

## Covid-19: chiarezza sui dati

*Covid-19: advocating clear information*

Giorgio Alleva, Alberto Zuliani, Sapienza Università di Roma

### Keywords

Coronavirus, epidemia, dati statistici

### Jel codes

I10, C10

**La pandemia da Covid-19 si presenta nei diversi paesi e nei diversi territori dell'Italia con caratteristiche molto differenti, cosicché non è facile trovare risposte alle numerose domande che i cittadini e gli stessi esperti si pongono. L'informazione resa disponibile dà conto della parte emersa dei contagi e degli esiti ma non consente di controllare né l'effettiva diffusione, né la reale letalità. Per capire di più è necessario un approccio interdisciplinare che coinvolga virologi, immunologi, epidemiologi, responsabili sanitari, fisici, matematici, statistici ed economisti e una forte collaborazione inter-istituzionale. Indagini campionarie sulla diffusione del contagio rappresentano un tassello importante di un sistema di osservazione e monitoraggio. La presenza di decisioni improvvisate e non sostenute dal conforto scientifico, l'allergia per gli aspetti organizzativi delle soluzioni prospettate, l'insofferenza dei comparti produttivi (e della popolazione) lasciano grande incertezza.**

*The Covid-19 pandemic is affecting areas all over Italy and nations worldwide in different ways, which makes it difficult to find answers to the many questions that citizens and experts have asked themselves. The information made available provides an account of the emergence and spread of infections, with various outcomes, but this data does not allow us to control the pandemic, nor its actual lethality. To understand more, an interdisciplinary approach involving virologists, immunologists, epidemiologists, physicists, mathematicians, statisticians and economists and a strong inter-institutional collaboration is needed. Sample surveys on the spread of contagion represent an important part of an observation and monitoring system. The presence of improvised decisions not supported by scientific evidence, the allergy to the organizational aspects of proposed solutions, the impatience of production sectors (and of the population) leave us with strong uncertainty.*

## I. Introduzione

Coronavirus, un termine che è diventato familiare. Ne conosciamo l'immagine poco rassicurante diffusa dai quotidiani, dai telegiornali e dai social media. Vorremmo saperne di più.<sup>1</sup> Gran parte delle persone si sono poste e in qualche caso ancora si pongono domande come queste: in quali modi si trasmette il contagio? Entro quanti giorni si presentano i sintomi? La guarigione dà immunità oppure ci si può infettare di nuovo? E dunque, ci sono le basi scientifiche per dare un «patentino di immunità»? Quale test risulta

più affidabile nelle diverse fasi dell'epidemia? La virulenza diminuirà con il sopraggiungere della stagione calda? Quali dispositivi proteggono meglio? Qual è il mix di misure restrittive più efficaci? Quali le terapie? Quali apparati e organi sono aggrediti oltre a quello respiratorio e ai polmoni? Il virus subirà mutazioni? Quanto durerà l'epidemia? Dobbiamo aspettarci ondate successive?

Virologi, immunologi, epidemiologi stanno dando via via risposte ad alcune di esse, risposte quasi mai conclusive e, spesso, differenti da quelle fornite inizialmente.<sup>2</sup> Alcuni confessano correttamente una certa dose di incertezza, altri si

<sup>1</sup> I dati auditel riferiti a sabato 28 marzo 2020, in piena emergenza, mettevano il Tg1 della sera al primo posto per share (31,29%), in particolare nei minuti dedicati all'informazione sull'epidemia Covid-19 (Grasso, 2020).

<sup>2</sup> Una interessante rassegna sui modelli di diffusione dell'epidemia è stata pubblicata da David Adam sulla rivista *Nature* il 3 aprile (Adam, 2020). La rassegna cita ripetutamente, fra gli altri, Neil Ferguson, epidemiologo matematico dell'Imperial College di Londra, il quale ha ispirato largamente le decisioni del governo britannico.

esprimono in modo perentorio; alcuni partono da una effettiva competenza, altri improvvisano. Qualche volta si tratta di opinioni, anche autorevoli; non abbastanza spesso di evidenza scientifica. Molte delle domande rimarranno senza risposte fondate ancora per molto tempo. D'altronde, l'esperienza accumulata fin qui è troppo breve e non uniforme. In alcuni casi, i dati riferiti sono in contraddizione perché i canoni di rilevazione sono diseguali; forse, almeno inizialmente e in qualche caso ancora oggi, c'è stata reti-

enza; inoltre, il numero di contagi accertato è correlato al numero di test effettuati che è diseguale tra i diversi paesi e, da noi, fra le regioni e province autonome. Infine, a una certa data, gli stadi e le dinamiche dell'epidemia possono risultare differenti, soprattutto fra paesi diversi. Tuttavia, pure al momento del picco, il rapporto contagiati/popolazione dei vari paesi resta molto variabile (tavola 1); così anche fra le regioni e province autonome in Italia (tavola 2)<sup>3</sup>. Anche i tassi di letalità (morti/contagiati per Covid-19) presentano

Tavola 1

**Contagi accertati in complesso, per 1.000 residenti e per 100 test; test effettuati in complesso e per 100 residenti in alcuni paesi** (dati al 28 maggio 2020a, ordinamento secondo il numero di contagi accertati)

Paesi	Contagi accertati	Contagi accertati per 1.000 residenti		Test effettuati	Contagi accertati per 100 test effettuati	Test effettuati per 100 residenti
		Al 28 maggio 2020	Al momento della massima diffusione del contagio			
Stati Uniti	1.699.933	5,14	6,20 <sup>b</sup>	15.192.481	11,2	4,6
Brasile	411.821	2,10	2,34 (31 mag)	132.467	-	0,06
Russia	370.680	2,60	1,52 (12 mag)	9.701.280	3,8	6,7
Regno Unito	267.240	3,94	1,16 (12 apr)	2.144.626	12,5	3,2
Spagna	236.769	5,06	1,61 (27 mar)	2.221.497	10,7	4,8
Italia	231.732	3,86	0,89 (22 mar)	3.683.144	6,3	6,1
Germania	179.717	2,15	0,58 (28 mar)	3.952.971	4,5	4,7
Turchia	158.762	1,88	0,62 (12 apr)	1.928.209	8,2	2,3
India	158.333	0,11	0,19 <sup>b</sup>	3.362.136	4,7	0,24
Francia	145.746	2,23	0,80 (01 apr)	724.574	20,1	1,3
Iran	141.591	1,69	1,96 (05 giu)	837.090	16,9	1,0
Canada	87.508	2,32	1,57 (04 mag)	1.559.204	5,6	4,1
Cina	84.106	0,06	0,04 (13 feb)	n. d.	n. d.	0,0
Messico	78.023	0,61	0,82 (05 giu)	211.277	36,9	0,16
Belgio	57.592	4,97	2,89 (16 apr)	658.513	8,7	5,7
Indonesia	23.851	0,09	0,12 <sup>b</sup>	201.311	11,8	0,00
Romania	18.594	0,97	0,31 (12 apr)	385.728	4,8	2,0
Giappone	16.651	0,13	0,05 (12 apr)	281.293	5,9	0,22
Austria	16.515	1,83	0,78 (27 mar)	427.372	3,9	4,8
Corea del Sud	11.344	0,22	0,06 (29 feb)	868.666	1,3	1,7
Nigeria	8.733	0,04	0,06 <sup>b</sup>	49.966	17,5	0,00
Australia	7.139	0,28	0,28 (23 mar)	1.367.444	0,5	5,4
<i>Coefficiente di variazione (Cv)</i>		<i>0,88</i>	<i>1,20</i>			

FONTE: ELABORAZIONE SU OUR WORLD IN DATA, PROGRAMME ON GLOBAL DEVELOPMENT, UNIVERSITY OF OXFORD

a I dati sui casi accertati in Spagna si riferiscono al 27 maggio. I dati sui test negli Stati Uniti si riferiscono al 27 maggio; in Belgio, Iran, Messico e Romania al 26 maggio; in Germania al 24 maggio; nel Regno Unito al 22 maggio; in Spagna al 21 maggio; in Francia al 28 aprile; in Brasile al 20 aprile.

b Dati riferito al 10 giugno 2020; il picco non era stato ancora raggiunto.

<sup>3</sup> L'evoluzione dell'epidemia è estremamente rapida e i dati nuovi modificano, in alcuni casi in misura anche importante, le tendenze acquisite in precedenza e le valutazioni che ne sono state fatte. Per quelli riportati in questo lavoro viene indicata la data dell'aggiornamento.

