagnero, Vito Andrea Dell'Anna, Rossana Bagna, Silvia Maria Borgione, Pier Claudio Murru, Claudio Plazzotta.

REFERENCES

1. Institute of Medicine (US) Committee on Understanding Premature Birth and Assuring Healthy Outcomes. The National Academies Collection: Reports funded by National Institutes of Health. In: Behrman RE, and Butler AS, editors. *Preterm Birth: Causes, Consequences, and Prevention.* Washington, DC: National Academies Press (US) (2007). p. 1–4. Available online at: http://www.ncbi.nlm.nih.gov/books/NBK11362/
2. Sweet DG, Carnielli V, Greisen G, Hallman M, Ozek E, Plavka R, et al. European Consensus Guidelines on the Management of Respiratory Distress Syndrome - 2016 Update. *Neonatology.* (2017) 111:107–25. doi: 10.1159/000448985
3. Higgins RD, Jobe AH, Koso-Thomas M, Bancalari E, Viscardi RM, Hartert TV, et al. Bronchopulmonary dysplasia: executive summary of a workshop. *J Pediatr.* (2018) 197:300–8. doi: 10.1016/j.jpeds.2018.01.043
4. Battersby C, Santhalingam T, Costeloe K, Modi N. Incidence of neonatal necrotising enterocolitis in high-income countries: a systematic review. *Arch Dis Child Fetal Neonatal Ed.* (2018) 103:F182–9. doi: 10.1136/archdischild-2017-313880
5. Poryo M, Boeckh JC, Gortner L, Zemlin M, Duppré P, Ebrahimi-Fakhari D, et al. Ante-, peri- and postnatal factors associated with intraventricular hemorrhage in very premature infants. *Early Hum Dev.* (2018) 116:1–8. doi: 10.1016/j.earlhumdev.2017.08.010
6. Hellström A, Smith LEH, Dammann O. Retinopathy of prematurity. *Lancet Lond Engl.* (2013) 382:1445–57. doi: 10.1016/S0140-6736(13)60178-6
7. Rolland A, Shankar-Aguilera S, Diomandé D, Zupan-Simunek V, Boileau P. Natural evolution of patent ductus arteriosus in the extremely preterm infant. *Arch Dis Child Fetal Neonatal Ed.* (2015) 100:F55–8. doi: 10.1136/archdischild-2014-306339
8. Shane AL, Sánchez PJ, Stoll BJ. Neonatal sepsis. *Lancet Lond Engl.* (2017) 390:1770–80. doi: 10.1016/S0140-6736(17)31002-4
9. Underwood MA, Danielsen B, Gilbert WM. Cost, causes and rates of rehospitalization of preterm infants. *J Perinatol.* (2007) 27:614–9. doi: 10.1038/sj.jp.7211801
10. Kuzniewicz MW, Parker SJ, Schnake-Mahl A, Escobar GJ. Hospital readmissions and emergency department visits in moderate preterm, late preterm, and early term infants. *Clin Perinatol.* (2013) 40:753–75. doi: 10.1016/j.clp.2013.07.008
11. Merritt TA, Pillers D, Prows SL. Early NICU discharge of very low birth weight infants: a critical review and analysis. *Semin Neonatol.* (2003) 8:95–115. doi: 10.1016/S1084-2756(02)00219-1
12. American Academy of Pediatrics Committee on Fetus and Newborn. Hospital discharge of the high-risk neonate. *Pediatrics.* (2008) 122:1119–26. doi: 10.1542/peds.2008-2174
13. Jefferies AL, Canadian Paediatric Society, Fetus and Newborn Committee. Going home: Facilitating discharge of the preterm infant. *Paediatr Child Health.* (2014) 19:31–42. doi: 10.1093/pch/19.1.31
14. American Academy of Pediatrics. Committee on Fetus and Newborn. Hospital stay for healthy term newborns. *Pediatrics.* (2010) 125:405–9. doi: 10.1542/peds.2009-3119
15. Cresi F, Martinelli D, Maggiora E, Locatelli E, Liguori SA, Baldassarre ME, et al. Cardiorespiratory events in infants with gastroesophageal reflux symptoms: is there any association? *Neurogastroenterol Motil.* (2017) 30:e13278. doi: 10.1111/nmo.13278
16. Darnall RA, Kattwinkel J, Nattie C, Robinson M. Margin of safety for discharge after apnea in preterm infants. *Pediatrics.* (1997) 100:795–801. doi: 10.1542/peds.100.5.795
17. Zupancic JAF, Richardson DK, O'Brien BJ, Eichenwald EC, Weinstein MC. Cost-effectiveness analysis of predischarge monitoring for apnea of prematurity. *Pediatrics.* (2003) 111:146–52. doi: 10.1542/peds.111.1.146
18. Lorch SA, Srinivasan L, Escobar GJ. Epidemiology of apnea and bradycardia resolution in premature infants. *Pediatrics.* (2011) 128:e366–73. doi: 10.1542/peds.2010-1567
19. Amin SB, Burnell E. Monitoring apnea of prematurity: validity of nursing documentation and bedside cardiorespiratory monitor. *Am J Perinatol.* (2013) 30:643–8. doi: 10.1055/s-0032-1329694
20. Vergales BD, Paget-Brown AO, Lee H, Guin LE, Smoot TJ, Rusin CG, et al. Accurate automated apnea analysis in preterm infants. *Am J Perinatol.* (2014) 31:157–62. doi: 10.1055/s-0033-1343769
21. Narvey MR, Canadian Paediatric Society, Fetus and Newborn Committee. Assessment of cardiorespiratory stability using the infant car seat challenge before discharge in preterm infants (<37 weeks' gestational age). *Paediatr Child Health.* (2016) 21:155–62. doi: 10.1093/pch/21.3.155
22. Davis NL. Screening for cardiopulmonary events in neonates: a review of the infant car seat challenge. *J Perinatol.* (2015) 35:235–40. doi: 10.1038/jp.2015.1
23. Wellington G, Elder D, Campbell A. 24-hour oxygen saturation recordings in preterm infants: editing artefact. *Acta Paediatr.* (2018) 107:1362–9. doi: 10.1111/apa.14293
24. Roberts T, Campbell A, Larsen P, Elder D. Preterm infants at discharge: nap polysomnography versus 24-hour oximetry. *Acta Paediatr.* (2017) 106:1754–9. doi: 10.1111/apa.13900
25. Wellington G, Campbell AJ, Elder DE. Oximetry for preterm infants at neonatal discharge: what is current practice in New Zealand and Australia? *J Paediatr Child Health.* (2016) 52:333–7. doi: 10.1111/jpc.13079
26. Elder DE, Whale J, Galletly D, Campbell AJ. Respiratory events in preterm infants prior to discharge: with and without clinically concerning apnoea. *Sleep Breath Schlaf Atm.* (2011) 15:867–73. doi: 10.1007/s11325-010-0457-x

27. Subhani M, Katz S, DeCristofaro JD. Prediction of postdischarge complications by predischarge event recordings in infants with apnea of prematurity. *J Perinatol.* (2000) 20:92–5. doi: 10.1038/sj.jp.7200320
28. Razi NM, Humphreys J, Pandit PB, Stahl GE. Predischarge monitoring of preterm infants. *Pediatr Pulmonol.* (1999) 27:113–6. doi: 10.1002/(SICI)1099-0496(199902)27:2<113::AID-PPUL7>3.0.CO;2-O
29. Lamarche-Vadel A, Blondel B, Truffer P, Burguet A, Cambonie G, Selton D, et al. Re-hospitalization in infants younger than 29 weeks' gestation in the EPIPAGE cohort. *Acta Paediatr*. (2004) 93:1340–5. doi: 10.1111/j.1651-2227.2004.tb02934.x
30. Chandrasekharan P, Rawat M, Reynolds AM, Phillips K, Lakshminrusimha S. Apnea, bradycardia and desaturation spells in premature infants: impact of a protocol for the duration of 'spell-free' observation on interprovider variability and readmission rates. *J Perinatol.* (2018) 38:86–91. doi: 10.1038/jp.2017.174
31. Ramanathan R, Corwin MJ, Hunt CE, Lister G, Tinsley LR, Baird T, et al. Cardiorespiratory events recorded on home monitors: Comparison of healthy infants with those at increased risk for SIDS. *JAMA*. (2001) 285:2199–207. doi: 10.1001/jama.285.17.2199
32. Di Fiore JM, Arko MK, Miller MJ, Krauss A, Betkerur A, Zadell A, et al. Cardiorespiratory events in preterm infants referred for apnea monitoring studies. *Pediatrics.* (2001) 108:1304–8. doi: 10.1542/peds.108.6.1304
33. Corvaglia L, Martini S, Aceti A, Capretti MG, Galletti S, Faldella G. Cardiorespiratory events with bolus versus continuous enteral feeding in healthy preterm infants. *J Pediatr*. (2014) 165:1255–7. doi: 10.1016/j.jpeds.2014.07.043
34. Di Fiore JM, Poets CF, Gauda E, Martin RJ, MacFarlane P. Cardiorespiratory events in preterm infants: etiology and monitoring technologies. *J Perinatol.* (2016) 36:165–71. doi: 10.1038/jp.2015.164
35. Poets CF, Stebbens VA, Richard D, Southall DP. Prolonged episodes of hypoxemia in preterm infants undetectable by cardiorespiratory monitors. *Pediatrics.* (1995) 95:860–3.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Cresi, Cocchi, Maggiora, Pirra, Logrippo, Ariotti, Peila, Bertino and Coscia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Journal of the Neurological Sciences 415 (2020) 116940

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

journal homepage: www.elsevier.com/locate/jns

Clinical short communication

Valproate impact and sex-dependent seizure remission in patients with idiopathic generalized epilepsy

Emanuele Cerulli Irelli[a], Enrico Cocchi[b], Alessandra Morano[a], Sara Casciato[c], Martina Fanella[a], Mariarita Albini[c], Giacomo Fisco[a], Francesca A. Barone[a], Biagio Orlando[a], Addolorata Mascia[c], Mario Manfredi[a], Jinane Fattouch[a], Anna T. Giallonardo[a], Giancarlo Di Gennaro[c], Carlo Di Bonaventura[a],*

[a] Epilepsy Unit, Department of Human Neurosciences, Policlinico Umberto I, "Sapienza" University, Rome, Italy
[b] Department of Public Health and Pediatrics, University of Turin, Turin, Italy
[c] IRCCS "NEUROMED", Pozzilli, IS, Italy

ARTICLE INFO

Keywords:
Gender medicine
Genetic generalized epilepsy (GGE)
Valproate restriction
IGE prognosis
Childbearing age

ABSTRACT

Background: Idiopathic Generalized Epilepsy (IGE) management has become increasingly challenging due to the restricted use of Valproate (VPA) in females. The aim of the study was to assess possible differences in terms of seizure outcome between men and women suffering from IGE.
Methods: A cohort of IGE patients (age range: 13–50 years) followed from 1980 to 2018 were included. Their medical history was retrospectively reviewed to investigate possible factors influencing seizure outcome. Seizure Remission (SR) was defined as the absence of any seizure type over 18 months prior to the last medical observation. The primary outcome was to evaluate sex differences in terms of SR at last observation. A multivariable logistic regression model was elaborated using SR as dependent variable.
Results: Three-hundred and sixty patients were included, 204 (56.7%) of whom were women. The median age at the end of follow-up was 30. At last medical observation, fewer women were receiving VPA compared with men (females 39.7% vs males 79.5%, $p < .001$). Overall SR was 70.6%. SR was significantly different according to sex (females 62.3% vs males 81.4%, $p < .001$). Multivariable logistic regression model showed that female sex (Odds Ratios [OR] = 0.52, 95% Confidence Interval [CI] = 0.29–0.94; $p = .03$), VPA treatment at last observation (OR = 0.44, 95% CI = 0.25–0.76; $p = .003$) and epilepsy syndrome ($p < .001$) were the factors independently associated with SR.
Conclusions: Recent modifications in VPA prescribing patterns may have determined a worse seizure control among IGE female patients. Comparative clinical trials assessing the best therapeutic options for women with childbearing potential are urgently needed.

1. Introduction

Idiopathic Generalized Epilepsy (IGE) represents one of the most common form of epilepsy and it is generally associated with a favorable prognosis [1]. Valproate (VPA) has always been considered the cornerstone of IGE treatment [2,3]; however, over the past years its use has been dramatically limited in women of childbearing potential due to its well-known teratogenic effects [4]. Consequently, ethical concerns about the prognostic implications of VPA avoidance in IGE female patients have been raised by several authors and expert task forces [5]. A recent work from our group focusing on IGE female patients during their reproductive age found that those not receiving VPA were less likely to achieve seizure control [6]. The aim of this study was to assess possible sex-related differences in seizure outcome in a population of IGE patients.

2. Methods

This is a cohort retrospective study performed according to STROBE guidelines. We retrospectively collected clinical data of patients followed at the Policlinico 'Umberto I' epilepsy outpatient clinic and the epilepsy unit of 'Neuromed' Institute from 1980 to 2018, through a

* Corresponding author at: Epilepsy Unit, Department of Human Neurosciences, Policlinico "Umberto I", "Sapienza" University, Viale dell'Università 30, 00185 Rome, Italy.
E-mail address: c_dibonaventura@yahoo.it (C. Di Bonaventura).

https://doi.org/10.1016/j.jns.2020.116940
Received 27 March 2020; Received in revised form 6 May 2020; Accepted 19 May 2020
Available online 30 May 2020
0022-510X/ © 2020 Elsevier B.V. All rights reserved.

careful revision of medical charts and a computerized database. Female patients during their reproductive age (15–50 years old) and age-matched male patients were considered eligible for the study. Then we enrolled the subjects who met the following inclusion criteria: 1) diagnosis of IGE according to ILAE criteria; 2) availability of a complete clinical documentation and ≥ 1 EEG recording; 3) a follow-up duration ≥18 months. For each patient we collected demographic data, family history of epilepsy in first- or second-degree relatives, history of febrile seizures, psychiatric comorbidities, age of epilepsy onset, follow-up duration, seizure type, possible triggering factors, EEG features (including photosensitivity) and prior/concomitant drug regimens. According to commonly accepted criteria, the diagnosis of IGE was confirmed by three trained epileptologists (CIE, DBC, GAT) who independently revised the patients' clinical and EEG data. According to the ILAE classification, the following epileptic syndromes were identified: Childhood Absence Epilepsy (CAE), Juvenile Absence Epilepsy, Juvenile Myoclonic Epilepsy, IGE with generalized tonic-clonic seizures only (IGE-GTCS) [7]. In order to achieve a more detailed clinical characterization we also considered other clinical entities not yet included in ILAE classification: [8] namely, Eyelid myoclonia with absences (accordingly with the criteria proposed by Striano et al.) [9] and IGE-undefined (the latter definition was only applied to not otherwise classified cases).

We defined Seizure Remission (SR) as the absence of any seizure type (i.e. myoclonic, absence and generalized tonic-clonic seizures - GTCS) over 18 months prior to the last medical observation.

The primary outcome of the study was to investigate sex-related differences in terms of SR. The possible influence of other demographic and clinical variables, with particular regard to VPA use at last medical observation, was subsequently analyzed. The study was approved by the Policlinico Umberto I Research Ethics Committee.

2.1. Statistical analysis

Data were tested for normal distribution using the Shapiro–Wilk test. Normal continuous data are presented as mean ± standard deviation (SD) and non-normal continuous data are presented as median [interquartile range (IQR)]. Continuous normal distributed data were compared across relevant groups using the Student *t*-test for normal distributed variables and Mann–Whitney *U* test or Wilcoxon signed-rank test for the non-normal distributed ones. Categorical variables were presented as frequency (count) and compared through the Fisher exact test. Group tests were two-sided with $P < .05$ considered statistically significant. In order to assess the independent effect on seizure outcome of sex and VPA treatment, and to assess their independence with respect to possible confounders, a multivariable logistic regression model was elaborated. SR at final observation was used as dependent variable and VPA treatment at the end of the follow-up, sex, history of febrile seizures, psychiatric comorbidity, photosensitivity and family history of epilepsy (1st or 2nd relative degree) were used as covariates and considered as nominal dichotomous variables. Age was included as continuous variable. Specific epileptic sub-syndromes were included in the model, considering undefined-IGE as reference. Categorized age at epilepsy onset (1–5 years vs 6–19 years vs ≥20 years) was also included in the model, with age 6–19 years as reference.

We therefore decided to deepen the statistical analysis elaborating a second multivariable logistic regression model (hereinafter referred to as "Combined Regression Model") in order to assess the combined effect of the variables "Sex" and "VPA treatment at last medical observation" on SR, with the above mentioned covariates. We obtained 4 possible categories: VPA(−)/Female sex, VPA(+)/Female sex, VPA(−)/Male sex, VPA(+)/Male sex. This Combined Regression Model gave us the chance to observe whether the combination of these variables could predict the outcome in a stronger way than their single independent effect.

Analyses were performed and figures generated using R 3.5.1 (R Project for Statistical Computing, Vienna, Austria).

Table 1
Demographic and clinical characteristics according to sex.

	Male (156 pts)	Female (204 pts)	*p* value
Age, years, median (IQR)	30 (18)	30 (18)	0.78
Follow-up duration, years, median (IQR)	16 (20)	16 (19)	0.73
Psychiatric Comorbidity, n, %	26, 12.7%	11, 7,1%	0.08
History of Febrile Seizures, n, %	18, 8.8%	19, 12.2%	0.3
Familiarity for epilepsy in 1st or 2nd relative	37, 23.7%	58, 28.4%	0.3
Type of epilepsy			
Childhood absence epilepsy, n, %	22, 14.1%	11, 5.4%	0.005*
Juvenile absence epilepsy, n, %	12,7.7%	25, 12.3%	0.16
Juvenile myoclonic epilepsy, n, %	50, 32.1%	88, 43.1%	0.03*
Eyelid myoclonia with absences, n, %	12, 7.7%	20, 9.8%	0.57
IGE with generalized tonic-clonic seizures only, %	37, 23.7%	43, 21.1%	0.61
IGE-Undefined, n, %	23, 14.7%	17, 8.3%	0.06
EEG features			
Photoparoxysmal response, n, %	61, 29.9%	44, 28.2%	0.49
AED regimen			
AED attempted during follow-up, median (IQR), n	2 (2)	2 (2)	0.08
AED at final observation, median (IQR), n	1(1)	1(1)	0.45
Patients on monotherapy, n, %	122, 59.8%	97, 62.2%	0.66

3. Results

We retrospectively reviewed data from 550 IGE patients, 360 of whom were enrolled in the study according to the inclusion criteria. Two-hundred-four of them (56.7%) were women. The year at first observation was 1980–1989 in 53 patients, 1990–1999 in 79 subjects, 2000–2009 in 93 and 2010–2018 in the remaining 135 cases. The last follow-up visit was performed during 2018 in 310/360 subjects (86.1%), and no patient ended follow-up before 2014. The population median age at the last visit was 30 (IQR 18). Further demographic, clinical, EEG and therapeutic data are summarized in the table. (Table 1)

At last medical observation, the number of patients who were taking VPA was significantly different in relation to sex (81/204 women −39.7%- vs 124/156 men −79.5%-, $p < .001$). In female patients still receiving VPA, its daily mean dose was lower compared with men (793.13 mg Standard Deviation -SD- ± 276.36 vs 1242.75 mg SD ± 414.35, $p < .001$). VPA blood levels were available in 300/360 patients at last observation, and were also significantly different according to sex (men 67.37 SD ± 21.79 vs women 54.28 SD ± 16.46, $p = .01$)

Among subjects not treated with VPA at last observation, the most used AEDs were Levetiracetam (79/155 patients, 50.1%), Lamotrigine (49/155 patients, 31.6%), Topiramate (19/155 patients, 12.3%), Clonazepam (18/155, 11.6%), Phenobarbital (14/155, 9%), Zonisamide (13/155, 8.4%), Perampanel (10/155, 6.4%), Ethosuximide (10/155, 6.4%), others AEDs (13/155, 8.4%).

The overall SR was 70.6% (254/360 patients). Among the subjects who continued experiencing seizures at last medical observation, 41/106 (38.7%) had GTCS. The rate of SR showed significant differences between sexes (127/204 women −62.3%- vs 127/156 men −81.4%-, $p < .001$), as well as the rate of freedom from GTCS over the last 18 months at last medical observation (174/204 women −85.3%- vs 145/156 men −92.9%-, $p = .03$).

The multivariable logistic regression model showed that sex (Odds Ratios [OR] = 0.52, 95% Confidence Interval [CI] = 0.29–0.94; p = .03), VPA treatment at final observation (OR = 0.44, 95% CI = 0.25–0.76; $p = .003$) and epileptic syndrome were the only

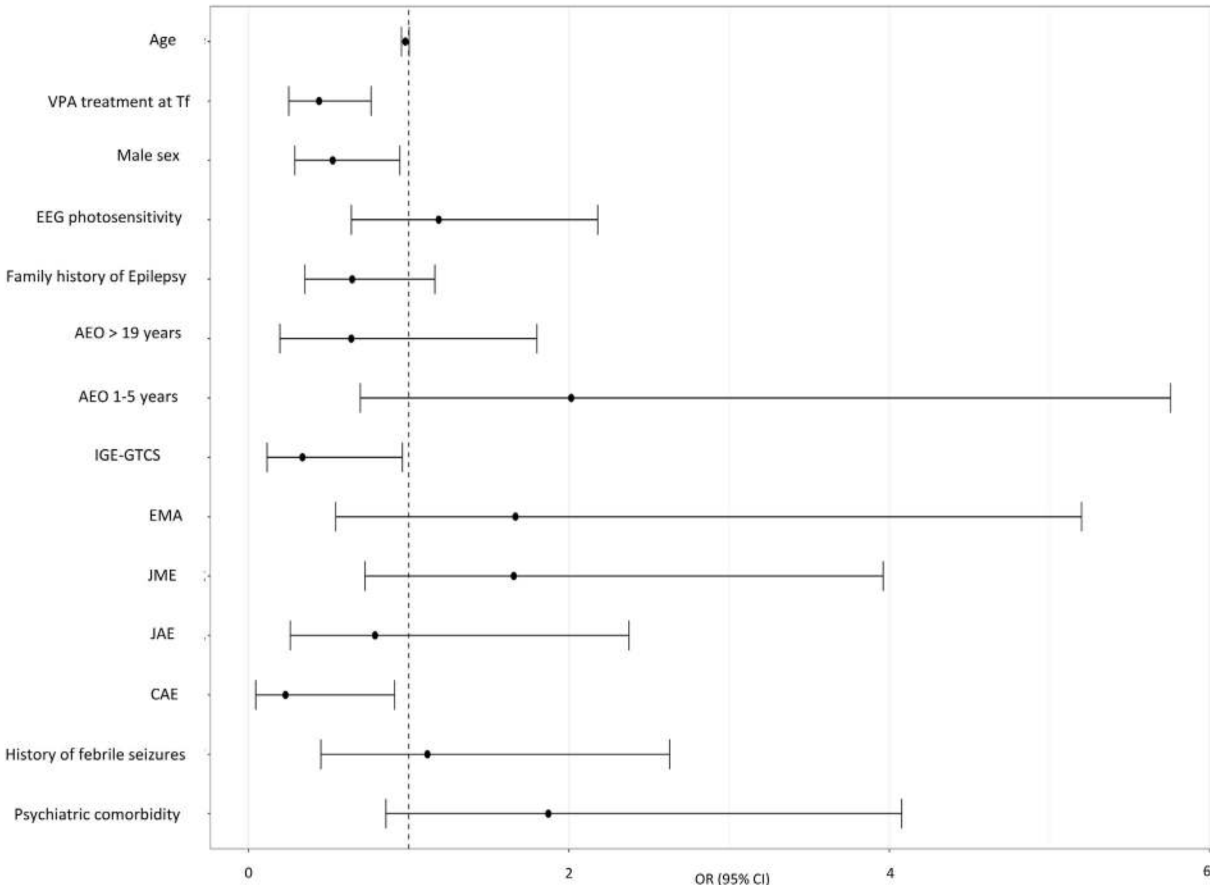


Fig. 1. Forest plot summary of logistic regression analysis for seizure remission at the end of follow up. This figure is a graphical representation of the Odds Ratio (OR) and relative 95% Confidence Interval (CI) of all independent variables included in the logistic regression model. OR (black dots) and the corresponding 95% CIs (black lines) for each covariate are shown. Abbreviations: AEO = Age of Epilepsy Onset; CAE = Childhood Absence Epilepsy; EMA = Eyelid Myoclonia with Absences; IGE-GTCS = Idiopathic Generalized Epilepsy with Tonic-Clonic Seizures Only; JAE = Juvenile Absence Epilepsy; JME = Juvenile Myoclonic Epilepsy.

factors which significantly affected SR. As to this, IGE-GTCS syndrome (OR = 0.33, 95% CI = 0.12–0.96; p = .042) and CAE (OR = 0.23, 95% CI = 0.05–0.91; p = .049) were found to be associated with a better outcome compared with the reference syndrome. The results of the logistic regression model are illustrated in Fig. 1. The results of the Combined Regression Model, showing the combined effect on SR of the variables “sex” and “VPA treatment at last medical observation”, are illustrated in Fig. 2.

4. Discussion

To the best of our knowledge, this is the first study exploring possible sex-related differences in seizure outcome after the recent changes in VPA prescribing pattern. The overall prognosis observed in our population was comparable to that reported in previous papers [10]. We found that seizure control at last medical observation was significantly worse among female patients of child-bearing potential compared with age-matched men. According to the logistic regression analysis, the most relevant clinical factor underlying such difference was the far more limited use of VPA among women with respect to men. This finding confirms that the choice to avoid VPA in IGE has a remarkable impact on the patients' chance of achieving seizure freedom [3,4]. However, female sex was found to be slightly associated with a worse outcome independently of the use of VPA. Such evidence could be partially explained by the lower dose of VPA, matched by lower serum levels, used by female patients still taking it in order to minimize endocrinological, teratogenic and cosmetic adverse effects [11,12]. Furthermore, it cannot be excluded a possible influence of sex-specific genetic or hormonal factors on seizure recurrence (e.g. catamenial seizures). Indeed, a small and non-significant trend towards a worse seizure control among women still taking VPA has been previously reported in literature [13].

We also found that IGE GTCS-only and CAE were associated with a more favorable clinical outcome. Such difference with myoclonic syndromes (in particular JME), aside from the possible intrinsic prognostic differences among IGE syndromes [1,14], might be explained by the more critical role of VPA in controlling myoclonic seizures than absence seizures and GTCS, which, in fact, might benefit from a broader range of alternative AEDs [3].

Our study has several limitations, first of all, its retrospective design, which particularly implies selection and recall biases. Besides, the wide study timespan prevented us from thoroughly assessing the effects of recent regulatory agencies' recommendations, not to mention the impact of new prescribing patterns including the latest AEDs.

Despite these limitations, our data confirmed the crucial role of VPA in the management of IGE patients, as demonstrated by the sex-dependent differences in SR. Consequently, our results highlighted once more the urgent need for comparative clinical trials tailored to IGE female patients with the aim to evaluate the most effective and tolerable therapeutic alternatives to VPA. Moreover, our study supports the necessity to extend our knowledge about the safety of the newest AEDs during pregnancy.

Funding

None.

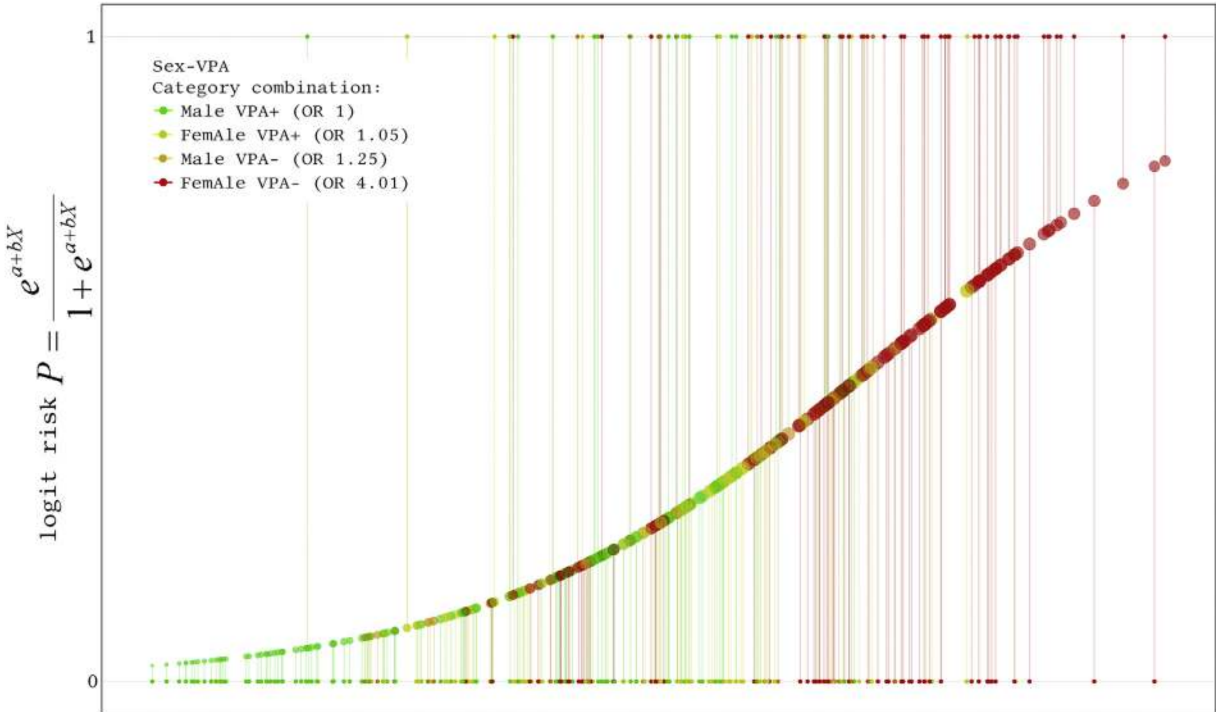


Fig. 2. The combination of Sex and VPA treatment is strongly predictive of seizure outcome at last observation.
The figure shows a graphical representation of the Combined Regression Model elaborated to assess the strength of the combined effect on seizure remission (SR) of the variables "sex" and "Valproate (VPA) treatment at final observation (Tf)". Four possible categories were obtained [VPA(−)/Female sex, VPA(+)/Female sex, VPA(−)/Male sex, VPA(+)/Male sex] and the dots are colored according to sex and VPA treatment at Tf. The dots lying on the horizontal line at level 1 of y-axis correspond to the patients still experiencing seizures, those lying on level 0 of y-axis are the SR patients. The dots aligned along the S-shaped curve between 0 and 1 (on vertical axis) represent the estimated risk of each subject based on the predictor variable values, transformed by the link logit function (represented on the Y axis) that characterizes binomial distribution. Each combination of predictor variables defines an estimated relative risk for each patient that, thanks to a logit transformation, lies between 0 (no risk) and 1 (full risk) assuming an asymptotic distribution. The vertical lines on the graph connect the real status of each patient with its respective estimated risk, representing the residual of each combination. As clearly visible, the model correctly predicts that the "VPA-/Female sex" combination is associated with the highest risk of uncontrolled seizures at Tf (being the closest to 1 on y-axis), as testified by an Odd Ratio [OR] = 4.01 95% Confidence Interval [CI] = 2.15–7.69. Groups with intermediate estimated risk tend to be either VPA(+)/Male (OR = 1.25, 95% CI = 0.58–2.67) or VPA(+)/Female (OR = 1.05, 95% CI = 0.34–2.94). The VPA(+)/Male group is instead the closest to 0 on y-axis (lowest risk).

