Development of a Prediction Score to Avoid Confirmatory Testing in Patients With Suspected Primary Aldosteronism

Development of a prediction score to avoid confirmatory testing in patients with suspected primary aldosteronism.

Jacopo Burrello MD[1*], Martina Amongero MS[2*], Fabrizio Buffolo MD[1], Elisa Sconfienza MD[1], Vittorio Forestiero MD[1], Alessio Burrello MS[3], Christian Adolf MD[4], Laura Handgriff MD[4], Martin Reincke MD[4], Franco Veglio MD[1], Tracy Ann Williams PhD[1,4], Silvia Monticone MD, PhD[1#], Paolo Mulatero MD[1#].

(1) Division of Internal Medicine and Hypertension, Department of Medical Sciences, University of Torino, Italy. (2) Department of Mathematical Sciences G. L. Lagrange, Polytechnic University of Torino, Italy. (3) Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi" (DEI), University of Bologna, Italy. (4) Medizinische Klinik und Poliklinik IV, Klinikum der Universität, Ludwig-Maximilians-Universität München, Munich, Germany.

* Contributed equally and should be considered as joint first authors.

# Contributed equally and should be considered as joint last authors.

Short title: Diagnostic modelling to confirm PA diagnosis

Keywords: aldosterone; primary aldosteronism; confirmatory testing; machine learning.

Corresponding author and person to whom reprints should be addressed: Prof. Paolo Mulatero Division of Internal Medicine and Hypertension Unit, Department of Medical Sciences, University of Torino, Città della Salute e della Scienza, Via Genova 3, 10126 Torino, Italy. Telephone/Fax number: +39.011.633.6959 / +39.011.633.6931. E-mail: paolo.mulatero@unito.it

Source(s) of Funding: This research did not receive any specific grant in Torino. The German Conn-Registry-Else-Kröner Hyperaldosteronism Registry is supported by the Else Kröner-Fresenius Stiftung (2013_A182, 2015_A171 and 2019_A104). CA and MR are supported by the Deutsche

Forschungsgemeinschaft (DFG, within the CRC/Transregio 205/1 “The Adrenal: Central Relay in Health and Disease”).

Conflict(s) of Interest/Disclosure(s): the authors have nothing to disclose.

ABSTRACT

Context. The diagnostic work-up of primary aldosteronism (PA) includes screening and confirmation steps. Case confirmation is time-consuming, expensive, and there is no consensus on tests and thresholds to be used. Diagnostic algorithms to avoid confirmatory testing may be useful for the management of patients with PA.

Objective. Development and validation of diagnostic models to confirm or exclude PA diagnosis in patients with a positive screening test.

Design, Patients and Setting. We evaluated 1,024 patients who underwent confirmatory testing for PA. The diagnostic models were developed in a training cohort (n=522), and then tested on an internal validation cohort (n=174) and on an independent external prospective cohort (n=328).

Main outcome measure. Different diagnostic models and a 16-point score were developed by machine learning and regression analysis to discriminate patients with a confirmed diagnosis of PA.

Results. Male sex, antihypertensive medication, plasma renin activity, aldosterone, potassium levels and presence of organ damage were associated with a confirmed diagnosis of PA. Machine learning based models displayed an accuracy of 72.9-83.9%. The Primary Aldosteronism Confirmatory Testing (PACT) score correctly classified 84.1% at training and 83.9% or 81.1% at internal and external validation, respectively. A flow chart employing the PACT score to select patients for confirmatory testing, correctly managed all patients, and resulted in a 22.8% reduction in the number of confirmatory tests.

Conclusions. The integration of diagnostic modelling algorithms in clinical practice may improve the management of patients with PA by circumventing unnecessary confirmatory testing.

INTRODUCTION

Primary aldosteronism (PA) represents the most frequent cause of secondary hypertension, with a prevalence reaching 29.8% in referral centers [1,2]. According to the Endocrine Society (ES), high risk groups accounting for up to 50% of patients with hypertension should be screened for PA by measurement of the aldosterone-to-renin ratio (ARR) [1,2]. Guidelines do not recommend an ARR cut-off for a positive screening test. Nonetheless, an ARR value ranging between 20 and 40 (as [ng/dL]/[ng/mL/h]) is suggested, depending on assay used for aldosterone and renin measurements and inter-center variability [1,2]. The cut-off used by the majority of referral centers is 30 [ng/dL]/[ng/mL/h], which maximizes sensitivity, but may lead to many false positive results [3]. For this reason, confirmatory testing, to either confirm or exclude the diagnosis of PA, is recommended in patients with a positive case detection. In particular, the ES guideline recommends one of the following: saline infusion test, captopril challenge test, oral sodium loading test, or fludrocortisone suppression test [1,2]. The confirmatory test could be avoided for patients who display spontaneous hypokalemia, suppressed renin levels and an aldosterone at screening greater than 20 ng/dL [1,2]. The Japanese Endocrine Society recommends the performance of at least 2 different confirmatory tests for all patients with a positive screening test [4].

Confirmatory testing aims to identify false positives at screening to avoid subsequent costly and invasive investigations, including adrenal vein sampling. However, these tests are time-consuming and there is no consensus on the best test or on the thresholds that should be used for confirmation or exclusion of PA. Studies that assessed between-test comparability suffer from several limitations, including sample size, retrospective design, and selection bias [1,2].