Tavola 2

**Contagi accertati in complesso, per 1.000 residenti e per 100 test; test effettuati in complesso e per 100 residenti nelle regioni e province autonome** (dati al 28 maggio 2020, ordinamento secondo il numero di contagi accertati)

Regioni e province autonome	Contagi accertati	Contagi accertati per 1.000 residenti	Tamponi effettuati	Contagi accertati per 100 tamponi effettuati	Tamponi effettuati per 100 residenti
Lombardia	88.183	8,77	713.068	12,4	7,1
Piemonte	30.445	6,99	304.358	10,0	7,0
Emilia Romagna	27.701	6,21	310.696	8,9	7,0
Veneto	19.125	3,90	632.124	3,0	12,9
Toscana	10.086	2,70	242.687	4,2	6,5
Liguria	9.605	6,19	100.625	9,5	6,5
Lazio	7.693	1,31	245.993	3,1	4,2
Marche	6.719	4,41	100.110	6,7	6,6
Campania	4.777	0,82	189.068	2,5	3,3
Puglia	4.481	1,11	111.946	4,0	2,8
P.A. Trento	4.425	8,18	83.018	5,3	15,3
Sicilia	3.438	0,69	142.516	2,4	2,9
Friuli Venezia Giulia	3.262	2,68	126.015	2,6	10,4
Abruzzo	3.237	2,47	71.788	4,5	5,5
P.A. Bolzano	2.595	4,89	63.289	4,1	11,9
Umbria	1.431	1,62	67.458	2,1	7,6
Sardegna	1.355	0,83	54.587	2,5	3,3
Valle d'Aosta	1.182	9,41	14.653	8,1	11,7
Calabria	1.158	0,59	66.956	1,7	3,4
Molise	435	9,81	13.847	3,1	31,2
Basilicata	399	0,71	28.342	1,4	5,0
Italia	231.732	3,86	3.683.144	6,3	6,1
Coefficiente di variazione (Cv)		0,78		0,63	

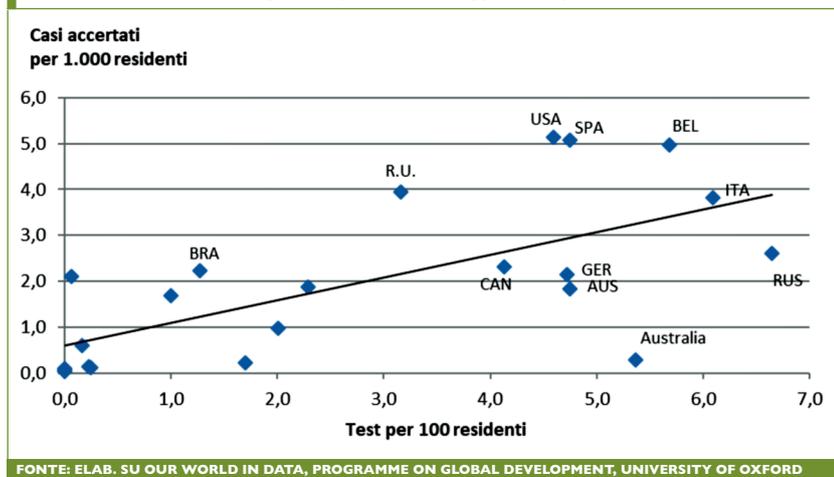
FONTE: ELABORAZIONE SU DATI PROTEZIONE CIVILE, [HTTPS://GITHUB.COM/PCM-DPC/COVID-19/TREE/MASTER/DATI-REGIONI](https://github.com/PCM-DPC/COVID-19/TREE/MASTER/DATI-REGIONI)

una variabilità non facilmente spiegabile, determinata – oltre che per i motivi precedenti – anche dalla sottostima, in proporzioni molto differenti, sia del numeratore sia del denominatore nei diversi domini territoriali; lo stesso accade per i tassi di guarigione.

**I principali paesi.** Tra i paesi più popolosi, al 28 maggio 2020, gli Stati Uniti risulterebbero molto più colpiti rispetto a Cina, India, Indonesia, Nigeria, Brasile e Russia, detenendo il record non soltanto in termini assoluti (1.700.000 contagi accertati) ma anche di probabilità di contagio, pari a 5,14 per 1.000 residenti. I casi registrati nei grandi paesi europei considerati nella tavola 1 presentano una variabilità non molto elevata; il numero di contagi risulta compreso tra 145.000 in Francia e 268.000 nel Regno Unito; la pro-

babilità di contagio per i residenti è invece molto diversa, passando dai valori 2,15 e 2,23 per 1.000 di Germania e Francia a 5,06 per 1.000 della Spagna; anche il Belgio presenta un valore molto elevato, pari a 4,97, derivante dalla scelta – unico tra i Paesi europei – di classificare come decessi connessi con il Covid-19 anche quelli di persone non sottoposte a esami ma che presentavano sintomi della patologia. La variabilità fra i tassi di contagio al momento di massima diffusione dell'epidemia è più elevata di quella che si rileva alla data del 28 maggio 2020, a testimonianza della diversa violenza con la quale il coronavirus ha colpito le aree geografiche del pianeta. I coefficienti di variazione risultano pari a 0,88 per le evidenze al 28 maggio e 1,20 nella fase di picco.

Figura 1

**Casi accertati per 1.000 residenti e test effettuati per 100 residenti in alcuni paesi** (dati al 28 maggio 2020)


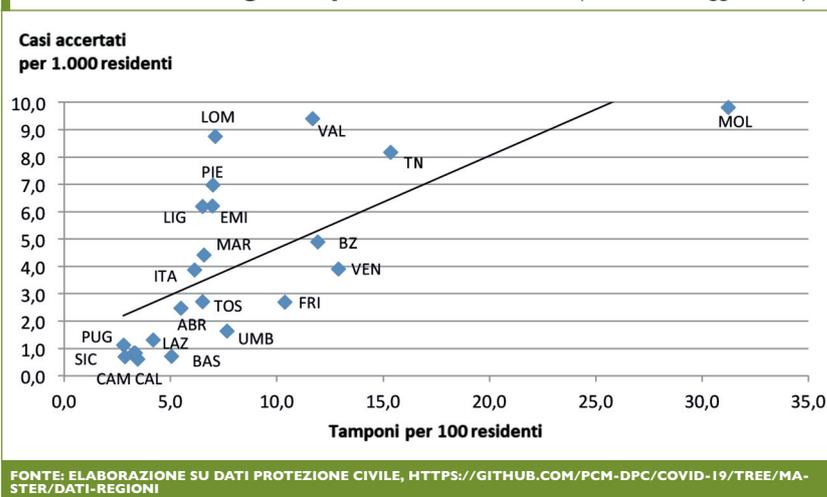
A riprova della sottostima generalizzata dei contagi, il loro numero è influenzato da quello dei tamponi o di altri test effettuati, che risulta molto diverso tra i paesi presi in considerazione per il confronto (figura 1). La relazione positiva che si può intravedere non è però nitida. La variabilità è molto ampia. In qualche caso, come in Italia, Belgio e, in misura minore, Spagna e Stati Uniti, l'elevato tasso di accertamenti attraverso test diagnostici è chiaramente connesso con la virulenza dell'epidemia. Una elevata frazione di popolazione sottoposta a tampone si osserva in Europa pure in paesi meno colpiti. È il caso di Austria e Germania, e soprattutto della Russia, che hanno contenuto anche in questo modo la diffusione e la letalità dell'epidemia<sup>4</sup>. L'Australia, al quarto posto tra i paesi considerati per la frazione di test effettuati, presenta uno dei valori più bassi per la diffusione dei contagi, pari a 0,28 per 1.000 residenti. Il Regno Unito registra contemporaneamente elevati tassi di contagio e bassa incidenza della popolazione sottoposta a test.

**Le regioni italiane.** La distribuzione dei contagi nelle regioni italiane mostra una forte concentrazione in Lombardia e in ge-

nerale nelle regioni del Nord. I tassi di contagio nella popolazione presentano i livelli massimi oltre che in Lombardia (8,8 contagi per 1.000 residenti), nelle piccole regioni della Val d'Aosta e del Molise e nella provincia autonoma di Trento, con valori compresi tra 8,2 e 9,8. I minimi si registrano in Campania, Basilicata, Calabria e nelle due isole maggiori, con valori compresi tra 0,59 e 0,82 per 1.000 residenti. Come si può osservare dalla figura 2, in diversi casi, i tassi di contagio registrati risultano correlati alla quota di popolazione sottoposta al tampone. L'associazione può derivare da un'attività più intensa di testing nelle aree più colpite<sup>5</sup>.

La relazione non è comunque univoca. Ad esempio, il minor tasso di contagi del Veneto rispetto alla Lombardia, come anche la minore letalità, sono associati a un livello molto più elevato di popolazione sottoposta al tampone (quasi doppio in Veneto). È possibile che la comparativamente migliore situazione in questa regione derivi anche da una migliore gestione dell'isolamento della popolazione dopo l'identificazione della positività, in particolare degli infetti pauci sintomatici e asintomatici. Com'era da atten-

Figura 2

**Casi accertati per 1.000 residenti e tamponi effettuati per 100 residenti nelle regioni e province autonome** (dati al 28 maggio 2020)


FONTE: ELABORAZIONE SU DATI PROTEZIONE CIVILE, [HTTPS://GITHUB.COM/PCM-DPC/COVID-19/TREE/MASTER/DATI-REGIONI](https://github.com/PCM-DPC/COVID-19/TREE/MASTER/DATI-REGIONI)

<sup>4</sup> Come si mostrerà nell'Appendice 2, la Germania è uno dei pochi paesi in cui i tamponi sono stati effettuati anche attraverso un campionamento statistico, volto alla stima del tasso di contagio nella popolazione.

<sup>5</sup> D'altronde, il test è l'unico modo per tentare di individuare e registrare i casi di contagio dei cosiddetti pauci sintomatici e degli asintomatici.

dersi la variabilità dei rapporti contagi per 100 tamponi è più contenuta rispetto a quella dei contagi per 1.000 residenti; i coefficienti di variazione per le due serie regionali valgono rispettivamente 0,63 e 0,78.

I dati statistici diffusi quotidianamente sull'evoluzione dell'epidemia in Italia non sono, a nostro parere, soddisfacenti: non aiutano a capire; alcuni di essi sono enfatizzati ingenuamente, come l'aumento del numero di guariti, destinato comunque a crescere; in qualche caso risultano fuorvianti per via del ritardo nella comunicazione degli aggiornamenti che dà luogo a fluttuazioni delle serie regionali e a letture di inattese accelerazioni o di agognati rallentamenti. Per diradare un poco la nebbia ci siamo poste tre questioni: perché vogliamo disporre di dati? Di quali dati vorremmo disporre e con quali caratteristiche? Come possiamo produrli? Nel prossimi tre paragrafi tentiamo di dare qualche risposta.