Data availability statement

Data are available upon reasonable request. Individual patient data will not be shared to conform with the privacy statement signed by the participants. Pseudononymised data may be shared upon request with the corresponding author.

Declaration of Competing Interest

None.

Acknowledgments

None.

References

[1] U. Seneviratne, M. Cook, W. D'Souza, The prognosis of idiopathic generalized epilepsy, Epilepsia. 53 (12) (2012) 2079–2090.
[2] A.G. Marson, A.M. Al-Kharusi, M. Alwaidh, et al., The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial, Lancet. 369 (9566) (2007) 1016–1026.
[3] T. Tomson, D. Battino, E. Perucca, Valproic acid after five decades of use in epilepsy: time to reconsider the indications of a time-honoured drug, Lancet Neurol. 15 (2) (2016) 210–218.
[4] K. Puteikis, I. Medžiaušaitė, R. Mameniškienė, Valproate utilisation trends among girls and women from 2013 to 2018, Seizure. 70 (2019) 77–81.
[5] T. Tomson, A. Marson, P. Boon, et al., Valproate in the treatment of epilepsy in girls and women of childbearing potential, Epilepsia. 56 (7) (2015) 1006–1019.
[6] E. Cerulli Irelli, A. Morano, E. Cocchi, et al., Doing without valproate in women of childbearing potential with idiopathic generalized epilepsy: implications on seizure outcome, Epilepsia. 61 (1) (2020) 107–114.
[7] I.E. Scheffer, S. Berkovic, G. Capovilla, et al., ILAE classification of the epilepsies: position paper of the ILAE commission for classification and terminology, Epilepsia. 58 (4) (2017) 512–521.
[8] C.P. Panayiotopoulos, Syndromes of idiopathic generalized epilepsies not recognized by the international league against epilepsy, Epilepsia. 46 (Suppl. 9) (2005) 57–66.
[9] S. Striano, P. Striano, C. Nocerino, et al., Eyelid myoclonia with absences: an overlooked epileptic syndrome? Neurophysiol. Clin. 32 (5) (2002) 287–296.
[10] F. Semah, M.C. Picot, C. Adam, et al., Is the underlying cause of epilepsy a major prognostic factor for recurrence? Neurology. 51 (5) (1998) 1256–1262.
[11] R1. Gürer, Ş2. Aydın, Ç. Özkara, Outcomes of low-dose valproic acid treatment in patients with juvenile myoclonic epilepsy, Seizure. 70 (2019) 43–48.
[12] J. Miró, M. Aiguabella, M. Veciana, et al., Low-dose sodium valproate in the treatment of idiopathic generalized epilepsies, Acta Neurol. Scand. 129 (5) (2014) e20–e23.
[13] W.F. Arts, O.F. Brouwer, A.C. Peters, et al., Course and prognosis of childhood epilepsy: 5-year follow-up of the Dutch study of epilepsy in childhood, Brain 127 (8) (2004) 1774–1784.
[14] A.T. Berg, S. Shinnar, S.R. Levy, et al., Two-year remission and subsequent relapse in children with newly diagnosed epilepsy, Epilepsia. 42 (12) (2001) 1553–1562.

Am J Physiol Renal Physiol 318: F486–F495, 2020.
First published December 23, 2019; doi:10.1152/ajprenal.00404.2019.

RESEARCH ARTICLE

Acute and chronic glomerular damage is associated with reduced CD133 expression in urinary extracellular vesicles

Veronica Dimuccio,[1] Licia Peruzzi,[2] Maria Felice Brizzi,[3] Enrico Cocchi,[3] Fabrizio Fop,[4] Alberto Boido,[1] Maddalena Gili,[3] Sara Gallo,[3] Luigi Biancone,[3,4] Giovanni Camussi,[3] and Benedetta Bussolati[1]

[1]*Department of Molecular Biotechnology and Health Sciences, University of Turin, Turin, Italy;* [2]*Pediatric Nephrology Unit, Regina Margherita Children's Hospital, Città della Salute e della Scienza di Torino, Turin, Italy;* [3]*Department of Medical Sciences, University of Turin, Turin, Italy; and* [4]*Division of Nephrology Dialysis and Transplantation, Città della Salute e della Scienza di Torino, Turin, Italy*

Submitted 30 August 2019; accepted in final form 17 December 2019

Dimuccio V, Peruzzi L, Brizzi MF, Cocchi E, Fop F, Boido A, Gili M, Gallo S, Biancone L, Camussi G, Bussolati B. Acute and chronic glomerular damage is associated with reduced CD133 expression in urinary extracellular vesicles. *Am J Physiol Renal Physiol* 318: F486–F495, 2020. First published December 23, 2019; doi:10.1152/ajprenal.00404.2019.—Extracellular vesicles released into urine (uEVs) can represent interesting biomarkers of renal cell damage. CD133, a stem/progenitor cell marker expressed by renal progenitor cells, is highly expressed in uEVs of healthy individuals. In the present study, we evaluated the level of CD133 in the uEVs of patients with acute and chronic glomerular damage by cytofluorimetric analysis. The level of CD133$^+$ uEVs was significantly decreased in pediatric patients with acute glomerulonephritis during the acute phase of renal damage, while it was restored after the subsequent recovery. A similar decrease was also observed in patients with chronic glomerulonephritis. Moreover, CD133$^+$ uEVs significantly declined in patients with type 2 diabetes, used as validation group, with the lowest levels in patients with albuminuria with diabetic nephropathy. Indeed, receiver-operating characteristic curve analysis indicates the ability of CD133$^+$ uEV values to discriminate the health condition from that of glomerular disease. In parallel, a significant decrease of CD133 in renal progenitor cells and in their derived EVs was observed in vitro after cell treatment with a combination of glucose and albumin overload, mimicking the diabetic condition. These data indicate that the level of CD133$^+$ uEVs may represent an easily accessible marker of renal normal physiology and could provide information on the "reservoir" of regenerating cells within tubules.

CD133; extracellular vesicles; prominin; regeneration; renal damage; urine

INTRODUCTION

The assessment and diagnosis of renal function during acute and chronic pathologies affecting the renal compartments are, at present, mainly based on the levels of retention markers such as creatinine and urea. In addition, alternative blood and urinary biomarkers appear to be precious tools that can refer not only to the degree of renal damage but also to the pathophysiological state of the kidney. In particular, urine detection markers already used in the clinical practice (13), such as neutrophil gelatinase-associated lipoprotein and cystatin C, appear as the most promising ones as they directly derive from the renal compartment and from cells of the urinary tract. In addition, urine is a readily available sample easy to collect without the need of any invasive procedure.

Extracellular vesicles (EVs) are round vesicles released in biological fluids by almost all types of cells, and they contain proteins and other molecules directly correlated with the tissue from which they originate. Particularly, urinary EVs (uEVs) carry molecules that are characteristic of the epithelial cells present in the whole length of the urinary tract (23, 25). Accordingly, uEVs are attracting increasing interest as potential urinary biomarkers, since they may be a precious source of information related to the pathophysiological state of renal tissue. Indeed, data have been so far provided on the role of uEVs as biomarkers of glomerular diseases, including diabetes, tubular damage, and renal fibrosis (4, 22, 27, 28).

The CD133 molecule represents a marker of a population of tubular cells, intercalated between other epithelial cells, with the ability to survive after damage and proliferate in response to cell injury, suggesting a role of this protein as marker of renal regenerative capability (3). CD133 itself may play a role in the increased proliferation of regenerating clonal tubular cells in acutely damaged kidney through the permissive effect on the activation of the Wnt/β-catenin pathway (2). Interestingly, we have previously reported that CD133, which is expressed at high level by the uEVs of healthy individuals, was reduced in uEVs of transplanted patients with acute tubular necrosis occurring after a renal transplantation and increased in the first week following transplant (7). These results suggest that the level of CD133$^+$ uEVs can mirror the regenerative state of the renal tubular compartment in healthy and pathological conditions. However, no data are currently available on CD133$^+$ uEV levels in patients with acute and chronic glomerular damage.

It is well established that glomerular damage affects the functionality of the tubular compartment, proteinuria being a main prognostic factor for interstitial fibrosis (11, 12). Indeed, tubular atrophy together with proteinuria represents the most reliable predictors of renal function in glomerular diseases (18, 30, 31).

Therefore, the aim of the present study was to verify if variation of CD133$^+$ uEV levels, previously observed in a

Address for reprint requests and other correspondence: B. Bussolati, Molecular Biotechnology Centre, Univ. of Torino, via Nizza 52, Torino 10126, Italy (e-mail: benedetta.bussolati@unito.it).

 1931-857X/20 Copyright © 2020 the American Physiological Society

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

Table 1. *Clinical characteristics of patients with chronic and acute glomerular nephropathy*

Patient Group	n	Sex (Male/Female)	Median Age, yr	Clinical Features	Hematuria Degree	Timing of Urine Collection	Serum Creatinine, mg/dL	Urinary Protein on Creatinine Ratio, mg/mg	Number of Extracellular Vesicles per mL of Urine
Healthy	7	3/4	12 (9–13)		−		0.52 ± 0.06	0.13 ± 0.01	$1.24\ e^{+10} \pm 2.05\ e^{+9}$
Chronic glomerulonephritis	12	11/1	9 (5–11)	Alport syndrome (n = 2), Schonlein-Henoch syndrome (n = 5), SLE (n = 2), and Nefrosic syndrome (n = 3)	+/−		0.49 ± 0.07	0.40 ± 0.12	$3.04\ e^{+9} \pm 7.03\ e^{+8}$
Acute postinfective glomerulonephritis in the acute phase	18	12/6	5 (4–6)	Poststreptococcal glomerulonephritis	+++	At time of diagnosis	0.74 ± 0.12	0.27 ± 0.06	$1.13\ e^{+10} \pm 3.54\ e^{+9}$
Acute postinfective glomerulonephritis in the recovery phase*	8	6/2	5 (4–6)	Poststreptococcal glomerulonephritis	−	One year after diagnosis	0.43 ± 0.07	0.04 ± 0.02	$1.70\ e^{+10} \pm 8.07\ e^{+9}$

Data are reported as means ± SE. Age is reported as medians (Q1–Q3). *Eight patients with acute postinfective glomerulonephritis of the eighteen patients in the acute phase were resampled during the recovery phase (at 1 yr). SLE, systemic lupus erythematosus.

group of acutely tubular damaged patients (7), also occurs in acute and/or persistent diseases involving the glomerular compartment. In the present study, changes in CD133$^+$ uEV levels over time were assessed in pediatric patients with acute poststreptococcal glomerulonephritis (GN), characterized by alteration of the glomerular filtration barrier, proteinuria, and resolution by minimal medical approaches. In addition, two groups of chronic glomerular diseases were studied. The first group included patients with chronic GN (CGN) and proteinuria in the absence of renal function impairment, and the second group included patients with type 2 diabetes (T2DM) with different degrees of renal impairment as a validation group.

In both acute and chronic glomerular diseases, the potential role of CD133 as a urinary biomarker modulated during renal damage was investigated. Finally, in vitro experiments were performed to evaluate whether CD133$^+$ EVs could be released by renal CD133$^+$-expressing cells. To this end, CD133$^+$ EV release was evaluated in cells exposed to hyperglycemia and albumin overload.

METHODS

Study groups. All patients enrolled in the present study provided informed written consent for the study. The study protocol was approved by the Bioethics Committee of the A.O.U. Città della Salute e della Scienza Hospital (protocol no. 0021671). The study was conducted according to the principles expressed by the Declaration of Helsinki of 1975, as revised in 2013. The study group was composed of a total of 30 pediatric patients and 7 pediatric healthy control individuals. Urine was collected from 18 pediatric patients with poststreptococcal GN (Table 1) at diagnosis and before treatment. At the time of resampling, 1 yr after diagnosis, some patients (8 of 18 patients) were all treatment free and experienced no relapse of proteinuria and/or hematuria after first remission. Samples were also collected from 12 pediatric patients with CGN (Table 1) with an active proteinuria and/or hematuria for at least 12 mo. All patients were under minimum steroid treatment at the time of sample recovery without additive immunosuppressive drugs. No significant differences among serum creatinine and urinary protein on creatinine ratio distribution were observed in the different groups using the Kruskal-Wallis test.

The validation group was composed by 47 adult patients with type 2 diabetes admitted to the clinic (HbA1c > 48 mmol/mol) and 13 adult healthy individuals age and sex matched. Table 2 shows patient and control clinical data. In patients with diabetes, only serum creatinine values appeared to be normally distributed, and the one-way ANOVA showed significant differences among groups (P = 0.04). Kruskal-Wallis test for the other nonnormally distributed data evidenced a significant difference in glomerular filtration rate (P = 0.002) and albuminuria ($P = 2.18\ e^{-10}$) but not in glycated hemoglobin values. Group-to-group comparison was performed with a t test for serum creatinine and Mann-Whitney test for glomerular filtration rate and albuminuria (Tables 1 and 2).

Urine collection and EV isolation. Morning urine samples (~100 mL) were collected in sterile containers. In parallel biochemical analysis were performed by the clinical laboratory of the A.O.U. Città

Table 2. *Clinical characteristics of patients with diabetes*

Patient Group	n	Sex (Male/Female)	Median Age, yr	Treatment (Sartans + Angiotensin-Converting Enzyme Inhibitor/Others)	Serum Creatinine, mg/dL	Estimated Glomerular Filtration Rate, $mL \cdot min^{-1} \cdot 1.73\ m^{-2}$	Albuminuria, mg/dL	HbA1c, mmol/mol	Number of Extracellular Vesicles per mL of Urine
Healthy	13	6/7	60 (42–71)		0.79 ± 0.05	96.1 ± 4.4	0 ± 0.0	0 ± 0	$3.45\ e^{+9} \pm 6.65\ e^{+8}$
NAlb DN	17	11/6	68 (63–77)	7/10	1.16 ± 0.12	72.4 ± 6.6	0.1 ± 0.1	74.1 ± 6.2	$1.09\ e^{+10} \pm 5.28\ e^{+9}$
MiAlb DN	15	10/5	79 (72–85)	7/8	1.40 ± 0.24	62.4 ± 6.4[b]	22.4 ± 2.2[d]	58.7 ± 4.4	$8.83\ e^{+9} \pm 1.74\ e^{+9}$
MaAlb DN	15	9/6	82 (70–87)	6/9	1.57 ± 0.22[a]	45.1 ± 7.1[c]	129.3 ± 27.2[e]	58.8 ± 3.1	$1.95\ e^{+10} \pm 1.07\ e^{+10}$

Data are reported as means ± SE. Age is reported as medians (Q1–Q3). NAlb DN, patients with diabetic normoalbuminuria; MiAlb DN, patients with diabetic microalbuminuria; MaAlb DN, patients with diabetic macroalbuminuria. [a]Creatinine vs. healthy individuals and patients with NAlb DN (t test, P < 0.05); [b]estimated glomerular filtration rate vs. healthy individuals (Mann-Whitney test, P = 0.02); [c]estimated glomerular filtration rate vs. healthy individuals (Mann-Whitney test, P = 0.0006) and vs. patients with NAlb DN (Mann-Whitney test, P = 0.007); [d]albuminuria vs. patients with NAlb DN (Mann-Whitney test, $P = 2.25\ e^{-7}$); [e]albuminuria vs. patients with NAlb DN (Mann-Whitney test, $P = 1.44\ e^{-7}$) and vs. patients with MiAlb DN (Mann-Whitney test, $P = 1.18\ e^{-6}$).

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

della Salute e della Scienza Hospital. Urine samples were centrifuged at 3,000 rpm for 15 min to remove whole cells, large membrane fragments, and other debris. Protease Inhibitor (PI) Cocktail (Sigma-Aldrich, St. Louis, MO, 100 μL PI/100 mL urine) and NaN_3 (10 mM, Sigma-Aldrich) were added immediately to the remaining supernatant. After filtration through 0.8- and 0.45-μm filters (Merck Millipore, Burlington, MA), uEVs were collected from the samples through ultracentrifugation (Beckman Coulter, OPTIMA L-100 K Ultracentrifuge, Rotor Type 70-Ti, Brea, CA) at 100,000 *g* for 1 h at 4°C. The pellet was then resuspended in RPMI (Euroclone, Turin, Italy) + 1% DMSO (Sigma-Aldrich) and stored at −80°C until use.

EV quantification. EVs were quantified by nanoparticle tracking analysis using the NanoSight NS300 system (NanoSight, Salisbury, UK) configured with a blue 488-nm laser and a high-sensitivity digital camera system (OrcaFlash 2.8, Hamamatsu C1 1440, NanoSight). Briefly, EVs stored in −80°C were thawed, strongly vortexed, and properly diluted in physiological solution (Fresenius Kabi, Bad Homburg, Germany) previously filtered with a 0.1-μm filter (Merck Millipore). For each sample, three videos of a 30-s duration were recorded. The settings of acquisition and analysis were optimized and kept constant between (19).

Transmission electron microscopy. Transmission electron microscopy was performed using a Jeol JEM 1010 electron microscope (Jeol, Tokyo, Japan) on uEVs coated on 200-mesh nickel formvar carbon-coated grids (Electron Microscopy Science, Hatfield, PA). EVs were placed in adhesion for 20 min on the grids that were then incubated with 2.5% glutaraldehyde containing 2% sucrose. After being washed in distilled water, EVs were negatively stained with NanoVan (Nanoprobes, Yaphank, NK) and analyzed.

Cytofluorimetric analysis. EVs were bound to surfactant-free white aldehyde/sulfate latex beads 4% (wt/vol), 4-μm diameter (Molecular Probes, ThermoFisher, Waltham, MA) for the subsequent cytofluorimetric analysis. Thirty micrograms of EVs were incubated with 5 μL of beads for 30 min at room temperature and subsequently for 30 min at +4°C. Adsorbed EVs were placed in different vials and incubated with the antibodies for exosomal markers (CD63 and CD81), renal exosomal marker CD24 and with CD133, or with the control isotypes, with a final dilution of 1:50, for 15 min at +4°C (see Supplemental Table S1; Supplemental Data for the article are available online at https://doi.org/10.6084/m9.figshare.9736400.v1). The adsorbed EVs were then washed and analyzed with a FACSCalibur and CellQuest software (Becton Dickinson Bioscience PharMingen, Franklyn Lake, NJ). The anti-aquaporin-1 (AQP1) antibody (Santa Cruz Biotechnology, Dallas, TX) was conjugated to Alexa Fluor 488 dye through the Alexa Fuor Antibody Labeling Kit (Molecular Probes) following the manufacturer's instructions. During the cytofluorimetric acquisition, the gating strategy was set on the physical parameters dot plot. Controls corresponded to EVs adsorbed on beads and marked with FITC-, phycoerythrin (PE)-, and allophycocyanine (APC)-conjugated mouse IgG1 isotypes (all purchased by Miltenyi, Bergisch Gladbach, Germany).

Floating purification of uEVs. To avoid protein contamination and to increase the purity of uEVs, a floating protocol through a sucrose gradient was applied as previously described (5, 16). Briefly, uEVs were washed with an ultracentrifuge and resuspended in 1.35 mL of buffer [0.25 M sucrose, 10 mM Tris (pH 8), and 1 mM EDTA, all products purchased from Sigma-Aldrich], transferred to a SW55Ti rotor tube (Beckman Coulter), and mixed with 60% stock solution of Optiprep (Sigma-Aldrich) at a 1:1 ratio. Next, 1.2 mL Optiprep 20% solution was layered on top followed by 1.1 mL of Optiprep 10% solution. The tubes were then ultracentrifuged at 350,000 *g* for 1 h at +4°C with "no brake" deceleration. The next day, five fractions of 1 mL each were collected from the top of the tubes, and each fraction was diluted in 25 mL PBS (Lonza, Basel, Switzerland) and washed with an ultracentrifugation at 100,000 *g* for 2 h at +4°C to pellet EVs. The highest fraction, containing pure EVs, as assessed by NanoSight analysis, was then resuspended in 200 μL of serum-free DMEM (Lonza) with 1% penicillin-streptomycin and 1% DMSO (both products purchased from Sigma-Aldrich) to allow freezing storage in −80°C until use.

Cell culture. CD133^{+} renal progenitor cells were obtained from biopsies of normal tissue of a human surgically removed kidney for polar carcinoma performed (after the approval of the Ethical Committee for the Use of Human Tissue of Molinette Hospital, no. 168/2004), as previously described (2). In particular, the outer medullary tissue at the opposite pole of the tumor was used. The absence of tumor-infiltrating cells was evaluated by pathologists. Tissue samples were cut to obtain 3- to 5-mm^3 fragments, digested in 0.1% collagenase type I (Sigma-Aldrich) for 30 min at 37°C, and subsequently forced through a graded series of meshes for the separation of cell components from stroma and aggregates. The filtrate was then pelleted by centrifugation. To recover CD133^{+} cells, the single cell suspension underwent magnetic separation for CD133/1 antigen (Miltenyi). CD133^{+} cells (>98% as evaluated by cytofluorimetric analysis) were resuspended in expansion medium (Endothelial Cell Growth Basal Medium Plus Supplement Kit, Lonza) and plated at density 1×10^4 viable cells/cm^2.

From *passages 2−4*, cells were treated for 48 h with human serum albumin (HAS Kedrion, Wilmington, DE) at different concentrations (0.1, 1, and 10 mg/mL), glucose (Sigma-Aldrich) at different concentrations (5.5 and 27.5 mM), or a combination of HSA (10 mg/mL) plus glucose (27.5 mM). After 48 h of treatment and overnight cell starvation, EVs were recovered by ultracentrifugation (2 h, 100,000 *g* at +4°C).

Western blot analysis. Proteins from cells were extracted using RIPA buffer containing 1% of phosphatase inhibitor cocktails, 1% protease inhibitors, and 1% PMSF (all purchased from Sigma-Aldrich), while proteins from EVs were obtained using the same lysis buffer on the ultracentrifuge pellet. Total protein concentration was estimated by Bradford (Bio-Rad, Hercules, CA) quantification. Ten micrograms of EV lysates and 50 μg of cells lysates were loaded on Mini-PROTEAN TGX precast electrophoresis gels (Bio-Rad). Proteins were subsequently transferred on iBlot nitrocellulose membranes (Invitrogen, Carlsbad, CA) and blotted with antibodies against CD133 (Miltenyi), actin (Santa Cruz Biotechnology), heat shock protein (HSP)70/heat shock complex (HSC)70 (Santa Cruz Biotechnology), and Rab5 (Santa Cruz Biotechnology) as provided in Supplemental Table S1 (available online at https://doi.org/10.6084/m9.figshare.9736400.v1). Chemiluminescent signals were detected using the ECL substrate (Bio-Rad).

Statistical analysis. The normality of clinical data was assessed through a Shapiro-Wilk test and graphical evaluation for serum creatinine and urinary protein on creatinine ratio in pediatric patients and of serum creatinine, glomerular filtration rate, albuminuria and glycated hemoglobin in patients with diabetes. Normally distributed variables are presented as means ± SD and were compared among multiple groups with one-way ANOVA test and a Student's *t* test for group-to-group comparison. Non-normally distributed data are presented as medians and interquartile ranges and were assessed for comparison among multiple groups with a nonparametric Kruskal-Wallis test and Mann-Whitney test for group-to-group comparison. Cytofluorimetric data followed a non-normal distribution, as assessed by a Kolmogorov-Smirnov test; therefore, nonparametric tests, such as Mann-Whitney or Kruskal-Wallis tests, were performed. Western blot experiments were analyzed with a Student's *t* test and with ordinary one-way ANOVA with Dunnett's multicomparison test. *P* values of <0.05 were considered significant. The sensitivity and specificity of CD133 as a marker of renal damage and 95% confidence intervals (CIs) were calculated using receiver-operating characteristic (ROC) curves. The accuracy of the test among different groups of patients was done by evaluating areas under the curve (AUC) and *P* values.

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

RESULTS

Characterization of uEVs from healthy pediatric subjects. EVs were isolated from urine obtained from healthy pediatric subjects using a protocol based on filtration and sequential centrifugations. Transmission electron microscopy images showed a heterogeneous population of rounded in shape vesicles (Fig. 1*A*). uEVs were also analyzed for number and size by Nanosight (mean diameter: 265.1 ± 69 nm; Fig. 1*B*).

Cytofluorimetric analysis was performed using uEV adsorption on latex beads, as this method allows the analysis of small vesicles that cannot be analyzed using direct cytofluorimetric analysis. The gating strategy of the evaluated beads and the setting of the negative threshold is shown in Fig. 1*C*. The results showed that uEVs from pediatric healthy subjects, as those obtained from healthy adult subjects (7), were positive for CD24 and AQP1 renal markers, for the exosomal marker CD81, and for the progenitor/stem cell marker CD133. Urinary EVs did not express the vascular marker VEGF receptor 2 or the podocyte marker podocalyxin (PDX) (see Supplemental Fig. S1*A*, available online at https://doi.org/10.6084/m9.figshare.9736400.v1). As a negative control, beads were incubated with antibodies in the absence of EVs, and they did not show positive fluorescence (see Supplemental Fig. S1*B*, available online at https://doi.org/10.6084/m9.figshare.9736400.v1).

Characterization of uEVs from pediatric patients with acute and chronic GN. uEVs from four different groups of patients suffering of GN were compared: healthy individuals, CGN, acute postinfective GN (AGN), AGN at diagnosis (AGN T1), and AGN on *day 365*, corresponding to the recovery phase (AGN T365). The Nanosight analysis showed no significant differences in the mean and mode size of uEVs among the four groups (Fig. 2, *A* and *B*). At variance, EV count revealed that the amount of uEVs per milliliter of urine was significantly decreased in patients with CGN compared with healthy individuals (Fig. 2*C*).

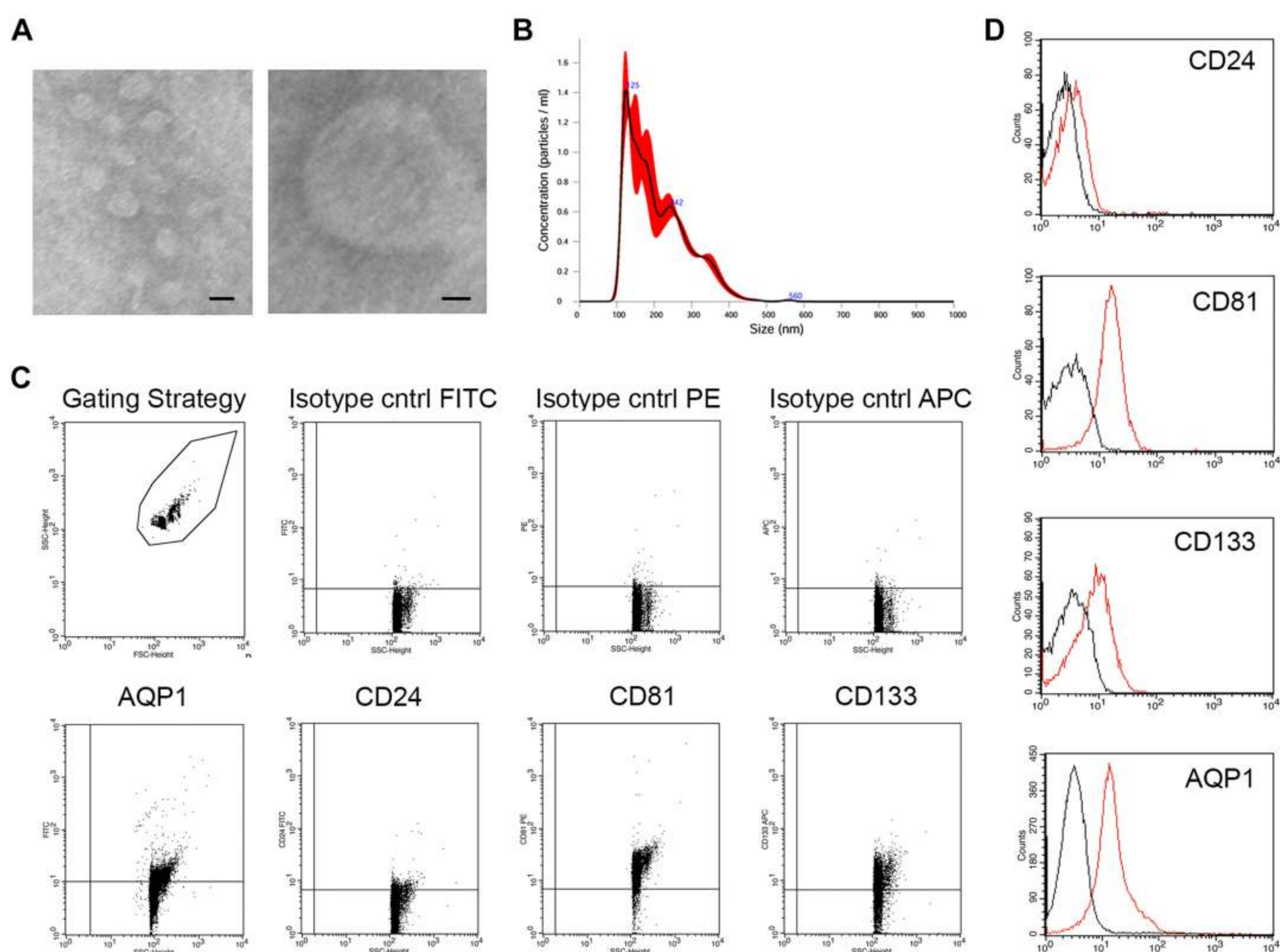


Fig. 1. Characterization of urinary extracellular vesicles (uEVs) in healthy individuals. *A*: representative transmission electron microscopy image of uEVs isolated from healthy individuals (scale bar = 150 nm) and an enlargement on one single uEV (scale bar = 20 nm). *B*: nanoparticle tracking analysis of a representative pool of uEVs showing their size distribution profile. *C* and *D*: cytofluorimetric analysis of surface markers expressed by uEVs from healthy individuals were represented as dot plot (the gating strategy and negative fluorescence controls are reported) (*C*) and histograms with isotype control (black line) plotted on the marker signal (red line) (*D*). Ten different preparations of different healthy subjects were tested with similar results. PE, phycoerythrin; APC, allophycocyanine; AQP1, aquaporin-1.