Diagnostic algorithms that employ clinical and biochemical parameters at screening would be useful to identify patients who can bypass confirmatory testing and proceed directly to subtype differentiation due to a high likelihood of PA and those with such a low likelihood that confirmatory

testing is unnecessary. The aim of the present study was to develop and validate computational models to confirm or exclude the diagnosis of PA in patients with a positive screening test. We propose different diagnostic algorithms based on machine learning techniques, and a flow chart for patient management which integrates the Primary Aldosteronism Confirmatory Testing (PACT) score to stratify patients according to their likelihood of PA.

METHODS

Single patient data extracted during the present study are not publicly available but are available from the corresponding author on reasonable request. Supplemental Data [5] are available at the link: https://github.com/CentroIpertenUnito/PACT-score.

Data extraction and study cohorts

For the developmental cohort, we retrospectively assessed data from 696 patients referred to the tertiary hypertension unit of Torino in whom confirmatory testing had been performed for a suspected diagnosis of PA. Inclusion criteria were: (1) a positive screening test for PA (see below); (2) a diagnosis of confirmed or not confirmed PA by confirmatory testing. Patients were excluded in case of autonomous cortisol secretion. Eligible patients from the developmental cohort were randomized to a training cohort (n=522) or to an internal validation cohort (n=174). An independent prospective cohort of 328 patients consecutively recruited from the Munich Klinikum der Universität was used for external validation. All the patients included in the present retrospective analysis gave extended written consent for the use of their personal data, according to Helsinki declaration. The study was approved by local ethical committees.

PA was diagnosed in accordance with the Endocrine Society (ES) guideline [1] and ESH consensus [2,6]. Screening was performed by measurement of the aldosterone concentration (AC) to plasma renin activity ratio (ARR) in the developmental cohort and by AC to direct renin concentration (DRC) in the external validation cohort. Interfering drugs were withdrawn according to guidelines [1,2]. The

screening test was considered positive if ARR was higher than 30 ng/dL/ng*mL^{-1}*h^{-1} and AC higher than 10 ng/dL. Patients with suspected PA underwent confirmatory testing by either an intravenous saline loading test or a captopril challenge test [2]. The cut-offs for a positive confirmatory test were a post-test AC > 5 ng/dL for intravenous saline loading, or a post-test ARR > 30 ng/dL/ng*mL^{-1}*h^{-1} for a captopril challenge test. After PA confirmation, subtype diagnosis was defined by computed tomography and adrenal venous sampling [6].

Statistics and machine learning analyses

Kolmogorov–Smirnov test was used to evaluate the distribution of patient parameters. Normally distributed parameters were expressed as mean ± standard deviation and analyzed by student t-test. Non-normally distributed parameters were expressed as median [interquartile range] and analyzed by Mann-Whitney's test. Categorical parameters were expressed as absolute number and percentage distribution and analyzed by Chi-square test. A *P*-value of less than 0.05 was considered significant. Univariate and multivariate logistic regression was used to assess the odds ratios (ORs). An OR greater than 1 was associated with an increased likelihood of a confirmed diagnosis of PA, an OR less than 1, a decreased likelihood.

The machine learning models, and the PACT score were built on the training cohort and then tested in the internal and external validation cohorts. Patients from the internal and external validation cohorts were distinct from the training cohort, in which all the models were built.

Supervised machine learning algorithms are used to formulate prediction about a select outcome on the base of a pre-defined set of labeled paired input-output data [7]. We used different models, including linear discriminant analysis (LDA), random forest (RF) classification algorithms, and support vector machine (SVM) with different kernels (linear and gaussian radial basis function). LDA employs linear combination of parameters to maximize the separation between groups by increasing precision estimates by variance reduction. The predicted diagnosis is derived from the following equation: Confirmed PA diagnosis = $LDAcoeff_1$*$Variable_1$ + $LDAcoeff_2$*$Variable_2$ + … +

LDAcoeff$_n$*Variable$_n$ > tested thresholds. The RF algorithm creates 30 classification trees with a maximum number of 7 splits for each tree. The predicted diagnosis resulted from the outcome of each classification tree of the forest; if at least 16 of 30 trees of the RF confirm PA, then the diagnosis of PA will be confirmed. Linear SVM builds a classification model to assign patients to their diagnosis given a linear boundary. The model finds out the plane which best separates groups of patients (i.e. confirmed *vs.* not confirmed diagnosis of PA), maximizing the distances between them. Patients are classified according to the following equation: SVMcoeff$_0$ + SVMcoeff$_1$*Variable1 + SVMcoeff$_2$*variable$_2$ + …. + SVMcoeff$_n$*Variable$_n$. Gaussian SVM allows to divide patients using a non-linear boundary. The corresponding equation is: SVMcoeff$_0$ + SVMcoeff$_1$*f(Variable1) + SVMcoeff$_2$*f(variable$_2$) + …. + SVMcoeff$_n$*f(Variable$_n$), where "f" is an exponential function coefficient.

The diagnostic performance of the PACT score was assessed by analysis of receiver operating characteristics (ROC) curves. The area under the curve (AUC) was evaluated to define the best cut-off by Youden Index (J = sensitivity + specificity - 1). Overfitting effect was defined as the difference between the accuracy at the training of the models and the accuracy at validation. A free-downloadable tool was developed to calculate the score and the predicted diagnosis (available at: https://github.com/CentroIpertenUnito/PACT-score/raw/master/PACT%20Score%20Calculator.xlsm).

