

Sincronie e asincronie negli sviluppi della pandemia di COVID-19 sul territorio

Synchronies and asynchronies in the development of COVID-19 pandemic in Italy

Maria Teresa Giraudo^{1,2}, Manuele Falcone^{2,3}, Cesare Cislighi^{2,4}, Francesca Cordero⁵, Simone Pernice⁵, Roberta Sirovich⁶

¹ Dipartimento di Scienze Cliniche e Biologiche, Università di Torino, Torino

² Gruppo MADE, Associazione Italiana di Epidemiologia

³ Azienda Regionale di Sanità della Toscana, Firenze

⁴ Epieconomista, Roma

⁵ Dipartimento di Informatica, Università di Torino, Torino

⁶ Dipartimento di Matematica "Giuseppe Peano", Università di Torino, Torino

Corrispondenza: Maria Teresa Giraudo; mariateresa.giraudo@unito.it

Riassunto

Introduzione: lo studio dei possibili determinanti dei periodi di crescita e decrescita dei contagi può essere di grande rilevanza, come si è sperimentato durante la pandemia di COVID-19. Uno dei metodi per comprendere se i determinanti siano contemporanei o si sviluppino per contiguità tra aree diverse è lo studio dell'indice di replicazione diagnostica RD_t tra le Regioni.

Obiettivi: presentare l'analisi della variabilità dell' RD_t e la successiva applicazione di un metodo di *clustering funzionale* recentemente introdotto come procedimenti di grande utilità per riconoscere la presenza di raggruppamenti con trend analogo nelle curve epidemiche.

Disegno: all'interno del periodo considerato, si analizzano in dettaglio gli andamenti degli RD_t regionali in quattro diversi intervalli di tempo.

Setting e partecipanti: per esemplificare questa metodologia, può essere di interesse lo studio della variabilità nel periodo tra la fine del 2021 e l'inizio del 2022.

Principali misure di outcome: la variabilità negli indici RD_t regionali è valutata tramite il coefficiente di correlazione pesato rispetto alle popolazioni delle singole Regioni. Il procedimento di *clustering* è applicato alle serie temporali dei valori assoluti dell' RD_t .

Risultati: i periodi di aumento della variabilità negli RD_t corrispondono alle fasi di crescita o decrescita iniziali dei contagi, mentre il *clustering funzionale* individua le macroaree in cui le curve epidemiche hanno avuto andamenti simili. Ciò che ha fatto crescere i contagi sembra riferito a un fattore non specifico di alcune aree, con l'apporto in alcuni casi di una dinamica di contagiosità tra aree adiacenti.

Conclusioni: la variabilità nell'andamento degli indici di replicazione diagnostica regionali, che vengono calcolati con pochissimi giorni di ritardo, è un ulteriore indicatore per l'*early detection* delle fasi di modifiche importanti nell'andamento delle curve epidemiche. La clusterizzazione delle curve degli indici epidemici può essere utile per stabilire se i determinanti agiscano contemporaneamente o per contiguità tra aree adiacenti.

Parole chiave: indice di riproduzione diagnostica, coefficiente di variazione, *clustering funzionale*, diffusione del contagio, determinante

Abstract

Background: the study of the possible determinants of the rise and fall of infections can be of great relevance, as

Cosa si sapeva già

■ L'andamento dell'incidenza nazionale è una misura assoluta dell'evoluzione dell'epidemia, ma non consente di rilevare i cambiamenti nelle dinamiche di trasmissione del contagio; è quindi utile considerare in modo comparativo gli andamenti regionali.

■ Lo studio dei possibili determinanti dei periodi di crescita e decrescita dei contagi è normalmente condotto in termini virologici, prendendo in considerazione l'insorgenza di possibili nuove varianti, o immunologici, considerando che la suscettibilità nella popolazione si modifica con il progredire dell'epidemia e il contagio di parte degli individui a rischio.

■ Nella prima parte dell'autunno del 2021 e, con minore enfasi, nel primo periodo del 2022 l'andamento dell'epidemia di COVID-19 in Italia sembrava avviata verso la remissione, ma il rinnovato incremento dei nuovi positivi immediatamente successivo ha richiesto di continuare il monitoraggio e di innalzare il livello di attenzione nei confronti del contagio.

Cosa si aggiunge di nuovo

■ L'indice di replicazione diagnostica RD_t consente di monitorare in tempo reale l'andamento dell'epidemia anche a livello delle singole Regioni; l'analisi dettagliata della sua variabilità regionale permette di individuare omogeneità o differenziazioni nella diffusione dei contagi in uno specifico periodo temporale.

■ Con il metodo di *clustering funzionale* si possono individuare raggruppamenti nei *pattern* di variabilità degli indici RD_t regionali, in particolare nelle fasi di innalzamento o di decrescita dell'incidenza a livello nazionale, e si può ipotizzare se lo sviluppo dell'epidemia sia simultaneo sul territorio o sia dilazionato per contiguità tra aree diverse.

■ La rilettura della curva epidemica nelle Regioni tra settembre 2021 e aprile 2022 in termini di variabilità dell'indice RD_t consente di mettere in corrispondenza la diffusione territoriale con gli eventi concomitanti e con le misure di contenimento e di ipotizzarne i determinanti principali; lo stesso metodo potrebbe essere impiegato in altri contesti.

was experienced during the COVID-19 pandemic. One of the methods to understand whether determinants are simultaneous or develop through contiguity between different areas is the study of the diagnostic replication index RD_t among regions.

Objectives: to introduce the analysis of RD_t variability and the subsequent application of a recently introduced functional clustering method as highly useful procedures for rec-

RASSEGNE E ARTICOLI

ognizing the presence of clusters with similar trends in epidemic curves.

Design: within the considered period, trends in regional RD_t are analyzed in detail over four different time intervals.

Setting and participants: to exemplify this methodology, the study of variability in the period from the end of 2021 to the beginning of 2022 may be of interest.