## 2. Comprendere l'evoluzione dell'epidemia

Dati statistici affidabili sono essenziali per almeno quattro finalità:

1. stimare le modalità di propagazione dell'epidemia;
2. stimare il numero effettivo dei contagiati, sintomatici, pauci sintomatici e asintomatici, dei guariti, dei decessi e degli immuni in modo da poter calcolare correttamente i tassi di contagio, i tassi di guarigione, letalità e immunizzazione e altri parametri indicativi della diffusione dell'epidemia (ad esempio,  $R_t$ )<sup>6</sup>; seguirne l'evoluzione nel tempo e fornire previsioni, costruendo scenari anche differenti e fornendo la relativa probabilità di accadimento;
3. valutare gli effetti delle misure via via introdotte in modo da poter orientarne inasprimento, attenuazione e rimozione nel corso del tempo;
4. prepararsi ad affrontare in modo informato le prossime prevedibili epidemie e pandemie ed eventuali nuove ondate di quella attuale.

Quanto alla prima finalità, è necessario che gli statistici lavorino insieme agli epidemiologi. Si possono dare risposte

riguardo al tempo di raddoppio dei contagi; al tasso  $R_t$ ; al rischio di contagio e di letalità per popolazioni selezionate e, in primo luogo, per persone con molte reti sociali ed elevato numero di contatti rispetto a quelle che conducono vita riservata con limitate uscite da casa e scarsi rapporti interpersonali, e inoltre per regione, genere, classi di età. Occorrono dati statisticamente affidabili riguardo al numero complessivo di contagi, guarigioni, decessi e immunizzazioni e non soltanto alla parte emersa, riferita a coloro che si sono rivolti alle strutture sanitarie, che è il dato attualmente disponibile. Le tecniche di campionamento per popolazioni elusive, i metodi calibrati sugli aspetti spaziali del contagio – il campionamento indiretto, quello adattivo, il campionamento spaziale bilanciato – consentono di progettare sistemi di rilevazione e monitoraggio finalizzati a letture affidabili dell'epidemia e di costruire modelli previsionali che forniscano anche misure di accuratezza.

Quanto alla seconda finalità. Il numero di contagi come viene valutato attualmente è ampiamente sottostimato; mancano quasi completamente gli asintomatici<sup>7</sup>. Di volta in volta vengono ipotizzati moltiplicatori: cinque volte, dieci volte rispetto al numero accertato sul campione di convenienza che si affaccia alle strutture sanitarie o che occasionalmente viene sottoposto a tampone o ad altro accertamento diagnostico. Nessuna di queste stime è fondata su un'osservazione statistica rigorosa che consenta di assegnare un grado probabilistico di affidabilità<sup>8</sup>.

Neppure il numero dei guariti e quello dei deceduti sono veritieri. Per i guariti è facile capirlo. Se fra i contagi non vengono compresi – o sono compresi in misura minima – i pauci sintomatici e gli asintomatici, non sarà possibile valutare in modo attendibile il numero dei guariti.

Più complessa è la stima del numero dei morti per Covid-19. Per comprenderne la letalità specifica sarebbe importante conoscere la presenza di patologie pregresse nelle persone decedute, risultate contagiate. D'altra parte, la precisione del tasso di letalità dipende anche dal denominatore e la forte sottostima del numero di contagi induce una corrispondentemente forte sovrastima del tasso di letalità, in particolare nei paesi che non effettuino molti test e, comun-

<sup>6</sup>  $R_t$  valutato al momento 0, cioè  $R_0$ , rappresenta il numero medio di infezioni secondarie prodotte da ciascun individuo infetto in una popolazione completamente suscettibile, mai venuta a contatto con l'agente patogeno, cioè il numero di persone che mediamente ogni malato può contagiare in tale situazione. È il criterio con il quale si misura la potenziale trasmissibilità quando non siano state adottate misure di contenimento, come avveniva da noi a metà febbraio. Ora (inizio giugno) deve essere considerato il tasso di riproduzione effettivo  $R_t$ .

<sup>7</sup> Sfuggono anche le forme lievi e, purtroppo, anche quelle gravi e spesso letali che non sono riuscite ad arrivare al pronto soccorso e alle strutture specializ-

zate. Una delle maggiori carenze portate allo scoperto dall'epidemia è relativa alla medicina territoriale.

<sup>8</sup> Giuseppe Arbia (2020) ha raccolto e messo in fila alcune stime della frazione di infetti privi di sintomi o con sintomi lievi e spesso non diagnosticati: per l'Università di Sapporo e Osaka 1/4, secondo il Governo cinese e Iaria Capua 1/3, per il World Data Forum 6/10, in base a un esperimento a Vo' Euganeo 6,2/10, per l'Oms 8/10, secondo la rivista *Science* 8,6/10, secondo una dichiarazione del responsabile della Protezione civile Angelo Borrelli 9/10 per l'Italia. Basandosi anche sulle stime precedenti, il numero di infetti potrebbe oscillare attualmente in Italia fra 56.000 (il numero dei positivi il 31 maggio 2020 era di 42.075

Tavola 3

**Morti, tassi di letalità e di mortalità per Covid-19 in alcuni paesi**  
(dati al 28 maggio 2020a, ordinamento secondo il numero di morti)

Paesi	Morti	Tasso di letalità (morti per 100 casi positivi accertati)	Tasso di mortalità (morti per 10.000 residenti)
Stati Uniti	100.442	5,91	3,03
Regno Unito	37.460	14,02	5,52
Italia	33.142	14,30	5,51
Francia	28.596	19,62	4,38
Spagna	27.118	11,45	5,80
Brasile	25.598	6,22	1,26
Belgio	9.364	16,26	8,08
Messico	8.597	11,02	0,67
Germania	8.411	4,68	1,00
Iran	7.564	5,34	0,90
Canada	6.765	7,73	1,79
Cina	4.638	5,51	0,03
India	4.531	2,86	0,03
Turchia	4.397	2,77	0,52
Russia	3.968	1,07	0,28
Indonesia	1.473	6,18	0,03
Romania	1.219	6,56	0,63
Giappone	858	5,15	0,07
Austria	645	3,91	0,72
Corea del Sud	269	2,37	0,05
Nigeria	254	2,91	0,01
Australia	103	1,44	0,04
<i>Coefficiente di variazione (Cv)</i>		<i>0,70</i>	<i>1,29</i>

FONTE: ELAB. SU OUR WORLD IN DATA, PROGRAMME ON GLOBAL DEVELOPMENT, UNIVERSITY OF OXFORD

a I dati sui decessi e i casi accertati della Spagna si riferiscono al 27 maggio.

que, se non si stimi la numerosità della popolazione infetta mediante indagini campionarie.

Come si può osservare dai dati riportati nella tavola 3 riferiti a diversi paesi fra i quali l'Italia, sia i tassi di letalità sia soprattutto quelli di mortalità per Covid-19 risultano fortemente variabili (con coefficienti di variazione rispettivamente pari a 0,7 e 1,3).

Per l'Italia i dati sui decessi diffusi dalla Protezione civile sono largamente approssimati per difetto. Come avviene per il confronto fra paesi, anche tra le regioni italiane (tavola 4)

unità) e 421.000, un margine di incertezza troppo elevato per poter orientarsi.

il tasso di mortalità presenta una variabilità nettamente superiore rispetto a quello di letalità (i rispettivi coefficienti di variazione sono pari a 0,9 e 0,3 e, secondo le aspettative, risultano ambedue inferiori a quelli corrispondenti calcolati per i diversi paesi nella tavola 3).

In Lombardia, oltre al maggior numero di decessi, si registrano anche i più elevati tassi di letalità e di mortalità. Seguono, per numero di decessi, Emilia Romagna e Piemonte che però presentano tassi di letalità e mortalità non molto diversi da quelli di altre regioni: Marche, Liguria, Valle d'Aosta e Abruzzo per il tasso di letalità; Valle d'Aosta, Liguria e provincia autonoma di Trento per il tasso di mortalità. Il Veneto, al quarto posto per numero di vittime, presenta un tasso di mortalità inferiore a quelli registrati in ben altre nove regioni o province autonome.

Tutte le regioni del Mezzogiorno, oltre Umbria e Lazio, presentano tassi di letalità inferiori a 10 su 100 casi positivi accertati; Lombardia, Emilia Romagna, Liguria e Marche superiori a 14 su 100 (in Lombardia oltre 18). Il tasso di mortalità non raggiunge 1 decesso su 10.000 residenti in alcuna regione del Mezzogiorno; supera 5 su 10.000 in tutte le regioni del Nord, a eccezione del Veneto, e inoltre nelle Marche.

È verosimile che lo stress al quale sono state sottoposte le strutture sanitarie nelle regioni più colpite possa avere inciso negativamente sulla capacità di registrazione dei decessi riconducibili a Covid-19; questo fenomeno sembra essersi attenuato nei mesi di aprile e soprattutto maggio.