Python 3.5 (library, scikit-learn) and IBM SPSS Statistics 26 (*IBM Corp., Armonk, New York, USA*) were used for analysis.

RESULTS

Clinical and biochemical characteristics

In the present study, we evaluated the data from 1,024 patients including a developmental cohort from Torino (n=696) and a prospective independent validation cohort from Munich (n=328). Patients from the developmental cohort included 421 subjects with a confirmed diagnosis of PA, and 275 for whom PA was not confirmed (Table 1). Patients with confirmed PA were more frequently male, with

higher systolic blood pressure, and defined daily dose (DDD) for antihypertensive medications, compared with patients in whom PA was not confirmed (p<0.01 for all comparisons). Aldosterone levels were higher, and plasma renin activity and potassium levels were lower in patients with confirmed *vs.* not confirmed PA (p<0.001). The prevalence of organ damage and cardiovascular events was higher in patients with a confirmed diagnosis of PA.

Univariate logistic regression analysis confirmed a relevant association of sex, systolic blood pressure, DDD, PRA and aldosterone levels at screening, lowest potassium, and prevalence of organ damage and cardiovascular events with a confirmed diagnosis of PA (Table S1). Multivariate regression analysis confirmed female sex (OR 0.41), DDD (OR 1.20), PRA (OR 0.07), aldosterone (OR 1.08), lowest potassium (OR 0.14), and presence of organ damage (OR 2.63) as independent predictors of a confirmed diagnosis of PA (Table 2 and Table S2).

Diagnostic modelling

Patients from the developmental cohort were randomly assigned to a training cohort (n=522) and to an internal validation cohort (n=174). No differences were found between the 2 cohorts with regard to clinical and biochemical parameters (Table S3). All diagnostic models were developed in the training cohort and tested in the internal validation cohort. A second external cohort of patients was used for an independent validation (n=328). Patients from the external validation cohort compared with the developmental cohort, displayed lower aldosterone levels, lower potassium, lower DDD, shorter duration of hypertension, and lower systolic blood pressure (p<0.01 for all comparisons). The prevalence of PA was also lower than in the developmental cohort (52.7 *vs.* 60.5%; p=0.019; Table S4).

The 6 parameters selected by regression analysis (Table 2) were used for the development of machine learning models. Their diagnostic performance in the discrimination of patients with a confirmed diagnosis of PA on the combined developmental cohort is shown in Figure S1. Predictor importance coefficients are reported in Figure S2. The best predictor was lowest potassium in all the models.

182 The linear combination of parameters by LDA is shown in the canonical plot (Figure S1A). In
183 particular, 551 of 696 patients (accuracy 79.2%) were correctly classified, with a sensitivity and
184 specificity for PA detection of 84.1% and 71.6%, respectively.
185 The RF classification algorithm reached a higher performance by correctly discriminating 571 of 696
186 patients (accuracy 82%) with a sensitivity and specificity of 87.9% and 73.1%, respectively. The RF
187 was composed by 30 trees; the first tree of the series is reported in Figure S1B.
188 Finally, SVM models (linear and gaussian kernel) displayed similar performance, with an accuracy
189 of 80.0% and 81.6%, respectively. The linear SVM was able to correctly classify 354 of 421 patients
190 with a confirmed diagnosis of PA (sensitivity 84.1%) and 203 of 275 patients with not confirmed PA
191 (specificity 73.8%). The gaussian SVM correctly classified 365 of 421 patients with a confirmed
192 diagnosis of PA (sensitivity 86.7%), at the same specificity of the linear kernel (73.8%).
193 Representative plots for main discriminants of SVM models are shown in Figure S1C and S1D.
194 Table S5 reports the confusion matrix and diagnostic performance of machine learning based models
195 at training and internal validation on the developmental cohort. The overfitting effect was low for all
196 the models (between 2.1% and 9.2%), thus suggesting an acceptable generalizability of the models.
197 The performance at external validation on the independent cohort was still very high, ranging between
198 72.9% and 78.7% (Table S6).
199
200 ***Development and validation of the PACT score***
201 The same 6 parameters employed for the machine learning models were used to develop a 16-point
202 scoring system. As for the diagnostic algorithms described above, the PACT score was built on the
203 training cohort and then tested in the internal and external validation cohorts (Table S7).
204 Categorization and points assignment are shown in Figure 1A and 1E.
205 The PACT score was directly correlated with the proportion of patients with a confirmed diagnosis
206 of PA (Figure 1B) and with pre-/post- confirmatory testing aldosterone and aldosterone-to-renin ratio
207 (Pearson's R ranging between 0.247 and 0.479; $p<0.001$; Table S8). Noteworthy, for all patients with

a score greater than 12, PA diagnosis was confirmed, whereas for all patients with a score lower than 5, PA diagnosis was not confirmed (Table S9). The analysis of ROC curves demonstrated a reliable performance of the score both at screening and internal validation (AUC of 0.879 and 0.877, respectively; Figure 1C and 1D). The cut-off with the highest accuracy was 8 points.

In the combined developmental cohort, a score equal or greater than 8 correctly discriminated a confirmed diagnosis of PA in 388 of 421 patients (sensitivity 92.2%), whereas a score lower than 8 identified those with not confirmed PA in 197 of 275 cases (specificity 71.6%), with an overall accuracy of 84.1%. A cut-off of equal or greater than 5 reached 100% sensitivity, correctly classifying all the patients with a confirmed diagnosis of PA, both in the training and validation cohorts. On the other side, a score lower than 13 correctly identified all the patients for which a diagnosis of PA was not confirmed at training and at validation of the score. Confusion matrix and diagnostic performance of the PACT score at training and internal validation are reported in Table S7.