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

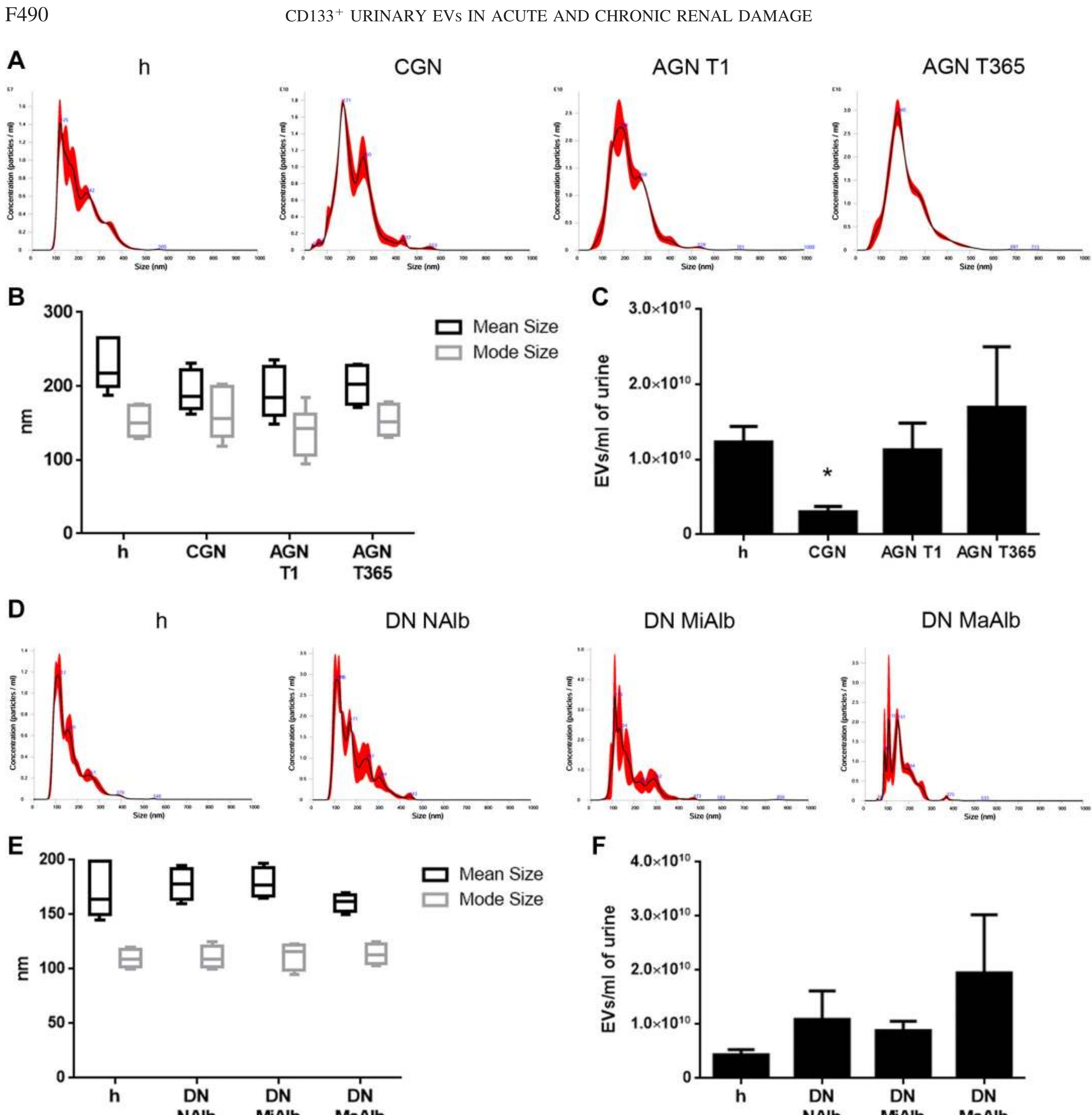
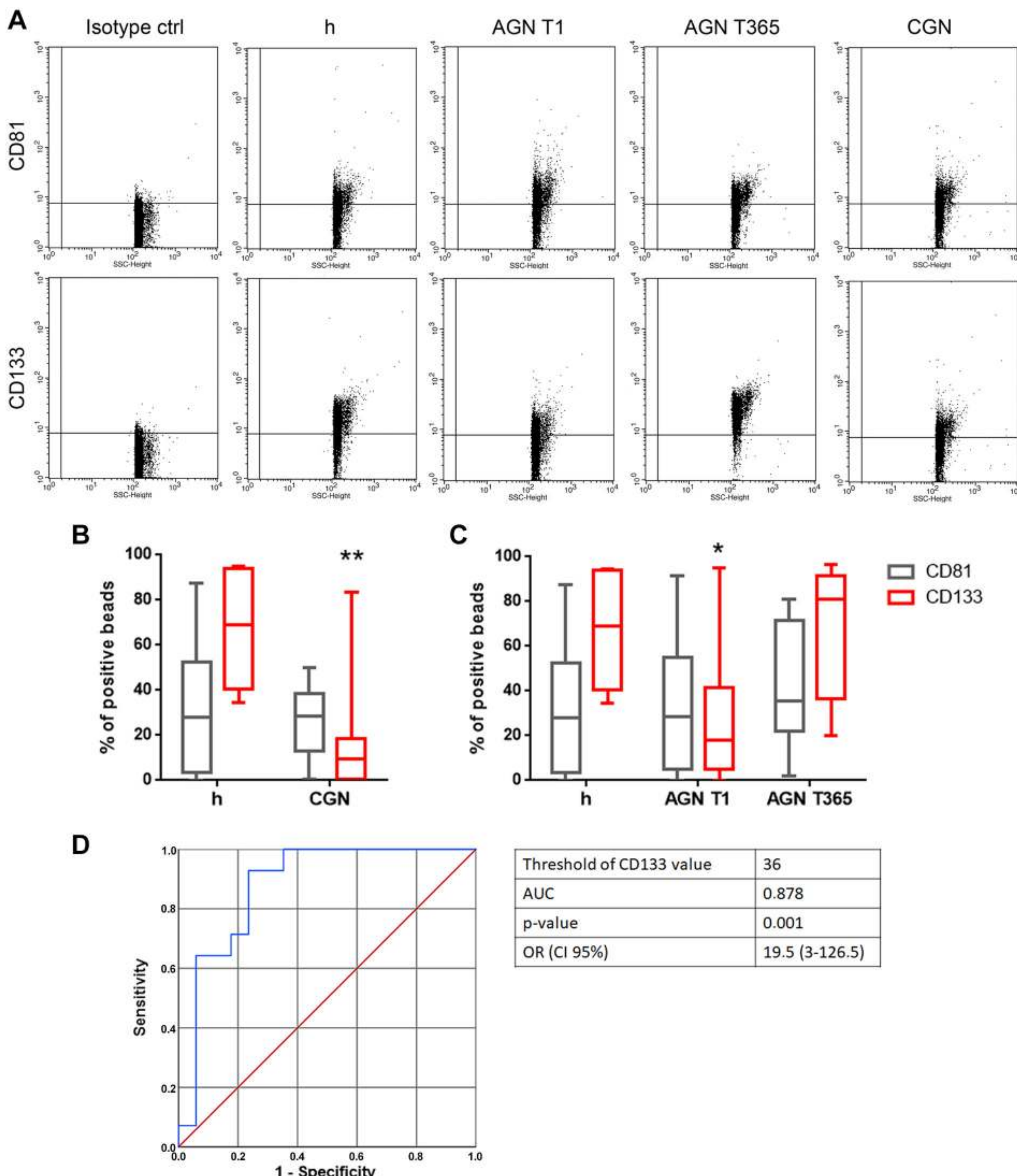


Fig. 2. Characterization of urinary extracellular vesicles (uEVs) recovered from patients with acute and chronic glomerular damage. *A* and *B*: representative nanoparticle tracking analysis (NTA) of uEVs from each group of pediatric patients (*A*) and box plot representation of mean and mode values of EV size (nm) of all patients in the study (*B*). *C*: evaluation of EV concentration in urine of the different groups of pediatric patients. *D* and *E*: representative NTA of uEVs from each group of patients with type 2 diabetes mellitus (T2DM; *D*) and box plot representation of mean and mode values of EV size (nm) of all patients in the study (*E*). *F*: evaluation of EV concentration in the urine of the different groups of patients with T2DM. Data in *C* and *F* are expressed as means ± SE of all patients in the study, reported in Table 1. CGN, chronic glomerulonephritis; h, healthy individuals; AGN, acute postinfective glomerulonephritis; AGN T1, AGN at *day 1*; AGN T365, AGN at *year 1*; NAlb, patients with diabetic normoalbuminuria; MiAlb, patients with diabetic microalbuminuria; MaAlb, patients with diabetic macroalbuminuria. $^*P < 0.05$ by an unpaired *t* test.

Cytofluorimetric analysis was performed to evaluate the relative amount of $CD133^+$ and $CD81^+$ uEVs, testing the same quantity of EVs previously shown to saturate the latex beads surface (7). Indeed, as shown in Fig. 3, the exosomal marker CD81 did not significantly vary among the different groups, indicating an equal EV loading. At variance, we observed a significant decrease of CD133 in uEVs of patients with CGN compared with healthy individuals. Moreover, $CD133^+$ uEVs were also significantly decreased in patients with AGN during the acute phase of the disease and returned to basal level on *day*

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.



Threshold of CD133 value	36
AUC	0.878
p-value	0.001
OR (CI 95%)	19.5 (3-126.5)

Fig. 3. Characterization of urinary extracellular vesicles (uEVs) obtained from acute damaged patients. *A*: dot plot cytofluorimetric analysis of surface markers in uEVs derived from healthy individuals and pediatric patients. The gating strategy and negative fluorescence controls are also reported. *B* and *C*: box plot representation of cytofluorimetric values of CD81 (gray plot) and CD133 (red plot) percent comparisons between healthy individuals (h) and patients with chronic glomerulonephritis (CGN; *B*) or between healthy individuals and patients with acute glomerulonephritis (AGN) (*C*) at *day 1* (T1) and *year 1* (T365) after the damage. Data from all patients in the study, described in Table 1, are reported. *D*: receiver-operating characteristic analysis for the detection of renal damage in different groups of acutely damaged patients using CD133. The table provides CD133 threshold value, the area under the curve (AUC) value, the *P* value, and the odd ratio (OR) with confidence interval (CI).

365. All these results are reported in dot plot graphs (Fig. 3*A*) and in the relative box plots (Fig. 3, *C* and *D*). These data indicate that $CD133^+$ uEVs, which mainly derived from renal tubular cells (23), were decreased after glomerular damage in both acute and chronic conditions.

Moreover, ROC curves were performed, using the values determined by the cytofluorimetric analysis, to investigate the potential utility of CD133 as a predictive biomarker of renal damage. When healthy individuals were compared with patients with AGN (AGN T1), the AUC value for CD133 was

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

0.878, with an odds ratio (OR) of 19.5 (95% CI: 3–126.5) and P = 0.001 (Fig. 3*D*). In the context of this analysis, the percentage of CD133^{+} uEVs considered as a threshold able to discriminate healthy individuals and patients with AGN was 36%.

Characterization of uEVs from patients with diabetes. We subsequently aimed to validate the reduction of CD133^{+} uEVs observed after glomerular damage in pediatric patients affected by T2DM-associated glomerular disease. Based on the degree of their renal impairment, patients were divided into the following three groups: patients with normoalbuminuric diabetes (NAlb DN), patients with microalbuminuric diabetes (MiAlb DN), and patients with macroalbuminuric diabetes (MaAlb DN) and were compared with age- and sex-matched healthy individuals. Nanosight analysis showed that the mean and mode uEV dimension were similar among groups. Similarly, no differences in the number of uEVs per milliliters of urine were detected (Fig. 2, *D–F*). The presence of CD133^{+} uEVs in patients with T2DM was subsequently compared with that of healthy individuals by cytofluorimetric analysis. As shown in Fig. 4, *B* and *C*, the level of CD81 did not change between healthy individuals and patients with diabetes, while CD133^{+} uEVs were significantly decreased in patients with DN. Moreover, when the analysis of CD133^{+} uEVs was based on their level of albuminuria, a further decrease of CD133^{+} uEVs was detected in patients with macroalbuminuria (Fig. 4*D*). To evaluate whether proteins in the urine of patients with albuminuric T2DM could affect the results of the cytofluorimetric analysis, three uEV pools belonging to the three different T2DM groups were purified from contaminating proteins by applying a glucose gradient floating protocol (5, 16). Floated uEVs were analyzed and compared with a pool of healthy uEVs subjected to the same protocol. As shown in Fig. 4*E*, the loading of 10 μg proteins from each sample showed a significant decrease in CD133 in all patients with T2DM, confirming the results of the cytofluorimetric analysis.

ROC curves were also performed on the results obtained from the cytofluorimetric analysis to confirm CD133 as a potential predictive biomarker of renal damage. When healthy individuals were compared with all patients with T2DM, the AUC value for CD133 was 0.775, with an OR of 9.8 (95% CI: 2.2–42.7) and P = 0.001 (Fig. 3*D*). In the context of this analysis, the percentage of CD133^{+} uEVs considered as threshold able to discriminate healthy individuals and patients with diabetes was 37%.

Modulation of CD133^{+} EV release by CD133^{+} cells subjected to albumin and glucose overload. To evaluate the contribution of renal progenitor cells in the release of CD133^{+} uEVs, CD133^{+} renal cells were treated with different concentrations of albumin and glucose for 48 h (Fig. 5*A*). As shown in Fig. 5*B*, albumin (0.1–10 mg/mL) but not glucose (5.5–27.5 mM) treatment induced a dose-dependent loss of CD133 expression by CD133^{+} cells (Fig. 5*B*). Likewise, albumin load induced the loss of CD133^{+} in CD133^{+} renal cell-derived EVs (Fig. 5*C*). To mimic diabetes in vitro, CD133^{+} cells were treated with a high albumin concentration (10 mg/mL) together with a high glucose concentration (27.5 mM) for 48 h. Western blot revealed a strong decrease in CD133 expression both in progenitor renal cells and their derived EVs (Fig. 5, *D* and *E*). These data support the results obtained by analyzing patients' urine and indicate that proteinuria and glucose may modify tubular progenitors cells and promote CD133 loss in cells and in their derived uEVs.

DISCUSSION

EVs are constantly released by virtually all cells, and their characteristics reflect the state of the cell of origin, so that they can mirror tissue health and/or disease. In the present study, we show that the levels of CD133^{+} uEVs, differently from normal individuals, declined in the urine of pediatric and adult patients with glomerular disease, suggesting that CD133^{+} uEVs may represent a marker of renal normal physiology. Indeed, ROC curve analysis indicated the ability of CD133^{+} uEV values to discriminate the health condition from that of glomerular disease.

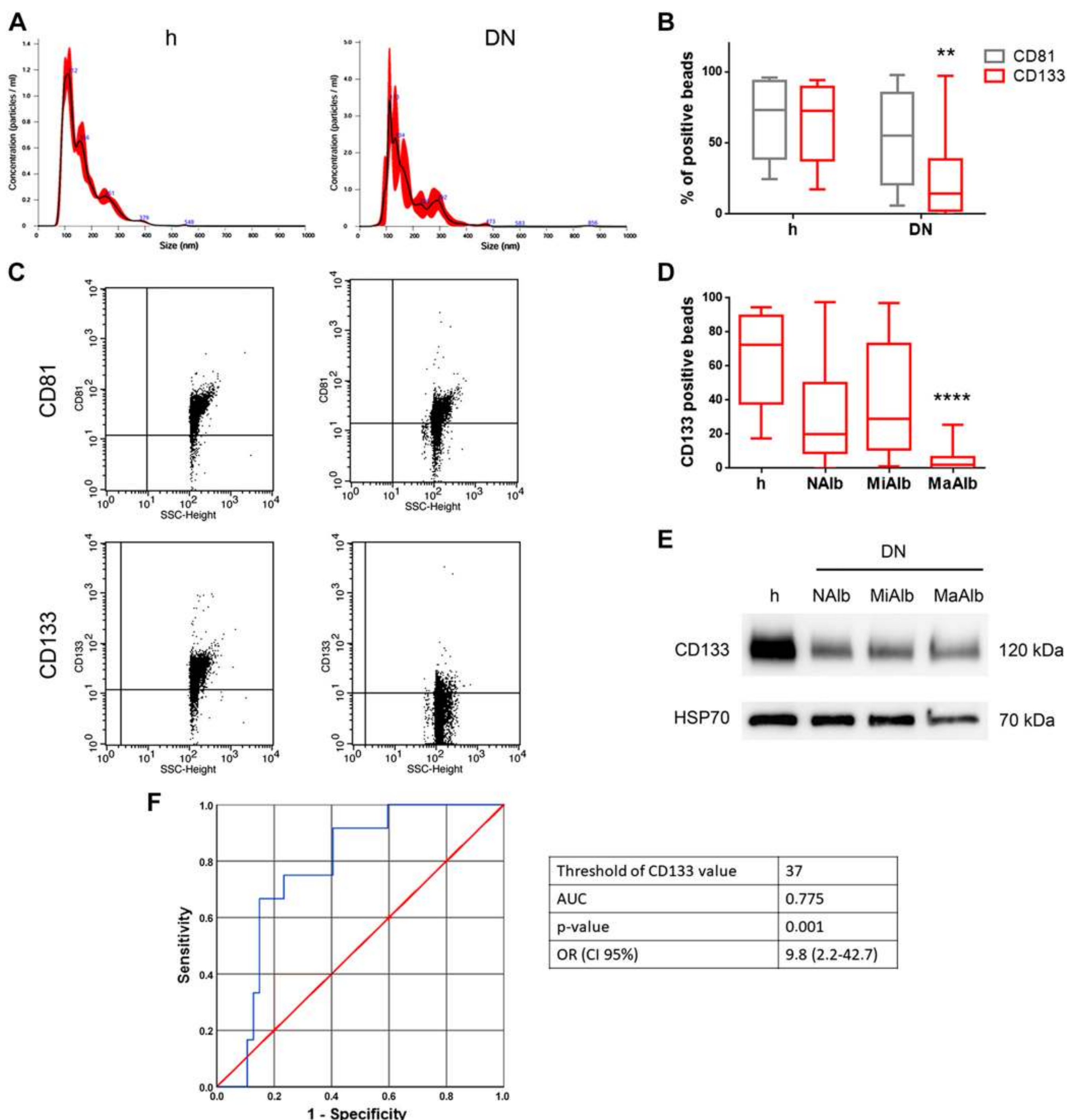
Since uEVs released by renal tissue could be detected in human urine, EVs may be exploited as markers for the physiological/pathological state of the renal tissue. Indeed, proteomic analysis of uEVs from healthy individuals demonstrated the presence of renal proteins, such as uromodulin, AQP1, and AQP2, indicating that the majority of EVs derived from tubular segments (9). During glomerular damage, podocyte-released EVs may be detected within urine (17), and podocyte-derived proteins within urinary exosomes were reported as specific markers of glomerular injury and chronic renal damage. In particular, exosomal Wilms' tumor 1 was described as a marker of focal segmental glomerulosclerosis in murine and human studies (29), while CD2-associated protein mRNA was reported as a marker of chronic kidney disease (20). The nature of urine vesicles is heterogeneous in type and dimension, including larger vesicles, so-called microparticles. Indeed, the number of podocyte-derived microparticles by itself was increased in patients with type 1 diabetes and could predict early renal damage (21).

In addition to being considered as biomarkers, EVs may also play a role in cell-to-cell communication along the urinary lumen. Indeed, proximal tubule-derived EVs were shown to be internalized by distal tubular cells (8). Furthermore, the increase of water flow in recipient cells proved the functional transfer of AQP2 from EVs deriving from collecting ducts (26). During glomerular disease such as diabetes, podocyte-derived EVs may promote fibrosis of tubular cells via detrimental glomerular-tubular signals (24). These data suggest that uEVs may orchestrate the trafficking of different renal messages occurring during physiological or pathological conditions.

In the present study, we focused on CD133, a tubular EV marker, previously shown to be highly abundant in normal urine (7). Indeed, CD133^{+} uEV sorting clearly showed the coexpression of proximal tubular markers (mainly aminopeptidase and AQP1) indicating its origin from cells of the proximal tubule (7). Moreover, we herein confirmed that CD133^{+} EVs are released, at least in vitro, by renal CD133^{+} cells. Similarly, CD133^{+} EVs have been detected in the spinocerebral fluid in association with pathological conditions (14, 15) and in cultures of CD133^{+} stem/progenitor cells of nervous and hematopoietic origins (1, 6). Therefore, it could be speculated that the levels of urinary CD133^{+} EVs may reflect the number of CD133^{+} cells lining the lumen of renal tubules.

Interestingly, in pediatric patients with AGN, CD133^{+} uEVs were downregulated during the acute phase to return to basal

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.



Threshold of CD133 value	37
AUC	0.775
p-value	0.001
OR (CI 95%)	9.8 (2.2-42.7)

Fig. 4. Characterization of urinary extracellular vesicles (uEVs) recovered from patients with chronic damage. *A*: representative nanoparticle tracking analysis of uEVs recovered from healthy individuals (h) and patients with diabetes (DN). *B* and *C*: cytofluorimetric results are reported as dot plot (*C*) and box plot (*B*) of CD81 (gray plot) and CD133 (red plot) in healthy individuals and patients with DN. $**P < 0.01$ by a Mann-Whitney comparison test. *D*: cytofluorimetric evaluation of CD133 in uEVs from healthy individuals and patients with diabetic normoalbuminuria (NAlb), diabetic microalbuminuria (MiAlb), and diabetic macroalbuminuria (MaAlb). Data from all patients in the study, described in Table 2, are reported. $****P < 0.0001$ by a Kruskal-Wallis multiple comparison test. *E* and *F*: Western blot analysis (*E*) of CD133 in floating purified uEVs. *F*: receiver-operating characteristic analysis for the detection of renal damage in different groups of patients with diabetes using CD133. The table provides CD133 threshold value, the area under the curve (AUC) value, the *P* value, and the odd ratio (OR) with confidence interval (CI).

level after 1 yr, as the result of tissue repair. Similarly, the $CD133^+$ uEV level was significantly low in pediatric patients with chronic renal disease. This observation suggests that CD133 could be considered a general marker of renal tissue damage that may reflect the state of the functional tubular compartment. This is consistent with data previously demonstrated in transplanted patients with slow graft function (7). Our results were further validated in patients with diabetes. Indeed, we found that the reduction of CD133 levels in patients with diabetes was even more evident in the subgroup of

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

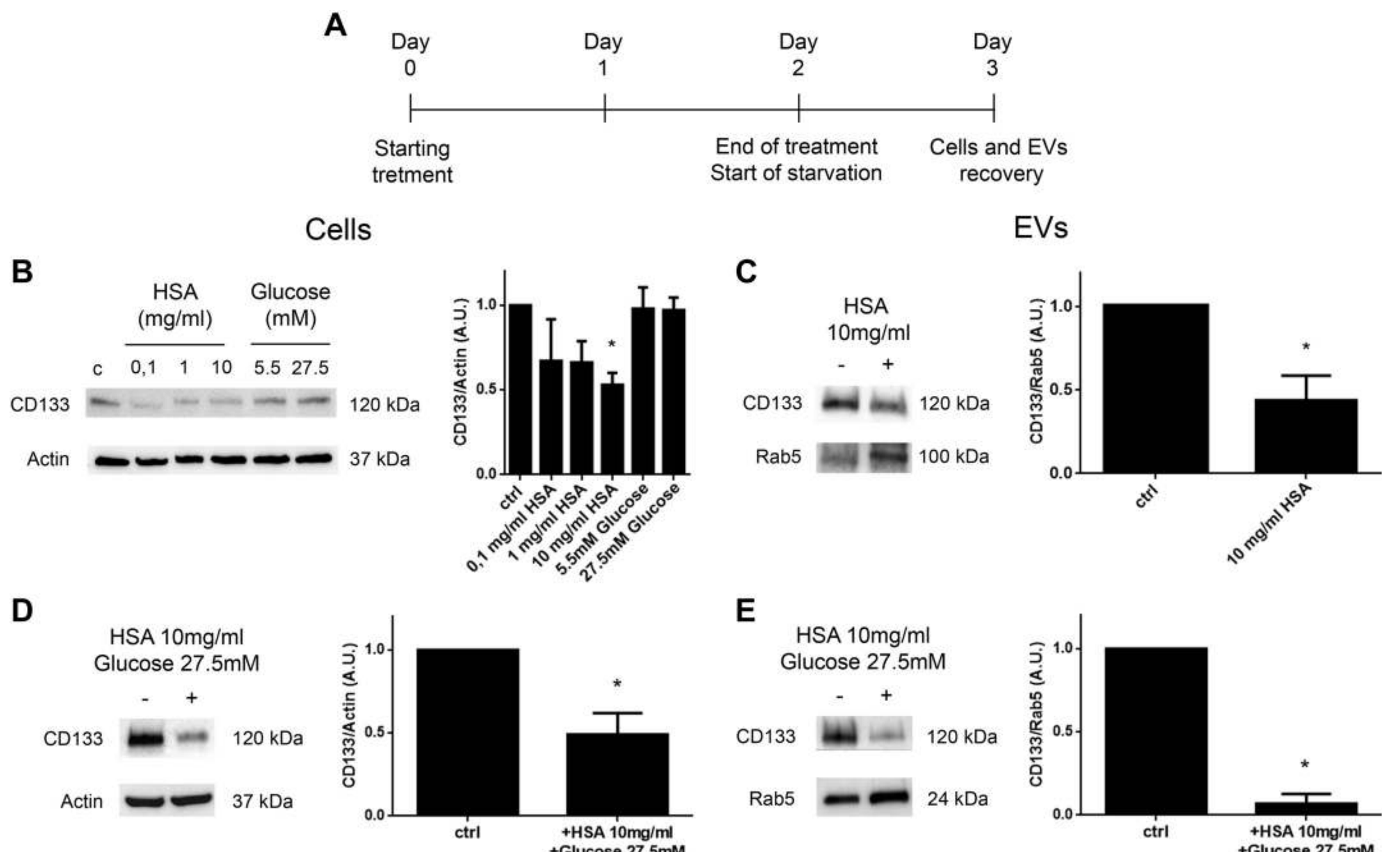


Fig. 5. Decrease of CD133 upon human serum albumin (HSA) and glucose treatment. *A*: timetable of the experiment performed on CD133$^+$ renal cells. *B* and *D*: CD133 levels in CD133$^+$ cells upon treatment with increasing concentration of HSA (0.1, 1, and 10 mg/mL) or glucose (5.5–27.5 mM) (*B*) or upon HSA (10 mg/mL) and glucose (27.5 mM) combination treatment (*D*). *C* and *E*: CD133 levels in CD133$^+$ cell-derived EVs upon HSA treatment (10 mg/mL; *C*) or upon HSA (10 mg/mL) and glucose (27.5 mM) combination treatment (*E*). A.U., arbitrary units. Data in *B–E* are expressed as means ± SE of at least 3 different experiments. *$P < 0.05$ by an unpaired *t* test.

patients with macroalbuminuria. However, although the CD133 decrease paralleled the decrease in renal function in these patients, no statistical correlation between CD133 levels and glomerular filtration rate was found.

These data suggest that the reduction of CD133$^+$ uEVs in patients with acute and chronic glomerular damage reflects the tubular damage occurring in the tubular compartment as a consequence of glomerular injury and protein overload. Consistently, in vitro experiments showed that the albumin overload decreased the expression of CD133 in cultured cells as well as in their derived EVs. Indeed, it is well established that proteinuria may accelerate kidney disease progression (10).

The analysis of ROC curves confirmed the ability of CD133$^+$ uEV levels to discriminate between healthy individuals and patients with renal disease. Moreover, the restoration of CD133$^+$ uEV levels in patients 1 yr after recovery from glomerular damage further supports the correlation of this marker with the renal pathophysiological condition. Considering the role of CD133 cells in renal repair and regeneration, it can be also inferred that CD133$^+$ uEVs may represent an easily accessible marker of the functional status of the renal tubular compartment and of the presence of cells with proliferative and repairing activity within tubules.

In conclusion, it can be speculated that the assessment of CD133$^+$ uEV levels in urine might provide a number of benefits. The use of a surface marker may allow a direct and quick evaluation compared with intravesicle protein, mRNA, or microRNA evaluation. Moreover, this assessment, being a percentage of total events, is independent from the number of total EVs in urine and from their concentration. Finally, the high level of CD133$^+$ uEVs in normal individuals guarantees a very good discrimination between healthy individuals and patients. This marker may therefore provide information on the renal parenchyma status and chances associated with tissue recovery after injury without invasive procedures, which represents an important step toward biopsy overcoming and patient's safety. The limit of our study relies on the use of a semiquantitative cytofluorimetric analysis. In fact, the absorption onto latex beads was an instrumental requirement for the assessment of small EVs. However, it should be considered that several new cytofluorimetric techniques, which are currently under evaluation, would allow a direct assessment of the small EV population. This implies that future studies may benefit of an easier and direct measurement of uEV number and markers.