La sottostima del numero di decessi nelle regioni del Nord sta emergendo nitidamente dal confronto tra il numero di morti per tutte le cause registrato dalle anagrafi nei primi mesi di quest'anno e nello stesso periodo negli anni precedenti. Su questo aspetto, l'Istat ha pubblicato dati in di-

Tavola 4

**Morti, tassi di letalità e di mortalità per Covid-19 nelle regioni e province autonome** (dati al 28 maggio 2020, ordinamento secondo il numero di morti)

Regioni e province autonome	Morti	Tasso di letalità (morti per 100 casi positivi accertati)	Tasso di mortalità (morti per 10.000 residenti)
Lombardia	16.112	18,11	16,01
Emilia Romagna	4.114	14,80	9,23
Piemonte	3.867	12,62	8,88
Veneto	1.918	10,01	3,91
Liguria	1.465	15,16	9,45
Toscana	1.041	10,30	2,79
Marche	987	14,67	6,47
Lazio	735	9,51	1,25
Puglia	504	11,21	1,25
P.A. Trento	462	10,43	8,54
Campania	412	8,58	0,71
Abruzzo	405	12,48	3,09
Friuli Venezia Giulia	333	10,17	2,74
P.A. Bolzano	291	11,21	5,48
Sicilia	274	7,96	0,55
Valle d'Aosta	143	12,08	11,38
Sardegna	130	9,59	0,79
Calabria	97	8,38	0,50
Umbria	76	5,31	0,86
Basilicata	27	6,77	0,48
Molise	22	5,05	4,96
Italia	33.415	14,34	5,56
Coefficiente di variazione (Cv)		0,30	0,91

FONTE: ELABORAZIONE SU DATI PROTEZIONE CIVILE, [HTTPS://GITHUB.COM/PCM-DPC/COVID-19/TREE/MA-STER/DATI-REGIONI](https://github.com/PCM-DPC/COVID-19/TREE/MA-STER/DATI-REGIONI)

verse riprese<sup>9</sup>. A partire dagli ultimi resi disponibili (Istituto superiore di sanità e Istat, 2020), riferiti a 7.270 di 7.904 comuni italiani, con una copertura in termini di popolazione residente elevata, pari a 93,5%, saranno presentate nel paragrafo successivo stime dell'eccesso di mortalità a varie scale territoriali fino al livello provinciale e di sistema del lavoro e una loro lettura rispetto ad alcuni caratteri demografici.

Una elaborazione nella stessa direzione è stata condotta da Giorgio Parisi e altri ricercatori con riferimento ad alcune

<sup>9</sup> L'Istituto ha pubblicato il 4 aprile il numero di decessi in 1.969 comuni. Il 4 maggio ha reso disponibile il numero di decessi fino al 15 aprile 2020 per 4.428 comuni e fino al 31 marzo 2020 per altri 2.432 comuni, mentre risultavano mancanti i dati per i restanti 1.038 comuni. Lo scorso 4 giugno sono stati pubblicati i dati sui decessi fino al 30 aprile per 7.270 comuni (Istituto superiore di sanità e Istat, 2020); non sono disponibili ancora quelli per i rimanenti 634 comuni. Si deve tenere a mente che il decesso viene registrato nell'anagrafe del comune di residenza e non in quello in cui è avvenuto.

regioni italiane (Bucci et al., 2020). Partendo dai dati relativi a 1.969 comuni resi disponibili dall'Istat il 4 aprile, hanno confrontato i dati ufficiali delle vittime da Covid-19 nei primi tre mesi dell'anno 2020 con l'eccedenza complessiva dei morti per tutte le cause nello stesso periodo, definita come differenza rispetto al numero di decessi registrati dalle anagrafi nel primo trimestre del 2016, anno considerato paragonabile. Gli autori hanno stimato in Lombardia circa 7.000 decessi in più rispetto a quelli certificati (figura 3) e circa 1.000 in più in Emilia Romagna. Nelle province maggiormente colpite, come Bergamo, hanno valutato che il numero reale dei decessi per Covid-19 sia stato più del doppio di quello riportato dalle autorità sanitarie. Nella discussione dei risultati definiscono due componenti alla base della sottostima: persone decedute, contagiate dal Coronavirus alle quali non sono stati fatti i tamponi e che quindi non risultavano registrate come casi di contagio; decessi dovuti indirettamente al Covid-19, ossia non provocati dal virus ma che non sarebbero avvenuti in condizioni normali, conseguenza ipotizzata a motivo delle condizioni critiche in cui ha dovuto operare il sistema sanitario in alcune regioni. Per distinguere le due componenti viene sfruttata

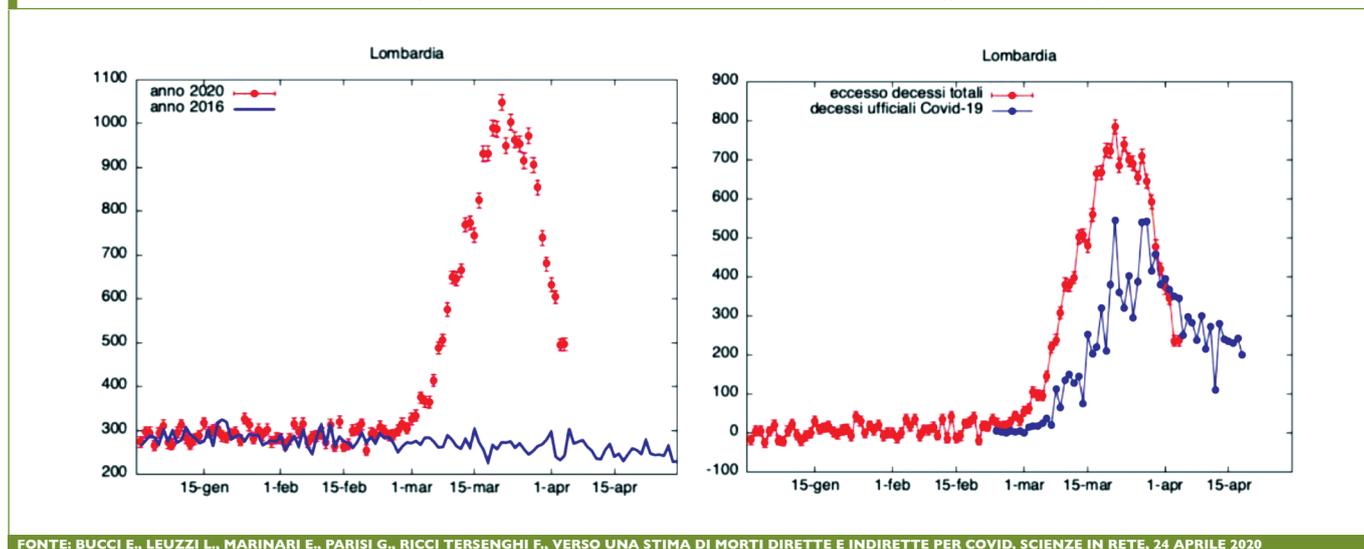
l'informazione relativa alla mortalità differenziale per genere da Covid-19<sup>10</sup>. La stima della frazione di decessi non Covid-19 sull'eccedenza in Lombardia sostituirebbe compresa fra 0,429 e 0,437. Il totale dei morti per Covid-19 nella regione, sommando quelli registrati come tali (8.905 al 5 aprile) a quelli stimati sull'eccedenza (fra 6.695 e 6.895), risulterebbero pari a 15.600-15.800 nell'intervallo di tempo tra il 23 febbraio e il 4 aprile<sup>11</sup>. Tale valore corrisponde alla cumulata dell'eccedenza giornaliera descritta nel grafico a destra della figura 3.

<sup>10</sup> Il tentativo è molto suggestivo; tuttavia, viene introdotta una stima esogena relativamente alla frazione femminile dei decessi per Covid-19 ( $f_i$ ).

<sup>11</sup> Gli autori sono consapevoli della necessità di ulteriori approfondimenti anche per la limitata numerosità dei comuni considerati.

Figura 3

**Decessi giornalieri totali in Lombardia: confronto tra il 2016 e il 2020** (grafico a sinistra). **Decessi in eccesso e numero di decessi certificati positivi al Covid-19 in Lombardia nel 2020** (grafico a destra)



Quanto alla terza finalità fra quelle indicate all’inizio del paragrafo, occorre considerare che, finora, le misure via via adottate sono state generalizzate all’intero Paese; soltanto recentemente alcune Regioni stanno assumendo iniziative correttive. Metodologie consolidate come, ad esempio, le stime difference-in-difference, consentono la valutazione di eventuali effetti differenziali, ma è necessario che ne venga progettato il disegno insieme alle misure introdotte e che i protocolli di rilevazione e analisi siano rigorosi. L’osservazione degli effetti di politiche o di interventi non disegnata e non realizzata scientificamente non è in grado di fornire informazioni solide.

Infine, e veniamo alla quarta finalità, conosciamo ancora poco del Sars-CoV-2; non sappiamo se il contagio e la successiva guarigione diano l’immunizzazione e se essa sia permanente. È possibile che l’epidemia si ripresenti con ondate successive. È anche probabile che altri virus, auguriamoci meno violenti di quello attuale, compaiano negli anni prossimi. È bene, quindi, collaudare un apparato di osservazione e di sorveglianza in grado di offrire un’informazione affidabile e utile per gestire nel modo migliore le epidemie venture.

### 3. Quali informazioni statistiche sono necessarie

**I contagi.** Nel corso degli ultimi mesi, a partire da febbraio, siamo stati aggiornati sul numero giornaliero e cumulato dei contagi diagnosticati dai presidi sanitari; sul numero giornaliero e cumulato delle persone contagiate in terapia presso le strutture sanitarie, con dettaglio per quella intensiva; lo stesso per quelle in quarantena presso la propria abitazione perché asintomatiche o con patologia lieve e comunque diagnosticate; sul numero giornaliero e cumulato delle persone guarite e di quelle decedute fra le diagnosticate. Si conosce anche lo stock giornaliero dei contagiati in trattamento. Si tratta di informazioni sulla capacità del sistema sanitario di fronteggiare l’epidemia, certamente molto utili ma che non fanno capire quale sia la reale diffusione e gravità del contagio. Il numero oscuro – è stato più volte accennato in precedenza – è rappresentato dai così detti asintomatici che non siano stati occasionalmente testati; di questo aggregato non si hanno stime affidabili. I dati completi dovrebbero essere dettagliati, oltre che territorialmente secondo alcuni principali caratteri (senz’al-

tro regione e densità demografica del luogo di residenza), anche per genere, età e abitudini di vita, un'informazione quest'ultima estremamente utile per decidere modalità, intensità e durata delle misure di contenimento.