Main outcomes measures: the variability in the regional RD_t indices is assessed by means of the correlation coefficient weighted with respect to the populations of the individual regions. The clustering procedure is applied to the time series of absolute RD_t values.

Results: it emerges that the periods of increasing variability in the RD_t correspond to the initial growth or decrease in the number of infections, while functional clustering iden-

tifies macro-areas in which the epidemic curves have had similar trends. What caused contagions to increase seems to relate to a factor that is not specific to certain areas, with the contribution in some cases of a contagion dynamic between adjacent areas.

Conclusions: the variability in the trend of regional diagnostic replication indices, which are calculated with only a few days delay, is a further indicator for the early detection of major changes in the trend of epidemic curves. The clustering of epidemic index curves may be useful to determine whether determinants act simultaneously or by contiguity between adjacent areas.

Keywords: diagnostic reproduction index, coefficient of variation, functional clustering, infection spread, determinant

Introduzione

La pura analisi dello sviluppo temporale degli indici di riproduttività di un'epidemia come quella di COVID-19, anche se condotta separatamente per ogni Regione, non consente di individuare eventuali discrepanze nell'accelerazione o decelerazione dei contagi in aree diverse sul territorio nazionale. La ricerca di corrispondenze tra discrepanze e aggregazioni a livello geografico o eventi a livello cronologico potrebbe suggerire l'emergere di *pattern* di determinanti ed essere utile per dare un ulteriore contributo alla comprensione dei fenomeni epidemici. L'individuazione di omogeneità o disomogeneità nelle variazioni dell' RD_t tra la fine del 2021 e l'inizio del 2022 può aiutare a fornire una risposta alla domanda se l'andamento non monotono delle curve epidemiche regionali sia da attribuirsi a modifiche nelle misure di contenimento adottate o alla diffusione di ulteriori varianti del virus. In questo modo, si disporrebbe di un ulteriore strumento per capire se i determinanti siano stati contemporanei o si siano sviluppati per contiguità tra aree diverse.

L'utilizzo della variabilità dell'indice RD_t a livello regionale richiede metodi specifici per la sua valutazione, vista la grande disparità di distribuzione della popolazione tra le unità in esame. In questo articolo sono stati, quindi, calcolati come indicatori medie, deviazioni standard e coefficienti di variazione pesati per l'indice RD_t lungo tutto l'intervallo di tempo considerato.

Dopo l'individuazione di periodi in cui analizzare in dettaglio gli andamenti dell'indice, è stato utilizzato un metodo innovativo di *clustering funzionale* in grado di individuare raggruppamenti di curve che condividono caratteristiche simili in un insieme più vasto di curve della stessa tipologia. Il confronto dei rispettivi trend di incidenza ha consentito, infine, di ipotizzare le dinamiche di trasmissione.

Obiettivi

Obiettivo del lavoro è di utilizzare la tecnica di indagine proposta per analizzare la variabilità inter-regionale dell'indice RD_t e la sua relazione con le curve epidemiche regionali nei periodi delle ondate di COVID-19 tra l'estate del 2021 e i primi mesi del 2022 per distinguere eventuali macroaree dove le dinamiche si sono differenziate. Sono state poste due ipotesi: se le accelerazioni o decelerazioni risultano contemporanee, si può ipotizzare una dinamica dovuta essenzialmente ai comportamenti adottati nel Paese, che in quel periodo sono stati piuttosto uniformi; se si identificano dei *cluster* molto specifici, si può pensare alla possibile diffusione per contagio di nuove varianti.

L'indice di replicazione diagnostica RD_t è stato introdotto in un contributo precedente¹ e valutato e reso disponibile a cadenza giornaliera sul sito MADE (Monitoraggio e Analisi dei Dati dell'Epidemia) a cura dell'Associazione Italiana di Epidemiologia.² È il rapporto tra la media mobile centrata non pesata dei valori settimanali di incidenza e la stessa media mobile valutata per la settimana precedente. Rispetto all'indice di riproduzione R_t , di cui può essere considerato una versione semplificata anche dal punto di vista computazionale, l' RD_t presenta il vantaggio di poter essere valutato in tempo quasi reale. Mentre l' R_t utilizza la data dell'inizio dei sintomi per la registrazione dei nuovi casi, l' RD_t utilizza la data di registrazione della positività. I due indici assumono solitamente valori molto simili e le discrepanze che si osservano non conducono perlopiù a diversità interpretative del fenomeno (per ulteriori dettagli, si veda il contributo precedentemente citato¹). L'indice RD_t fornisce un'indicazione dell'accelerazione o decelerazione della diffusione dei contagi, segnalate da valori rispettivamente superiori o inferiori a 1.

Metodi

Lo studio è stato condotto sulla base dei dati resi pubblicamente disponibili sul sito della Protezione Civile per tutto il periodo dal 02.03.2020 al 28.10.2022. Il flusso di dati, comunicati giornalmente dalle singole Regioni e dalle province autonome di Trento e Bolzano, comprendeva il numero di nuovi contagi nelle 24 ore precedenti, a cui si aggiungevano i ricoveri, le guarigioni, i decessi, il numero di tamponi, il numero delle persone in isolamento domiciliare e la prevalenza di casi. Quest'ultima veniva fornita anche per le singole province. Come considerato in dettaglio in Giraud et al.,¹ i dati di incidenza e prevalenza così ottenuti risentono di una certa approssimazione sia per l'inclusione di nuovi casi solo a seguito di un tampone ufficialmente registrato sia per l'utilizzo della data di diagnosi e non di contagio per la determinazione degli indici epidemici. Sebbene i dati disponibili presentino un livello non ottimale di qualità dovuta anche alla gestione dell'operatività durante il corso del fenomeno epidemico, la loro validità a livello informativo e interpretativo è indubbia.

L'andamento della curva epidemica in Italia e dell'indice RD_t dal settembre 2021 al marzo 2022 sono rappresentati in figura 1. Sono riportati sia i valori giornalieri (in blu) sia le medie mobili settimanali (in rosso). Per tutte le analisi effettuate si è utilizzato il software R versione 4.2.3 rilasciata il 15.03.2023.³ Per analizzare le differenze nello sviluppo dell'epidemia a seconda della Regione, è stata considerata la variabilità negli indici RD_t specifici.