Accuracy at internal validation was 83.9%. The overfitting effect was minimum (0.2%), and consistently the performance at external validation confirmed the very high generalizability of our score system with an accuracy of 81.1% (Table S6). The PACT score displayed a sensitivity ranging between 78.6% and 91.9% and specificity between 73.3% and 83.9%, at internal and external validation. Of note, this performance was similar and even higher than that of machine learning models (accuracy at external validation: 78.4%, 72.9%, 78.4%, and 78.7%, for the LDA, the RF algorithm, linear and gaussian SVM, respectively).

Management of patients with PA

The PACT score was implemented in a flow chart for the management of patients with PA (Figure 2). Patients with a positive screening test were stratified for their likelihood of PA diagnosis according to our score-system (developmental cohort and external validation cohort; n=1,024). For patients with a score less than 5, PA diagnosis was excluded without a confirmation test (n=107); instead, for patients with a score equal or greater than 13, PA diagnosis was confirmed without further tests

(n=126). All the remaining patients (n=791; score between 5 and 12) should undergo confirmatory test and be allocated according with subsequent investigations. This approach resulted in the correct management of all patients (accuracy 100%), with the reduction of 22.8% of unnecessary confirmatory tests (233 of 1,024 procedures).

A second model for the management of patients with PA was developed by the use of stricter cut-offs (Figure S3A). In this case, a PACT score lower than 6 correctly excluded PA diagnosis in 170 of 191 patients; 21 patients with a confirmed diagnosis of PA were missed (none of them displayed a diagnosis of aldosterone producing adenoma). A cut off of equal or greater than 11 correctly confirmed the diagnosis of PA in 258 of 277 patients; 19 patients with low-renin hypertension would undergo to inappropriate adrenal venous sampling. At this regard, we would underline that for 11 of the 19 misclassified low-renin hypertensive patients, the confirmatory test was performed before 2014 in recumbent position, thus representing potentially false negative patients [8]. The second flow chart displayed an accuracy of 96.1% (96.5% and 95.6% of sensitivity and specificity, respectively), reducing the number of necessary confirmatory tests of 45.7% (Figure S3B).

Finally, we performed a sub-analysis on patients with unilateral PA. The performance of all the proposed diagnostic models is reported in Table S10: the accuracy ranged between 95 and 100%. In particular, the two flow charts for patient management correctly classified as confirmed PA all the patients with unilateral disease.

DISCUSSION

We used supervised machine-learning algorithms and regression analysis to develop prediction models and a simple scoring system, to discriminate patients with confirmed PA. We provided evidence that confirmatory testing could be avoided in a subset of patients selected by their clinical and biochemical parameters at screening. We developed a flow-chart integrating the PACT score to stratify patients with suspected PA and a positive screening test according to their likelihood of PA diagnosis. This model provides options to clinicians, who may propose a strict follow-up for patients

with a low likelihood of PA (PACT score less than 5), or directly proceed to subtype diagnosis with CT scanning and adrenal venous sampling for patients with a high likelihood of PA (PACT score equal or greater than 13). Patients with an intermediate risk should follow the diagnostic flow-chart recommended by the guidelines and undergo confirmatory testing. This approach results in the correct management of all patients and reduces the number of confirmatory tests by 23%. An online downloadable tool facilitates the application of the PACT score in clinical practice.

The growing awareness of the scientific community on the importance of screening for secondary causes of hypertension [9], will lead to an increase in the number of patients with a positive screening test for PA, and therefore an increased requirement of confirmatory tests.

The systematic confirmation of all patients with a positive screening test determines an increase in costs, time, risks and complexity in the management of patients with PA [10], thus contributing to the underdiagnosis of PA [11]. The PACT score may simplify the diagnostic work-up for a subset of selected patients at low or high likelihood of PA, resulting in an increased availability of resources to be allocated for subtype diagnosis, and targeted therapy, which are efficacious and cost-effective [12].

According to the ES guideline and ESH consensus, patients with hypokalemia with suppressed renin and high aldosterone levels could skip confirmatory testing [1,2]; in the overall population included in the study (developmental and validation cohorts), 45 patients (4.4%) displayed these characteristics and could directly undergo subtype diagnosis. On the other side, 126 patients (12.3%) could directly undergo subtype diagnosis following the PACT score, and in another 10.4% of cases, the diagnosis of PA could be excluded without further testing, resulting in a net advantage on confirmatory testing reduction (23% with the PACT score, compared to 4.4% with the ES recommendations). Noteworthy, using stricter cut-offs the PACT score could allow the reduction of 46% of confirmatory tests maintaining an overall accuracy of 96.1%.

Avoiding potentially unnecessary confirmatory tests could have an impact not only on the reduction of costs, but also on patient management. Even if minimally invasive, confirmatory testing is

associated with side effects, including hypertensive or hypotensive episodes, arrhythmias, vertigo, headache, dyspnea, and neurological symptoms [10]. A reduction of the number of tests will also reduce the incidence of clinical complications, related to intravascular volume expansion (in case of saline loading test) or hypotension (after captopril administration).