**I test.** Come accertare la presenza dell'infezione, attuale o pregressa, e lo stato di immunità delle persone? Tramite il test molecolare (attraverso il tampone naso-faringeo o salivare) o il test sierologico (con prelievo del sangue)? O ambedue? Senza entrare nel merito di un tema prettamente sanitario, sembrerebbe che, in una fase ascendente, iniziale o di seconda ondata, l'esito del tampone (anche ripetuto) sia fondamentale per decidere se isolare una persona e identificarne i contatti nei quattordici giorni precedenti (durata stimata della contagiosità); nella fase discendente dell'epidemia, invece, l'esito dell'esame sierologico può consentire di apprezzare la dimensione della popolazione non a rischio in determinate aree geografiche o luoghi di lavoro e dunque di valutare i tempi per il riavvio delle attività. Si tratta di un tema rilevante nella progettazione di un sistema di sorveglianza, perché ha effetti su tempi, costi e organizzazione (con la possibilità di robotizzare alcune fasi di somministrazione e analisi dei test). L'aspetto di carattere statistico riguarda l'affidabilità dei risultati, un dato fondamentale non soltanto per la scelta dello specifico kit da adottare, ma soprattutto per le decisioni da prendere a valle dell'esito dell'esame. Occorre tenere conto dei due errori che possono essere commessi in sede di test: i cosiddetti falsi-positivi e falsi-negativi. Nel trade-off tra sensibilità e specificità di un test, occorre decidere quale aspetto fra i due tenere sotto controllo in misura maggiore. È più importante un'elevata sensibilità del test in modo da minimizzare i falsi negativi? Oppure un'elevata specificità, in modo da minimizzare i falsi positivi?<sup>12</sup> La scelta dipende dalla fase e dall'intensità dell'epidemia, dai diversi costi sociali e dall'appartenenza delle persone a categorie a maggiore o minore rischio. Massimo Fuggetta discute il tema in modo approfondito, arrivando alla conclusione che la parola d'ordine debba essere comunque «test, test, test», durante tutte le fasi dell'epidemia (Fuggetta, 2020).

**I guariti.** L'incertezza relativa ai contagi si riflette inevitabilmente

sui dati delle guarigioni. Con un tasso di letalità intorno al 14,3% dei contagiati conclamati, valutato al 31 maggio 2020 sulla base dei dati cumulati diffusi dalla Protezione civile, il numero complessivo di guariti è destinato a crescere con il passare dei giorni. Se si disponesse di dati attendibili sui contagiati asintomatici e sui pauci sintomatici, il numero di guarigioni e il relativo tasso risulterebbero molto elevati, tenuto conto che per questi due gruppi l'esito favorevole è pressoché generale.

**I decessi.** Ogni morte per coronavirus è accompagnata da grande sofferenza e dallo strazio delle famiglie; ed è motivo di apprensione per tutti. Il numero di decessi è sottostimato, lo abbiamo già accennato, ma non si sa di quanto. Il quadro dei decessi per Covid-19 aggiornato al 31 maggio 2020 per mese e regione è descritto nella tavola 5.

Le prime morti registrate per Sars-CoV-2 vengono accertate alla fine del mese di febbraio in Lombardia (23), Emilia Romagna (4) e Veneto (2). La Lombardia raggiunge il picco a marzo con 7.176 morti (57,9% del totale nazionale), un mese prima delle altre regioni, e mantiene un livello elevato (6.573 decessi, pari al 42% del totale) nel mese successivo. In aprile la mortalità si intensifica in Piemonte, Veneto, Liguria, Emilia Romagna e Toscana. Fra aprile e maggio la diminuzione di mortalità è generalizzata, costantemente al di sopra del 50% e, in generale del 70-80%. Fa eccezione il solo Lazio che registra un ulteriore aumento del 5,3%, passando da 162 morti a marzo a 279 ad aprile e a 294 a maggio<sup>13</sup>. Nel complesso, i morti per Covid-19 registrati dalla Protezione civile nei primi cinque mesi del 2020 sono stati 33.415, 64,6% dei quali nelle regioni del Nord-Ovest e 21,3% in quelle del Nord-Est. Le regioni centrali presentano 2.839 decessi (8,5% del totale); il Sud 1.467 decessi (4,4%) e le due isole maggiori 404 (1,2%). Se si considera anche il tasso di mortalità calcolato nella tavola 4, dopo quattro mesi dall'inizio dalle prime avvisaglie dell'epidemia, si configurano tre realtà territoriali molto diverse: il Nord, con situazioni più gravi in Lombardia, Valle d'Aosta, Liguria, Emilia Romagna e Piemonte; il Centro, con una mortalità molto più contenuta e la situazione più grave nelle Marche; il Sud e specialmente le Isole con un numero decisamente

<sup>12</sup> A testimonianza della rilevanza di questo tema, nel bando di gara per la scelta del test sierologico da effettuare su un campione di 150.000 residenti (indagine avviata a fine maggio) è stato indicato che il prodotto doveva assicurare una sensibilità almeno pari al 90% e una specificità almeno pari al 95%.

<sup>13</sup> Peraltro, il bilancio del Lazio è forse ancora peggiore, secondo la revisione dei dati Protezione civile effettuato da parte dell'Istituto superiore di sanità.

Tavola 5

**Morti per Covid-19 per mese e regione nel periodo 28 febbraio<sup>1</sup>-31 maggio e nel periodo 28 febbraio<sup>1</sup>-30 aprile 2020**

Domini territoriali	Decessi					Decessi nel periodo febbraio-aprile			
	Febbraio	Marzo	Aprile	Maggio	Totale	Fonte Prot. civ.	Fonte Ist. sup. san.	Differenza	Differenza %
Italia	29	12.399	15.539	5.448	33.415	27.967	27.846	-121	-0,4
<b>Ripartizioni geografiche</b>									
Nord-Ovest	23	8.514	9.605	3.445	21.587	18.142	18.017	-125	-0,7
Nord-Est	6	2.468	3.518	1.126	7.118	5.992	6.039	47	0,8
Centro	-	895	1.361	583	2.839	2.256	2.280	24	1,1
Sud	-	410	816	241	1.467	1.226	1.165	-61	-5,0
Isole	-	112	239	53	404	351	345	-6	-1,7
<b>Regioni</b>									
Piemonte	-	854	2.212	801	3.867	3.066	2.220	-846	-27,6
Valle d'Aosta	-	56	81	6	143	137	127	-10	-7,3
Lombardia	23	7.176	6.573	2.340	16.112	13.772	14.480	708	5,1
Trentino Alto Adige	-	240	453	60	753	693	666	-27	-3,9
Veneto	2	475	982	459	1.918	1.459	1.513	54	3,7
Friuli Venezia Giulia	-	113	176	44	333	289	300	11	3,8
Liguria	-	428	739	298	1.465	1.167	1.190	23	2,0
Emilia-Romagna	4	1.640	1.907	563	4.114	3.551	3.560	9	0,3
Toscana	-	244	598	199	1.041	842	856	14	1,7
Umbria	-	37	30	9	76	67	61	-6	-9,0
Marche	-	452	454	81	987	906	840	-66	-7,3
Lazio	-	162	279	294	735	441	523	82	18,6
Abruzzo	-	115	205	85	405	320	304	-16	-5,0
Molise	-	9	12	1	22	21	18	-3	-14,3
Campania	-	133	226	53	412	359	338	-21	-5,8
Puglia	-	110	305	89	504	415	399	-16	-3,9
Basilicata	-	7	18	2	27	25	16	-9	-36,0
Calabria	-	36	50	11	97	86	90	4	4,7
Sicilia	-	81	154	39	274	235	220	-15	-6,4
Sardegna	-	31	85	14	130	116	125	9	7,8

**Fonte:** ELABORAZIONE SU DATI PROTEZIONE CIVILE E ISTITUTO SUPERIORE DI SANITÀ

<sup>1</sup> Data dei primi decessi per Covid-19.

basso di decessi, in assoluto e in rapporto alla popolazione residente, con il tasso di mortalità per sei regioni su sette inferiore a 1 ogni 10.000 residenti (soltanto in Puglia 1,25 per 10.000). Possono aver contribuito a questa accentuata diversificazione territoriale le misure di distanziamento sociale adottate nelle regioni del Centro e del Mezzogiorno

all'esordio del contagio la cui espansione è stata quindi contenuta. Ma si dovrà approfondire ancora molto in sede epidemiologica per fornire analisi e nessi causali convincenti, ammesso che i dati prodotti si rivelino adeguati per questo scopo.

Ancora nella tavola 5 è riportato il confronto dei dati sui

decessi resi disponibili dalla Protezione civile e dall'Istituto superiore di sanità. La fonte è la medesima: le comunicazioni giornaliere delle regioni. Tuttavia, si presentano alcuni disallineamenti al livello territoriale disaggregato: per il Piemonte l'Iss stima 2.220 morti rispetto ai 3.066 di Protezione civile (-27,6%); per la Lombardia 14.408 rispetto a 13.772 (+5,1%); nel Veneto 54 in più (+3,7%); nel Lazio 82 in più (+18,6%); nelle Marche 66 in meno (-7,3%). Le differenze vengono giustificate come disallineamenti temporali. Ambedue le serie di dati tengono conto del luogo dove è avvenuto il decesso. D'altronde, il focus delle analisi regionali (fino al livello comunale) sono i contagi, le curve epidemiche, la trasmissibilità; quindi se le morti fossero registrate per luogo di residenza sarebbe introdotta una distorsione<sup>14</sup>. Però, allorché si proceda al confronto con i dati anagrafici, si dovrebbe tenere conto di questa disarmonia e ricostruire serie omogenee.