A questo scopo, è stata calcolata dapprima la media aritmetica pesata, utilizzata quando i valori considerati non contribuiscono tutti allo stesso modo alla media finale, ed è stato assegnato loro un peso commisurato alla rilevanza relativa. In generale, la media pesata \bar{x} (o \bar{x}_{wtd}) di n valori x_i , ognuno con un peso w_i , è

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

Gli indici RD_t possono essere influenzati dalla numerosità della popolazione per cui si valutano, dal momento che per popolazioni di piccole dimensioni subiscono oscillazioni molto più ampie. Quindi, sono stati considerati come pesi le numerosità delle popolazioni delle singole Regioni tratte dai dati del Censimento Istat 2021.⁴ Anche per calcolare la varianza è necessario tenere conto dell'utilizzo di medie pesate in luogo di medie ordinarie. Così facendo, si ovvia alla perdita di gradi di libertà indotta dalla presenza di alcuni pesi preponderanti introducendo il valore "numero effettivo di misurazioni" n_{eff} :

$$n_{eff} = \frac{(\sum_{i=1}^n w_i)^2}{\sum_{i=1}^n w_i^2}$$

La varianza pesata $Var(x)_{wtd}$ è

$$Var(x)_{wtd} = (s_x^2)_{wtd} = \frac{\sum_{i=1}^n w_i (x_i - \bar{x}_{wtd})^2}{\sum_{i=1}^n w_i} \frac{n_{eff}}{n_{eff} - 1}$$

$$= \left(\frac{\sum_{i=1}^n w_i x_i^2}{\sum_{i=1}^n w_i} - (\bar{x}_{wtd})^2 \right) \frac{(\sum_{i=1}^n w_i)^2}{(\sum_{i=1}^n w_i)^2 - (\sum_{i=1}^n w_i^2)}$$

Il coefficiente di variazione pesato CV_{wtd} è il rapporto tra la deviazione standard pesata $(s_x)_{wtd}$ e la media pesata.

Le curve giornaliere degli RD_t regionali sono state poi clusterizzate con una procedura di *clustering funzionale*, sviluppata per il raggruppamento di dati che non rappresentano un punto in uno spazio anche di grande dimensione, ma sono costituiti dal campionamento a istanti discreti di una funzione del tempo. Il metodo utilizzato è stato proposto da James et al.⁵ e successivamente incluso nel tool CONNECTOR, arricchito di strumenti per la visualizzazione, l'ottimizzazione dei parametri liberi e la valutazione della qualità del *clustering* finale in Pernice et al.⁶

Il *clustering funzionale* proposto da James et al.⁵ è particolarmente efficace quando le curve sono campionate in modo irregolare e sparso, ovvero quando gli istanti di campionamento sono pochi e non permettono l'utilizzo di tecniche di *filtering* o *regolarizzazione*. Il metodo è *model-based*: le curve vengono rappresentate usando una base di *spline* cubiche naturali i cui coefficienti vengono ottenuti con un modello a effetti aleatori che include un termine di appartenenza al *cluster* (incognito). Il modello viene stimato mediante un algoritmo *expectation-maximization* che tratta i parametri di appartenenza al *cluster* come mancanti. Vengono, quindi, stimati la curva per tutti gli istanti dell'intervallo temporale analizzato e i relativi intervalli di confidenza e di predizione. CONNECTOR guida l'utente nella selezione dei parametri liberi del modello, ovvero la dimensione della base di *spline* e il numero di *cluster*, tramite l'introduzione di procedure dedicate. In particolare, la dimensione della base di *spline* viene fissata con l'osservazione della *cross-validated likelihood*. Il numero di *cluster* viene scelto usando tre indici. Il primo è un indice introdotto in Pernice et al.⁶ con il nome di *functional DB index*, che è l'estensione funzionale dell'indice DB (Davies-Bouldin), che misura la bontà del *clustering* privilegiando gruppi di curve vicine tra di loro

RASSEGNE E ARTICOLI

e con centri lontani dai centri degli altri *cluster*. Effettua, inoltre, la stima dei parametri del modello di *clustering* e restituisce la visualizzazione delle curve raggruppate per *cluster* di appartenenza. Il secondo è un indice di *total tightness*, che misura la somma delle distanze delle curve dai centri dei *cluster*. Il terzo è un indice di stabilità della clusterizzazione finale, che dipende dall'inizializzazione dell'algoritmo di *fitting*.

L'applicazione di questo metodo consente di suddividere le curve degli RD_t regionali in gruppi distinti, all'interno dei quali le dinamiche della variabilità nelle curve epidemiche sono simili.

Risultati

L'andamento delle medie pesate giornaliere degli RD_t regionali e dei coefficienti di variazione è riportato in figura 2.

I periodi durante i quali la media pesata degli RD_t registra un incremento o un decremento rilevante corrispondono alle fasi di accelerazione o decelerazione dell'estensione dei contagi. Si sono verificati nel periodo intorno alla fine del 2021 e all'inizio del 2022, in concomitanza con l'inizio e la conclusione delle festività natalizie. Rivestono un interesse particolare, però, le fasi di innalzamento e successiva decrescita del coefficiente di variazione degli indici, in quanto denotano una disomogeneità tra le Regioni.

Per esemplificare l'applicazione del metodo di *clustering*, sono stati presi in considerazione tre periodi temporali:

- i giorni dal 16 al 25 dicembre 2021;
- i giorni dal 26 dicembre 2021 al 4 gennaio 2022, a cavallo di Capodanno;
- i giorni dal 12 al 18 gennaio 2022.