GRANTS

This work was supported by Unicyte, StemKidney Research grant (to B. Bussolati, M. F. Brizzi, and G. Camussi).

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

DISCLOSURES

G. Camussi is component of the Unicyte Scientific Board. V. Dimuccio, G. Camussi, and B. Bussolati are inventors in a related patent application.

AUTHOR CONTRIBUTIONS

B.B. conceived and designed research; V.D., E.C., A.B., M.G., S.G., and G.C. performed experiments; V.D., L.P., M.F.B., E.C., F.F., L.B., and B.B. analyzed data; V.D., L.P., G.C., and B.B. interpreted results of experiments; V.D. prepared figures; V.D. and B.B. drafted manuscript; M.F.B., L.P., and B.B. edited and revised manuscript; V.D., L.P., M.F.B., E.C., F.F., A.B., M.G., S.G., L.B., G.C., and B.B. approved final version of manuscript.

REFERENCES

1. **Bauer N, Wilsch-Bräuninger M, Karbanová J, Fonseca AV, Strauss D, Freund D, Thiele C, Huttner WB, Bornhäuser M, Corbeil D.** Haematopoietic stem cell differentiation promotes the release of prominin-1/CD133-containing membrane vesicles—a role of the endocytic-exocytic pathway. *EMBO Mol Med* 3: 398–409, 2011. doi:10.1002/emmm.201100147.
2. **Brossa A, Papadimitriou E, Collino F, Incarnato D, Oliviero S, Camussi G, Bussolati B.** Role of CD133 molecule in Wnt response and renal repair. *Stem Cells Transl Med* 7: 283–294, 2018. doi:10.1002/sctm.17-0158.
3. **Bussolati B, Collino F, Camussi G.** CD133+ cells as a therapeutic target for kidney diseases. *Expert Opin Ther Targets* 16: 157–165, 2012. doi:10.1517/14728222.2012.661417.
4. **Campion CG, Sanchez-Ferras O, Batchu SN.** Potential role of serum and urinary biomarkers in diagnosis and prognosis of diabetic nephropathy. *Can J Kidney Health Dis* 4: 1–54, 2017. doi:10.1177/2054358117705371.
5. **Colombo M, Raposo G, Théry C.** Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. *Annu Rev Cell Dev Biol* 30: 255–289, 2014. doi:10.1146/annurev-cellbio-101512-122326.
6. **Corbeil D, Marzesco AM, Wilsch-Bräuninger M, Huttner WB.** The intriguing links between prominin-1 (CD133), cholesterol-based membrane microdomains, remodeling of apical plasma membrane protrusions, extracellular membrane particles, and (neuro)epithelial cell differentiation. *FEBS Lett* 584: 1659–1664, 2010. doi:10.1016/j.febslet.2010.01.050.
7. **Dimuccio V, Ranghino A, Praticò Barbato L, Fop F, Biancone L, Camussi G, Bussolati B.** Urinary CD133+ extracellular vesicles are decreased in kidney transplanted patients with slow graft function and vascular damage. *PLoS One* 9: e104490, 2014. doi:10.1371/journal.pone.0104490.
8. **Gildea JJ, Seaton JE, Victor KG, Reyes CM, Bigler Wang D, Pettigrew AC, Courtner CE, Shah N, Tran HT, Van Sciver RE, Carlson JM, Felder RA.** Exosomal transfer from human renal proximal tubule cells to distal tubule and collecting duct cells. *Clin Biochem* 47: 89–94, 2014. doi:10.1016/j.clinbiochem.2014.06.018.
9. **Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, Star RA, Kleta R, Wang NS, Knepper MA.** Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol* 20: 363–379, 2009. doi:10.1681/ASN.2008040406.
10. **Gorriz JL, Martinez-Castelao A.** Proteinuria: detection and role in native renal disease progression. *Transplant Rev (Orlando)* 26: 3–13, 2012. doi:10.1016/j.trre.2011.10.002.
11. **Guo JK, Marlier A, Shi H, Shan A, Ardito TA, Du Z-P, Kashgarian M, Krause DS, Biemesderfer D, Cantley LG.** Increased tubular proliferation as an adaptive response to glomerular albuminuria. *J Am Soc Nephrol* 23: 429–437, 2012. doi:10.1681/ASN.2011040396.
12. **Han Y, Ly ND, Tesch GH, Poronnik P, Nikolic-Paterson DJ.** Reduced tubular degradation of glomerular filtered plasma albumin is a common feature in acute and chronic kidney disease. *Clin Exp Pharmacol Physiol* 45: 241–249, 2018. doi:10.1111/1440-1681.12878.
13. **Holzscheiter L, Beck C, Rutz S, Manuilova E, Domke I, Guder WG, Hofmann W.** NGAL, L-FABP, and KIM-1 in comparison to established markers of renal dysfunction. *Clin Chem Lab Med* 52: 537–546, 2014. doi:10.1515/cclm-2013-0693.
14. **Huttner HB, Corbeil D, Thirmeyer C, Coras R, Köhrmann M, Mauer C, Kuramatsu JB, Kloska SP, Doerfler A, Weigel D, Klucken J, Winkler J, Pauli E, Schwab S, Hamer HM, Kasper BS.** Increased membrane shedding–indicated by an elevation of CD133-enriched membrane particles–into the CSF in partial epilepsy. *Epilepsy Res* 99: 101–106, 2012. doi:10.1016/j.eplepsyres.2011.10.029.
15. **Huttner HB, Janich P, Köhrmann M, Jászai J, Siebzehnrubl F, Blümcke I, Suttorp M, Gahr M, Kuhnt D, Nimsky C, Krex D, Schackert G, Löwenbrück K, Reichmann H, Jüttler E, Hacke W, Schellinger PD, Schwab S, Wilsch-Bräuninger M, Marzesco AM, Corbeil D.** The stem cell marker prominin-1/CD133 on membrane particles in human cerebrospinal fluid offers novel approaches for studying central nervous system disease. *Stem Cells* 26: 698–705, 2008. doi:10.1634/stemcells.2007-0639.
16. **Kowal J, Arras G, Colombo M, Jouve M, Morath JP, Primdal-Bengtson B, Dingli F, Loew D, Tkach M, Théry C.** Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc Natl Acad Sci USA* 113: E968–E977, 2016. doi:10.1073/pnas.1521230113.
17. **Kwon SH, Woollard JR, Saad A, Garovic VD, Zand L, Jordan KL, Textor SC, Lerman LO.** Elevated urinary podocyte-derived extracellular microvesicles in renovascular hypertensive patients. *Nephrol Dial Transplant* 32: 800–807, 2017. doi:10.1093/ndt/gfw077.
18. **Leh S, Hultström M, Rosenberger C, Iversen BM.** Afferent arteriolopathy and glomerular collapse but not segmental sclerosis induce tubular atrophy in old spontaneously hypertensive rats. *Virchows Arch* 459: 99–108, 2011. doi:10.1007/s00428-011-1100-3.
19. **Lopatina T, Grange C, Fonsato V, Tapparo M, Brossa A, Fallo S, Pitino A, Herrera-Sanchez MB, Kholia S, Camussi G, Bussolati B.** Extracellular vesicles from human liver stem cells inhibit tumor angiogenesis. *Int J Cancer* 144: 322–333, 2019. doi:10.1002/ijc.31796.
20. **Lv LL, Cao YH, Pan MM, Liu H, Tang RN, Ma KL, Chen PS, Liu BC.** CD2AP mRNA in urinary exosome as biomarker of kidney disease. *Clin Chim Acta* 428: 26–31, 2014. doi:10.1016/j.cca.2013.10.003.
21. **Lytvyn Y, Xiao F, Kennedy CRJ, Perkins BA, Reich HN, Scholey JW, Cherney DZ, Burger D.** Assessment of urinary microparticles in normotensive patients with type 1 diabetes. *Diabetologia* 60: 581–584, 2017. doi:10.1007/s00125-016-4190-2.
22. **Min QH, Chen XM, Zou YQ, Zhang J, Li J, Wang Y, Li SQ, Gao QF, Sun F, Liu J, Xu YM, Lin J, Huang LF, Huang B, Wang XZ.** Differential expression of urinary exosomal microRNAs in IgA nephropathy. *J Clin Lab Anal* 32: e22226, 2018. doi:10.1002/jcla.22226.
23. **Miranda KC, Bond DT, McKee M, Skog J, Păunescu TG, Da Silva N, Brown D, Russo LM.** Nucleic acids within urinary exosomes/microvesicles are potential biomarkers for renal disease. *Kidney Int* 78: 191–199, 2010. doi:10.1038/ki.2010.106.
24. **Munkonda MN, Akbari S, Landry C, Sun S, Xiao F, Turner M, Holterman CE, Nasrallah R, Hébert RL, Kennedy CRJ, Burger D.** Podocyte-derived microparticles promote proximal tubule fibrotic signaling via p38 MAPK and CD36. *J Extracell Vesicles* 7: 1432206, 2018. doi:10.1080/20013078.2018.1432206.
25. **Pisitkun T, Shen RF, Knepper MA.** Identification and proteomic profiling of exosomes in human urine. *Proc Natl Acad Sci USA* 101: 13368–13373, 2004. doi:10.1073/pnas.0403453101.
26. **Street JM, Birkhoff W, Menzies RI, Webb DJ, Bailey MA, Dear JW.** Exosomal transmission of functional aquaporin 2 in kidney cortical collecting duct cells. *J Physiol* 589: 6119–6127, 2011. doi:10.1113/jphysiol.2011.220277.
27. **Sun H, Yao W, Tang Y, Zhuang W, Wu D, Huang S, Sheng H.** Urinary exosomes as a novel biomarker for evaluation of α-lipoic acid's protective effect in early diabetic nephropathy. *J Clin Lab Anal* 31: e22129, 2017. doi:10.1002/jcla.22129.
28. **Yu Y, Bai F, Qin N, Liu W, Sun Q, Zhou Y, Yang J.** Non-proximal renal tubule-derived urinary exosomal miR-200b as a biomarker of renal fibrosis. *Nephron* 139: 269–282, 2018. doi:10.1159/000487104.
29. **Zhou H, Cheruvanky A, Hu X, Matsumoto T, Hiramatsu N, Cho ME, Berger A, Leelahavanichkul A, Doi K, Chawla LS, Illei GG, Kopp JB, Balow JE, Austin HA 3rd, Yuen PS, Star RA.** Urinary exosomal transcription factors, a new class of biomarkers for renal disease. *Kidney Int* 74: 613–621, 2008. doi:10.1038/ki.2008.206.
30. **Zhu X, Li H, Liu Y, You J, Qu Z, Yuan S, Peng Y, Liu F, Liu H.** Tubular atrophy/interstitial fibrosis scores of Oxford classification combinded with proteinuria level at biopsy provides earlier risk prediction in IgA nephropathy. *Sci Rep* 7: 1100, 2017. doi:10.1038/s41598-017-01223-3.
31. **Żyłka A, Dumnicka P, Kuśnierz-Cabala B, Gala-Błądzińska A, Ceranowicz P, Kucharz J, Ząbek-Adamska A, Maziarz B, Drożdż R, Kuźniewski M.** Markers of glomerular and tubular damage in the early stage of kidney disease in type 2 diabetic patients. *Mediators Inflamm* 2018: 1, 2018. doi:10.1155/2018/7659243.

Downloaded from journals.physiology.org/journal/ajprenal (005.090.062.111) on February 24, 2023.

Check for updates

Received: 14 August 2019 | Revised: 14 November 2019 | Accepted: 14 November 2019

DOI: 10.1111/epi.16407

FULL-LENGTH ORIGINAL RESEARCH

Epilepsia

Doing without valproate in women of childbearing potential with idiopathic generalized epilepsy: Implications on seizure outcome

Emanuele Cerulli Irelli[1] | Alessandra Morano[1] | Enrico Cocchi[2] | Sara Casciato[3] | Martina Fanella[1] | Mariarita Albini[1] | Federica Avorio[1] | Luca M. Basili[1] | Giacomo Fisco[1] | Francesca A. Barone[1] | Addolarata Mascia[3] | Alfredo D'Aniello[3] | Mario Manfredi[1] | Jinane Fattouch[1] | Pierpaolo Quarato[3] | Anna Teresa Giallonardo[1] | Giancarlo Di Gennaro[3] | Carlo Di Bonaventura[1]

[1]Epilepsy Unit, Department of Human Neurosciences, Policlinico "Umberto I", "Sapienza" University, Rome, Italy

[2]Department of Public Health and Pediatrics, University of Turin, Turin, Italy

[3]IRCCS "NEUROMED", Pozzilli, Italy

Correspondence
Carlo Di Bonaventura, Department of Human Neurosciences, "Sapienza" University of Rome, Viale dell'Università, 30, 00185 Rome, Italy.
Email: c_dibonaventura@yahoo.it

Abstract

Objective: Valproate (VPA) use in women with idiopathic generalized epilepsy (IGE) who are of reproductive age has been a matter of concern and debate, which eventually led to the recent restrictions by regulatory agencies. The aim of our study was to investigate the relationship between VPA avoidance/switch and seizure outcome in women of childbearing potential.

Methods: We retrospectively reviewed data from female patients with IGE, 13-50 years of age, followed since 1980. We evaluated the prescription habits, and the rate of VPA switch for other antiepileptic drugs (AEDs) and its prognostic implications. Seizure remission (SR) was defined as the absence of any seizure type more than 18 months before the last medical observation. The main aim of the study was to assess (a) possible changes in seizure outcome related to VPA switch for other AEDs, especially in patients planning a pregnancy; and (b) possible differences in SR based on the presence/absence of VPA at last observation.

Results: One hundred ninety-eight patients were included in the study. Overall SR at last medical observation was 62.7%. SR significantly differed between subjects taking and those not taking VPA ($P < .001$) at last visit. Multiple regression models showed that taking VPA at last medical observation was strongly associated with SR in both the general population ($P < .001$) and the juvenile myoclonic epilepsy (JME) group ($P < .001$). Thirty-six (70.6%) of 51 patients who switched from VPA during follow-up experienced a clinical worsening. Switching back to VPA was more frequently associated with SR at last observation ($P < .001$). In those patients who substituted VPA in view of a pregnancy, SR and drug burden (monotherapy vs polytherapy) differed significantly before and after the switch.

Significance: Our study suggests that VPA avoidance/switch might be associated with unsatisfactory seizure control in women with IGE who are of childbearing potential. Our findings further highlight the complexity of the therapeutic management of female patients of reproductive age.

© 2019 International League Against Epilepsy

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

KEYWORDS
idiopathic generalized epilepsy, pregnancy, reproductive age, restriction, valproate, women

1 | INTRODUCTION

Idiopathic generalized epilepsy (IGE) is a well-described form of epilepsy that is believed to have strong genetic bases.[1] It accounts for approximately 20% of the adult subjects who attend epilepsy outpatient clinics,[2] and is more common in women.[3] The heterogeneous IGE spectrum includes several syndromes that differ greatly in terms of clinical features and prognosis[4,5]; nevertheless, the goal of clinical remission can be achieved in up to 70%-80% of the cases thanks to an appropriate therapy.[6] Valproic acid (or valproate [VPA]) has always been considered the gold standard for the treatment of most IGE syndromes,[7,8] given its extraordinary efficacy for all types of generalized seizures (absences, myoclonic, and tonic-clonic seizures).[9] However, over the past 10 years, several studies have shown that VPA administration in pregnant women is associated with a considerably higher risk of major congenital malformations and impaired postnatal motor, behavioral, and cognitive development.[10–16] Based on emerging evidence, worldwide regulatory agencies have strongly discouraged prescribing of VPA to women of reproductive age, and have recommended that those already taking VPA switch to other anti-epileptic drugs (AEDs).[17] However, to date, none of the other broad-spectrum AEDs, not even the most recently marketed ones, have proved to be as effective as VPA on all generalized seizures.[18–20] According to the International League Against Epilepsy (ILAE) recommendations, VPA should still be considered as first-line treatment when it is supposed to be the most effective medication for a specific epilepsy syndrome and once established that a future pregnancy is extremely unlikely.[21] Despite the clinical relevance of this topic, very few randomized controlled trials comparing the efficacy of different AEDs in IGE are currently available, and they are mostly limited to specific syndromes.[7,22] Moreover, no clinical study has thoroughly evaluated the relationship between seizure outcome and VPA avoidance/switch in women of childbearing potential since the recommendations of regulatory agencies were released.

The aim of our retrospective study was to assess the prescription habits and the clinical implications of VPA avoidance/replacement in terms of seizure outcome in a population of women with IGE in their reproductive age.

Key Points

- The treatment of female patients with IGE during reproductive age has always been challenging, especially after the restrictions of regulatory authorities on valproate (VPA) use
- Many patients switched from VPA to other antiepileptic drugs (AEDs) during follow-up for childbearing potential issues
- VPA switching often led patients, including those substituting VPA in view of a pregnancy, to clinical worsening
- In this specific population, taking VPA at last medical observation was strongly associated with seizure remission

2 | METHODS

In this retrospective multicenter study, we reviewed data from patients followed at Policlinico Umberto I and Neuromed Epilepsy Unit from 1980 to 2018. The study was developed in accordance with strengthening the reporting of observational studies in epidemiology (STROBE) guidelines. Relevant data were obtained through the review of both clinical charts and a computerized database. Subjects were enrolled according to the following inclusion criteria: (a) female gender; (b) age ranging from 13 to 50 years; (c) diagnosis of IGE; (d) availability of a complete clinical documentation and ≥1 electroencephalography (EEG) recording; and (e) follow-up duration of at least 18 months.

For each patient, we collected demographic data, family history of epilepsy in first- or second-degree relatives, history of febrile seizures, psychiatric comorbidities, age at onset, seizure type, possible triggering factors, EEG features (including photosensitivity), and prior/concomitant drug regimens. According to commonly accepted criteria, the diagnosis of IGE was confirmed by three trained epileptologists (CIE, DBC, GAT) who independently revised the patients' electroclinical findings. According to the ILAE classification,[23] we further identified specific epileptic syndromes: childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME), and IGE with generalized tonic-clonic seizures only (IGE-GTCS). However, to achieve a more accurate characterization, other clinical entities already included in the latest classification proposal,

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

were considered as well: namely, eyelid myoclonia with absences (EMA)[24,25] and IGE-undefined (the latter definition was only applied to not otherwise classified cases).

The main aim of the study was to assess the impact on seizure outcome of VPA presence (VPA+) or absence (VPA–) in the patients' drug regimen. We considered different clinical situations including (a) patients who avoided VPA at treatment beginning; (b) patients who switched from VPA to another AED, especially in view of a future pregnancy; (c) patients who switched back to VPA; and (d) patients who received VPA (as either add-on treatment or sequential monotherapy) during follow-up. We also evaluated the prescribing pattern of AEDs alternative to VPA and the reasons for VPA substitution.

Seizure remission (SR) was defined as the absence of any seizure type (myoclonic, absence, and generalized tonic-clonic seizures [GTCS]) over the 18 months prior to the last medical observation. For methodologic purposes, we defined absences and myoclonic seizures as "minor seizures."

2.1 | Statistical analysis

Data were tested for normal distribution using the Shapiro-Wilk test, resulting in generalized nonnormal distribution. Data were therefore presented as median (interquartile range [IQR]), and comparison across relevant groups was performed through Mann-Whitney U test or Wilcoxon signed-rank test. Categorical variables were presented as frequency (count) and compared across relevant groups through the Fisher exact test. Group tests were two-sided, with $P < .05$ considered statistically significant. To study the effect of VPA therapy on seizure outcome, multiple regression models were elaborated. In the first model (M1) we used VPA treatment at the end of the follow-up and ineffectiveness of any first-line monotherapy as potential predictors. All models were corrected for age, follow-up duration, photosensitivity, psychiatric comorbidity, history of febrile seizures, and specific syndrome (covariates). M1 model was applied to analyze three different outcomes: SR (M1a), persistence of GTCS (M1b), and polytherapy (>1 AED) (M1c) at final observation. We also elaborated a model (M2) in which we used VPA ineffectiveness at any time during follow-up as a possible predictor, with the same above-mentioned covariates. Analyses were performed and figures generated using R 3.5.1 (R Project for Statistical Computing).

3 | RESULTS

3.1 | General characteristics of the patient population

We reviewed data from 260 female patients diagnosed with IGE. Sixty-two subjects were excluded because they did not meet the inclusion criteria; therefore 198 patients were actually considered for analysis. The median age of the study participants was 29.84 years (IQR 22.79-40.25) and the median follow-up was 11.01 years (IQR 5.84-24.85). Follow-up duration was comparable among different epileptic syndromes except for CAE ($P = .021$). The timing of the first visit for all patients is illustrated in Figure 1, whereas 176 patients (88.9%) performed the last one in 2018. The most common syndrome was JME, diagnosed in 81/198 women (40.9%). Demographic and clinical data are summarized in Table 1.

3.2 | AED prescribing pattern at first observation

Valproate was the first-line treatment in 90 patients (45.5%), levetiracetam (LEV) in 51 (25.8%), lamotrigine (LTG) in

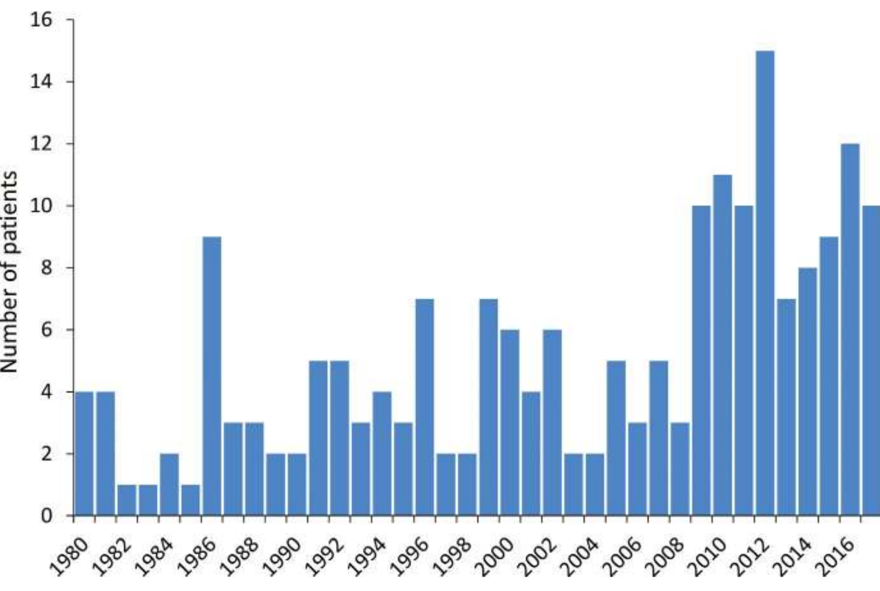


FIGURE 1 Distribution of patients according to the year of enrollment

TABLE 1 Demographic and clinical characteristics

Age, y, median (IQR)	29.84 (22.79-40.25)
Follow-up, y, median (IQR)	11.01 (5.84-24.85)
Family history in first and/or second relative degree, n, %	81, 40.9%
Psychiatric comorbidity, n, %	26, 13.1%
History of febrile seizures, n, %	18, 9.1%
Type of epilepsy	
Childhood absence epilepsy (CAE), n, %	11, 5.6%
Juvenile absence epilepsy (JAE) n, %	25, 12.6%
Juvenile myoclonic epilepsy (JME), n, %	81, 40.9%
Eyelid myoclonia with absences (EMA), n, %	20, 10.1%
IGE with generalized tonic-clonic seizures only (IGE-GCTS), n, %	43, 21.7%
IGE-undefined, n, %	18, 9.1%
EEG features	
Photoparoxysmal response (PPR), n, %	59, 29.8%

20 (10.1%), phenobarbital (PB) in 12 (6.1%), ethosuximide (ETS) in 4, and topiramate (TPM) in 2 (among the other AEDs, carbamazepine was used in 6 cases and clonazepam in 5).

3.3 | Drug regimen changes during follow-up

One-third (30/90) of patients receiving VPA as first-line AED continued to take it until the last medical observation, whereas 60 subjects (66.7%) switched from VPA to another medication during follow-up. Clinical reasons for switching are summarized in Figure 2, the most common (28/60) being the planning of a pregnancy. LEV and LTG were the most frequent alternatives to VPA after its discontinuation (in 45.1% and 25.5% of cases, respectively). The other AEDs chosen by physicians to replace VPA are shown in Table 2. Among 108 patients taking first-line AEDs other than VPA, 34 added VPA at some point during follow-up due to poor seizure control.

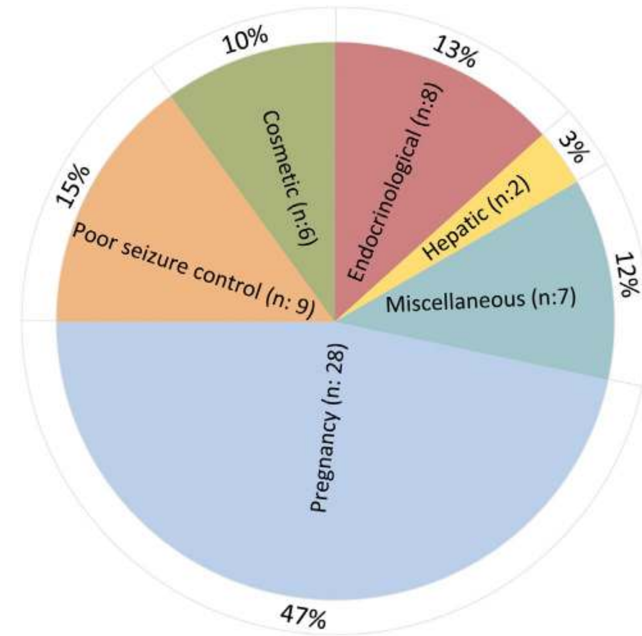


FIGURE 2 Reasons for switching from VPA to other AEDs in the overall population

3.4 | Clinical characteristics and outcome of patients who switched VPA during follow-up

Of 51 patients discontinuing VPA for causes other than poor seizure control, clinical worsening was documented in 36 (70.6%), 13 of whom (36.1%) switched back to VPA later on during follow-up. Patients who experienced a clinical worsening after VPA substitution and those who remained stable were comparable in terms of demographics, clinical features, and AED regimen. The rate of SR at last medical observation differed significantly between patients who switched back to VPA and those who did not (10/13%-76.9%- vs 3/23%-13%-, $P < .001$).

3.5 | Clinical characteristics and outcome of patients switching VPA in view of pregnancy

As previously stated, 28 patients switched VPA for another AED in consideration of a future pregnancy. In more than a half of cases (13/28, 53.6%) the switch occurred after 2014, with a median follow-up of 4 years (IQR 3-13.5). Before the switch, 23/28 patients (82.1%) were taking VPA as a monotherapy; 24/28 of patients were on SR while 4/28 were still experiencing "minor" seizures. Twenty of 28 patients (71.4%) presented a clinical worsening after switching (in most of them, seizures relapsed within 2-3 months), and 11

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 AEDs used in substitution of VPA during switch

	Total of patients N = 51 pts	Worsening after VPA switch N = 36 pts	Stable after VPA switch N = 15 pts
Levetiracetam, n (%)	23 (45.1%)	16/23 (69.6%)	7/23 (30.4%)
Lamotrigine, n (%)	13 (25.5%)	9/13 (69.2%)	4/13 (30.8%)
Phenobarbital, n (%)	10 (19.6%)	7/10 (70%)	3/10 (30%)
Clonazepam, n (%)	2 (3.9%)	1/2 (50%)	1/2 (50%)
Topiramate, n (%)	2 (3.9%)	2/2 (100%)	0
Ethosuximide, n (%)	1 (2%)	1/1 (100%)	0

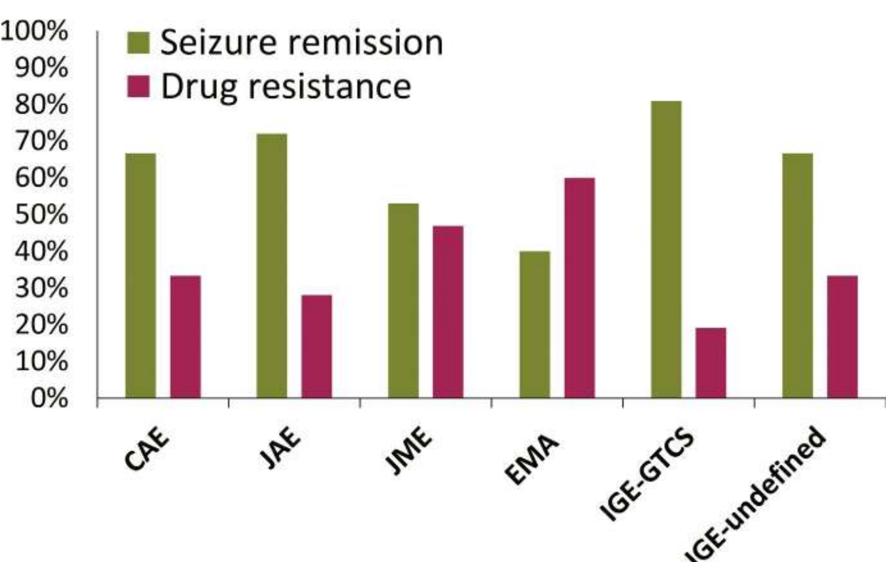


FIGURE 3 Seizure outcome at last medical observation according to different epilepsy syndromes

patients (39.3%) switched back to VPA during follow-up. As far as the drug burden is concerned, at last medical observation the number of patients taking monotherapy was significantly decreased when compared with the period before the switch (14/28 vs 23/28, $P = .02$). Considering seizure outcome, we observed a statistically significant difference in overall SR rate (15/28 vs 24/28, $P = .01$) and GTCS occurrence (5/28 vs 0/28, $P = .05$) at last medical observation when compared with the time before the switch.