Previous studies proposed different cut-offs for AC and ARR to avoid confirmatory tests. Nanba et al. first demonstrated that most patients evaluated with saline infusion test, captopril challenge test and/or furosemide upright test displayed a confirmed diagnosis of PA when ARR was equal or greater than 100 [ng/dL]/[ng/mL/h] in the presence of an AC of at least 25 ng/dL [13]. Maiolino et al. observed that the progressive increase of the ARR at screening was associated with an increased specificity for the diagnosis of an aldosterone producing adenoma [14]. Recommendations from the French Endocrinology and Hypertension societies suggest the avoidance of confirmatory testing in the presence of an AC above 20 ng/dL, high ARR, with or without hypokalemia. On the contrary, PA diagnosis could be ruled out if AC at screening is less than 9 ng/dL in two different occasions [15]. This approach resulted in a positive predictive value of 93% in a recent study on a cohort of 173 hypertensive patients referred to a single hypertension center [16]. Finally, two recent reports observed that an AC above 30 ng/dL, or above 20 ng/dL in the presence of hypokalemia, rendered confirmatory testing unnecessary [17,18]. Main limitations of all these studies are represented by the general applicability of the proposed cut-offs, the absence of an independent validation, and the relatively low sensitivity and negative predictive value.

To our knowledge, one single study combined different clinical parameters to develop a scoring system to skip confirmatory testing [19]. Using age, BMI, number of antihypertensive medications, sodium, potassium, and presence of diabetes, the authors reported 100% specificity and positive predictive value for PA diagnosis, with a reduction of 42.2% of confirmatory test. However, this was a single-center retrospective study with a limited sensitivity (52%) not validated in an independent cohort [19]. In the PACT score, we combined parameters which are easily available for patients who

underwent screening for PA: sex, DDD, PRA and aldosterone values, potassium, and the presence of organ damage. In line with previous reports, hypokalemia, aldosterone and PRA at screening were the main discriminants [17,18,20], with lowest potassium being the best predictor in all machine learning models. In particular, an AC of at least 28 ng/dL and a PRA less than 0.2 mg/mL/h, in presence of history of hypokalemia (lowest potassium < 3.7 mEq/L) resulted in a score of 11, corresponding to a 77.8% likelihood of PA. Male patients had an increase probability of a confirmed diagnosis of PA and female sex has been associated with a false positive result at screening test [21,22]. Antihypertensive medications expressed as DDD and target organ damage completed the score.

Some study limitations should be acknowledged. First of all, the generalizability of our approach for patient management could be limited by bias related to the cohort characteristics and assays used for AC and PRA measurement. However, we tested our diagnostic models in an external validation cohort which differed significantly for 5 of the 6 included parameters, and where patients were screened using DRC instead of PRA and displayed lower median aldosterone levels. Moreover, we built our score in a retrospective developmental cohort, but the validation was provided in an independent external prospective cohort of patients enrolled consecutively. Finally, the PACT score has been designed to select patients at screening test to avoid a proportion of confirmatory tests, but it was not tested in patients who had PRA and aldosterone measured under interfering medications. In this regard, it should be noted that the effects of interfering drugs on hormonal measurements cannot be standardized, which is a major issue when developing a prediction model. The assessment of the diagnostic performance of our models in a cohort of patients screened for PA under interfering drugs should be evaluated in future studies.

This is the first study which implemented machine learning algorithms and regression analysis to develop and validate prediction models to discriminate patients with a confirmed diagnosis of PA,

using screening parameters, in a large cohort of patients from two specialized referral centers. Internal and external validation of the models demonstrated a reliable generalizability with a low overfitting effect. The performance of the proposed diagnostic algorithms was higher than all previously proposed approaches, with a reduction of the number of confirmatory tests of up to 23%. The algorithm appears to be a significant improvement compared with recommendations of international guidelines, which would have bypassed confirmatory testing in 4.4% of patients. This approach may have a high potential impact on the management of PA with a reduction of costs and simplification of the diagnostic work-up of patients with hypertension.

CONCLUSIONS

Combining different clinical parameters in the PACT score, we discriminate with high accuracy patients with a confirmed diagnosis of PA, in a large cohort of patients with a positive screening test. This approach could result in a significant reduction of unnecessary confirmatory tests. The integration of diagnostic modelling algorithms in clinical practice will increase the detection rate of PA and improve the management of these patients.

Source(s) of Funding: This research did not receive any specific grant in Torino. The German Conn-Registry-Else-Kröner Hyperaldosteronism Registry is supported by the Else Kröner-Fresenius Stiftung (2013_A182, 2015_A171 and 2019_A104). CA and MR are supported by the Deutsche Forschungsgemeinschaft (DFG, within the CRC/Transregio 205/1 “The Adrenal: Central Relay in Health and Disease”).

Conflict(s) of Interest/Disclosure(s): the authors have nothing to disclose.