L'analisi di *clustering funzionale* degli RD_t regionali per il primo periodo è riportata nella figura 3. In figura 3A sono rappresentati gli andamenti dei singoli indici per i 4 *cluster* risultanti dalla suddivisione ottimale, mentre in figura 3B le Regioni sulla mappa dell'Italia sono evidenziate con colori diversi a seconda della loro appartenenza ai *cluster*. I primi due *cluster* (A e B) corrispondono a *plateau* negli andamenti dell' RD_t , mentre per gli altri due si riscontra un andamento crescente. La prima situazione è verificata in quasi tutte le Regioni della zona adriatica, dalle province autonome di Trento e Bolzano, dalle isole e da due Regioni del Sud. Le Regioni più estese che appartengono alla seconda coppia di *cluster* sono Piemonte, Lombardia e Toscana.

La caratterizzazione dei *cluster* individuati corrisponde ad andamenti specifici nelle incidenze corrispondenti. Si rilevano, in particolare, situazioni di stallo nella prima coppia di *cluster* e picchi verso il termine del periodo per la seconda.

L'analisi relativa al secondo periodo è riportata nella

figura 4. Si individuano 3 *cluster* (figura 4A), evidenziati con colori diversi sulla mappa riportata in figura 4B. Il primo *cluster* raggruppa sostanzialmente le stesse Regioni della zona adriatica già individuate per il periodo precedente e corrisponde a un andamento a U rovesciata degli RD_t . Al secondo e al terzo *cluster* appartengono le Regioni per cui l' RD_t , pur avendo un andamento di crescita-decrescita simile al precedente, raggiunge un valore di picco più alto o molto più alto. In questi due gruppi, il valore massimo dell'indice viene raggiunto leggermente in ritardo rispetto alle altre Regioni. Soprattutto negli ultimi due casi, l'andamento corrisponde a un picco nelle incidenze verso la fine dell'anno, posticipato nel secondo rispetto al primo. I rapporti tra i valori delle incidenze nei periodi di crescita e di decrescita risultano più elevati quando la curva rappresentativa dell'andamento del *cluster* assume valori maggiori.

La figura 5 riporta l'analisi effettuata per il terzo periodo, per cui sono stati individuati 4 *cluster*. Il primo comprende la Valle d'Aosta, la Basilicata, le isole e la fascia centrale di Regioni dalla Lombardia al Lazio, a eccezione delle Marche. In questo caso, l'andamento dell' RD_t è stato abbastanza stabile, non discostandosi troppo dal valore 1. Al secondo gruppo appartengono il Piemonte, il Nord-Est e la Calabria, con un andamento leggermente decrescente dell'indice. In Liguria, Campania e Molise i valori dell' RD_t sono risultati decrescenti, mentre in Marche e Puglia l'indice, prima in crescita, ha iniziato a decrescere solo al termine del periodo.

Per quanto riguarda le incidenze, le curve rappresentative dei *cluster* rivelano una chiara decrescita per il terzo gruppo fino al giorno 17 e una crescita fino al giorno 15 per il quarto gruppo, seguita da una diminuzione repentina.

Discussione

La tecnica di *clustering funzionale* consente di individuare *cluster* nei trend di quantità monitorate a istanti di tempo discreti anche per intervalli di breve durata. L'uso di questa metodologia per gli indici di replicazione diagnostica può consentire di far luce sulla modalità di propagazione dei contagi in atto. Una dinamica abbastanza omogenea tra le Regioni, che corrisponde a una diffusione sostanzialmente contemporanea dei contagi, riconduce a un ruolo preponderante dell'introduzione di norme di contenimento a livello nazionale. Un andamento molto differenziato degli indici tra zone ricondurrebbe, invece, a un'alterazione differenziata nella contagiosità del virus, dovuta verosimilmente all'emergere di nuovi varianti. Un trend casuale, con incrementi e decrementi istantanei che si verificano ovunque, porta a pensare alla sovrapposizione di fattori di vario genere riconducibili a matri-

RASSEGNE E ARTICOLI

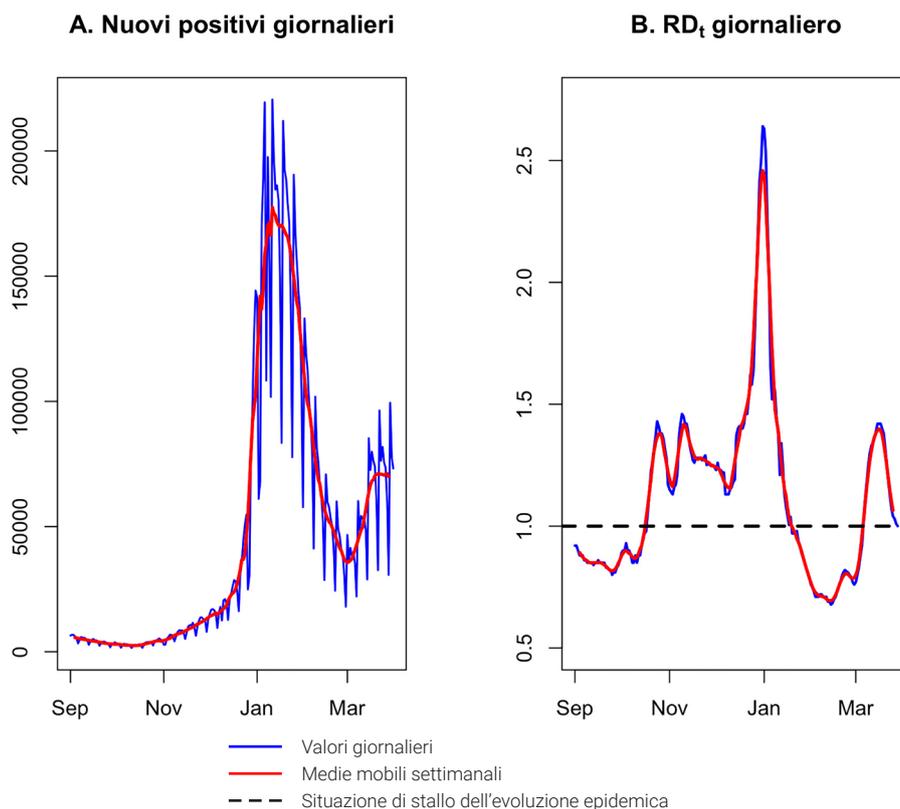


Figura 1. (A) Andamento dei contagi giornalieri e (B) andamento dell'indice di replicazione diagnostica RD_t , con la rispettiva media mobile settimanale nel periodo da settembre 2021 a marzo 2022.