**Il confronto con gli anni precedenti.** Sulla base della comparazione della mortalità nei primi quattro mesi del 2020 e di quella media nello stesso periodo nei cinque anni precedenti, dati resi disponibili recentemente (Istituto superiore di sanità e Istat, 2020), vengono presentate nel seguito le stime di alcuni aggregati e parametri rilevanti con riferimento a specifici intervalli temporali e domini territoriali di interesse e se ne dà una rappresentazione grafica a livello comunale e di sistemi locali<sup>15</sup>.

Una prima stima effettuata è quella dei decessi per tutte le cause nel corso del primo quadrimestre del 2020 nei 634 comuni per i quali l'Istat non ha ancora prodotto i dati. La stima ha utilizzato i tassi di mortalità specifici per età e genere riferiti ai comuni appartenenti al medesimo sistema locale. Si è quindi ipotizzato che la diffusione del virus per cluster territoriali possa essere stimata sulla base della contiguità geografica e degli spostamenti quotidiani per motivi di lavoro. La stima per i 634 comuni ammonta a 15.423 decessi e porta il totale per l'Italia a 270.020 unità. Considerato che nel quinquennio 2015-2019 i decessi sono stati mediamente 235.690, si sarebbe dunque registrato un aumento pari a 34.330 unità (+14,6%)<sup>16</sup>. Tale dinamica a livello nazionale è fortemente influenzata dalla significativa

riduzione della mortalità registrata in tutto il Paese a gennaio e febbraio 2020 e dalla concentrazione dei decessi per Covid-19 in alcune aree del Nord.

L'eccedenza di morti può essere ricondotta a tre tipologie: i decessi dovuti al Sars-CoV-2 in assenza di altre patologie pregresse; quelli di malati cronici con patologie gravi (in gran parte anziani) per i quali Covid-19 ha rappresentato la causa scatenante della morte; infine, quelli dovuti alla mancata o inadeguata risposta da parte del Servizio sanitario nazionale relativamente a patologie differenti, o al mancato ricorso a esso da parte di tali malati per timore di contagio da Coronavirus<sup>17</sup>. C'è da aspettarsi nei prossimi mesi un rimbalzo negativo dovuto alle morti anticipate del secondo e del terzo gruppo. In ogni caso, i primi due gruppi devono essere ascritti al Covid-19. Il terzo dovrebbe avere una numerosità limitata.

Venendo ai dati e con riferimento al primo quadrimestre del 2020, il diverso bilancio in Italia e in Lombardia tra comuni con eccedenza (o diminuzione) del numero di decessi rispetto agli anni precedenti è illustrato nei due grafici della figura 4.

In Lombardia la frazione dei comuni dove si registra un'eccedenza percentuale di decessi è sensibilmente maggiore in confronto a quella riferita alla restante parte del Paese (85,8% rispetto a 49,2%). In tutti i comuni italiani con eccedenza di mortalità risiedono 36.377.372 persone, delle quali 9.655.310 (26,5%) nei comuni della Lombardia. Nel primo trimestre l'eccedenza percentuale dei decessi in Lombardia era pari a 55,1, più che doppia rispetto al 23,3% per l'intero Paese. L'eccedenza che fino a marzo si apprezzava localmente nelle zone maggiormente colpite ora, aggiungendo il mese di aprile, si manifesta anche a livello nazionale.

Gli effetti letali del Covid-19 si sono concentrati nei mesi di marzo e aprile, attenuati a maggio, e protratti successivamente. Per questo motivo, si presentano anche le stime dei decessi riferite alla fase più acuta (1° marzo-30 aprile 2020). Pure in questo caso, si è proceduto a integrare i dati per i 634 comuni per i quali non stati ancora resi disponibili, con un procedimento identico a quello già illustrato in prece-

<sup>14</sup> Quale denominatore dei tassi di mortalità dovrebbe quindi essere utilizzato più correttamente l'ammontare della popolazione presente, anziché di quella residente; ma si dispone soltanto di quest'ultima.

<sup>15</sup> I sistemi locali (SI) determinati dall'Istat rappresentano una griglia territoriale le cui partizioni, indipendenti dall'articolazione amministrativa del territorio, sono definite utilizzando i flussi degli spostamenti giornalieri casa/lavoro (pendolarismo) rilevati in occasione dei Censimenti generali della popolazione e delle abitazioni. [https://www.istat.it/it/files/2014/12/nota-metodologica\\_SII2011\\_rev20150205.pdf](https://www.istat.it/it/files/2014/12/nota-metodologica_SII2011_rev20150205.pdf).

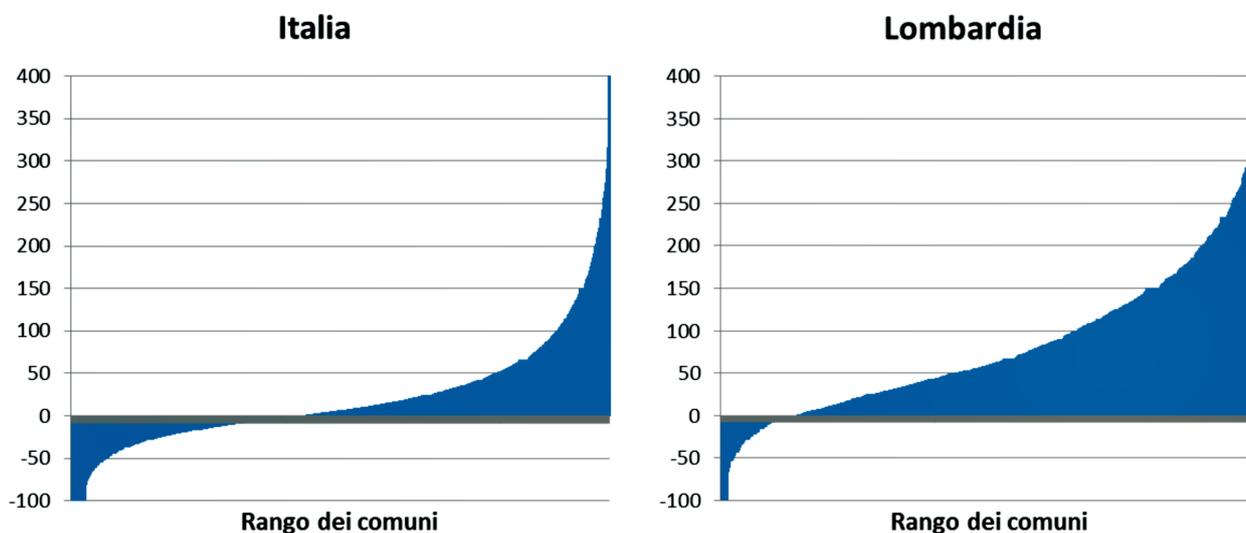
<sup>16</sup> Il confronto relativo ai decessi del primo trimestre effettuato con i dati al-

lora disponibili aveva fornito una variazione negativa: anziché un'eccedenza di mortalità: una diminuzione di 43.229 unità (-17,7%), un risultato in contrasto con le aspettative.

<sup>17</sup> In effetti, il sistema sanitario ha fronteggiato prioritariamente l'epidemia e quindi, prevedibilmente, è diminuita la capacità di risposta ad altre esigenze. La Società italiana di anestesia, analgesia, rianimazione e terapia intensiva (Siaarti), nella fase di massima diffusione del Covid-19, ha elaborato tra gli strumenti per gli anestesisti-rianimatori impegnati le Raccomandazioni di etica clinica per l'ammissione a trattamenti intensivi e per la loro sospensione, in condizioni eccezionali di squilibrio tra necessità e risorse disponibili. Nel documento si af-

Figura 4

**Diminuzione o eccedenza del numero di decessi nei comuni italiani e in quelli lombardi nel primo quadrimestre 2020 rispetto alla media per lo stesso periodo negli anni 2015-2019<sup>1</sup>** (diminuzione o eccedenza percentuale in ordinata; rango dei comuni ordinati secondo la variazione percentuale in ascissa)

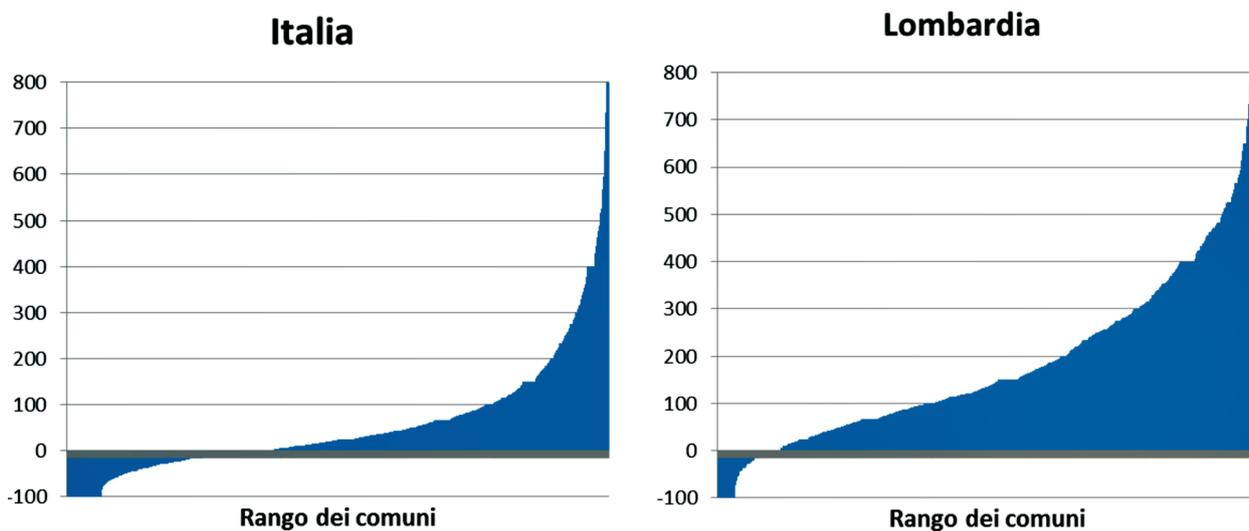


FONTE: NOSTRE STIME SU DATI ISTAT

<sup>1</sup> Sono esclusi alcuni piccoli comuni che presentano un'eccedenza superiore al 400% (13 in Italia dei quali 7 in Lombardia).