3.6 | AED regimen and prognostic factors affecting seizure outcome at the last observation

On the final follow-up visit, 124 patients (62.6%) achieved SR, whereas 74 (37.4%) had persistent seizures. In the latter group, 32 patients (43.2%) still presented GTCS. Considering the different IGE syndromes, the lowest SR rate was observed in EMA (40%), and the highest in IGE-GTCS only (80.9%) (Figure 3). At last observation, 118 patients were on monotherapy, 61 were receiving two AEDs, 18 were taking three, and only one patient was on four medications. SR was observed in 63/80 subjects (78.7%) taking VPA and in 61/118 (51.7%) not taking it ($P < .001$) (Figure 4). Clinical and demographic characteristics of patients taking or not taking VPA at final observation are summarized in Table 3.

M1a showed that only a drug regimen including VPA at the end of follow-up was associated with SR (odds ratio [OR] 0.26, 95% confidence interval [CI] 0.12-0.53; <0.001). No single epileptic syndrome appeared significantly associated with a worse outcome in this model. However, when considering all syndromes with predominant myoclonic seizures, namely EMA and JME, as a single group, a trend to a worse prognosis was found (OR 2.36, 95% CI 0.93-6.82; $P = .08$). Moreover, significant differences in SR were documented in JME patients according to their "VPA status" (83.3% in the VPA+ group vs 35.3% in VPA- group; Fisher exact test $P < .001$) (Figure 4). Indeed, once restricted to JME patients, the M1a model showed that taking VPA at the final observation was strongly associated with SR (OR 0.09, 95% CI 0.02-0.31; $P < .001$).

M1b demonstrated that the ineffectiveness of any first-line monotherapy was associated with an increased risk of persistence of GTCS (OR 2.91, 95% CI 1.28-6.83; $P = .0124$), whereas age (OR 0.9, 95% CI 0.82-0.98; $P = .04$) and treatment with VPA appeared to be associated with GTCS freedom (OR 0.26, 95% CI 0.08-0.68; $P = .008$). Once again, when

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

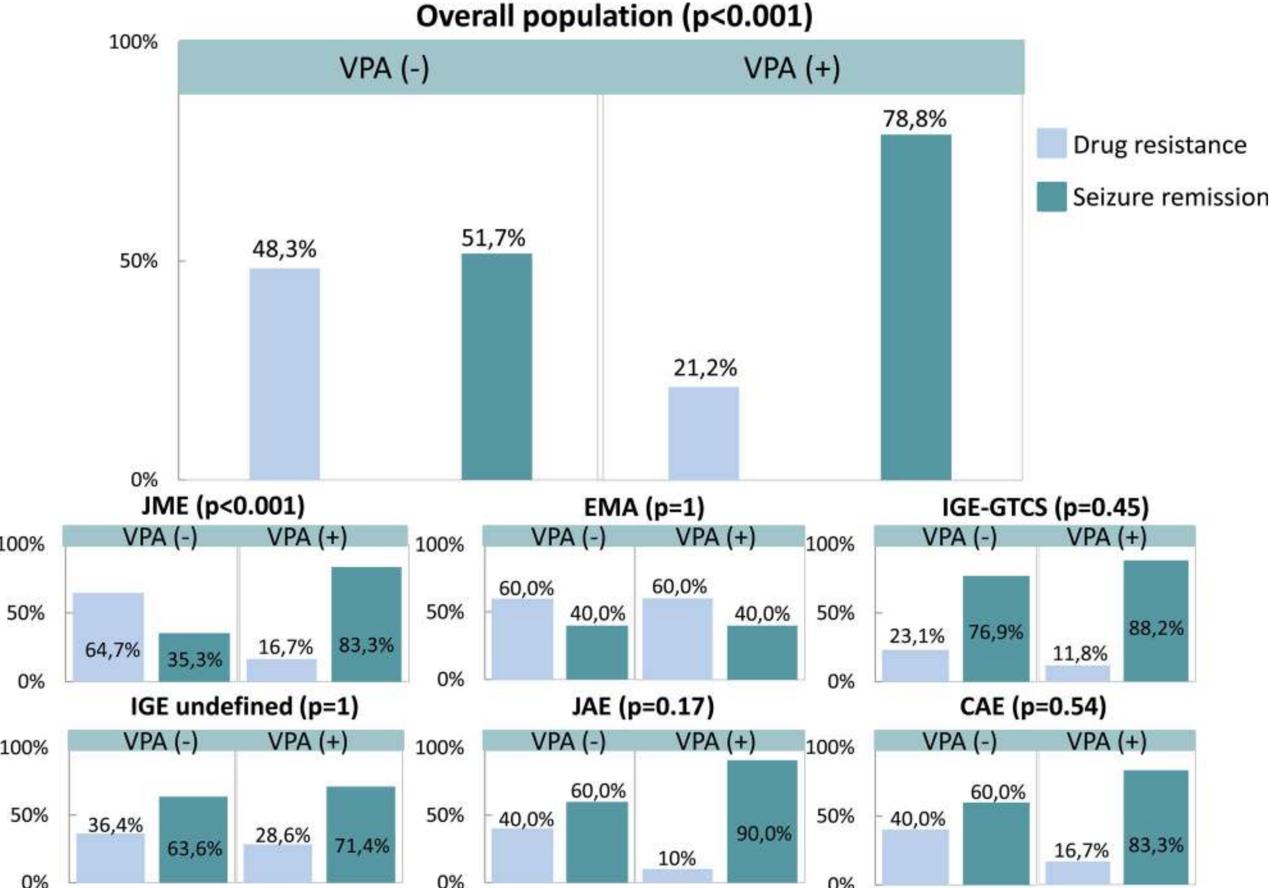


FIGURE 4 Seizure remission rate at last medical observation in the overall population (upper panels) and in specific epilepsy syndromes (lower panels) according to VPA status (VPA+ vs VPA-)

restricted to JME alone, M1b showed that VPA resulted to be associated with GTCS remission (OR 0.06, 95% CI 0.003-0.39; $P = .01$). Besides, M1c showed that the ineffectiveness of any first-line monotherapy predicted the need for polytherapy (OR 12.57, 95% CI 6.24-26.77; $P < .001$). No significant correlation was documented between either specific epileptic syndromes or clinical features and the necessity of a combination therapy.

Finally, M2 demonstrated that the VPA treatment ineffectiveness at any time during follow-up was significantly associated with polytherapy compared with the failure of other AEDs (OR 9.69, 95% CI 2.15-73.54; $P = .009$).

4 | DISCUSSION

To the best of our knowledge, this is the first study to explore the possible prognostic implications of VPA avoidance in women of childbearing potential. Despite the growing concerns about the possible consequences of restrictions on VPA use among female patients,[26–28] no study has yet assessed the potential impact of VPA avoidance/switch on seizure control.

As regards seizure outcome, persisting seizures were reported at the last observation by 37.4% of our patients, almost half of whom were still experiencing GTCS. The rate of remission in our population was lower than expected,[4,6] probably because this study included mainly patients with IGE persisting in adult life, who are more likely to achieve unsatisfactory seizure control.[29] Notwithstanding, in this special population, the avoidance/withdrawal of VPA could have been an additional determinant. Indeed, in our study, the SR rate at final observation differed according to "VPA status" (78.7% in the VPA+ group vs 51.7% in the VPA- group). Moreover, the regression models confirmed that the presence of VPA in the patients' drug regimen at final observation was the only factor strongly associated with SR (OR 0.26). Although age and follow-up duration were significantly different between VPA+ and VPA- patients, their potential influence on seizure outcome was ruled out by statistical analysis. VPA treatment at final observation was also found to be associated with freedom from GTCS, both in the overall population and in the JME group.

In addition to this, VPA ineffectiveness at any time (compared with other drugs) predicted the need for polytherapy at the end of follow-up (OR 9.69), suggesting an increased risk of unsatisfactory seizure control with alternative monotherapies.[29]

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Demographic and clinical characteristics of patients taking (VPA+) or not (VPA-) VPA at final observation

	VPA+ (80 pts)	VPA− (118 pts)	*P* value
Age, y, median (IQR)	31.6 (23.3-43.2)	26.2 (20.4-35.2)	**.004**
Follow-up, y, median (IQR)	17.6 (8.6-26.8)	7.9 (4.1-18.7)	**<.001**
EEG photosensitivity, n (%)	28	31	.2
Psychiatric comorbidity, n (%)	12	14	.5
Febrile seizures, n (%)	10	8	.2
Number of AEDs, n, mean	1.56	1.47	.32
Number of patients on polytherapy, n (%)	36 (46.15%)	44 (37.3%)	.27
CAE, n (%)	6 (7.5%)	5 (4.2%)	.29
JAE, n (%)	10 (12.5%)	15 (12.7%)	.96
JME, n (%)	30 (37.5%)	52 (44.1%)	.35
EMA, n (%)	10 (12.5%)	10 (8.5%)	.35
IGE-GTCS, n (%)	17 (21.2%)	26 (22%)	.89
IGE-undefined, n (%)	7 (8.8%)	10 (8.5%)	.94

Note: Bold indicates the statistical significance values.

In our work, the rate of SR in the VPA- group (51.7%) was dramatically lower than that reported in previous studies focusing on IGE prognosis.[4,6] This discrepancy suggests that VPA avoidance might determine a less favorable seizure control in female patients. However, this observation should be cautiously interpreted, in light of the several limitations of this study such as its retrospective nature and the possible selection bias.

Our data appear to be in line with the criticism provided by the recent ILAE position paper about the ban of VPA in childbearing age, especially in the management of more insidious epileptic syndromes. These considerations seem to be further supported by our results about the switch from VPA to other AEDs (and vice versa). Indeed, in the group of patients who switched VPA for other AEDs, the proportion experiencing clinical worsening was remarkable. Accordingly, a better seizure control was documented in the cohort of patients who switched back to VPA, when compared with patients who did not (76.9% vs 13%). The evidence of a large proportion of patients experiencing clinical worsening after VPA substitution was already reported by studies investigating the same issue in young women who withdrew VPA during pregnancy.[30]

In our work, the main reason for VPA substitution lay in the planning of a pregnancy. When focusing on these patients, we observed a lower SR rate and a higher drug burden (ie, polytherapy) at last observation with respect to the time preceding VPA switch. In the same group, poor seizure control justified VPA reintroduction in almost 40% of the cases. This observation seems to provide additional information about the debated topic of VPA use in this specific population,[31] highlighting the need for an accurate counselling, which especially deals with the risk of seizure worsening and the potential increase of drug burden.

Although our study provided some interesting results, it was limited by several factors including (a) the intrinsic bias depending on its retrospective design; (b) the challenging interpretation of the patients' outcomes due to the differences in follow-up duration within the study population; (c) a selection bias related to the enrollment of adult subjects followed in a tertiary epilepsy center, where a higher rate of drug resistance is usually expected; and (d) the poor availability of data concerning novel AEDs that were necessarily underprescribed in our cohort.

In conclusion, our study seems to suggest that doing without VPA can expose women with IGE to an increased risk of unsatisfactory seizure control. In light of these findings, clinicians should thoroughly discuss with their patients about potential risks and benefits of VPA avoidance, especially in cases of specific syndromic contexts (eg, JME). Overall, our data further highlight the need for alternative drugs that may ensure both effectiveness and safety in women of reproductive age with IGE. However, prospective studies on larger populations are warranted to provide more solid evidence.

CONFLICT OF INTERESTS

None of the authors has any conflict of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Carlo Di Bonaventura https://orcid.org/0000-0003-1890-5409

REFERENCES

1. Mullen SA, Berkovic SF, ILAE Genetics Commission. Genetic generalized epilepsies. Epilepsia. 2018;59:1148–53.
2. Jallon P, Latour P. Epidemiology of idiopathic generalized epilepsies. Epilepsia. 2005;46(suppl 9):10–4.
3. McHugh JC, Delanty N. Epidemiology and classification of epilepsy: gender comparisons. Int Rev Neurobiol. 2008;83:11–26.

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4. Seneviratne U, Cook M, D'Souza W. The prognosis of idiopathic generalized epilepsy. Epilepsia. 2012;53:2079–90.
5. Sullivan JE, Dlugos DJ. Idiopathic generalized epilepsy. Curr Treat Options Neurol. 2004;6:231–42.
6. Semah F, Picot MC, Adam C, Broglin D, Arzimanoglou A, Bazin B, et al. Is the underlying cause of epilepsy a major prognostic factor for recurrence? Neurology. 1998;51:1256–62.
7. Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, et al. The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. Lancet. 2007;369:1016–26.
8. Bourgeois B, Beaumanoir A, Blajev B, de la Cruz N, Despland PA, Egli M, et al. Monotherapy with valproate in primary generalized epilepsies. Epilepsia. 1987;28(suppl 2):S8–11.
9. Tomson T, Battino D, Perucca E. Valproic acid after five decades of use in epilepsy: time to reconsider the indications of a time-honoured drug. Lancet Neurol. 2016;15:210–8.
10. Meador KJ, Baker GA, Browning N, Cohen MJ, Bromley RL, Clayton-Smith J, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. Lancet Neurol. 2013;12:244–52.
11. Tomson T, Battino D, Bonizzoni E, Craig J, Lindhout D, Perucca E, et al. Comparative risk of major congenital malformations with eight different antiepileptic drugs: a prospective cohort study of the EURAP registry. Lancet Neurol. 2018;17:530–8.
12. Cohen MJ, Meador KJ, Browning N, May R, Baker GA, Clayton-Smith J, et al. Fetal antiepileptic drug exposure: adaptive and emotional/behavioral functioning at age 6 years. Epilepsy Behav. 2013;29:308–15.
13. Meador KJ, Baker GA, Browning N, Clayton-Smith J, Combs-Cantrell DT, Cohen M, et al. Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs. N Engl J Med. 2009;360:1597–605.
14. Christensen J, Grønborg TK, Sørensen MJ, Schendel D, Parner ET, Pedersen LH, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. JAMA. 2013;309:1696–703.
15. Hernández-Díaz S, Smith CR, Shen A, Mittendorf R, Hauser WA, Yerby M, et al. Comparative safety of antiepileptic drugs during pregnancy. Neurology. 2012;78:1692–9.
16. Campbell E, Kennedy F, Russell A, Smithson WH, Parsons L, Morrison PJ, et al. Malformation risks of antiepileptic drug monotherapies in pregnancy: updated results from the UK and Ireland Epilepsy and Pregnancy Registers. J Neurol Neurosurg Psychiatry. 2014;85:1029–34.
17. European Medicines Agency. Assessment report. Procedure under Article 31 of Directive 2001/83/EC resulting from pharmacovigilance data, 2014. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Referrals_document/Valproate_and_related_substances_31/Recommendation_provided_by_Pharmacovigilance_Risk_Assessment_Committee/WC500177352.pdf. Accessed October 9, 2014.
18. Biraben A, Allain H, Scarabin JM, Schück S, Edan G. Exacerbation of juvenile myoclonic epilepsy with lamotrigine. Neurology. 2000;55:1758.
19. Auvin S, Chhun S, Berquin P, Ponchel E, Delanoë C, Chiron C. Aggravation of absence seizure related to levetiracetam. Eur J Paediatr Neurol. 2011;15:508–11.
20. Perucca E, Tomson T. The pharmacological treatment of epilepsy in adults. Lancet Neurol. 2011;10:446–56.
21. Tomson T, Marson A, Boon P, Canevini MP, Covanis A, Gaily E, et al. Valproate in the treatment of epilepsy in girls and women of childbearing potential. Epilepsia. 2015;56:1006–19.
22. Glauser TA, Cnaan A, Shinnar S, Hirtz DG, Dlugos D, Masur D, et al. Ethosuximide, valproic acid, and lamotrigine in childhood absence epilepsy. N Engl J Med. 2010;362:790–9.
23. Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, et al. ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017;58:512–21.
24. Appleton RE, Panayiotopoulos CP, Acomb BA, Beirne M. Eyelid myoclonia with typical absences: an epilepsy syndrome. J Neurol Neurosurg Psychiatry. 1993;56:1312–6.
25. Joshi CN, Patrick J. Eyelid myoclonia with absences: Routine EEG is sufficient to make a diagnosis. Seizure. 2007;16:254–60.
26. Lawthom C. Valproate and epilepsy: for women as well as men. Pract Neurol. 2018;18:222–3.
27. Craig JJ. Women and valproate: what should neurologists do? Pract Neurol. 2018;18:219–21.
28. Sisodiya SM. Valproate and childbearing potential: new regulations. Pract Neurol. 2018;18:176–8.
29. Nicolson A, Appleton RE, Chadwick DW, Smith DF. The relationship between treatment with valproate, lamotrigine, and topiramate and the prognosis of the idiopathic generalised epilepsies. J Neurol Neurosurg Psychiatry. 2004;75:75–9.
30. Tomson T, Battino D, Bonizzoni E, Craig J, Lindhout D, Perucca E, et al. Withdrawal of valproic acid treatment during pregnancy and seizure outcome: Observations from EURAP. Epilepsia. 2016;57:e173–177.
31. Serafini A, Gerard E, Genton P, Crespel A, Gelisse P. Treatment of Juvenile Myoclonic epilepsy in patients of child-bearing potential. CNS Drugs. 2019;33:195–208.

How to cite this article: Cerulli Irelli E, Morano A, Cocchi E, et al. Doing without valproate in women of childbearing potential with Idiopathic generalized epilepsy: Implications on seizure outcome. *Epilepsia*. 2020;61:107–114. https://doi.org/10.1111/epi.16407

15281167, 2020, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/epi.16407 by Columbia University Libraries, Wiley Online Library on [24/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Authors: Elliott, MD; Marasa, M; Cocchi, E; Vena, N; Zhang, JY; Khan, A; Murthy, SK; Bheda, S; Rasouly, HM; Povysil, G; Kiryluk, K; Gharavi, AG

Title: Clinical and Genetic Characteristics of CKD Patients with High-risk APOL1 Genotypes

Running title: Features of CKD in high-risk APOL1

Manuscript Type: Original Article - Clinical Research

Manuscript Category:

Funders:
Funding Institutions: Department of Defence Research Award, (Grant / Award Number: 'PR201425'). Dr. Khan is supported by grants from the Foundation for the National Institutes of Health (Grant/Award Numbers: K25DK128563 and UL1TR001873). Dr. Elliott was supported by a Helios post-fellowship training award from the University of Calgary.

Financial Disclosure:

Dr. Gharavi reports receiving research grants from the National Institutes of Health and the Renal Research Institute. A. Gharavi also reports Consultancy: Astra Zeneca Center for genomics research, Goldfinch Bio: Actio biosciences, Novartis: Travere; Ownership Interest: Actio; Research Funding: Natera; Honoraria: Sanofi, Alnylam; and Advisory or Leadership Role: Editorial board: JASN and Journal of Nephrology. Dr. Povysil is an employee of and has ownership interest in Waypoint Bio. E. Cocchi reports Research Funding: American Society of Nephrology. H. Milo Rasouly reports Research Funding: Natera. K. Kiryluk reports Consultancy: Calvariate, HiBio; and Research Funding: AstraZeneca, Vanda, Bioporto, Aevi Genomics, and Visterra.
Because Ali Gharavi is an Associate Editor of the Journal of the American Society of Nephrology, he was not involved in the peer review process for this manuscript. A guest editor oversaw the peer review and decision-making process for this manuscript.
Study Group/Organization Name:
Study Group Members' Names:
Total number of words: 2900

ACCEPTED

Abstract:

Background: *APOL1* genotype has significant effects on kidney disease development and progression that vary among specific causes of kidney disease, suggesting the presence of effect modifiers.

Methods: We assessed the risk of kidney failure and eGFR decline rate in patients with chronic kidney disease (CKD) carrying high-risk (N=239) and genetically matched low-risk (N=1187) *APOL1* genotypes. Exome sequencing revealed monogenic kidney diseases. Exome-wide association studies and gene-based and gene-set based collapsing analyses evaluated genetic modifiers of the effect of *APOL1*

Copyright © 2023 by the American Society of Nephrology

genotype on CKD.

Results: Compared to genetic ancestry-matched CKD cases with low-risk *APOL1* genotypes, those with high-risk *APOL1* genotypes had a higher risk of kidney failure (HR= 1.58), a higher decline in eGFR (6.55 vs 3.63 mL/min/1.73m^2/year), and were younger at time of kidney failure (45.1 vs 53.6 years), with the G1/G1 genotype demonstrating the highest risk. The rate for monogenic kidney disorders was lower among CKD patients with high-risk *APOL1* genotypes (2.5%) compared to those with low-risk genotypes (6.7%). Gene-set analysis identified an enrichment of rare missense variants in the inflammasome pathway in individuals with high-risk *APOL1* genotypes and CKD (OR = 1.90).

Conclusions: In this genetically matched cohort, high-risk *APOL1* genotypes were associated with an increased risk of kidney failure and eGFR decline rate, with a graded risk between specific high-risk genotypes and a lower rate of monogenic kidney disease. Rare missense variants in the inflammasome pathway may act as genetic modifiers of *APOL1* effect on kidney disease.

Significance Statement

APOL1 high-risk genotypes confer a significant risk of kidney disease, but variability in patient outcomes suggests the presence of modifiers of the *APOL1* effect. We show that a diverse population of CKD patients with high-risk *APOL1* genotypes have an increased lifetime risk of kidney failure and higher eGFR decline rates, with a graded risk among specific high-risk genotypes. CKD patients with high-risk *APOL1* genotypes have a lower diagnostic yield for monogenic kidney disease. Exome sequencing revealed enrichment of rare missense variants within the inflammasome pathway modifying the effect of *APOL1* risk genotypes, which may explain some clinical heterogeneity.

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Journal of the American Society of Nephrology Publish Ahead of Print
DOI: 10.1681/ASN.0000000000000095

Clinical and Genetic Characteristics of CKD Patients with High-risk *APOL1* Genotypes

Mark D. Elliott[1,2,3,4], Maddalena Marasa[1,2], Enrico Cocchi[1,2,5], Natalie Vena[2,3], Jun Y. Zhang[1], Atlas Khan[1,2], Sarath Krishna Murthy[2], Shiraz Bheda[2], Hila Milo Rasouly[1,2], Gundula Povysil[3], Krzysztof Kiryluk[1,2,3], Ali G. Gharavi[1,2,3]

1 Columbia University Vagelos College of Physicians and Surgeons, Division of Nephrology, Department of Medicine

2 Columbia University Vagelos College of Physicians and Surgeons, Center for Precision Medicine and Genomics, Department of Medicine

3 Columbia University Vagelos College of Physicians and Surgeons, Institute for Genomic Medicine

4 University of Calgary, Division of Nephrology, Department of Medicine

5 Universita' degli Studi di Torino, Department of Pediatrics

Short running title: Features of CKD in high-risk *APOL1*

Word Count Abstract: 240 Word

Count Text: 2670

Address for correspondence Ali G.

Gharavi, MD

1150 St. Nicholas Ave., Russ Berrie Pavilion 413, New York, NY 10032 Email:

ag2239@columbia.edu

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Abstract:

Background: *APOL1* genotype has significant effects on kidney disease development and progression that vary among specific causes of kidney disease, suggesting the presence of effect modifiers.

Methods: We assessed the risk of kidney failure and eGFR decline rate in patients with chronic kidney disease (CKD) carrying high-risk (N=239) and genetically matched low-risk (N=1187) *APOL1* genotypes. Exome sequencing revealed monogenic kidney diseases. Exome-wide association studies and gene-based and gene-set based collapsing analyses evaluated genetic modifiers of the effect of *APOL1* genotype on CKD.

Results: Compared to genetic ancestry-matched CKD cases with low-risk *APOL1* genotypes, those with high-risk *APOL1* genotypes had a higher risk of kidney failure (HR= 1.58), a higher decline in eGFR (6.55 vs 3.63 mL/min/1.73m^2/year), and were younger at time of kidney failure (45.1 vs 53.6 years), with the G1/G1 genotype demonstrating the highest risk. The rate for monogenic kidney disorders was lower among CKD patients with high-risk *APOL1* genotypes (2.5%) compared to those with low-risk genotypes (6.7%). Gene-set analysis identified an enrichment of rare missense variants in the inflammasome pathway in individuals with high-risk *APOL1* genotypes and CKD (OR = 1.90).

Conclusions: In this genetically matched cohort, high-risk *APOL1* genotypes were associated with an increased risk of kidney failure and eGFR decline rate, with a graded risk between specific high-risk genotypes and a lower rate of monogenic kidney disease. Rare missense variants in the inflammasome pathway may act as genetic modifiers of *APOL1* effect on kidney disease.

Introduction:

Individuals of African ancestry are at an increased risk of developing chronic kidney disease (CKD) and in the United States of America (USA) carry a three-fold increased risk of incident kidney failure compared to Americans of European ancestry.[1–3] This is attributable to interacting environmental, social, and genetic factors.[4] The G1 and G2 *APOL1* alleles are unique because they are strongly associated with kidney disease while being common in populations of sub- Saharan African ancestry and therefore account for a significant portion of these ancestry- based differences.[5–7] In the USA, the G1 and G2 alleles are most common in Americans of sub- Saharan African ancestry, where up to 15% have high-risk genotypes and are very rare in Americans of European ancestry.[8–10]

In individuals with and without prevalent kidney disease, high-risk *APOL1* genotypes are associated with an increased rate of decline in estimated glomerular filtration rate (eGFR), and an increased risk of incident proteinuria and kidney failure. [5,6,11–13] Despite this, only 15 to 20% of individuals with high-risk *APOL1* genotypes develop kidney disease in their lifetime.[11,12] Additionally, the effect of *APOL1* risk genotype is highly variable between specific kidney diseases and self-declared Black ancestry is associated with more rapid progression independent of *APOL1* genotype, suggesting the presence of

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

modifiers.[14,16,17] One known modifier is elevations in interferon levels, which induce *APOL1* expression in podocytes, potentially explaining the HIV and SARS-CoV-2 associated collapsing glomerulopathy seen in individuals with high-risk *APOL1* genotypes.[5,10,13–15] Recent studies have implicated the *NLRP3* inflammasome and pyroptosis in interferon induced *APOL1* cytotoxicity, suggesting these pathways as potential modifiers for *APOL1*-associated phenotypes.[16,17]

Several studies have attempted to identify genetic modifiers of the effect of *APOL1* risk genotype on kidney disease but no strong signals have been identified and replicated. A gene- based collapsing analysis applied to a smaller sample from the same biobank used for this study identified a suggestive enrichment of rare, non-benign variants in the *AHDC1* gene in individuals with high-risk *APOL1* genotypes and CKD.[23] Admixture mapping in individuals with high-risk *APOL1* genotypes and FSGS found an increase in African ancestry at the *UBD* locus that correlated with lower *UBD* expression and increased *APOL1*-mediated cytotoxicity suggesting a role as a modifier.[14] A relatively small genome-wide association study (GWAS) of common variants did not identify any significant associations or interactions with *APOL1* risk genotype and did not replicate *AHDC1* or *UBD* leaving us with a limited understanding of the genetic architecture of *APOL1*-associated kidney disease.[24]

Here we studied the effect of *APOL1* risk genotype on the risk of kidney failure and the rate of eGFR decline in a large, ethnically diverse, cohort with all cause CKD. We used exome sequencing data to genetically match individuals with high- and low- risk *APOL1* genotypes and compared the impact of genetically determined and self-declared ancestry on outcomes.
Exome sequencing was also used to evaluate individuals for monogenic kidney disease and novel *APOL1* variants. Finally, genetic modifiers of kidney disease in individuals with high-risk

APOL1 genotypes were evaluated using an exome-wide association study (ExWAS) and gene- based and gene-set based collapsing analyses.

ACCEPTED

METHODS:

Study Population and Design:

Individuals with CKD at New York Presbyterian Hospital/Columbia University Irving Medical Center in New York City, USA, were recruited between October 2013 and July 2021 into the Genetic Studies of Chronic Kidney Disease biobank with Columbia University Institutional Review approval. All participants provided informed consent. So far 4388 individuals from the biobank had undergone exome sequencing following standard protocols at the Institute for Genomic Medicine (IGM) at Columbia University (supplementary methods).[23,25] Retrieval, annotation, and filtering of variants were done using analysis tool for annotated variants (ATAV).[26]

APOL1 risk genotypes were defined by the presence of the G1 ($G1_G$ 22-36661906-A-G (rs73885319) and $G1_M$ 22-36662034-T-G (rs60910145)) and G2 (22-36662041-AATAATT-A (rs71785313)) alleles. Within the biobank 239 individuals were identified with high-risk *APOL1*

Copyright © 2023 by the American Society of Nephrology

genotypes (genotypes G1/G1, G2/G2, or G1/G2) along with a comparison cohort of 1187 genetic ancestry-matched individuals without high-risk *APOL1* genotypes, including 336 single risk allele carriers that have either the G1/G0 or G2/G0 genotypes, selected by Louvain clustering of 12,400 informative genetic ancestry markers (Figure 1, Supplementary Figure 1, supplementary methods).[26,27] Louvain clustering was used because it does not require the number of ground truth clusters to be known and does not make assumptions about cluster size, density or shape.[28] Clinical data were collected at recruitment and updated from the electronic health record (EHR). Descriptive statistics were performed in R 4.0.3 (R Core Team, Vienna Austria).[2929]

Analysis of Clinical Diagnosis:

The association of *APOL1* risk alleles with common primary causes of kidney disease was evaluated using logistic regression. CKD cases with a diagnosis of either autosomal dominant polycystic kidney disease (ADPKD) or congenital anomalies of the kidney and urinary tract (CAKUT) were used as a reference cohort as these disorders are not known to be associated by *APOL1* genotype. This analysis was adjusted for sex, diabetes mellitus, hypertension, genetic ancestry cluster, quintile of the median income of their home ZIP code as an area-based surrogate of socioeconomic status, and ICD-10 based Swiss AHQI Elixhauser co-morbidity score.[30–32] Missing Elixhauser co-morbidity score and ZIP code based median income data were imputed using multiple imputations by chained equations (supplementary methods).[3333]

Time to Event Analyses:

Kaplan-Meier, Cox proportional hazard, and Fine-Gray competing risk models were used to evaluate the onset of kidney failure (defined as chronic dialysis or kidney transplantation) with a median of 3.5 years of follow-up data.[34,35] Cox proportional hazard and Fine-Gray competing risk models were performed unadjusted and adjusted for sex, family history of kidney disease,presence of monogenic kidney disease diagnosis, diabetes mellitus, hypertension, genetic ancestry cluster, quintile of the median income of their home ZIP code as an area-based surrogate of socioeconomic status, and ICD-10 based Swiss AHQI Elixhauser co-morbidity score.[30–32] The proportional-hazard assumptions were met with the inclusion of time-varying covariates. The effects of specific high-risk *APOL1* genotypes, single allele carrier status, and the interaction of clinical diagnosis and *APOL1* genotype were evaluated in secondary analyses.

eGFR Decline Analysis:

CKD patients with 2 or more serum creatinine measurements collected at least 1 year apart were included. CKD-EPI (2021) without race term was used to estimate GFR.[36,37] Values collected at less than 14.2 years of age, where the CKD-EPI formula loses accuracy, and after the initiation of dialysis or transplantation were excluded.[38] 94 CKD patients (2216 creatinine values) with high-risk *APOL1* genotypes, were compared to 519 genetic ancestry-matched CKD patients (20869 creatinine values) without high-risk *APOL1* genotypes which included 162 (6916 creatinine values) single risk allele carriers. A median of 5.5 years of data and 21 creatinine values were available per individual. Linear

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

mixed-effects modeling evaluated the eGFR decline rate using a random slope and intercept for each subject, nested within their primary cause of kidney disease and fixed effects for *APOL1* risk genotype, family history of kidney disease, hypertension, diabetes mellitus, initial eGFR value, and genetic ancestry cluster.[39,40] The interaction term of *APOL1* risk genotype and time since the first creatinine value represented the decline in eGFR per unit time. Numeric variables were centered and scaled. As a secondary analysis the effects of specific high-risk *APOL1* genotypes and single risk allele carrier status were evaluated. Fixed effect confidence intervals were estimated using parametric bootstrapping.