Figure 1. (A) Trend of daily infections and (B) trend of the diagnostic replication index RD_t , with the respective weekly moving average from September 2021 to March 2022.

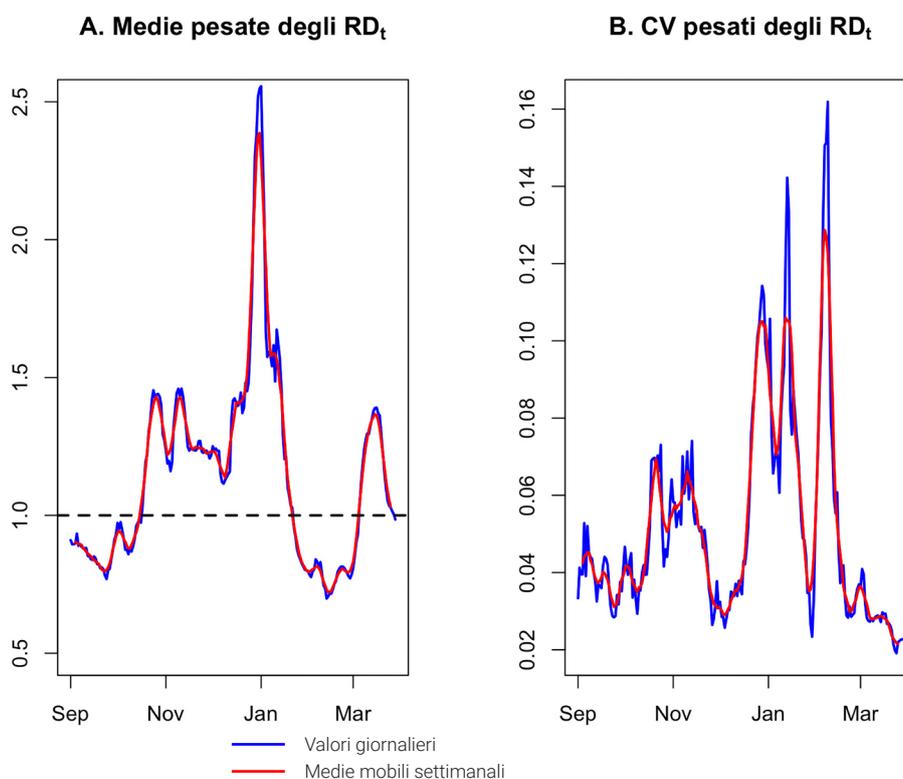


Figura 2. (A) Medie pesate giornaliere e (B) coefficienti di variazione pesati giornalieri degli indici di replicazione diagnostica RD_t regionali nel periodo da settembre 2021 a marzo 2022.

Figure 2. (A) Daily weighted averages and (B) daily weighted coefficients of variation of the regional diagnostic replication indices RD_t from September 2021 to March 2022.

RASSEGNE E ARTICOLI

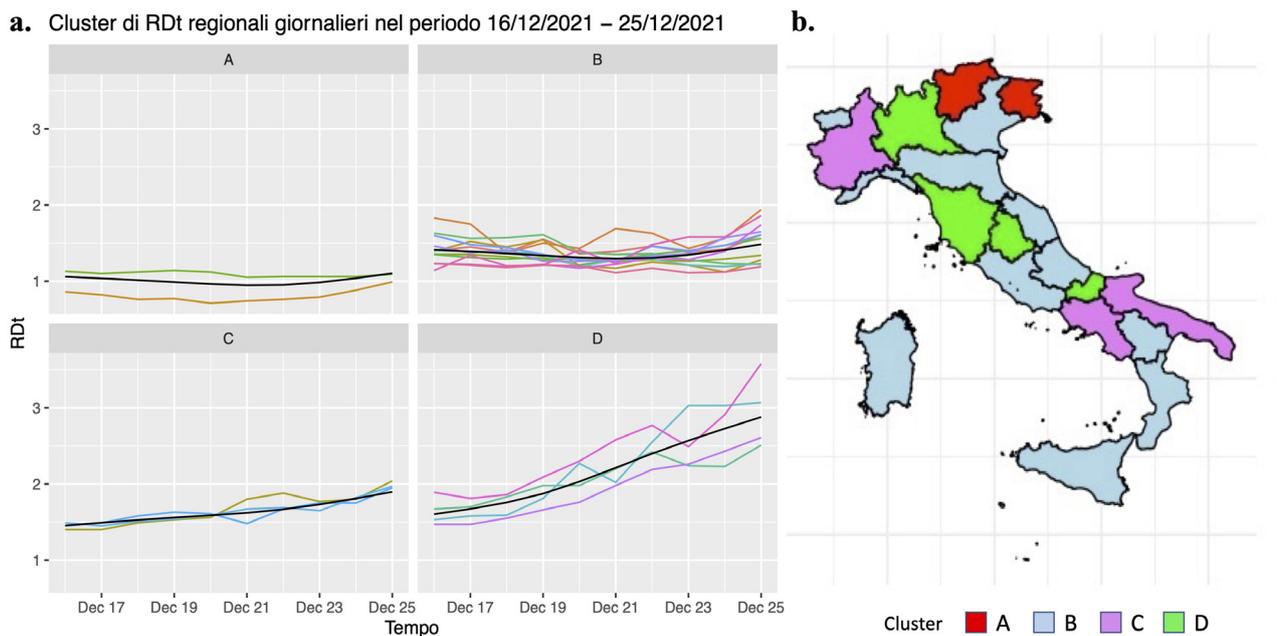


Figura 3. (A) Andamento degli indici RDt per i cluster di Regioni risultanti dalla suddivisione ottimale (una curva per ogni Regione) e (B) mappa geografica dei cluster nel periodo dal 16 al 25 dicembre 2021.
Figure 3. (A) Trend of the RDt indices for the clusters of regions resulting from an optimal subdivision (one curve for each region) and (B) geographical map of the clusters arising from 16 to 25 December 2021.