Figura 5

**Diminuzione o eccedenza del numero di decessi nei comuni italiani e in quelli lombardi tra il 1° marzo e il 30 aprile 2020 rispetto alla media per lo stesso periodo negli anni 2015-2019<sup>1</sup>** (diminuzione o eccedenza percentuale in ordinata; rango dei comuni ordinati secondo la variazione percentuale in ascissa)



FONTE: NOSTRE STIME SU DATI ISTAT

<sup>1</sup> Sono esclusi alcuni piccoli comuni che presentano un'eccedenza superiore all'800% (41 in Italia dei quali 20 in Lombardia).

ferma che «come estensione del principio di proporzionalità delle cure, l'allocatione in un contesto di grave carenza (shortage) delle risorse sanitarie deve puntare a garantire i trattamenti di carattere intensivo ai pazienti con maggiori possibilità di successo terapeutico: si tratta dunque di privilegiare la "maggior speranza di vita"».

denza. La rappresentazione grafica è stata quindi replicata (figura 5). Il confronto è eloquente. In Lombardia la frazione dei comuni dove si registra un'eccedenza di decessi risulta molto maggiore rispetto a quella riferita al resto del Paese (88,7% rispetto a 56,5%). In tutti i comuni italiani con eccedenza di mortalità risiedono 40.066.781 persone, delle quali 9.770.348 (24,4%) nei comuni della Lombardia. L'eccedenza per l'intero paese ha inizio dal 42-esimo percentile e si innalza sopra 100% dal 78-esimo; per la Lombardia dall'11-esimo e rispettivamente dal 40-esimo.

Sempre con riferimento al periodo acuto nella tavola 6 viene presentata una lettura dell'eccedenza di decessi secondo la dimensione demografica dei comuni. L'evidenza empirica è interessante perché mostra che l'eccedenza è particolarmente ampia nei comuni piccoli fino a 10.000 abi-

tanti: per il 23,0% è superiore al 100% e per l'11,3% superiore al 200%. Le stesse percentuali per i comuni con oltre 100.000 abitanti valgono rispettivamente 11,1% e 4,4%. Dei 131 comuni con una variazione superiore al 500% 127 hanno una popolazione non superiore a 10.000 abitanti, i restanti 4 sono compresi nella classe 10.000-50.000 (in ordine decrescente si tratta dei comuni di Nembro, Alzano Lombardo, Albino, Ponte San Pietro, tutti in Lombardia). Può aver svolto un ruolo importante l'intensità dei rapporti familiari, più forte nei comuni piccoli, tenendo conto che molti contagi sono avvenuti nell'ambiente familiare attraverso persone infette asintomatiche. Questa lettura è coerente con i riscontri per Vo' Euganeo (Crisanti et al., 2020). Nella tavola 7 l'eccedenza (o diminuzione) dei decessi rispetto alla mortalità normale nel periodo 1° marzo-30

Tavola 6

**Numero di comuni secondo la diminuzione o l'eccedenza dei decessi tra il 1° marzo e il 30 aprile 2020 in percentuale rispetto alla media per lo stesso periodo negli anni 2015-2019 e secondo la classe di ampiezza demografica dei comuni di residenza**

Diminuzione o eccedenza di decessi nel 2020 (%)	Classi di ampiezza demografica al 1° gennaio 2019				Totale
	Fino a 10.000	10.001-50.000	50.001-100.000	oltre 100.000	
Fino a -10	2.020	181	16		2.217
-10-0	556	133	23	9	721
0-10	264	107	17	6	395
10-25	655	165	16	11	846
25-50	659	194	14	10	877
50-75	534	78	6	2	621
75-100	430	70	4	2	506
100-150	546	72	3	1	623
150-200	233	32	1	2	268
200-300	314	30	1	2	347
300-500	311	22			333
500-800	86	4			90
800-1.000	28				28
Oltre 1.000	13				13
	19 <sup>1</sup>				19 <sup>1</sup>
<b>Totale</b>	<b>6.669</b>	<b>1.088</b>	<b>101</b>	<b>45</b>	<b>7.904</b>

FONTE: NOSTRE STIME SU DATI ISTAT

<sup>1</sup> Comuni con un numero molto piccolo di residenti che non hanno registrato morti fra il 1° marzo e il 30 aprile negli anni 2015-2019.

Tavola 7

**Diminuzione o eccedenza del numero dei decessi nei comuni italiani tra il 1° marzo e il 30 aprile 2020 in percentuale rispetto alla media per lo stesso periodo negli anni 2015-2019 e secondo la classe di età, il genere e la regione di residenza**

Regioni	Fino a 64 anni		65-74 anni		75-84 anni		85 anni e oltre		Totale	
	M	F	M	F	M	F	M	F	M	F
Piemonte	21,2	2,7	49,8	22,6	59,6	48,1	71,1	60,2	56,9	49,5
Valle d'Aosta	18,5	-22,4	28,0	64,6	88,7	45,6	79,4	85,0	63,3	64,7
Lombardia	94,1	42,7	184,0	108,9	190,3	136,1	167,3	146,1	168,0	132,2
Trentino-Alto Adige	10,2	41,2	45,9	57,8	64,7	74,5	93,1	77,2	63,6	72,4
Veneto	-0,6	-1,6	15,3	5,8	35,0	17,2	43,0	31,7	29,2	23,4
Friuli-Venezia Giulia	-13,3	-2,6	3,3	-2,7	29,7	13,4	19,3	18,3	15,4	13,8
Liguria	17,0	-0,6	62,8	44,8	67,5	56,4	65,3	57,7	60,2	52,6
Emilia-Romagna	51,7	19,7	82,5	56,9	83,0	52,3	66,5	52,0	72,6	50,4
Toscana	5,7	0,9	22,6	8,8	18,5	8,4	21,7	13,5	19,0	11,1
Marche	19,6	25,5	71,1	17,7	61,3	36,6	50,6	28,9	53,1	29,6
Umbria	-7,3	-6,1	-7,3	-4,5	-2,1	-5,4	4,5	9,9	-0,8	4,1
Lazio	-16,7	-17,5	-8,9	-8,9	-5,8	-13,0	0,2	-4,5	-6,0	-8,4
Abruzzo	15,3	22,7	4,6	26,9	2,8	2,9	16,2	14,3	9,8	12,9
Molise	-16,2	19,4	-20,1	-16,8	-11,5	10,4	5,6	2,2	-7,1	3,8
Campania	-5,4	-8,0	12,8	-1,8	-4,0	-14,5	4,0	-0,4	1,3	-5,5
Puglia	6,3	1,2	7,9	13,7	6,1	5,8	20,1	13,6	11,3	10,5
Basilicata	-17,1	-24,5	0,0	13,2	-6,6	-11,5	-2,0	10,3	-5,3	2,1
Calabria	-7,8	-17,4	6,2	10,5	-2,3	-7,8	11,4	6,8	3,1	0,9
Sicilia	-14,8	-8,0	2,9	-1,0	-6,6	-11,9	-0,2	2,5	-3,9	-2,9
Sardegna	-9,9	0,1	-1,2	2,5	10,2	3,2	11,7	23,2	5,1	13,5
<b>Totale</b>	<b>16,0</b>	<b>5,2</b>	<b>45,9</b>	<b>26,7</b>	<b>49,7</b>	<b>31,0</b>	<b>48,5</b>	<b>42,6</b>	<b>43,8</b>	<b>35,2</b>

Fonte: NOSTRE STIME SU DATI ISTAT

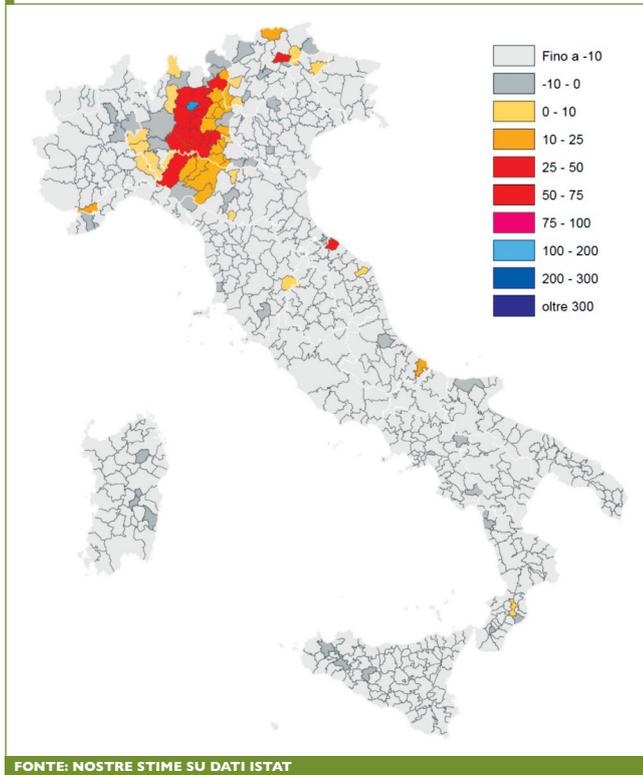
aprile 2020 è esposta in percentuale per genere e classe di età distintamente per regione. Al livello nazionale viene confermata la minore esposizione al virus delle donne: l'eccedenza risulta del 43,8% per gli uomini e del 35,2% per le donne. Il vantaggio femminile tende ad attenuarsi con l'aumentare dell'età: fino a 64 anni l'eccedenza è del 16,0% per gli uomini e del 5,2% per le donne; fra i 65 e i 74 anni le percentuali risultano rispettivamente 45,9% e 26,7%; si avvicinano (49,7% rispetto a 31,0%) per la classe 75-84 anni; il divario si riduce ulteriormente per gli 85-enni e oltre: 48,5% rispetto a 42,6%. Passando al dettaglio regionale, si deve considerare, in primo luogo, che le regioni piccole (Valle d'Aosta, Friuli Venezia Giulia, Umbria, Molise, Basi-

licata) esibiscono andamenti non regolari in ragione del limitato numero di decessi che le interessa. In quelle maggiormente colpite dall'epidemia (Lombardia, Piemonte, Emilia Romagna, Liguria) le differenze di genere diminuiscono molto. Lazio, Campania e Sicilia presentano, per molte classi di età e, salvo la Campania, per ambedue i generi, riduzioni di mortalità rispetto a quella considerata normale. Questo suggerisce di riferirsi a valori di base differenziati temporalmente per le diverse aree territoriali. È un esercizio che verrà proposto più avanti.