Diagnostic Analysis of Monogenic Kidney Disorders:

Sequencing data were analyzed using ATAV to identify pathogenic and likely pathogenic variants diagnostic of an individual's kidney disease using a curated list of 625 genes known to cause kidney disease. [41] Variant interpretation followed the American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP) guidelines with the implementation of criteria updates recommended by the ClinGen Sequence Variant Interpretation working group.[42] Diagnostic variants were reviewed at multidisciplinary genetic sign-out rounds and reached consensus interpretation and are considered returnable after Clinical Laboratory Improvement Amendments (CLIA) certified lab confirmation. *APOL1* high- risk genotypes are also considered returnable after CLIA confirmation. Variants that did not fit the individual's kidney disease phenotype were considered non-diagnostic. Novel *APOL1* coding variants were also analyzed. The diagnostic rate based on *APOL1* risk genotype was compared using logistic regression with self-described race and ethnicity and family history as covariates.

Exome based association studies:

The 239 CKD patients with high-risk *APOL1* genotypes were compared to a genetic ancestry- matched comparison cohort without CKD using Louvain clustering.[43] Following quality control (QC), coverage harmonization and kinship pruning 232 CKD cases with high-risk *APOL1* genotypes remained and were compared to 5482 genetic ancestry-matched controls without CKD from the IGM sequencing database (Figure 1, supplementary methods). In addition to the full cohort, two *a priori* subgroup analyses were performed, examining those who developed kidney failure (168 cases; 4413 matched controls) representing the most severely effected individuals, and those with FSGS (79 cases; 4389 matched controls) which has been strongly associated with *APOL1* risk genotype.[18]

To search for potential genetic modifiers of the effect of *APOL1* genotype on kidney disease, we first conducted a single variant association study using Regenie 2.0 (supplementary methods).[44] Sex and the first 10 principle components (PCs) were included as covariates. GCTA 1.93.3 was used to perform conditional analysis where loci P-values are recalculated by conditioning upon the top loci signal.[45,46] The conditional analysis was repeated by selecting the new top loci signal until no new signals reached the prespecified significance threshold. The approximate Firth likelihood ratio test was used to evaluate the association between single variants and disease status. The ExWAS significance threshold was adjusted for the total number of exome-wide variants tested across the three analyses using a q

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

value significance threshold of Q<0.05.

We next performed a gene-based rare-variant gene-based collapsing analysis, applying one recessive and 11 dominant nonsynonymous models for selecting qualifying variants (QVs) (Supplementary Table 1). The work flow, QC, and coverage harmonization followed prior published reports (supplementary methods). [23,47,48] 137 of the identified cases with high-risk *APOL1* genotypes included in this analysis were used in a published gene-based collapsing analysis of patients with CKD.[23]

Finally, we conducted gene-set based collapsing analysis with three predefined gene-sets: 1. the pyroptosis gene-set 2. the inflammasome gene-set and 3. known autosomal recessive kidney disease genes (Supplementary Table 2). These gene-sets were evaluated across all 12 models from the rare-variant gene-based collapsing analysis (Supplementary Table 1). Gene and gene-set analyses used the exact 2-sided Cochran-Mantel-Haenszel test of independence across ancestry clusters and a q value significance threshold of Q<0.05.[49] Once a suggestive signal was identified, we conducted two follow-up analyses for the pyroptosis and inflammasome gene-sets. The first compared CKD cases with high-risk *APOL1* genotypes (N =232) to controls with high-risk *APOL1* genotypes without CKD (N = 176) to assess whether enrichment was due to high-risk *APOL1* genotype independent of CKD status and is termed high-risk only (HRO). The second analysis compared CKD cases with 1 *APOL1* risk allele (N = 339) to controls with 0 *APOL1* risk alleles without CKD (N = 4563) to assess whether enrichment was instead associated with CKD in the absence of high-risk *APOL1* genotypes and is termed *APOL1* het.

RESULTS:

Study population:

Baseline characteristics of the 239 CKD cases with high-risk *APOL1* genotypes, and 1187 genetic ancestry-matched CKD cases with low-risk *APOL1* genotypes are summarized in Table 1. Novel rare variants in *APOL1* were identified in the low-risk genotype cohort, but not in the high-risk genotype cohort (Supplemental Table 3). A large proportion of individuals with high-risk and low-risk genotypes self-reported their race and ethnicity as Black or African American and not Hispanic or Latinx (61%, 22% respectively) or White and Hispanic or Latinx (17%, 36% respectively). CKD cases with high-risk *APOL1* genotypes were more likely to carry a primary diagnosis of FSGS and hypertension-attributed CKD (OR = 17.76, 3.91; P = 6.1×10^{-14}, 8.3×10^{-4} respectively; Figure 2, Table 2). In addition, we noted an increased risk of FSGS in those with a single *APOL1* risk allele (OR = 1.85; P = 0.041).

Lifetime risk of kidney failure and rate of eGFR decline

This analysis included 239 CKD cases with high-risk *APOL1* genotypes and 1187 genetic ancestry matched CKD cases with low-risk genotypes. CKD cases with high-risk *APOL1* genotypes were younger at kidney failure compared to those with low-risk genotypes (45.1 vs 53.6 years; log- rank P = 2.0×10^{-6}; Figure 3, Table 3). In the adjusted model, CKD cases with high-risk *APOL1* genotypes were at an

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

increased lifetime risk of developing kidney failure compared to those with low-risk genotypes (HR = 1.58; P = 1.5 x 10^{-6}; Table 3). The inclusion of the competing risk of death showed similar results (HR = 1.59; P = 2.7 x 10^{-6}; Table 3). Similarly, CKD cases with high-risk *APOL1* genotypes experienced a higher rate of eGFR decline compared to those with low-risk genotypes (6.55 vs 3.63 mL/min/1.73m^2/year; P = 6.9 x10^{-4}; Figure 4, Table 3). We also detected a significant interaction between *APOL1* risk genotype and lupus nephritis for the risk of kidney failure but not with other primary diagnoses (OR = 2.41; P = 0.027; Supplementary Figure 2). In this genetically matched cohort, only self-declared race and ancestry of Black/African American, Hispanic or Latinx was significantly associated with risk of progression (HR = 1.52; P = 0.049) while other self-declared race and ethnicities, and genetic ancestry clusters did not have an impact. None of these affected the impact of *APOL1* risk genotype on kidney outcomes.

In secondary analyses (Table 3, Supplementary Table 4), we noted a graded risk of kidney failure in CKD cases with high-risk *APOL1* genotypes. The G1/G1 genotype was associated with the highest risk of kidney failure while G1/G2 carried an intermediate risk. Consistent with these data, the G1/G1 genotype group had the highest rate of eGFR decline and developed kidney failure earliest, followed by the G1/G2 genotype, the G2/G2 genotype, and the low-risk genotype groups (Supplementary Figure 3, Supplementary figure 4). CKD cases carrying a single *APOL1* risk allele showed a nominally higher rate of eGFR decline compared to those with zero *APOL1* risk alleles but no difference in the risk of kidney failure(4.43 vs 3.23 mL/min/1.73m^2/year; P = 0.066; HR = 1.08; P = 0.40; Supplementary Figure 5).

Monogenic Kidney Diseases:

There was a lower rate of monogenic diagnoses in CKD cases with high-risk *APOL1* genotypes (6/239, 2.5%) compared to low-risk genotypes (79/1187, 6.7%; OR = 0.38; P = 0.032; Supplementary Table 5) despite similar rates of positive family history (34.7% vs 33.6%). Positive family history of kidney disease was associated with a diagnosis of monogenic kidney disease, but not a high-risk *APOL1* genotype (OR = 3.04, 1.11; P = 1.17 x 10^{-6}, 0.50, respectively). Removing those cases with a monogenic kidney disease did not impact the effect of *APOL1* risk genotype on the lifetime risk of kidney failure. There was no difference in diagnostic rate between the genetic ancestry clusters or between self-reported race and ethnicity. Additionally, we identified 6 individuals with sickle cell anemia and 83 individuals with sickle cell trait with no enrichment based on *APOL1* risk genotype.

Exome based association studies:

We analyzed exome sequencing data from CKD patients with high-risk *APOL1* genotypes and compared them with genetic ancestry-matched individuals without known kidney disease. Two prespecified subgroups of CKD patients with kidney failure and FSGS were also analyzed. As expected, the ExWAS identified three independent significant signals after step-wise conditional analysis; the *APOL1* $G1_M$ allele, (OR = 5.17, P = 4.52 x 10^{-78}; Q = 5.94 x 10^{-73}), the *APOL1* G2 allele (OR = 3.74, P =

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

2.48×10^{-33}; Q = 1.63×10^{-28}) and the synonymous variant 22-36661842-G-A which forms part of the G3 haplotype in *APOL1* (OR = 0.16, P = 1.90×10^{-7}; Q = 0.015).500 Within the gene-based collapsing analyses no genes reached the prespecified level of significance (Supplementary Table 6, Supplementary Table 7). The top two enrichment signals were within the subgroup of individuals with FSGS, in the genes *DHDH* (OR, 95% CI = 38.75, 7.42– 187; P = 2.33×10^{-5}; Q = 1.00) and *NLRP1* (OR, 95% CI = 13.00, 4.20 – 34.19; P = 2.33×10^{-5}; Q =1.00). We did not find any compelling signals in the previously reported *UBD* or *AHDC1* genes.

Gene-set based collapsing analyses demonstrated a significant enrichment of rare missense variants within the inflammasome gene-set in the full CKD case cohort and the subgroup with kidney failure (OR = 1.90, 2.03; P = 0.0005, 0.0008; Q = 0.038, 0.041 respectively; Supplementary Table 8). This was driven by 57 QVs in 45 CKD cases, none of which have been associated with disease in the literature (Supplementary Table 9). The enrichment remained significant when CKD cases with high-risk *APOL1* genotypes were compared to non-CKD controls with high-risk *APOL1* genotypes (OR = 3.31; P = 0.0005; Q = 0.038) but not when single risk allele *APOL1* carriers with CKD were compared to non-CKD cases with zero risk alleles (OR = 0.46; P = 0.0166; Q = 0.277), indicating specificity of the inflammasome signal to the CKD group with high risk *APOL1* genotypes. No significant signals were identified in the pyroptosis gene- set or within autosomal recessive kidney disease genes (Supplementary Table 8).

Discussion:

In this study of a diverse cohort of CKD patients, which included a large Hispanic and Latinx population, we confirmed that high-risk *APOL1* genotypes conferred an increased risk of developing kidney failure (HR = 1.58) and a higher rate of eGFR decline (difference of 2.92 mL/min/1.73m^2/year). We found a graded risk between high-risk genotypes, where G1/G1 were most affected. By including individuals with all causes of kidney disease, we found that those with high-risk *APOL1* genotypes were more likely to have a diagnosis of FSGS or hypertension-attributed CKD. There was also evidence that individuals with a single *APOL1* risk allele were more likely to have FSGS which had previously only been described in association with HIV associated Nephropathy.[18] The diagnostic rate of exome sequencing for monogenic kidney diseases was lower in individuals with high-risk *APOL1* genotypes. While the analysis of genetic modifiers of the effect of *APOL1* genotype on kidney disease did not yield any genome- wide significant genes or loci, there were suggestive signals in the inflammasome pathway in the gene-set analysis.

This report supports prior work showing an increased risk of kidney failure, higher rates of eGFR decline, and an enrichment of FSGS and hypertension-associated kidney disease in individuals with high-risk *APOL1* genotypes.[5,11,13,51] Compared to prior reports, we studied an ethnically diverse population where the high and low-risk genotype groups were genetically matched for ancestry and identified cases with a monogenic diagnoses of kidney disease. The availability of genetic information reduced the potential confounding effect of ancestry and concomitant monogenic disease on

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

progression. Our results also align with prior studies showing that CKD patients with high-risk *APOL1* genotypes develop kidney failure 7 to 9 years earlier than those without a high-risk genotype which has important prognostic implications for these individuals.[52,53]. Our study did identify a small residual effect of self-declared Black/African American, Hispanic or Latinx race and ethnicity on renal outcomes, after *APOL1* risk genotype and genetic ancestry was taken into account. Recent data also suggest that polygenic score for kidney disease may predict future risk of progression.[54] Hence, incorporation of genome-wide polygenic score may help in improved risk prediction in patients with high risk *APOL1* genotypes.

We identified a graded risk between specific high-risk *APOL1* genotypes, with the G1/G1 genotype associated with the highest risk, in concordance with allele specific effects that have been reported in assays of anti-trypanosomal activity and in model systems of *APOL1* associated kidney disease.[7,55–57] The *APOL1* G2 allele provided resistance to serum resistance- associated protein (SRA) producing *Trypanosoma brucei rhodesience* by abolishing the interaction of SRA with APOL1, however the G1 allele only had a moderate effect on the SRA- APOL1 interaction, suggesting it may lead to protection through other mechanisms, such as increased cytotoxicity.[7,57] In addition, individuals with the G1 allele who are infected by *T. b. gambiense,* which causes the majority of human diseases, more often develop a state of asymptomatic carriage, while those with the G2 allele are more likely to have clinically overt disease, further supporting allele specific effects.[55] These findings highlight the importance of separating specific high-risk *APOL1* genotypes in future analyses of observational and clinical trials, because of potential differences in prognosis. In addition, the nominal trend towards a higher rate of eGFR decline in subjects with a single *APOL1* risk allele is consistent with studies showing increased risk of kidney failure, CKD, cardiovascular disease and cardiomyopathy in additive models of *APOL1* risk variants, suggesting that heterozygotes carry some increased risk.[52,58,59]

The lower diagnostic rate of exome sequencing for monogenic kidney disease in CKD cases with high-risk *APOL1* genotypes suggests that dual genetic diagnoses are rarer in this patient population. Although a family history of kidney disease was predictive of a diagnosis of monogenic kidney disease, it was not predictive of *APOL1* risk genotype, implicating additional factors such as polygenic background or environmental effects as contributors to the effect of family history in this group.[25,60] These data are important for clinicians selecting patients for genetic testing.

The gene-based collapsing analyses did not identify significant modifiers of the effect of *APOL1* genotype, including the previously suggested modifiers *AHDC1* and *UBD*. [14,23] The suggestive signals in *NLRP1,* and the inflammasome gene-set are noteworthy because this pathway has been implicated in *APOL1* mediated cytotoxicity.[61] The cytotoxicity of *APOL1* risk variants may be due to increased expression following viral infections, inflammation and interferon signaling.[16,19] In model systems of *APOL1*-associated kidney disease, over-expression of *APOL1* risk variants was shown to stimulate both the simulator of interferon genes (STING) pathway and the inflammasome and ultimately lead to podocyte death by pyroptosis.[21] This cellular injury was attenuated by genetic depletion of either *NLRP3* or *GSDMD* within the inflammasome or the *STING* gene. Hence, genetic variants within these pathways may be plausible modifiers of the effect of *APOL1* risk genotype.

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

This study's limitations include the single center design which may introduce unmeasured center specific practices in the analyses, such as medication usage, the timing of initiation of kidney replacement therapy and transplantation referral. In the process of genetic ancestry- matching, only one African ancestry dominant cluster was identified, suggesting that the cohort did not not capture the full spectrum of genetic diversity in Africa.[8,9] The use of clinically collected creatinine values may produce variation due to differences in laboratory measurements. We included both inpatient and outpatient creatinine values in this analysis which could include creatinine values that are in flux as patients experience or recover from acute kidney injury, however the restriction of eGFR analyses to only those individuals with over one year of data aimed to reduce the effects of short-term creatinine changes, while the use of subject specific intercept and slopes within the eGFR slope modeling accounted for subject specific differences. The computation of co-morbidity scores from EHR ICD-10 codes is routinely performed but can introduce bias based on billing and diagnostic patterns. Collapsing analyses carry their own set of limitations including the identification of QVs that are biologically meaningful and sequencing data variability.[47] To address this we used a standardized QC and analysis pipeline that includes coverage harmonization along with measures of missense intolerance and multiple models to identify appropriate QV screening criteria. Power analyses suggest the rare-variant analyses performed were underpowered to identify many interactions with *APOL1* genotype, demonstrating the need for larger sample sizes to identify genetic modifiers and confirm the novel clinical associations detected in this study.

In conclusion, we confirmed that high-risk *APOL1* genotypes conferred an increased risk of developing kidney failure and a higher rate of eGFR decline, and identified a graded risk between high-risk genotypes, where G1/G1 were most affected. In single *APOL1* risk allele carriers we identified an increased risk of FSGS, and a trend towards a faster rate of eGFR decline. We found a lower diagnostic rate of exome sequencing for monogenic kidney diseases in individuals with high-risk *APOL1* genotypes, and identified suggestive signals in the inflammasome pathway in the gene-set analysis.

Author Contributions:

Mark Elliott: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing

Maddalena Marasa: Conceptualization, Data curation, Project administration, Writing – review & editing

Enrico Cocchi: Formal analysis, Software, Writing – review & editing

Natalie Vena: Data curation, Formal analysis, Writing – review & editing

Jun Zhang: Data curation, Writing – review & editing

Atlas Khan: Formal analysis, Software, Writing – review & editing

Sarathbabu Krishnamurthy: Formal analysis, Software, Writing – review & editing

Shiraz Bheda: Formal analysis, Software, Writing – review & editing

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Hila Milo-Rasouly: Methodology, Writing – review & editing

Gundula Povysil: Formal analysis, Software, Writing – review & editing

Krzysztof Kiryluk: Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing

Ali Gharavi: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – originaldraft, Writing – review & editing

Acknowledgments:

We thank all CUIMC participants and investigators who have been involved in the Genetic Studies of Chronic Kidney Disease biobank.

Disclosures:

Dr. Gharavi reports receiving research grants from the National Institutes of Health and the Renal Research Institute. A. Gharavi also reports Consultancy: Astra Zeneca Center for genomics research, Goldfinch Bio: Actio biosciences, Novartis: Travere; Ownership Interest: Actio; Research Funding: Natera; Honoraria: Sanofi, Alnylam; and Advisory or Leadership Role: Editorial board: JASN and Journal of Nephrology. Dr. Povysil is an employee of and has ownership interest in Waypoint Bio. E. Cocchi reports Research Funding: American Society of Nephrology. H. Milo Rasouly reports Research Funding: Natera. K. Kiryluk reports Consultancy: Calvariate, HiBio; and Research Funding: AstraZeneca, Vanda, Bioporto, Aevi Genomics, and Visterra.

Because Ali Gharavi is an Associate Editor of the Journal of the American Society of Nephrology, he was not involved in the peer review process for this manuscript. A guest editor oversaw the peer review and decision-making process for this manuscript.

Funding:

Funding Institutions: Department of Defence Research Award, (Grant / Award Number: 'PR201425'). Dr. Khan is supported by grants from the Foundation for the National Institutes of Health (Grant/Award Numbers: K25DK128563 and UL1TR001873). Dr. Elliott was supported by a Helios post-fellowship training award from the University of Calgary.

ACCEPTED

Data Sharing Statement:

Individual level data for 1426 participants with CKD is available via National Center for Biotechnology Information (NCBI) dbGaP database (phs001828.v2.p1, in process).

Supplementary Material Table of Contents:

Supplementary Table 1: Model characteristic used for gene- and gene-set based collapsing analyses.

Supplementary Table 2: Gene names used for gene-set analyses. Supplementary

Copyright © 2023 by the American Society of Nephrology

Table 3: Novel *APOL1* variants identified.

Supplementary Table 4: Expanded demographics and baseline data of included CKD patients, divided by APOL1 risk genotype, including single risk allele carriers and specific high-risk genotypes.

Supplementary Table 5: Monogenic kidney diseases identified within the study. Includes variant classification, and criteria asserted.

Supplementary Table 6: Top 15 ranked genes across the rare-variant collapsing analysis.

Supplementary Table 7: Top results of the common and rare variant secondary gene-based burden analysis.

Supplementary Table 8: Results from the gene-set based rare-variant collapsing analysis.

Supplementary Table 9: Qualifying variants identified within the inflammasome gene-set within the dominant rare missense model for the full cohort.

Supplementary Figure 1: Genetic ancestry and clustering using principal components analyses of exome ancestry markers on a UMAP projection.

Supplementary Figure 2: Lifetime risk of the development of kidney failure comparing subjects with specific high-risk APOL1 genotypes to low-risk APOL1 genotype.
Supplementary Figure 3: eGFR change over time in subjects by specific APOL1 genotype. eGFR decline rate calculated with covariates held at mean values.

Supplementary Figure 4: Lifetime risk of developing kidney failure for the interaction between APOL1 high-risk genotype and primary cause of kidney disease.

Supplementary Figure 5: eGFR change over time in subjects characterized by number of *APOL1* risk alleles.

Supplementary Methods.

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

References:

1. Saran, R. *et al.* US Renal Data System 2019 Annual Data Report: Epidemiology of Kidney Disease in the United States. *American Journal of Kidney Diseases* **75**, A6–A7 (2020).
2. Trinh, E., Na, Y., Sood, M. M., Chan, C. T. & Perl, J. Racial Differences in Home Dialysis Utilization and Outcomes in Canada. *Clin J Am Soc Nephrol* **12**, 1841–1851 (2017).
3. Mathur, R., Dreyer, G., Yaqoob, M. M. & Hull, S. A. Ethnic differences in the progression of chronic kidney disease and risk of death in a UK diabetic population: an observational cohort study. *BMJ Open* **8**, e020145 (2018).
4. Patzer, R. E. & McClellan, W. M. Influence of race, ethnicity and socioeconomic status on kidney disease. *Nat Rev Nephrol* **8**, 533–541 (2012).
5. Grams, M. E. *et al.* Race, APOL1 Risk, and eGFR Decline in the General Population. *J Am Soc Nephrol* **27**, 2842–2850 (2016).
6. Peralta, C. A. *et al.* APOL1 Genotype and Race Differences in Incident Albuminuria and Renal Function Decline. *J Am Soc Nephrol* **27**, 887–893 (2016).
7. Genovese, G. *et al.* Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans. *Science* **329**, 841–845 (2010).
8. Limou, S., Nelson, G. W., Kopp, J. B. & Winkler, C. A. APOL1 Kidney Risk Alleles: Population Genetics and Disease Associations. *Adv Chronic Kidney Dis* **21**, 426–433 (2014).
9. Nadkarni, G. N. *et al.* Worldwide Frequencies of Apolipoprotein L1 Renal Risk Variants. *N Engl J Med* **379**, 2571–2572 (2018).
10. Friedman, D. J. & Pollak, M. R. APOL1 Nephropathy: From Genetics to Clinical Applications. *CJASN* **16**, 294–303 (2021).
11. Foster, M. C. *et al.* APOL1 Variants Associate with Increased Risk of CKD among African Americans. *JASN* **24**, 1484–1491 (2013).
12. Chen, T. K. *et al.* APOL1 Risk Variants, Incident Proteinuria, and Subsequent eGFR Decline in Blacks with Hypertension-Attributed CKD. *Clin J Am Soc Nephrol* **12**, 1771–1777 (2017).

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

13. Parsa, A. *et al.* APOL1 Risk Variants, Race, and Progression of Chronic Kidney Disease. *N Engl J Med* **369**, 2183–2196 (2013).

14. Zhang, J.-Y. *et al.* UBD modifies APOL1-induced kidney disease risk. *Proc Natl Acad Sci U S A* **115**, 3446–3451 (2018).

15. Dummer, P. D. *et al.* APOL1 kidney disease risk variants – an evolving landscape. *Semin Nephrol* **35**, 222–236 (2015).

16. Ma, L., Divers, J. & Freedman, B. I. Mechanisms of injury in APOL1-associated kidney disease. *Transplantation* **103**, 487–492 (2019).

17. Paranjpe, I. *et al.* Association of APOL1 Risk Genotype and Air Pollution for Kidney Disease. *Clin J Am Soc Nephrol* **15**, 401–403 (2020).

18. Kopp, J. B. *et al.* APOL1 Genetic Variants in Focal Segmental Glomerulosclerosis and HIV-Associated Nephropathy. *J Am Soc Nephrol* **22**, 2129–2137 (2011).

19. Nichols, B. *et al.* Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1. *Kidney Int* **87**, 332–342 (2015).

20. May, R. M. *et al.* A multi-center retrospective cohort study defines the spectrum of kidney pathology in Coronavirus 2019 Disease (COVID-19). *Kidney International* **100**, 1303–1315 (2021).

21. Wu, J. *et al.* The key role of NLRP3 and STING in APOL1-associated podocytopathy. *J Clin Invest* **131**, (2021).

22. Jha, A. *et al.* Alterations in Plasma Membrane Ion Channel Structures Stimulate NLRP3 Inflammasomes Activation in APOL1 Risk Milieu. *FEBS J* **287**, 2000–2022 (2020).

23. Cameron-Christie, S. *et al.* Exome-Based Rare-Variant Analyses in CKD. *Journal of the American Society of Nephrology* **30**, 1109–1122 (2019).

24. Langefeld, C. D. *et al.* Genome-wide association studies suggest that APOL1-environment interactions more likely trigger kidney disease in African Americans with non-diabetic nephropathy than strong APOL1-second gene interactions. *Kidney Int* **94**, 599–607 (2018).

25. Groopman, E. E. *et al.* Diagnostic Utility of Exome Sequencing for Kidney Disease. *New England*

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Journal of Medicine NEJMoa1806891 (2018) doi:10.1056/NEJMoa1806891.

26. Ren, Z. *et al.* ATAV: a comprehensive platform for population-scale genomic analyses. *BMC Bioinformatics* **22**, 149 (2021).
27. Abraham, G. & Inouye, M. Fast Principal Component Analysis of Large-Scale Genome-Wide Data. *PLOS ONE* **9**, e93766 (2014).
28. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
29. R Core Team (2020). R: A language and environment for statistical computing.
30. Elixhauser, A., Steiner, C., Harris, D. R. & Coffey, R. M. Comorbidity Measures for Use with Administrative Data. *Medical Care* **36**, 8–27 (1998).
31. Berkowitz, S. A., Traore, C. Y., Singer, D. E. & Atlas, S. J. Evaluating Area-Based Socioeconomic Status Indicators for Monitoring Disparities within Health Care Systems: Results from a Primary Care Network. *Health Serv Res* **50**, 398–417 (2015).
32. Sharma, N., Schwendimann, R., Endrich, O., Ausserhofer, D. & Simon, M. Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. *BMC Health Services Research* **21**, 13 (2021).
33. White, I. R., Royston, P. & Wood, A. M. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* **30**, 377–399 (2011).
34. Fine, J. P. & Gray, R. J. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association* **94**, 496–509 (1999).
35. Ravani, P. *et al.* Influence of Mortality on Estimating the Risk of Kidney Failure in People with Stage 4 CKD. *JASN* **30**, 2219–2227 (2019).
36. Levey, A. S. *et al.* A New Equation to Estimate Glomerular Filtration Rate. *Ann Intern Med* **150**, 604–612 (2009).
37. Inker, L. A. *et al.* New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race. *New England Journal of Medicine* **385**, 1737–1749 (2021).

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

38. Azzi, A. *et al.* Is there an age cutoff to apply adult formulas for GFR estimation in children? *J Nephrol* **28**, 59–66 (2015).

39. Shou, H. *et al.* Analytic Considerations for Repeated Measures of eGFR in Cohort Studies of CKD. *CJASN* **12**, 1357–1365 (2017).

40. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48 (2015).

41. Rasouly, H. M. *et al.* The burden of Candidate pathogenic variants for kidney and genitourinary disorders emerging from exome sequencing. *Annals of Internal Medicine* **170**, 11–21 (2019).

42. Richards, S. *et al.* Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine* **17**, 405–424 (2015).

43. Povysil, G. *et al.* Assessing the Role of Rare Genetic Variation in Patients With Heart Failure. *JAMA Cardiol* **6**, 379–386 (2021).

44. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet* **53**, 1097–1103 (2021).

45. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am J Hum Genet* **88**, 76–82 (2011).

46. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369-S3 (2012).

47. Povysil, G. *et al.* Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nat Rev Genet* **20**, 747–759 (2019).

48. Auwera, G. A. V. der & O'Connor, B. D. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. (O'Reilly Media, 2020).

49. Storey, J. D. False Discovery Rate. in *International Encyclopedia of Statistical Science* (ed. Lovric, M.) 504–508 (Springer, 2011). doi:10.1007/978-3-642-04898-2_248.

50. Palmer, N. D. *et al.* Lack of Association of the APOL1 G3 Haplotype in African Americans with ESRD. *J*

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Am Soc Nephrol **26**, 1021–1025 (2015).