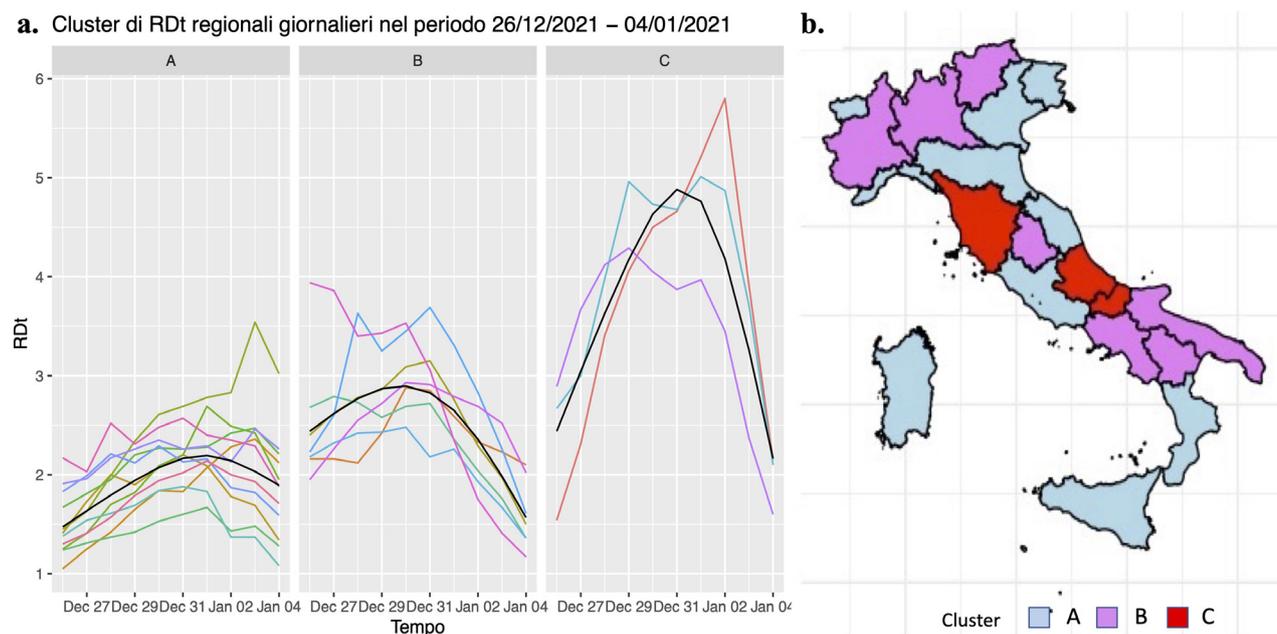


Figura 4. (A) Andamento degli indici RDt per i cluster di Regioni risultanti dalla suddivisione ottimale (una curva per ogni Regione) e (B) mappa geografica dei cluster nel periodo dal 26 dicembre 2021 al 4 gennaio 2022.
Figure 4. (A) Trend of the RDt indices for the clusters of regions resulting from an optimal subdivision (one curve for each region) and (B) geographical map of the clusters arising from 26 December 2021 to 4 January 2022.

ci comuni quali flussi rilevanti di spostamenti inter-regionali. Tra la fine del 2021 e l'inizio del 2022, l'analisi degli RD_t non ha consentito di riscontrare forti differenziazioni tra le Regioni. Nel periodo immediatamente precedente il Natale, sono stati individuati due tipi di cluster: uno corrispondente a un andamento costan-

te, un altro a un andamento crescente dell'RD_t. Per i giorni a cavallo di Capodanno 2022, un primo cluster presenta un andamento a U rovesciata degli RD_t, mantenuto negli altri due, ma con picchi più rilevanti e leggermente posticipati. Nel terzo periodo, all'incirca la seconda decina di giorni dell'anno, i valori dell'RD_t hanno oscillato intorno al valore 1 per mol-