REFERENCES

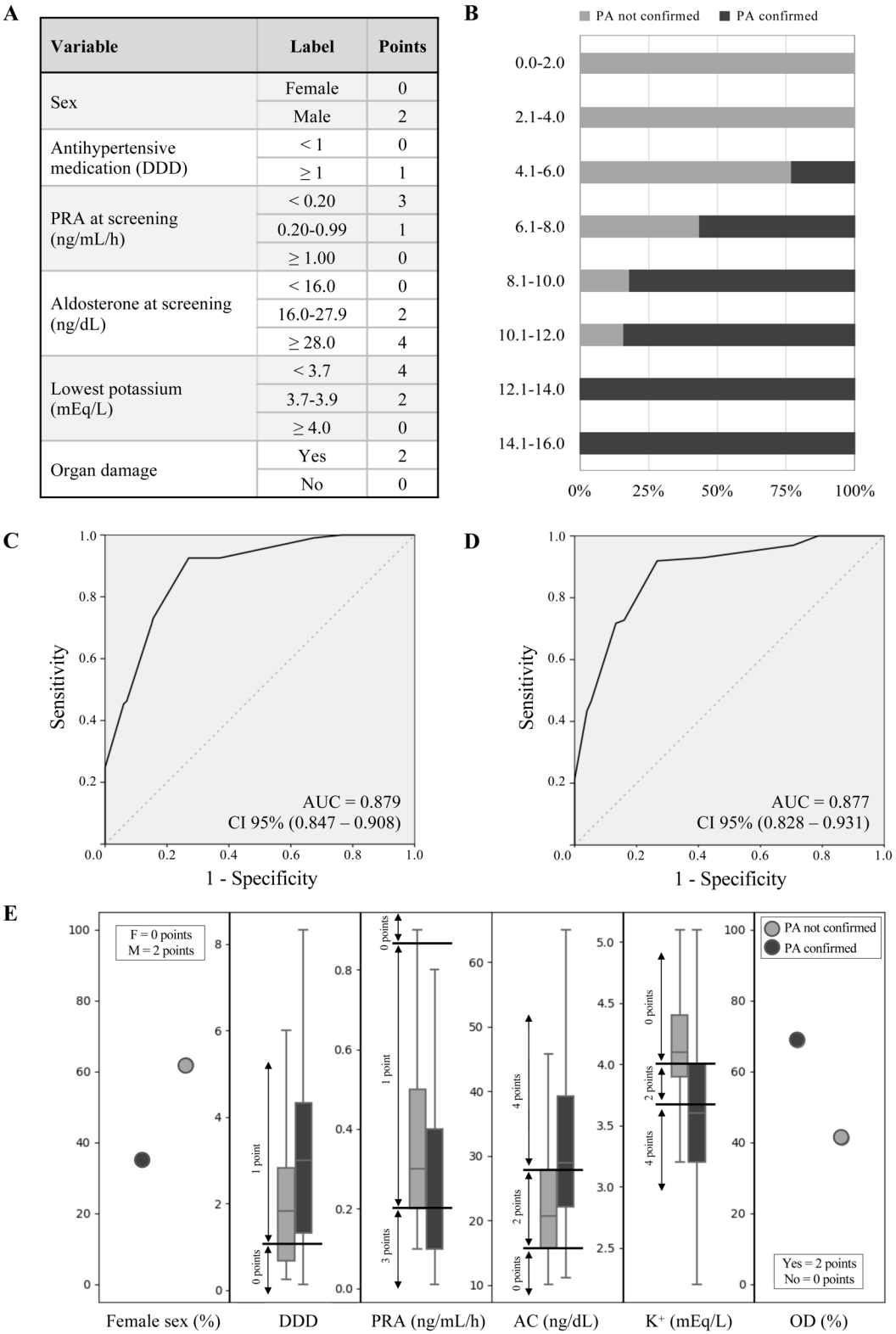
1. Funder JW, Carey RM, Mantero F, et al. The Management of Primary Aldosteronism: Case Detection, Diagnosis, and Treatment: An ES Clinical Practice Guideline. *J Clin Endocrinol Metab.* 2016;101:1889-916.
2. Mulatero P, Monticone S, Deinum J, et al. Genetics, prevalence, screening and confirmation of primary aldosteronism: a position statement and consensus of the Working Group on Endocrine Hypertension of The European Society of Hypertension. *J Hypertens* 2020;38:1919-28.
3. Monticone S, Burrello J, Tizzani D, et al. Prevalence and Clinical Manifestations of Primary Aldosteronism Encountered in Primary Care Practice. *J Am Coll Cardiol.* 2017;69:1811-20.
4. Nishikawa T, Omura M, Satoh F, et al; Task Force Committee on Primary Aldosteronism, The Japan Endocrine Society. Guidelines for the diagnosis and treatment of primary aldosteronism--the Japan Endocrine Society 2009. *Endocr J.* 2011;58:711-21.
5. Burrello J, Amongero A, Buffolo F, et al. Data from: Development and validation of a prediction score to avoid confirmatory testing in patients with suspected primary aldosteronism. *J Clin Endocrinol Metab* 2020. https://github.com/CentroIpertenUnito/PACT-score. Deposited 15 December 2020.
6. Mulatero P, Sechi LA, Williams TA, et al. Subtype diagnosis, treatment, complications and outcomes of primary aldosteronism and future direction of research: a position statement and consensus of the Working Group on Endocrine Hypertension of the European Society of Hypertension. *J Hypertens.* 2020;38:1929-36.
7. Burrello J, Burrello A, Stowasser M, et al. The Primary Aldosteronism Surgical Outcome Score for the Prediction of Clinical Outcomes After Adrenalectomy for Unilateral Primary Aldosteronism. *Ann Surg.* 2020;272:1125-32.
8. Ahmed AH, Cowley D, Wolley M, et al. Seated saline suppression testing for the diagnosis of primary aldosteronism: a preliminary study. *J Clin Endocrinol Metab.* 2014;99:2745-53.