51. Jagannathan, R. *et al.* Association Between APOL1 Genotype and Kidney Diseases and Annual Kidney Function Change: A Systematic Review and Meta-Analysis of the Prospective Studies. *IJNRD* **14**, 97–104 (2021).
52. Tzur, S., Rosset, S., Skorecki, K. & Wasser, W. G. APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease. *Nephrol Dial Transplant* **27**, 1498–1505 (2012).
53. Lee, B. T. *et al.* The APOL1 Genotype of African American Kidney Transplant Recipients Does Not Impact 5-Year Allograft Survival. *American Journal of Transplantation* **12**, 1924–1928 (2012).
54. Khan, A. *et al.* Genome-wide polygenic score to predict chronic kidney disease across ancestries. *Nat Med* **28**, 1412–1420 (2022).
55. Cooper, A. *et al.* APOL1 renal risk variants have contrasting resistance and susceptibility associations with African trypanosomiasis. *eLife* **6**, e25461 (2017).
56. Hayek, S. S. *et al.* A tripartite complex of suPAR, APOL1 risk variants and αvβ3 integrin on podocytes mediates chronic kidney disease. *Nat Med* **23**, 945–953 (2017).
57. Thomson, R. *et al.* Evolution of the primate trypanolytic factor APOL1. *Proc Natl Acad Sci U SA* **111**, E2130–E2139 (2014).
58. Hughson, M. D. *et al.* APOL1 Risk Variants Independently Associated With Early Cardiovascular Disease Death. *Kidney International Reports* **3**, 89–98 (2018).
59. Bajaj, A. *et al.* Phenome-wide association analysis suggests the APOL1 linked disease spectrum primarily drives kidney-specific pathways. *Kidney Int* **97**, 1032–1041 (2020).
60. Cocchi, E., Nestor, J. G. & Gharavi, A. G. Clinical Genetic Screening in Adult Patients with Kidney Disease. *Clinical Journal of the American Society of Nephrology* CJN.15141219 (2020) doi:10.2215/cjn.15141219.
61. Masters, S. L. *et al.* NLRP1 inflammasome activation induces pyroptosis of hematopoietic progenitor cells. *Immunity* **37**, 1009–1023 (2012).

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Figures and Tables:

Table 1. Demographics and baseline data of included CKD patients, characterized by *APOL1* risk genotype.

	Low-Risk Genotype	High-Risk Genotype
Number	1187	239
Biologic Sex, Female (%)	564 (48%)	105 (44%)
Age at recruitment, years median (IQR)	43.7 (27.9 – 57.8)	43.4 (30.2 – 54.6)
Age at last followup, years median (IQR)	48.4 (32.1 – 62.2)	48.8 (35.9 – 59.3)
Follow-up time, years median (IQR)	3.7 (0 – 6.1)	3.7 (0 – 7.1)
Pediatric at recruitment (%)	149 (13%)	4 (2%)
Developed Kidney Failure (%)	630 (53%)	168 (70%)
Age at Kidney Failure, years median (IQR)	43.3 (30.5 – 55.4)	40.6 (28.0 – 50.2)
Received Kidney Transplant (%)	538 (45%)	150 (63%)
Family History of Kidney Disease (%)	399 (34%)	83 (35%)
Initial eGFR, mL/min/1.73m^2 mean (IQR)	57.9 (33.2 – 97.6)	49.7 (34.5 – 81.7)
Self-Described Race and ethnicity		
Asian, Hispanic or Latinx	3 (0.2%)	0 (0%)
Asian, Not Hispanic or Latinx	5 (0.4%)	1 (0.4%)
Black/African American, Hispanic or Latinx	70 (6%)	14 (6%)
Black/African American, Not Hispanic or Latinx	294 (25%)	149 (61%)
Black/African American, Unspecified Ethnicity	5 (0.4%)	2 (1%)
Native American, Hispanic or Latinx	3 (0.2%)	0 (0%)
Unknown or preferred not to specify, Hispanic or Latinx	265 (22%)	16 (7%)
Unknown or preferred not to specify, Not Hispanic or Latinx	41 (3%)	1 (0.4%)
Unknown or preferred not to specify, Unspecified Ethnicity	4 (0.3%)	0 (0%)
White, Hispanic or Latinx	428 (36%)	41 (17%)
White, Not Hispanic or Latinx	62 (5%)	15 (6%)
White, Unspecified Ethnicity	7 (0.6%)	0 (0%)
Genetic Ancestry Cluster		
Majority Black/African American	364 (31%)	170 (71%)
Majority Hispanic or Latinx	195 (16%)	5 (2%)
Admixed connecting cluster	628 (53%)	64 (27%)
ZIP Code Median Income (by quintile)		
1 (Highest)	33 (3%)	3 (1%)
2	355 (30%)	71 (30%)
3	515 (43%)	102 (43%)
4	218 (18%)	52 (22%)
5 (lowest)	18 (2%)	3 (1%)
Missing	48 (4%)	8 (3%)
Hypertension	452 (38%)	108 (45%)
Diabetes Mellitus	183 (12%)	37 (15%)
Elixhauser Comorbidity Score, median (IQR)	5 (0 – 9)	7 (4 – 8)
Missing	450 (38%)	86 (36%)

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Table 2. Most common primary causes of kidney disease by *APOL1* risk genotype.

Clinical Diagnosis	Low-Risk *APOL1* Genotype	Single Risk Allele G1 or G2	High-Risk *APOL1* Genotype
FSGS	108 (9%)	35 (10%)	91 (38%)
FSGS developing kidney failure	42 (39%)	14 (40%)	51 (56%)
C1q Nephropathy	2 (2%)	0 (0%)	5 (5%)
H-CKD	103 (9%)	34 (10%)	33 (14%)
H-CKD developing kidney failure	79 (77%)	25 (74%)	33 (100%)
LN	103 (9%)	37 (11%)	10 (4%)
LN developing Kidney Failure	48 (47%)	23 (62%)	8 (80%)
CAKUT	148 (12%)	28 (8%)	11 (5%)
IgA Nephropathy	112 (9%)	23 (7%)	7 (3%)
DKD	159 (13%)	50 (15%)	17 (7%)
DKD developing Kidney Failure	136 (86%)	45 (90%)	15 (88%)
Membranous Nephropathy	62 (5%)	20 (6%)	8 (3%)
ADPKD	48 (4%)	17 (5%)	1 (0.4%)
Other	334 (28%)	96 (29%)	61 (26%)

Other includes: Membranoproliferative glomerulonephritis, ANCA vasculitis, anti-GBM disease, minimal change disease, chronic tubulointerstitial nephritis, Alport syndrome, thin basement membrane disease, nephrolitiasis, Gitelman syndrome, thrombotic microangiopathy, amyloidosis, nephronophthisis, medullary cystic kidney disease, sarcoidosis, Bartter syndrome, preeclampsia, scleroderma, and tuberous sclerosis.

Abbreviations: autosomal dominant polycystic kidney disease (ADPKD), congenital anomalies of kidney and urinary tract (CAKUT), diabetic kidney disease (DKD), focal segmental glomerulosclerosis (FSGS), Hypertension associated chronic kidney disease (H-CKD), lupus nephritis (LN)

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Table 3. Analyses of kidney failure outcomes by *APOL1* genotype. Kaplan-Meier and Cox-Proportional hazard modelling applied to development of kidney failure, competing risk analysis includes competing risk of death, eGFR decline modelled using linear mixed effects modelling. Time to event analyses adjusted for sex, ZIP code based median income, genetic ancestry cluster, family history of kidney disease, presence of monogenic kidney disease, Elixhauser comorbidity score, diabetes mellitus and hypertension. eGFR decline adjusted for initial eGFR measurement, genetic ancestry cluster, diabetes mellitus, hypertension, family history of kidney disease, and primary cause of kidney disease. Statistical testing was not corrected for multiple testing

	Number of Subjects	Age at Kidney Failure (Median, 95% CI, Years)	Adjusted HR of Kidney Failure (95% CI), P	Competing Risk HR of Kidney Failure (95% CI), P	eGFR Decline Rate (Mean, 95% CI, mL/min/1.73m^2/year), P
Low-risk	1187	53.6 (51.8 – 55.6)	1 (reference)	1 (reference)	3.63 (5.36 – 1.88), reference
High-risk	239	45.1 (42.0 – 48.7)	**1.58 (1.31 – 1.91), P=1.5 x10^{-6}**	**1.59 (1.31 – 1.92), P=2.7 x10^{-6}**	**6.55 (8.24 – 4.83), P=6.9 x10^{-4}**
G1/G1	98	42.1 (40.5 – 48.9)	**1.88 (1.45 – 2.44), P=1.9 x10^{-6}**	**1.86 (1.43 – 2.41), P=3.7 x10^{-6}**	**8.16 (10.83 – 5.56), P=8.9 x10^{-4}**
G1/G2	109	46.2 (42.0 – 50.0)	**1.53 (1.19 – 1.97), P=1.0 x10^{-3}**	**1.54 (1.17 – 2.02), P=1.8 x10^{-3}**	**6.05 (8.27 – 3.93), P=0.03**
G2/G2	32	53.4 (42.4 – 66.5)	1.15 (0.74 – 1.79), P=0.54	1.15 (0.74 – 1.80), P=0.53	4.82 (8.90 – 0.71), P=0.56

ACCEPTED

Copyright © 2023 by the American Society of Nephrology

Figure 1. Patient flow within the study.

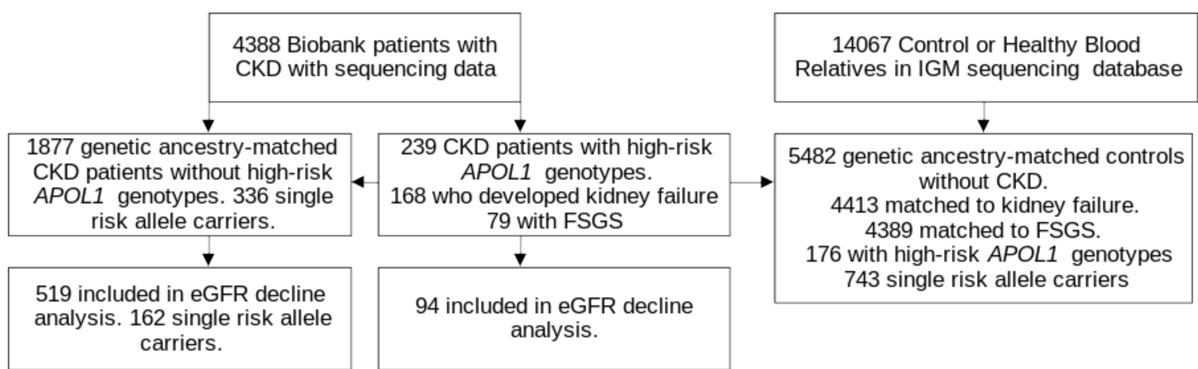


Figure 2. Risk of developing specific kidney diseases based on the number of *APOL1* risk alleles. The comparison group is composed of individuals with congenital anomalies of the kidney and urinary tract and autosomal dominant polycystic kidney disease. Adjusted for family history of kidney disease, sex, ZIP code based median income, genetic ancestry cluster, Elixhauser comorbidity score, diabetes mellitus and hypertension.

Abbreviations: FSGS = focal segmental glomerulosclerosis, HTN = hypertension-associated nephropathy, C3GN = C3 glomerulopathy, MN = membranous nephropathy, DKD = diabetic kidney disease, IgAN = IgA nephropathy, SLE = systemic lupus erythematosus.

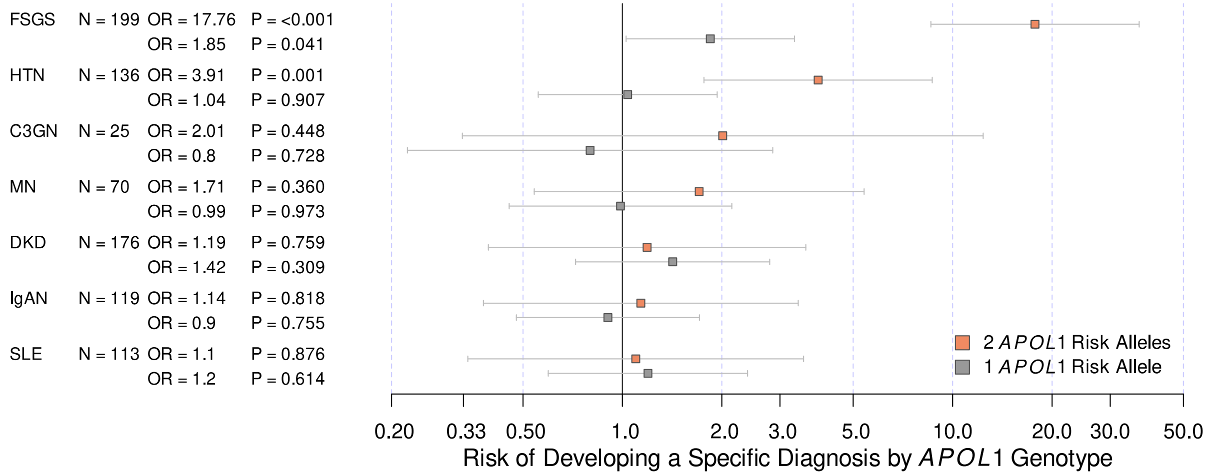


Figure 3. Lifetime risk of developing kidney failure comparing subjects with high-risk *APOL1* genotypes to low-risk *APOL1* genotypes. Median time to kidney failure denoted and Logrank P-value shown.

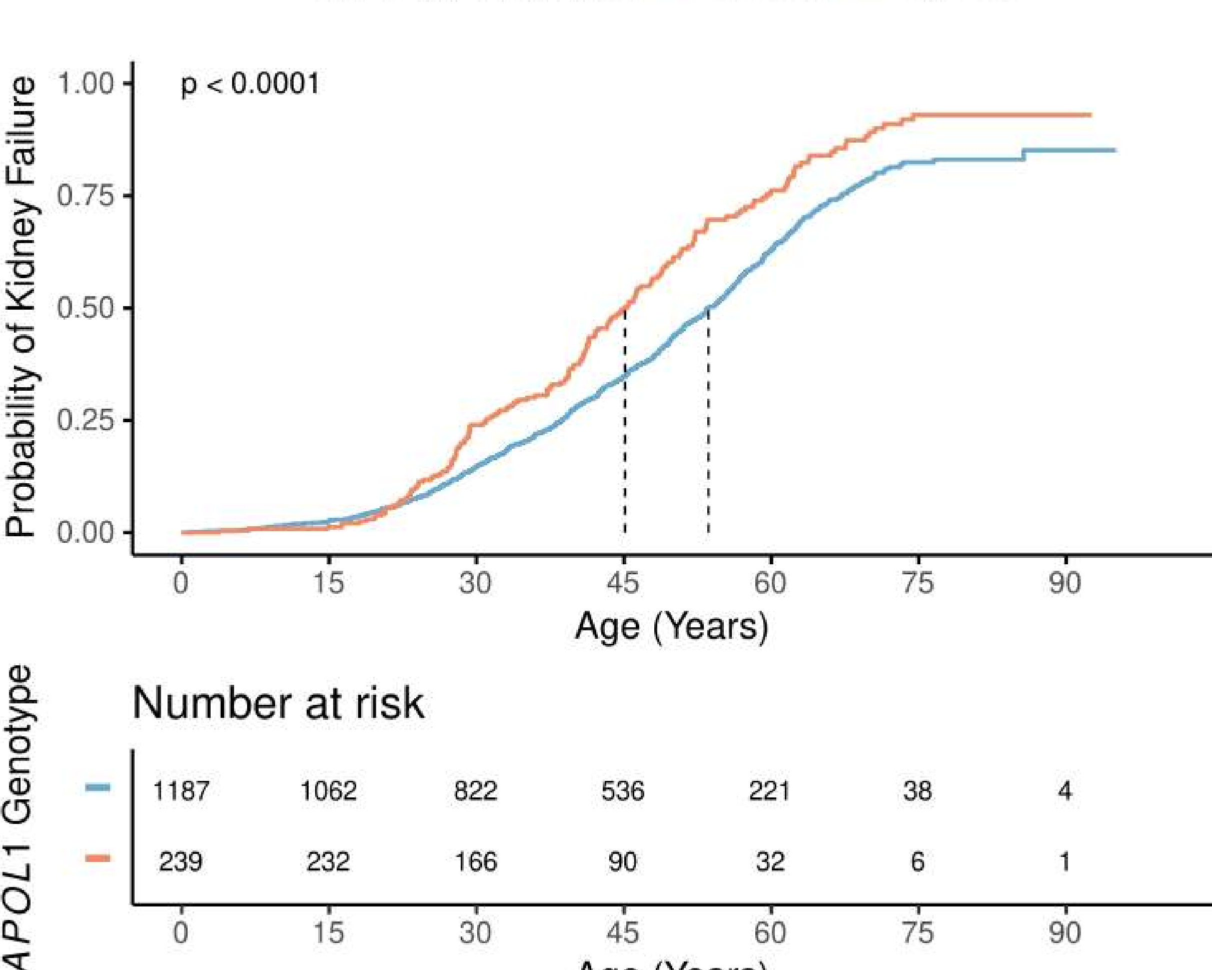
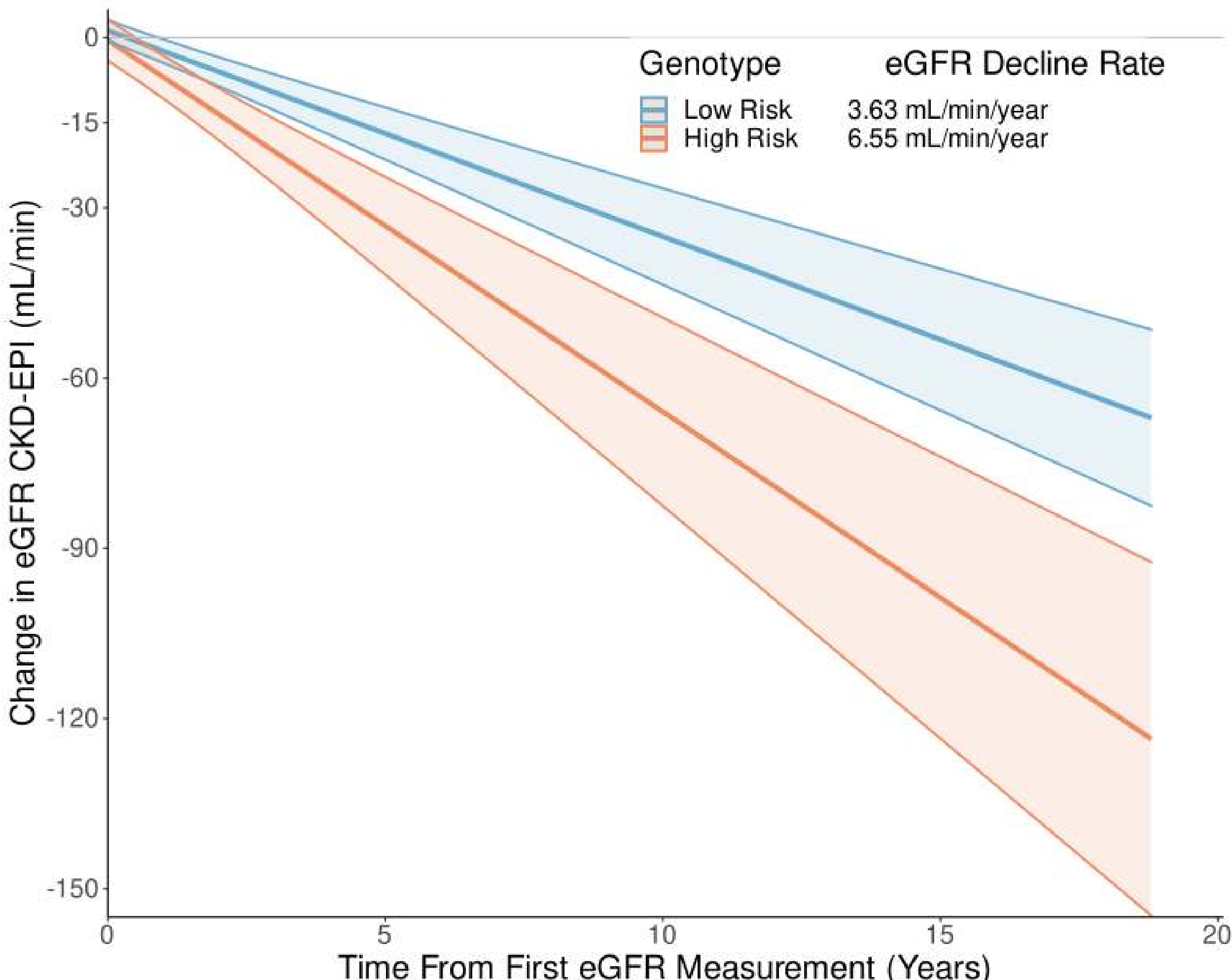


Figure 4. eGFR decline from baseline in subjects with high and low-risk *APOL1* genotypes. eGFR decline rate calculated with covariates held at mean values with 95% confidence intervals shown. Adjusted for primary cause of kidney disease, family history of kidney disease, diabetes mellitus, hypertension, genetic ancestry cluster, and initial measured eGFR.



Journal of Nephrology (2020) 33:849–857
https://doi.org/10.1007/s40620-019-00660-9

ORIGINAL ARTICLE

Post-transplant recurrence of steroid resistant nephrotic syndrome in children: the Italian experience

William Morello[1] · Sairaj Puvinathan[1] · Giuseppe Puccio[1] · Gian Marco Ghiggeri[2] · Luca Dello Strologo[3] · Licia Peruzzi[4] · Luisa Murer[5] · Michela Cioni[2] · Isabella Guzzo[3] · Enrico Cocchi[4] · Elisa Benetti[5] · Sara Testa[1] · Luciana Ghio[1] · Gianluca Caridi[2] · Massimo Cardillo[6] · Rosanna Torelli[6] · Giovanni Montini[1]

Received: 29 July 2019 / Accepted: 9 October 2019 / Published online: 15 October 2019
© The Author(s) 2019

Abstract

Background Steroid resistant nephrotic syndrome (SRNS) is a frequent cause of end stage renal disease in children and post-transplant disease recurrence is a major cause of graft loss.

Methods We identified all children with SRNS who underwent renal transplantation in Italy, between 2005 and 2017. Data were retrospectively collected for the presence of a causative gene mutation, sex, histology, duration of pre-transplant dialysis, age at onset and transplant, HLA matching, recurrence, therapy for recurrence, and graft survival.

Results 101 patients underwent a first and 22 a second renal transplant. After a median follow-up of 58.5 months, the disease recurred on the first renal transplant in 53.3% of patients with a non-genetic and none with a genetic SRNS. Age at transplant > 9 years and the presence of at least one HLA-AB match were independent risk factors for recurrence. Duration of dialysis was longer in children with relapse, but did not reach statistical significance. Overall, 24% of patients lost the first graft, with recurrence representing the commonest cause. Among 22 patients who underwent a second transplant, 5 suffered of SRNS recurrence. SRNS relapsed in 5/9 (55%) patients with disease recurrence in their first transplant and 2 of them lost the second graft.

Conclusions Absence of a causative mutation represents the major risk factor for post-transplant recurrence in children with SRNS, while transplant can be curative in genetic SRNS. A prolonged time spent on dialysis before transplantation has no protective effect on the risk of relapse and should not be encouraged. Retransplantation represents a second chance after graft loss for recurrence.

Keywords Steroid-resistant nephrotic syndrome · Kidney transplant · Post-transplant recurrence

✉ Giovanni Montini
giovanni.montini@unimi.it

[1] Pediatric Nephrology, Dialysis and Transplant Unit, Department of Clinical Sciences and Community Health, Fondazione IRCCS Ca' Granda-Ospedale Maggiore Policlinico, University of Milan, via della Commenda, 9, 20122 Milan, Italy

[2] Division of Nephrology, Dialysis, and Transplantation, Scientific Institute for Research and Health Care, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[3] Renal Transplant Unit, Bambino Gesù Children's Research Hospital IRCCS, Rome, Italy

[4] Pediatric Nephrology Unit, Regina Margherita Children's Hospital, Città della Salute e della Scienza di Torino, Turin, Italy

[5] Pediatric Nephrology, Dialysis and Transplant Unit, Department of Women's and Children's Health, Hospital-University of Padova, Padua, Italy

[6] North Italy Transplant program (NITp), UOC Coordinamento Trapianti, Fondazione IRCCS Ca' Granda-Ospedale Maggiore Policlinico, Milan, Italy

Introduction

Steroid resistant nephrotic syndrome (SRNS) is the most common acquired cause of end stage renal failure (ESRD) requiring transplantation in children. Advances in genetic screening have allowed the identification of a monogenic cause of SRNS in one-third of cases [1]. Genetic SRNS are associated with an underlying mutations in genes encoding podocyte associated proteins, resulting in structural or functional disruption of the glomerular filtration barrier [2]. The pathophysiology of SRNS without underlying mutations remains poorly explained and is thought to involve an unknown circulating permeability factor [3] which may also be implicated in the recurrence soon after transplantation [4].

Unfortunately, in up to 50% of patients, SRNS relapses after transplantation and disease recurrence is a major cause of graft loss [4–6]. Genetic SRNS have been reported to have a low rate of recurrences [7–9]. On the other hand, previous studies have suggested non-African race, rapid progression to ESRD (< 3 years) and previous recurrence after transplantation to be associated with SRNS relapse [8, 10, 11], but no established risk factors can actually predict the outcome. A longer time on dialysis before transplantation was believed to decrease the risk of relapse, but not confirmed by large reports [12, 13].

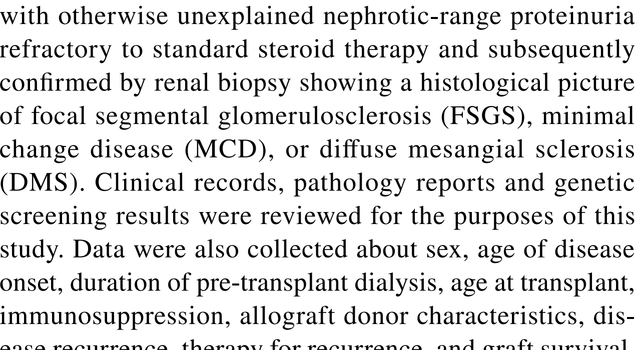
There is currently little consensus regarding the best management of post-transplant SRNS recurrence, which represents a devastating complication for families and physicians, and poses a significant threat to allograft survival. Plasma exchange (PE), steroids and rituximab are the most common strategies to treat the recurrence [14].

The objective of this study was to identify factors affecting the risk of recurrence and graft loss in children with SRNS, by stratifying the population according to their genetic status.

Methods

We performed a retrospective, multicentre, observational cohort study to address the long-term prognosis of renal graft, the risk factors for recurrence and the predictors for response to therapy following recurrence in paediatric patients undergoing renal transplantation because of a SRNS.

We identified patients who underwent renal transplantation at all five Italian paediatric transplant centres, between 2005 and 2017, with a primary diagnosis of SRNS and onset before 18 years. Patients were included if a clinical diagnosis of SRNS was made in an individual



with otherwise unexplained nephrotic-range proteinuria refractory to standard steroid therapy and subsequently confirmed by renal biopsy showing a histological picture of focal segmental glomerulosclerosis (FSGS), minimal change disease (MCD), or diffuse mesangial sclerosis (DMS). Clinical records, pathology reports and genetic screening results were reviewed for the purposes of this study. Data were also collected about sex, age of disease onset, duration of pre-transplant dialysis, age at transplant, immunosuppression, allograft donor characteristics, disease recurrence, therapy for recurrence, and graft survival.

Patients were divided in three groups: Group A (Genetic SRNS): patients with an identified causative genetic mutation and/or a first degree relative with SRNS and/or extra-renal disease manifestations pathognomonic of SRNS, Group B (Idiopathic SRNS): patients with a negative or heterozygous recessive genetic test result and without a first degree family history or associated extra-renal manifestations pathognomonic of SRNS, Group C (unknown genetic status): patients with no genetic analysis performed and no family history or extra-renal manifestations typical of SRNS.

Definitions

Nephrotic-range proteinuria, urine protein:creatinine ratio (uPr/uCr) ≥ 2 mg/mg. Age at disease onset, age at first clinical presentation of nephrotic syndrome. Steroid resistance, persistence of nephrotic range proteinuria following 4 weeks of daily 60 mg/sqm prednisone therapy. Post-transplant disease recurrence, an otherwise unexplained persistent nephrotic range proteinuria after renal transplantation, when rejection was excluded. Graft loss, functional failure of the renal allograft, necessitating renal replacement therapy. Remission after recurrence, complete resolution of proteinuria (uPr/uCr < 0.2 mg/mg). Partial remission after recurrence, persistent reduction of proteinuria (uPr/uCr < 2 mg/mg) with preserved renal function.

Statistical analysis

Categorical variables were compared using the Chi squared test for independence. The distribution of continuous variables in groups was compared using the Wilcoxon signed-rank test and the Kruskal–Wallis test. Linear regression models were used to compare continuous variables. For multivariate analysis, multiple logistic regression models were used. A p value < 0.05 was considered statistically significant. All statistical analyses were performed using the open source software R. (R Core Team, 2014. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria).

Results

Study cohort

During the study period, a total of 728 (618 deceased and 110 living donors) renal grafts were performed at the 5 Italians pediatric transplant centres, of whom 123 in patients with ESRD secondary to SRNS. 101 patients received a first renal allograft and 22 a second renal transplant (12 failures of the original cohort and 10 failures of a first transplant that occurred before the study period). The number of patients who received a first transplant at each center is as follows: Istituto G. Gaslini, Genova—31, Bambino Gesù Children's Hospital, IRCCS, Rome—23, Fondazione IRCCS Ca' Granda, Ospedale Maggiore, Milan—20, Regina Margherita Children's Hospital, Turin—15, University Hospital of Padua—12. The study cohort is summarized in Fig. 1.

First renal graft

101 patients (52.5% males) underwent a first renal transplant. The median age at onset was 2.8 years of age (range 0–17.2); 24 individuals (25.2%) presented with congenital SRNS, defined as onset of disease within the first 3 months of life. Renal histology was consistent with FSGS in 85 cases, MCD in 14 and DMS in 2. Main demographic and clinical characteristics are summarized in Table 1. At transplant all patients received an induction therapy with basiliximab and immunosuppression with steroids, calcineurin inhibitors and mofetil mycophenolate. Two patients were treated with plasmapheresis pre-transplantation.