RASSEGNE E ARTICOLI

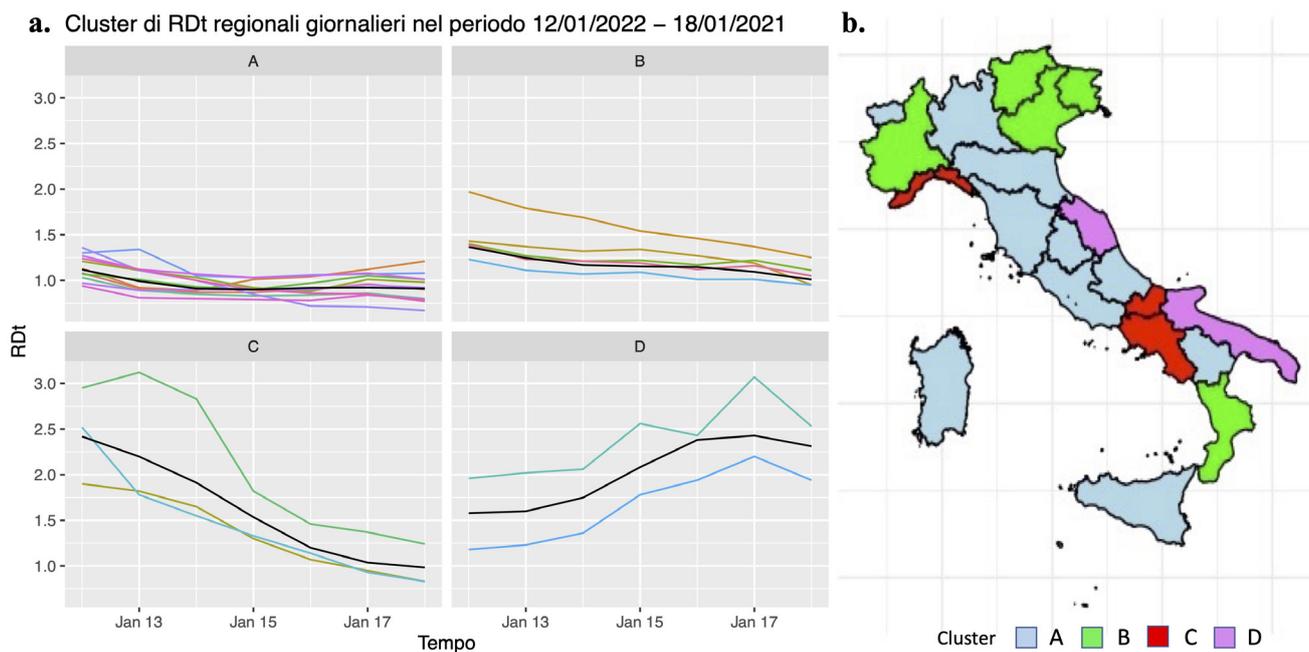


Figura 4. (A) Andamento degli indici RDt per i cluster di Regioni risultanti dalla suddivisione ottimale (una curva per ogni Regione) e (B) mappa geografica dei cluster nel periodo dal 12 al 18 gennaio 2022.

Figure 4. (A) Trend of the RDt indices for the clusters of regions resulting from an optimal subdivision (one curve for each region) and (B) geographical map of the clusters arising from 12 to 18 January 2022.

te Regioni, mentre hanno manifestato in tempi diversi una tendenza alla decrescita da valori più elevati per altre. Non si osservano, quindi, fenomeni di diffusione successiva del contagio tra Regioni o aree contigue né emergono forti asincronie tra aree geografiche specifiche.

Parallelamente, può essere utile considerare l'impatto della variante VOC Omicron (lineage B.1.1.529) durante lo stesso periodo, analizzato in dettaglio da Stefanelli et al.⁷ A seguito di quattro indagini campionarie genomiche nazionali condotte tra il 06.12.2021 e il 17.01.2022, la variante Omicron era identificata all'inizio di dicembre esclusivamente nel Nord Italia, con stime di proporzione al di sotto del 2%. Entro il 20 dicembre, Omicron risultava diffusa ovunque, con stime di proporzione intorno al 20%. Al 3 gennaio, Omicron costituiva la maggioranza delle varianti circolanti, con stime di proporzione intorno all'80%, mentre al 17 gennaio risultava la variante dominante con una proporzione variabile tra il 94,6% e il 97,8%. Il lieve anticipo nell'innalzamento dei valori dell' RD_t per alcune Regioni del Nord Italia riscontrato dall'analisi dei cluster nel periodo iniziale si potrebbe in minima parte ascrivere all'innesco della propagazione della variante rilevato da Stefanelli et al.⁷ Per il resto, la diffusione della nuova variante è risultata piuttosto omogenea nel periodo considerato. Il vantaggio selettivo di Omicron rispetto alla variante Delta, fino ad allora maggiormente diffusa, potrebbe essere ascritto a una maggiore trasmissibilità, responsabile anche

dell'innalzamento della curva epidemica a cavallo della fine dell'anno. La crescita in varie modalità nei valori dell'indice di riproduzione nel periodo a cavallo tra i due anni potrebbe essere dovuta alla rapida e capillare diffusione della variante. Nel contempo, va tenuto presente che i cambiamenti nei modelli di contatto dovuti al periodo natalizio potrebbero aver avuto un effetto difficile da quantificare.

In generale, ciò che ha fatto crescere i contagi sembra essere maggiormente riferito a un fattore non specifico di alcune aree, con una possibile lieve dinamica di contagiosità tra aree adiacenti, piuttosto che all'emergere e al successivo diffondersi asincrono di nuove varianti.

Un'ulteriore fonte di differenziazione nella propagazione dei contagi potrebbe risultare dall'eventuale varietà territoriale delle misure di contenimento. Per far fronte alla situazione emergente, le norme in vigore nel periodo esaminato sono state piuttosto omogenee su tutto il territorio nazionale,^{8,9} con tutte le Regioni e Province autonome in zona bianca o gialla eccetto la Valle d'Aosta, entrata in zona arancione soltanto dal 17.01.2022. Le restrizioni previste per le zone bianca e gialla erano in quel momento essenzialmente simili, mentre il loro impatto nelle zone arancioni e rosse era limitato agli spostamenti tra Comuni, ma in modo molto meno radicale che in precedenza. Con diversi Decreti-Legge emanati tra novembre e dicembre 2021, veniva regolamentata in modo più restrittivo la normativa relativa all'obbligo di *green pass*. È stato

RASSEGNE E ARTICOLI

introdotto l'obbligo di *green pass* rafforzato, la certificazione verde rilasciata a persone vaccinate o guarite, per accedere a tutta una serie di eventi. Era stato esteso l'obbligo di mascherina all'aperto anche in zona bianca e l'uso della mascherina FFP2 in una serie di luoghi al chiuso. La situazione nazionale è stata quasi del tutto uniforme, confermando l'assenza di ulteriori fattori specifici responsabili dell'andamento della curva epidemica.

Conclusioni

Diversi studi sono stati pubblicati sui *pattern* spazio-temporali di diffusione del COVID-19 in Italia¹⁰⁻¹⁵ o all'estero¹⁶, così come molti lavori sono stati dedicati all'identificazione o alla modellizzazione dei possibili determinanti della diffusione dell'epidemia¹⁷⁻²⁰, ma riferendosi soprattutto ai periodi iniziali e utilizzando tecniche che non consentono un'analisi in tempo pressoché reale. Il monitoraggio della variabilità nell'andamento degli indici RD_t , che vengono calcolati con pochissimi giorni di ritardo, sembra costituire un ulteriore indicatore per l'*early detection* delle fasi di modifiche importanti nell'andamento delle curve epidemiche.