9. Olsen MH, Angell SY, Asma S, et al. A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on hypertension. *Lancet.* 2016;388:2665-12.
10. Heinrich DA, Adolf C, Quinkler M, et al. Safety of medical adjustment and confirmatory testing in the diagnostic work-up of primary aldosteronism. *Eur J Endocrinol.* 2019;181:421-8.
11. Mulatero P, Monticone S, Burrello J, Veglio F, Williams TA, Funder J. Guidelines for primary aldosteronism: uptake by primary care physicians in Europe. *J Hypertens.* 2016;34:2253-7.
12. Lubitz CC, Economopoulos KP, Sy S, et al. Cost-Effectiveness of Screening for Primary Aldosteronism and Subtype Diagnosis in the Resistant Hypertensive Patients. *Circ Cardiovasc Qual Outcomes.* 2015;8:621-30.
13. Nanba K, Tamanaha T, Nakao K, et al. Confirmatory testing in primary aldosteronism. *J Clin Endocrinol Metab.* 2012;97:1688-94.
14. Maiolino G, Rossitto G, Bisogni V, et al; PAPY Study Investigators. Quantitative Value of Aldosterone-Renin Ratio for Detection of Aldosterone-Producing Adenoma: The Aldosterone-Renin Ratio for Primary Aldosteronism (AQUARR) Study. *J Am Heart Assoc.* 2017;6:e005574.
15. Reznik Y, Amar L, Tabarin A. SFE/SFHTA/AFCE consensus on primary aldosteronism, part 3: Confirmatory testing. *Ann Endocrinol (Paris).* 2016;77:202-7.
16. Vivien M, Deberles E, Morello R, Haddouche A, Guenet D, Reznik Y. Evaluation of Biochemical Conditions Allowing Bypass of Confirmatory Testing in The Workup of Primary Aldosteronism: A Retrospective Study in a French Hypertensive Population. *Horm Metab Res.* 2019;51:172-177.
17. Umakoshi H, Sakamoto R, Matsuda Y, et al. Role of Aldosterone and Potassium Levels in Sparing Confirmatory Tests in Primary Aldosteronism. *J Clin Endocrinol Metab.* 2020;105:dgz148.

18. Kawashima J, Araki E, Naruse M, et al. Baseline Plasma Aldosterone Level and Renin Activity Allowing Omission of Confirmatory Testing in Primary Aldosteronism. *J Clin Endocrinol Metab.* 2020;105:dgaa117.
19. Kietsiriroje N, Wonghirundecha R, Suntornlohanakul O, Murray RD. Construction of a predictive scoring system as a guide to screening and confirmation of the diagnosis of primary aldosteronism. *Clin Endocrinol (Oxf).* 2020;92:196-205.
20. Burrello J, Monticone S, Losano I, et al. Prevalence of Hypokalemia and Primary Aldosteronism in 5100 Patients Referred to a Tertiary Hypertension Unit. *Hypertension.* 2020;75:1025-33.
21. Pizzolo F, Raffaelli R, Memmo A, et al. Effects of female sex hormones and contraceptive pill on the diagnostic work-up for primary aldosteronism. *J Hypertens.* 2010;28:135-42.
22. Ahmed AH, Gordon RD, Taylor PJ, Ward G, Pimenta E, Stowasser M. Are women more at risk of false-positive primary aldosteronism screening and unnecessary suppression testing than men? *J Clin Endocrinol Metab.* 2011;96:E340-6.

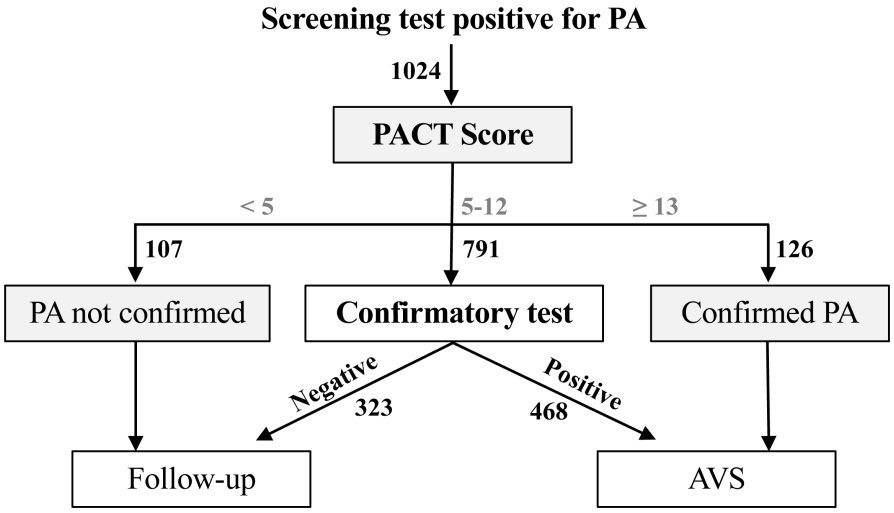
A

Variable	Label	Points
Sex	Female	0
	Male	2
Antihypertensive medication (DDD)	< 1	0
	≥ 1	1
PRA at screening (ng/mL/h)	< 0.20	3
	0.20-0.99	1
	≥ 1.00	0
Aldosterone at screening (ng/dL)	< 16.0	0
	16.0-27.9	2
	≥ 28.0	4
Lowest potassium (mEq/L)	< 3.7	4
	3.7-3.9	2
	≥ 4.0	0
Organ damage	Yes	2
	No	0