Genetic testing results were available for 76 individuals (75.2%) (Table 2): 39 had an autosomal dominant mutation or were homozygous for a recessive mutation, 8 were heterozygous carriers and 29 had a negative genetic test. Genetic results were unavailable for 25 patients, among whom we were able to identify 2 additional patients with genetic SRNS: one had a sibling with established genetic SRNS and another showed extra-renal disease manifestations

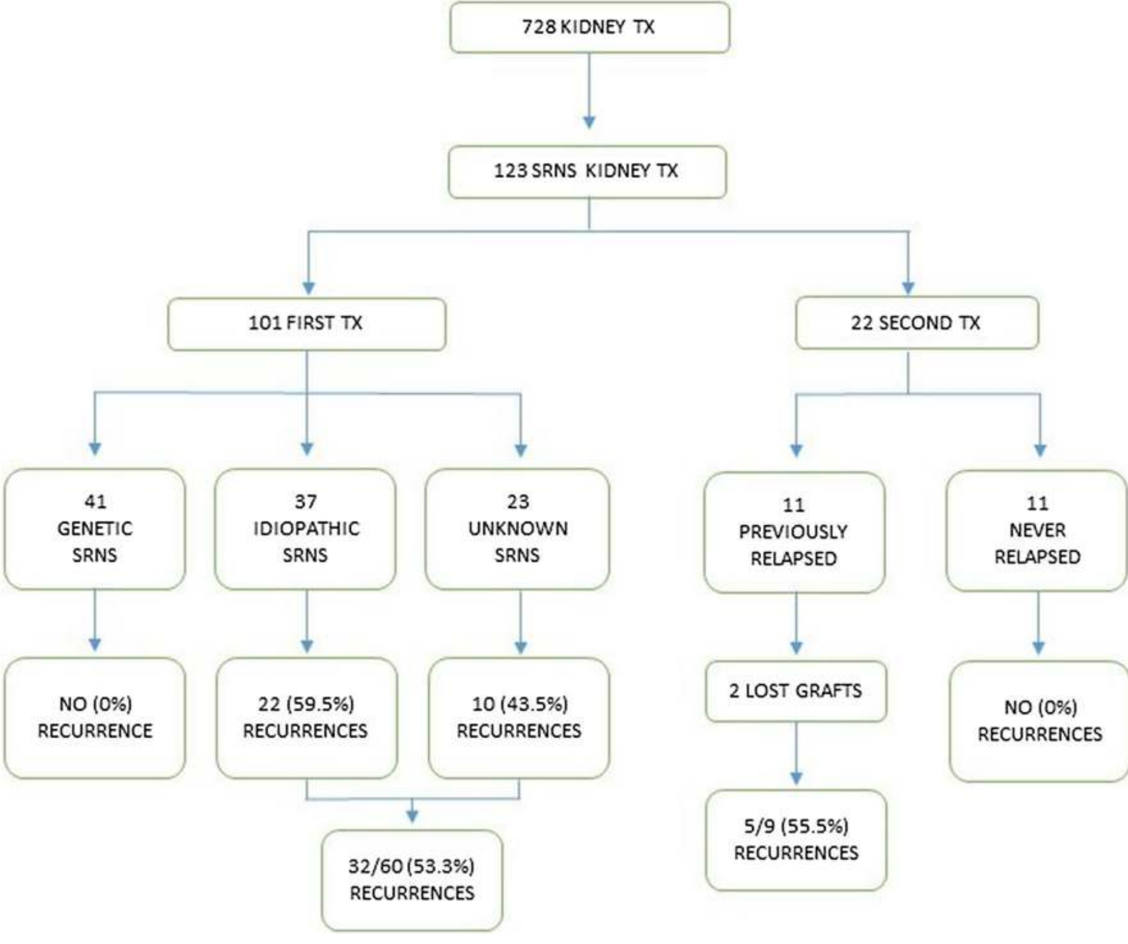


Fig. 1 Study cohort

Table 1 Main demographic and clinical characteristics of SRNS children transplanted between 2005 and 2017

Characteristics	First renal Tx	Second renal Tx
Total	101	22
Gender		
Male	53 (52.5%)	13 (59.1%)
Female	48 (47.5%)	9 (40.9%)
Genetic disease		
Yes	41 (40.6%)	5 (22.7%)
No	37 (36.6)	13 (56.6%)
Unknown	23 (22.8%)	4 (18.2%)
Age at onset (years) median (range)	2.8 (0–17.2)	4.45 (0–14.29)
Age at transplant (years) median (range)	11.8 (2.6–20.8)	16.71 (4.56–31.1)
Time to ESRD (years) median (range)	3.3 (1.7–14.3)	2.5 (0–7.5)
Time on dialysis before transplantation (years) median (range)	2 (0–9)	Not available
Donor type		
Living	6 (5.9%)	2 (9.1%)
Deceased	95 (94.1%)	20 (90.9%)
Follow-up (months) median (range)	58.5 (0.7–157.8)	40 (0–148)

Table 2 Prevalence of mutations among patients with available genetic results

Gene	Encoded protein	Mode of inheritance	Genetic tests n = 76
NPHS1	Nephrin	Recessive	13
WT1	Wilms tumour protein	Dominant	11
NPHS2	Podocin	Recessive	8
ACTN4	α-Actinin	Dominant	2
PLCE1	Phospholipase C	Recessive	1
COL4A5	Type IV collagen α5 chain	X-linked recessive	1
SMARCAL1	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily A-like protein 1	Recessive	1
LMX1B	LIM homeobox transcription factor 1 β	Dominant	1
COQ2	Coenzyme Q2	Recessive	1
Heterozygous carriers		Recessive	8
No mutations			29

suggesting a genetic disease. The patient cohort was therefore comprised of: Group A (genetic SRNS): 41 individuals (40.6%), Group B (idiopathic SRNS): 37 individuals (36.6%), and Group C (unknown genetic status): 23 individuals (22.8%). Age at the onset was similar between Group B (idiopathic) and Group C (unknown), while it was younger for patients with a genetic disease (Group A) ($p < 0.0001$).

NPHS1, encoding nephrin, was the single most commonly mutated gene and accounted for one-third (33.3%) of positive genetic results, followed by WT1, encoding Wilms tumour protein and NPHS2, encoding podocin. Mutations in these genes were responsible of 28.2% and 20.5% of genetic SRNS, respectively. Taken together, mutations in NPHS1, NPHS2, and WT1 accounted for 82% of identified genetic cases. Pathogenetic mutations were also identified in the following genes: PLCE1, ACTN4, COL4A5, SMARCAL1, LMX1B, COQ2 (Table 2). In eight cases renal disease was associated with a syndromic presentation, as follows: Denys Drash syndrome in four cases and one case each of Frasier syndrome, WAGR syndrome, Leopard syndrome and Schimke immuno-osseous dysplasia.

Post-transplant disease recurrence

Median follow-up is 58.5 months (range 0.7–157.8). SRNS recurred in 32 individuals (31.7%) after the first renal transplant, at a median time of 2 days post-transplantation. When stratified by genetic status, the incidence of post-transplant

disease recurrence was 59.5% in Group B (idiopathic SRNS) and 43.5% in Group C (unknown genetic status). No Group A (genetic SRNS) child experienced disease recurrence and this group was therefore excluded from further analysis (Table 3).

The difference in post-transplant disease recurrence between Group B and C, however, was not significant (p = 0.23). Risk factors for recurrence were evaluated in the remaining 60 patients (Group B and C). Overall, SRNS recurred in 32/60 (53.3%) non-genetic patients.

As all relapses except one (identified 10 years after transplantation) occurred within 8 months from transplant, the analysis was made at 8 months of follow-up and included all evaluable patients (54 patients). Age at transplant was categorized as ≥ 9 years, following a ROC analysis identifying it as the best cut-off for relapse prediction in our dataset (Fig. 2). Bivariate analysis was performed by Wilcoxon test for independent samples. Multivariate analysis was performed by a logistic regression model.

At bivariate analysis, the following variable were significantly associated to relapse:

- Age ≥ 9 years (p = 0.01823)
- At least one HLA AB match (p = 0.01752)
- At least one HLA DR match (p = 0.01763)

Gender, donor age and donor type (living or deceased) did not affect the risk of recurrence. Time to ESRD and duration of dialysis before transplant were not significantly associated with relapse; anyway, they were both longer in children with relapse (median = 4.6 vs 2.7 years, p = 0.2673 and 2.4 vs 1.8 years, p = 0.06582, respectively).

We were not able to assess the role of different induction schedules, since all patients were homogeneously treated with basiliximab and immunosuppressive therapy. However, among the two patients treated with pre-transplantation PE, one experienced relapse the day after transplant.

Multivariate analysis included all the variables associated with recurrence at the bivariate analysis with a p value < 0.1 (Table 4). Age at transplant > 9 years and HLA-AB match were the only independent risk factors for recurrence after transplant (p = 0.01017 and p = 0.02465, respectively). However, the best prediction model for relapse, characterized by the lowest residual deviance and lowest AIC, included also a longer duration of dialysis before transplant (Null deviance: 68.029 on 49 degrees of freedom; Residual deviance: 49.584 on 46 degrees of freedom; AIC: 57.584; p = 0.06994). The model including the aforementioned 3 variables has a likelihood ratio test p value of 0.000356, and a pseudo R squared value of 0.271136 (McFadden method).

The risk factors for recurrence in Group B and C are summarized in Table 5.

Disease recurrence was treated in all patients with PE with a median of 20 sessions (range 4–79). 22 were treated with rituximab and 9 with high dose steroids. The use of other therapeutic agents was as follows: ofatumumab (3), mesenchymal stromal cells (2), intravenous immunoglobulins (2), abatacept (2), cyclophosphamide (2), cyclosporin (1) and thymoglobulin (1). Overall, a complete or partial remission was achieved in 15 and 4 patients, respectively, 13 patients (40.6%) failed to achieve sustained disease remission, despite treatment and 11/13 subsequently lost the graft. Use of rituximab or high dose steroids did not influence the response rate (p = 0.3574). The remission was persistent with preserved renal function in 13/15 patients. One patient

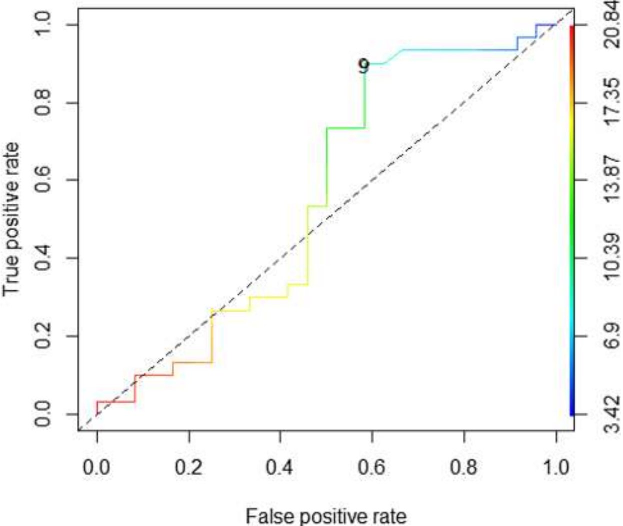


Fig. 2 ROC curve identifying the best cut-off for age at transplant with FPR = 0.58333333, TPR (sensitivity) = 0.86666667, Specificity = 0.41666667, p value = 0.01823

Table 3 Incidence of recurrence, stratified by genetic testing results

Characteristics	Total, n = 101	Recurrence, n = 32	No recurrence, n = 69
Genetic results *n*, %			
Negative	37	22 (59.5)	15 (40.5)
Unknown	23	10 (43.5)	13 (57.5)
Positive	41	0 (0.0)	41 (100.0)

Table 4 Variables included in the multivariate analysis

Variables
Age ≥ 9 years
At least one HLA AB match
At least one HLA DR match
Duration of dialysis before transplant

Table 5 Risk factors for post-transplant disease recurrence in Group B (idiopathic SRNS) and Group C (unknown genetic status) individuals, at 8 months of follow-up

Variables	Total	Recurrence, n (%)	No recurrence, n (%)	Univariate analysis *p* value	Multivariate analysis *p* value
Gender					
Male	30 (55.6%)	18 (60%)	12 (40%)	0.4624	0.26990
Female	24 (44.4%)	12 (50%)	12 (50%)		
Age at transplant, *yr*					
≥ 9	40 (74.0%)	26 (65%)	14 (35%)	**0.01823**	**0.01017**
< 9	14 (26.0%)	4 (28.6%)	10 (71.4%)		
HLA-AB matching				0 vs > 0	0 vs > 0
0	7 (13.5%)	1 (14.3%)	6 (65.7%)	**0.01752**	**0.02465**
1	18 (34.6%)	13 (72.2%)	5 (27.8%)		
2	17 (32.7%)	9 (52.9%)	8 (47.1%)		
3	10 (19.2%)	6 (60%)	4 (40%)		
HLA-DR matching					
0	16 (30.8%)	5 (31.2%)	11 (68.8%)	**0.01763**	0.46309
1	34 (65.4%)	24 (70.6%)	10 (29.4%)		
2	2 (3.8%)	0 (0%)	2 (100%)		
		Recurrence	No recurrence		
Duration of dialysis median, range		2.4, 0.6–9	1.8, 0.1–5	0.06582	0.06994
Time to ESRD median, range		4.6, 0–12	2.7, 0–13	0.2673	0.72323
Donor age median, range		14, 1–63	11, 1–56	0.1609	0.84874

Statistically significant values in bold

had second untreatable relapse 10 years after transplant and lost the kidney. One additional patient with partial remission, following experienced rejection and lost his graft.

Graft loss

24 patients (23.8%) experienced loss of a first renal graft. The causes of graft loss were as follows: disease recurrence in 12 (50%), rejection in 6 (25%), primary non-functioning graft in 3 (12.5%), thrombosis in 2 (8.3%) and chemotherapy toxicity for post-transplant thrombo-proliferative disease in one case (4.1%). In addition, death with a functional graft due to sepsis occurred in two patients.

Second renal transplant

During the study period, 22 SRNS patients received a second renal graft; 11 of them had had a recurrence in the first graft, while 11 lost their transplant for different reasons. Among patients with a previous recurrence, 2/11 patients lost their graft immediately after the transplantation for reasons different from relapse (death and surgery complications) and were not included in the following analysis. After a median follow-up of 40 months, five patients relapsed on the second transplant. All of them have had a recurrence in the first graft. Therefore, in our population, only 5 out 9 (55.5%) evaluable patients with a previous relapse experienced recurrence of proteinuria after the second kidney transplantation. Among five relapsed patients, only two subsequently lost the second graft. None of the 11 patients who lost the first transplant for different reasons suffered of relapse, with 4/11 having a genetic disease. Outside SRNS recurrence, two patients experienced a graft rejection and lost the second kidney graft and one patient died after a post-transplant lymphoproliferative disease.

Discussion

SRNS is a leading cause of ESRD in children. Post-transplant recurrence is a common complication, associated with an increased risk of graft loss. Many efforts have been made to identify the risk factors for recurrence in order to improve prevention and treatment strategies [5, 6, 15].

Our study gives a clear picture of the Italian experience with kidney transplantation in children with SRNS, during a period of over 10 years, encompassing all the recent acquisition regarding the etiopathology and therapeutic options for SRNS.

The overall incidence of post-transplant disease recurrence (53% of non-genetic patients) is consistent with the available scientific literature, stretching back almost three decades [16, 17]. Recurrence is confirmed to be a very early event, with a median time from transplant of 2 days and 30/32 events occurring within the first 2 months after transplantation. By stratifying the cohort according to genetic status, we have been able to confirm that genetic SRNS does not recur after transplant. While previous reports have identified the genetic status as an important risk factor for disease recurrence, most available retrospective studies are unable to account for the genetic status of the majority of their cohorts [17, 18]. To the best of our knowledge, no previous studies were able to assess the risk of recurrence in an equally characterized population, as regards genetic disease. In our study, genetic results were indeed available for 76/101 patients (75.2%), furthermore since the remaining individuals (Group C, unknown genetic status) closely resemble Group B (Idiopathic SRNS) in key clinical features, including similar age of onset and rate of recurrence (43.5%), it is likely that most of them also represent cases of idiopathic SRNS. Indeed, we believe that patients with early onset or congenital SRNS were more likely to be tested for a genetic disease, while genetic testing was less performed in older children and adolescents with a clinical picture of idiopathic SRNS.

Our observation is in line with previous studies that report no or very low relapse rate after transplantation in children with genetic SRNS [9, 17, 19]. Few old reports have suggested a risk of relapse for genetic SRNS, but they are almost all related to NPHS2 mutation, including heterozygous individuals [20–23]. The causative role of the variants included in these reports should be reconsidered, as exquisitely suggested in a recent review by Bierzynska [15].

The rate of recurrence in idiopathic SRNS (Group B) was 59.5%. The result is slightly superior than previously reported. When genetic patients are excluded, Ding [17] and Pelletier [19] found a relapse rate of 46.3% and 47%, respectively. The lower recurrence rate described by these groups could be justified by the presence of unknown genetic SRNS patients. Indeed, when both Group B and C are considered, the overall rate of recurrence was 53.3% in our cohort. Therefore, our data underline the importance of a genetic evaluation for SRNS genes in order to plan transplantation, as it represents the principle risk factor for recurrence.

Aside from absence of a genetic aetiology, our study identified age at transplant greater than 9 years and HLA-AB match as independent risk factors for recurrence. The best prediction model for recurrence included also a longer duration of dialysis.

Nehus et al. reported a higher rate of recurrence in younger children, among a cohort of 327 patients, though genetic results were not reported for any participants [24]. No significant difference in relapse rate according to the age at transplant were detected by Tejani and Stablein [12] and in the more recent studies by Ding et al. and Pelletier et al. [17, 19]. Again, unavailability of genetic testing for the majority of their patients could justify the different findings.

HLA-AB match was independently related to recurrence in our cohort, in contrast HLA AB or DR match did not influence the risk of relapse in the study by Tejani et al. [12] and did not affect transplant outcome in adolescent with SRNS in a retrospective study of the NAPRTCS registry [25].

Following the evidence that a circulating factor is responsible for recurrence, it has been suggested that a prolonged dialysis prior to renal transplantation would have a protective effect as far as the risk of relapse is concerned. The results of our study do not support this hypothesis. Indeed, in our study cohort, a longer duration of dialysis was associated with an increased risk of relapse. Even if this variable did not reach significance, its inclusion identifies the best prediction model for recurrence (R^2 0.271136). Among the few studies which investigated the association between duration of dialysis and disease recurrence, no significant differences were found in a single centre experience of 43 patients by Senggutuvan [16]. In a larger cohort of 132 paediatric renal transplants, found no relationship between disease recurrence and duration of dialysis was found [12]. Hence, since no protective effect was proven by others and our data show a longer duration of dialysis in patients with recurrence, even if not statistically significant, it is not justified to prolong the duration of dialysis before transplantation in children with SRNS.

Whether donor type (living vs deceased) is significantly associated with disease recurrence remains controversial. Data from old registries [26, 27] found no increased recurrence rates according to the type of donors. Other studies have suggested living donor as an independent risk factor for recurrence [17, 28, 29]. Our study cohort included only six living donor recipients (5.9%), reflecting the reluctance of paediatric nephrologists to use living donors in SRNS patients, due to the risk of recurrence and graft loss. On the other hand, since in our cohort no relapses occurred in genetic SRNS, another important clinical implication of our study is that living kidney donors can be safely used in genetic SRNS patients.

All patients from our cohort were treated with PE, following SRNS recurrence. Complete or partial remission was achieved in 19/32 (59.4%) children, with a functioning graft after a median follow-up of 39.5 months. Similar rates of response were previously reported. Kashgary et al. in their meta-analysis identified a remission rate of 70.2% in children treated with PE [14]. A lower response rate was reported by Pelletier [19], but remission information was available only

for 49/64 (77%) relapsed patients and the detailed immunosuppressive strategy is missing. According to our results, PE is confirmed as an effective treatment for recurrence. Even the small numbers, rituximab and high dose steroids did not influence the response rate in our cohort.

On the other hand, disease recurrence was the leading cause of graft loss in non-genetic SRNS and the rate of graft loss after relapse (34.3%) in our study is consistent with previous data [26].

Among the small number of retransplanted individuals included in our study, the overall incidence of relapse in a second renal graft after a first recurrence is not significantly different from the first transplant (55%). In 4/9 patients who experienced a relapse in their first transplant, proteinuria did not recur after the second transplantation. This contrasts with reported small cohorts in whom the incidence of recurrence approaches 100% once the first transplant was lost for recurrent SRNS [10, 12, 28, 30]. We are not able to identify the factors responsible of the different outcome, but according to our data retransplantation after relapse can be considered in children with SRNS.

Conclusions

Twelve years of the Italian experience with post-transplant SRNS recurrence allows us to reach different important conclusions. Firstly, the absence of underlying genetic mutations predicts a high risk of post-transplant recurrence, therefore genetic screening must be performed in all children with SRNS before transplantation in order to best plan their care in the post-transplant period. Age > 9 years is an independent risk factor for recurrence, while a prolonged time spent on dialysis before transplantation has no protective effect on the risk of relapse and should not be encouraged. Living donor did not influence the risk of relapse and can be safely used in genetic SRNS patients. PE based treatment strategies are effective in the majority of relapsed patients. Finally, in those who experience graft loss, even for recurrence, it is appropriate to consider retransplantation, as it maybe curative in the long term.

Acknowledgements We thank "La Nuova Speranza onlus" foundation for the management support.

Author Contributions WM designed the study, collected clinical data, performed data analysis and interpretation, drafted the article, and approved the final version of the manuscript. SP collected clinical data, performed data analysis, drafted the article, and approved the final version of the manuscript. GP performed the statistical analysis and data interpretation, drafted the article and approved the final version of the manuscript. GMG concepted the study, critically revised the article, and approved the final version of the manuscript. LDS collected clinical data, critically revised the article, and approved the final version of the manuscript. LP collected clinical data, critically revised the article, and approved the final version of the manuscript. LM collected clinical data, critically revised the article, and approved the final version of the manuscript. MC collected clinical data, critically revised the article, and approved the final version of the manuscript. IG collected clinical data, critically revised the article, and approved the final version of the manuscript. EC collected clinical data, critically revised the article, and approved the final version of the manuscript. EB collected clinical data, critically revised the article, and approved the final version of the manuscript. ST collected clinical data and approved the final version of the manuscript. LG collected clinical data and approved the final version of the manuscript. GC collected clinical data and approved the final version of the manuscript. MC collected clinical data and approved the final version of the manuscript. RT collected clinical data and approved the final version of the manuscript. GM concepted and designed the study, performed data interpretation, critically revised the article, and approved the final version of the manuscript.

Funding This study received no specific fund from any public, commercial, or not-for-profit agency.

Compliance with ethical standards

Conflict of interest WM has received a speaker honorarium from Sanofi-Genzyme. No other conflict of interest to declare.

Ethical statement The study was approved by the institutional review board of Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan.

Informed consent Informed consent was not required in view of the retrospective study design and the anonymity of the patient records reviewed.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Sadowski CE, Lovric S, Ashraf S et al (2015) A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. J Am Soc Nephrol 26:1279–1289. https://doi.org/10.1681/ASN.2014050489
2. Becherucci F, Mazzinghi B, Provenzano A et al (2016) Lessons from genetics: is it time to revise the therapeutic approach to children with steroid-resistant nephrotic syndrome? J Nephrol 29:543–550. https://doi.org/10.1007/s40620-016-0315-4
3. Irginia V, Avin JS, Am R et al (1996) Circulating factor associated with increased glomerular permeability to albumin in recurrent focal segmental glomerulosclerosis. N Engl J Med 334:878–883. https://doi.org/10.1056/NEJM199604043341402
4. Dall'Amico R, Ghiggeri G, Carraro M et al (1999) Prediction and treatment of recurrent focal segmental glomerulosclerosis after renal transplantation in children. Am J Kidney Dis 34:1048–1055. https://doi.org/10.1016/S0272-6386(99)70010-7
5. Weber S, Tönshoff B (2005) Recurrence of focal-segmental glomerulosclerosis in children after renal transplantation: clinical

and genetic aspects. Transplantation 80:128–134. https://doi.org/10.1097/01.tp.0000187110.25512.82
6. Mahesh S, Del Rio M, Feuerstein D et al (2008) Demographics and response to therapeutic plasma exchange in pediatric renal transplantation for focal glomerulosclerosis: a single center experience. Pediatr Transplant 12:682–688. https://doi.org/10.1111/j.1399-3046.2007.00880.x
7. Jungraithmayr TC, Hofer K, Cochat P et al (2011) Screening for NPHS2 mutations may help predict FSGS recurrence after transplantation. J Am Soc Nephrol 22:579–585. https://doi.org/10.1681/asn.2010010029
8. Vinai M, Waber P, Seikaly MG (2010) Recurrence of focal segmental glomerulosclerosis in renal allograft: an in-depth review. Pediatr Transplant 14:314–325. https://doi.org/10.1111/j.1399-3046.2009.01261.x
9. Weber S, Gribouval O, Esquivel EL et al (2004) NPHS2 mutation analysis shows genetic heterogeneity of steroid-resistant nephrotic syndrome and low post-transplant recurrence. Kidney Int 66:571–579. https://doi.org/10.1111/j.1523-1755.2004.00776.x
10. Fine RN (2007) Recurrence of nephrotic syndrome/focal segmental glomerulosclerosis following renal transplantation in children. Pediatr Nephrol 22:496–502. https://doi.org/10.1007/s00467-006-0361-6
11. Ulinski T (2010) Recurrence of focal segmental glomerulosclerosis after kidney transplantation: strategies and outcome. Curr Opin Organ Transplant 15:628–632. https://doi.org/10.1097/MOT.0b013e32833dee3a
12. Tejani A, Stablein D (1992) Recurrence of focal segmental glomerulosclerosis posttransplantation: a special report of the North American Pediatric Renal Transplant Cooperative Study. J Am Soc Neph 2:S258–S263. https://doi.org/10.1017/CBO9781107415324.004
13. Donckerwolcke RA, Broyer M, Brunner FP et al (1983) Combined report on regular dialysis and transplantation of children in Europe, XI, 1981. Proc Eur Dial Transplant Assoc 19:61–91
14. Kashgary A, Sontrop JM, Li L et al (2016) The role of plasma exchange in treating post-transplant focal segmental glomerulosclerosis: a systematic review and meta-analysis of 77 case-reports and case-series. BMC Nephrol 17:1–8. https://doi.org/10.1186/s12882-016-0322-7
15. Bierzynska A, Saleem MA (2018) Deriving and understanding the risk of post-transplant recurrence of nephrotic syndrome in the light of current molecular and genetic advances. Pediatr Nephrol 33:2027–2035. https://doi.org/10.1007/s00467-017-3793-2
16. Senggutuvan P, Cameron JS, Hartley RB et al (1990) Recurrence of focal segmental glomerulosclerosis in transplanted kidneys: analysis of incidence and risk factors in 59 allografts. Pediatr Nephrol 4:21–28
17. Ding WY, Koziell A, McCarthy HJ et al (2014) Initial steroid sensitivity in children with steroid-resistant nephrotic syndrome predicts post-transplant recurrence. J Am Soc Nephrol 25:1342–1348. https://doi.org/10.1681/asn.2013080852
18. Hubsch H, Montan B, Abitbol C et al (2005) Recurrent focal glomerulosclerosis in pediatric renal allografts: the Miami experience. Pediatr Nephrol 20:210–216. https://doi.org/10.1007/s00467-004-1706-7
19. Pelletier JH, Kumar KR, Engen R et al (2018) Recurrence of nephrotic syndrome following kidney transplantation is associated with initial native kidney biopsy findings. Pediatr Nephrol 34:539. https://doi.org/10.1007/s00467-018-4103-3
20. Bertelli R, Ginevri F, Caridi G et al (2003) Recurrence of focal segmental glomerulosclerosis after renal transplantation in patients with mutations of podocin. Am J Kidney Dis 41:1314–1321
21. Billing H, Müller D, Ruf R et al (2004) NPHS2 mutation associated with recurrence of proteinuria after transplantation. Pediatr Nephrol 19:561–564. https://doi.org/10.1007/s00467-003-1408-6
22. Becker-Cohen R, Bruschi M, Rinat C et al (2007) Recurrent nephrotic syndrome in homozygous truncating NPHS2 mutation is not due to anti-podocin antibodies. Am J Transplant 7:256–260. https://doi.org/10.1111/j.1600-6143.2006.01605.x
23. Caridi G, Bertelli R, Perfumo F, Ghiggeri GM (2004) Heterozygous NPHS1 or NPHS2 mutations in responsive nephrotic syndrome and the multifactorial origin of proteinuria. Kidney Int 66:1715–1716. https://doi.org/10.1111/j.1523-1755.2004.938_9.x
24. Nehus EJ, Goebel JW, Succop PS, Abraham EC (2013) Focal segmental glomerulosclerosis in children. Transplant J 96:550–554. https://doi.org/10.1097/TP.0b013e31829c2431
25. Baum MA, Ho M, Stablein D et al (2002) Outcome of renal transplantation in adolescents with focal segmental glomerulosclerosis. Pediatr Transplant 6:488–492
26. Abbott KC, Sawyers ES, Oliver JD et al (2001) Graft loss due to recurrent focal segmental glomerulosclerosis in renal transplant recipients in the United States. Am J Kidney Dis 37:366–373. https://doi.org/10.1053/ajkd.2001.21311
27. Hariharan S, Adams MB, Brennan DC et al (1999) Recurrent and de novo glomerular disease after renal transplantation: a report from Renal Allograft Disease Registry (RADR). Transplantation 68:635–641
28. First MR (1995) Living-related donor transplants should be performed with caution in patients with focal segmental glomerulosclerosis. Pediatr Nephrol 9(Suppl):S40–S42
29. Francis A, Trnka P, McTaggart SJ (2016) Long-term outcome of kidney transplantation in recipients with focal segmental glomerulosclerosis. Clin J Am Soc Nephrol 11:2041–2046. https://doi.org/10.2215/CJN.03060316
30. Striegel JE, Sibley RK, Fryd DS, Mauer SM (1986) Recurrence of focal segmental sclerosis in children following renal transplantation. Kidney Int Suppl 19:S44–S50

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.