La successiva clusterizzazione delle curve fa emergere chiaramente le macroaree in cui le curve epidemici-

che hanno avuto andamenti simili ed è quindi utile per stabilire se i determinanti agiscano, come appare più probabile nei periodi analizzati, senza un *pattern* di innesco e propagazione specifico.

Sebbene la suddivisione regionale abbia consentito di ipotizzare i determinanti della crescita dei contagi, le unità geografiche considerate erano per forza di cose ampie e predefinite. L'estensione a porzioni di aree geografiche più ristrette potrebbe condurre a risultati ancora più specifici. In particolare, si potrebbe pensare di ripetere le stesse analisi a livello di singole Province per alcune Regioni, in modo da mettere in luce possibili analogie di evoluzione dei contagi tra zone limitrofe che si distinguono da altre, per esempio, per connotazioni geomorfologiche o ambientali. Analogamente, potrebbe esistere una somiglianza tra città metropolitane che le differenzia da zone a diversa densità abitativa. Lo stesso tipo di analisi si potrebbe poi ripetere su categorie più specifiche di soggetti, suddividendoli, per esempio, per fasce di età.

Non da ultimo, il metodo di analisi proposto potrebbe essere impiegato per altri tipi di rilevazioni epidemiologiche di interesse per il monitoraggio e l'analisi nell'ambito della sanità pubblica.

Conflitti di interesse dichiarati: nessuno.

Bibliografia

- Giraud MT, Falcone M, Cadum E et al. R_t or RD_t , that is the question! *Epidemiol Prev* 2020;44(5-6):42-50.
- Monitoraggio e analisi dei dati dell'epidemia – MADE. Disponibile all'indirizzo: <https://epiprev.it/apps/made.php>
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2020. Disponibile all'indirizzo: <https://www.R-project.org>.
- Istat. Popolazione residente e dinamica demografica. Anno 2021. Disponibile all'indirizzo: <https://www.istat.it/it/files/2022/12/CENSIMENTO-E-DINAMICA-DEMOGRAFICA-2021.pdf>
- James GM, Sugar CA. Clustering for sparsely sampled functional data. *J Am Stat Assoc* 2003;98:397-408. doi: 10.1198/016214503000189
- Pernice S, Sirovich R, Grassi E et al. CONNECTOR, fitting and clustering of longitudinal data to reveal a new risk stratification system. *Bioinformatics* 2023;39(5):btad201. doi: 10.1093/bioinformatics/btad201
- Stefanelli P, Trentini F, Petrone D et al. Tracking the progressive spread of the SARS-CoV-2 Omicron variant in Italy, December 2021 to January 2022. *Euro Surveill* 2022;27(45):2200125. doi: 10.2807/1560-7917.ES.2022.27.45.2200125
- Gestione della pandemia di COVID-19 in Italia. Disponibile all'indirizzo: https://it.wikipedia.org/wiki/Gestione_della_pandemia_di_COVID-19_in_Italia#Aprile_2022_-_Maggio_2023
- Pandemia di COVID-19 in Italia. Disponibile all'indirizzo: https://it.wikipedia.org/wiki/Pandemia_di_COVID-19_in_Italia#Aprile_2022_-_Maggio_2023
- Azzolina D, Lorenzoni G, Silvestri L, Prosepe I, Berchiolla P, Gregori D. Regional differences in mortality rates during the COVID-19 epidemic in Italy. *Disaster Med Public Health Prep* 2022;16(4):1355-61. doi: 10.1017/dmp.2020.486
- D'Angelo N, Abbruzzo A, Adelfio G. Spatio-Temporal Spread Pattern of COVID-19 in Italy. *Mathematics* 2021;9:2454. doi: 10.3390/math9192454
- Cutrini E, Salvati L. Unraveling spatial patterns of COVID-19 in Italy: Global forces and local economic drivers. *Reg Sci Policy Pract* 2021;13(S1):73-108.
- Mattera R. A weighted approach for spatio-temporal clustering of COVID-19 spread in Italy. *Spat Spatiotemporal Epidemiol* 2022;41:100500.
- Cocco P, De Matteis S. The determinants of the changing speed of spread of COVID-19 across Italy. *Epidemiol Infect* 2022;150:1-26. doi: 10.1017/S095026882200084X
- Reno C, Sanmarchi F, Stoto MA, Fantini MP, Lenzi J, Golinelli D. The impact of health policies and vaccine rollout on the COVID-19 pandemic waves in Italy. *Health Policy Technol* 2022;11(2):100604. doi: 10.1016/j.hlpt.2022.100604
- Yie KY, Chien TW, Yeh YT, Chou W, Su SB. Using Social Network Analysis to Identify Spatiotemporal Spread Patterns of COVID-19 around the World: Online Dashboard Development. *Int J Environ Res Public Health* 2021;18(5):2461. doi: 10.3390/ijerph18052461
- Schuppert A, Polotzek K, Schmitt J, Busse R, Karschau J, Karagiannidis C. Different spreading dynamics throughout Germany during the second wave of the COVID-19 pandemic: a time series study based on national surveillance data. *Lancet Reg Health Eur* 2021;6:100151. doi: 10.1016/j.lanepe.2021.100151
- Schmidt P, Nelson EC, Kearney G, Kraft S, Oliver BJ. International, national and local trends in the spread of COVID-19: a geographic view of COVID-19 spread and the role to be played by coproduction. *Int J Qual Health Care* 2021;33(S2):iii71-77. doi: 10.1093/intqhc/mzab074
- Kapitsinis N. The underlying factors of the COVID-19 spatially uneven spread. Initial evidence from regions in nine EU countries. *Reg Sci Policy Pract* 2020;12(6):1027-45. doi: 10.1111/rsp3.12340.
- Amadou M, Arcuri G, Levratto N. Are regions equal in adversity? A spatial analysis of spread and dynamics of COVID-19 in Europe. *Eur J Health Econ* 2021;22(4):629-42. doi: 10.1007/s10198-021-01280-6