Legend to Figure 1 – **Development of the PACT score.** Univariate/multivariate regression analyses were used to assign points to each variable according to stratification level. The score was developed in the training cohort (n=522) and tested on the internal validation cohort from Torino (n=174). Data on training and validation of the score are reported in Table S7. (**A**) The table reports included variables and scoring-point system. If only direct renin concentration (DRC) is available, the following cut-offs could be used: DRC < 2.5 mU/L (3 points); DRC 2.5-12.3 mU/L (1 point); DRC ≥ 12.4 mU/L (0 point). (**B**) Histogram showing the proportion of patients (x-axis, %) for each diagnosis (PA confirmed, black; PA not confirmed, grey), stratified by score points (y-axis) on the

developmental cohort (n=696). The total number of patients (N) for each score level and their proportion (%) are reported in Table S9. (**C**, **D**) Receiver operating characteristics (ROC) curve to assess the area under the curve (AUC) in the training (n=522; left) and internal validation cohort from Torino (n=174; right). (**E**) Representation of variable categorization and assigned points (PA confirmed, black; PA not confirmed, grey); the bars indicate median and interquartile range.



Legend to Figure 2 – **Management of patients with suspected PA.** Flow chart for the management of patients with a positive screening test according to the PACT score (Developmental Cohort + External Validation Cohort; n=1,024). The number of patients is indicated in bold; cut-offs are indicated in grey. AVS, Adrenal Venous Sampling; PA, Primary Aldosteronism; PACT, Primary Aldosteronism Confirmatory Testing Score.

Table 1. Patient Characteristics

Variable	Confirmatory / Exclusion Test		*P*-value
	PA confirmed (n=421)	PA not confirmed (n=275)	
Age at diagnosis (years)	50 ± 10.2	51 ± 9.5	0.202
Female sex, n (%)	148 (35.2)	170 (61.8)	**<0.001**
Duration of HTN (months)	68 [24; 135]	54 [18; 125]	0.084
Systolic BP (mmHg)	157 ± 20.7	152 ± 19.4	**0.003**
Diastolic BP (mmHg)	95 ± 11.2	94 ± 10.5	0.051
Antihypertensive medication (DDD)	3.00 [1.33; 4.33]	1.83 [0.67; 2.83]	**<0.001**
BMI (Kg/sqm)	25.9 ± 4.48	25.5 ± 3.95	0.204
PRA at screening (ng/mL/h)	0.20 [0.10; 0.40]	0.30 [0.20; 0.50]	**<0.001**
Aldosterone at screening (ng/dL)	28.9 [22.1; 39.3]	20.7 [15.7; 27.9]	**<0.001**
Lowest Potassium (mEq/L)	3.6 ± 0.63	4.1 ± 0.41	**<0.001**
eGFR (mL/min)	91 ± 17.1	91 ± 16.9	0.666
Diabetes, n (%)	31 (7.4)	13 (4.7)	0.162
Organ damage, n (%)	290 (68.9)	114 (41.5)	**<0.001**
CV events, n (%)	61 (14.5)	20 (7.3)	**0.004**

Characteristics of patients included in the analysis: confirmed diagnosis of PA (n=421) vs. PA not confirmed (n=275). HTN, Hypertension; BP, Blood Pressure; DDD, Defined Daily Dose (average maintenance dose per day for a drug used for its main indication in adults); PRA, Plasma Renin Activity; eGFR, estimated Glomerular Filtration Rate; CV, Cardiovascular. Organ damage is defined as presence of left ventricular hypertrophy at echocardiography and/or microalbuminuria. Normally and non-normally distributed variables were reported as mean ± standard deviation or median [interquartile range], respectively. Categorical variables were reported as absolute number (n) and proportion (%).

Table 2. Regression analysis on discriminant for PA diagnosis

Variable (ref. PA confirmed)	Univariate analysis		Multivariate Analysis	
	OR (CI 95%)	*P*-value	OR (CI 95%)	*P*-value
Female sex, n (%)	0.34 (0.24-0.46)	**<0.001**	0.41 (0.28-0.62)	**<0.001**
Antihypertensive medication (DDD)	1.40 (1.27-1.54)	**<0.001**	1.20 (1.07-1.35)	**0.002**
PRA at screening (ng/mL/h)	0.28 (0.14-0.57)	**<0.001**	0.07 (0.03-0.20)	**<0.001**
Aldosterone at screening (ng/dL)	1.06 (1.04-1.08)	**<0.001**	1.08 (1.06-1.10)	**<0.001**
Lowest Potassium (mEq/L)	0.13 (0.09-0.19)	**<0.001**	0.14 (0.09-0.23)	**<0.001**
Organ damage, n (%)	3.13 (2.28-4.29)	**<0.001**	2.63 (1.75-3.95)	**<0.001**

Odds ratio (OR) and the 95% confidence interval (CI) were evaluated by univariate and multivariate logistic regression analysis, as indicated. An OR greater than 1 indicates an increased likelihood of confirmed PA, and an OR less than 1 a decreased likelihood (i.e., an OR of 1.06 for aldosterone levels means an increase of 6% in the likelihood of confirmed PA, for each 1 ng/dL increase of aldosterone; an OR of 0.07 for PRA means an increase of 43% in the likelihood of confirmed PA, for each 0.1 ng/mL/h decrease in PRA). Antihypertensive medication (expressed as DDD), PRA and aldosterone at screening, and lowest potassium were treated as continuous variables; sex and organ damage were treated as categorical variables.