Gli andamenti appena descritti possono essere ricondotti a una molteplicità di cause. Nelle regioni più colpite la popolazione è mediamente più anziana e, nelle età più avan-

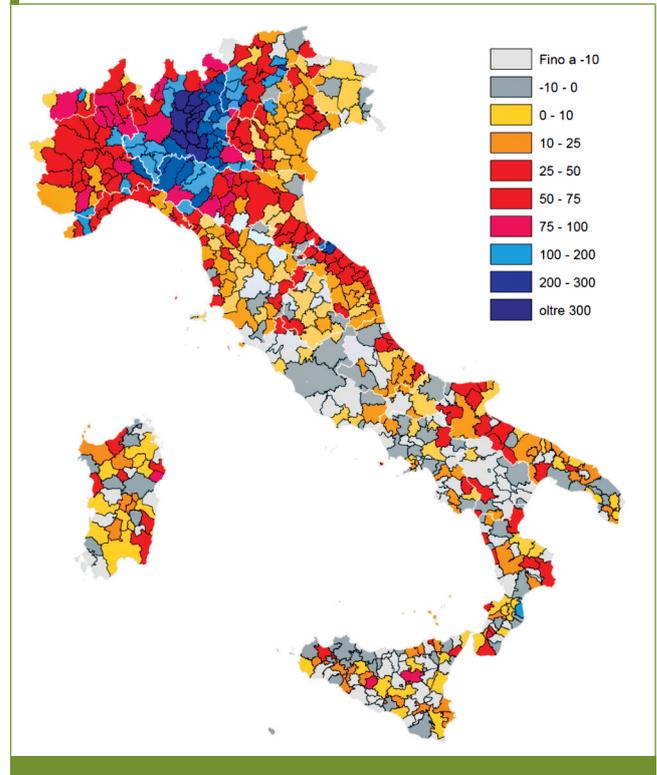
Cartogramma 1

**Diminuzione o eccedenza percentuale dei decessi per tutte le cause nel primo trimestre del 2020 rispetto alla media per lo stesso periodo negli anni 2015-2019 nei 610 sistemi locali**



Cartogramma 2

**Diminuzione o eccedenza percentuale dei decessi per tutte le cause tra il 1° marzo e il 15 aprile 2020 rispetto alla media per lo stesso periodo negli anni 2015-2019 nei 610 sistemi locali**



zate, la componente femminile è maggioritaria. Inoltre, molte morti per Covid-19, spesso non certificate come tali per mancanza dell'accertamento del contagio, sono avvenute nelle residenze sanitarie assistite nelle quali la presenza femminile è prevalente.

La proporzione per genere esibita per l'intero Paese dall'eccedenza di mortalità nel periodo marzo-aprile 2020 rispetto agli anni precedenti (uomini 56,7%; donne 43,3%) può essere confrontata con quella registrata per i decessi certificati positivi da Covid-19 (uomini 59,2%; donne 40,8%)<sup>18</sup>. La prossimità delle proporzioni conforta nel ricondurre l'eccedenza essenzialmente all'epidemia<sup>19</sup>.

Una lettura interessante della mortalità prevalentemente riconducibile a Covid-19 è offerta nei cartogrammi 1 e 2. Vengono considerate le partizioni funzionali del territorio

delineate dai sistemi locali. Le differenze fra primo trimestre e periodo di acuzie (marzo-metà aprile) sono eclatanti: la macchia cromatica esplose nel Nord-Ovest, in particolare in Lombardia e Piemonte, alle quali si aggiunge l'Emilia Romagna.

Nella tavola 8 vengono stimati, per vari domini territoriali, sia la mortalità certificata per Covid-19, sia l'eccedenza del numero di decessi nel bimestre marzo-aprile 2020 rispetto alla media nello stesso periodo dei cinque anni precedenti (2015-2019), sia i tassi di mortalità (decessi per 1.000.000 di residenti) riferiti alla mortalità certificata per Covid-19 e all'eccedenza di mortalità rispetto al livello normale, che può essere considerata un limite superiore di quella specifica<sup>20</sup>. I tassi non sono standardizzati per tenere conto della differente composizione per età delle popolazioni nei di-

<sup>18</sup> Dati aggiornati al 26 maggio 2020 (Istituto superiore di sanità, 2020).

<sup>19</sup> «Una misura più universale dell'impatto dell'epidemia sulla mortalità della popolazione è data dall'eccesso dei decessi per il complesso delle cause, risultanti dal confronto, a parità di periodo, del dato del 2020 con la media dei decessi del quinquennio precedente (2015-2019)» (Istituto Superiore di Sanità e Istat, 2020).

<sup>20</sup> Sulla stima dell'eccedenza così determinata può avere una qualche influenza la diminuzione del numero di morti per incidenti stradali, infortuni sul lavoro e altre cause sopite durante il periodo di lockdown.

Tavola 8

**Decessi certificati positivi da Covid-19 e diminuzione o eccedenza dei decessi per tutte le cause tra il 1° marzo e il 30 aprile 2020 rispetto alla media per lo stesso periodo negli anni 2015-2019 per alcuni domini territoriali di stima<sup>1</sup>**

Domini territoriali di stima	Decessi				Tassi di mortalità per Covid-19 (per 1.000.000 di residenti)	
	(1)	(2)	(3) = (2)-(1)	(4) = [(1)/(2)] x 100	(5) = [(1)/Pop] x 1.000.000	(6) = [(2)/Pop] x 1.000.000 (limite superiore)
	Certificati positivi da Covid-19	Diminuzione o eccedenza rispetto agli anni precedenti	Differenza	Certificati positivi da Covid-19 su eccedenza	Per i decessi certificati positivi da Covid-19	Per il totale stimato dei decessi da Covid-19
Italia	27.846	34.330	6.484	81,1	462	569
<i>Ripartizioni geografiche</i>						
Nord-Ovest	18.017	28.688	10.671	62,8	1.119	1.782
Nord-Est	6.039	7.679	1.640	78,6	518	659
Centro	2.280	-386	-	-	190	-
Sud	1.165	-725	-	-	83	-
Isole	345	-927	-	-	52	-
<i>Regioni (ordinate secondo l'eccedenza o diminuzione di decessi per tutte le cause rispetto agli anni precedenti)</i>						
Lombardia	13.749	23.491	9.742	58,5	1.366	2.334
Emilia-Romagna	3.547	4.593	1.046	77,2	795	1.030
Piemonte	3.066	3.588	522	85,5	704	824
Veneto	1.457	1.767	310	82,4	297	360
Liguria	1.167	1.477	310	79,0	753	953
Trentino-Alto Adige	693	1.087	394	63,8	646	1.014
Marche	906	1.032	126	87,8	594	676
Toscana	842	482	-360	174,7	226	129
Puglia	415	341	-	-	103	-
Friuli-Venezia Giulia	289	232	-	-	238	-
Sardegna	116	199	83	58,3	71	121
Abruzzo	320	157	-163	203,5	244	120
Valle d'Aosta	137	132	-	-	1.093	-
Basilicata	25	-50	-	-	44	-
Molise	21	-119	-	-	69	-
Umbria	67	-161	-	-	76	-
Calabria	86	-252	-	-	44	-
Campania	359	-803	-	-	62	-
Sicilia	235	-1.125	-	-	47	-
Lazio	441	-1.739	-	-	75	-

FONTE: NOSTRE STIME SU DATI ISTAT E ISTITUTO SUPERIORE DI SANITÀ

<sup>1</sup> Le stime dei decessi a livello nazionale, di ripartizione geografica e regionale riguardano la totalità dei comuni.

versi domini territoriali, procedimento utile per migliorare il confronto<sup>21</sup>. In effetti, quello che interessa qui soprattutto è valutare l'impatto dell'epidemia sulla mortalità e quindi stimare i tassi di mortalità nei differenti contesti territoriali. Ebbene, per l'intero Paese, il numero dei decessi aumenterebbe dai 27.846 certificati positivi a 34.330 (+23,3%); nella ripartizione Nord-Ovest l'aumento sarebbe del

59,2%; in Lombardia, la regione più colpita, del 70,9%; in l'Emilia Romagna del 29,5%. Per la ripartizione Sud e per alcune regioni si registra una diminuzione anziché un'eccedenza di decessi rispetto ai livelli considerati normali. Dati questi limiti di validità, che si tenterà di superare più avanti, il numero di decessi per 1.000.000 di abitanti residenti passerebbe per l'Italia da 452 a 569; quello per il

Tavola 9

**Decessi certificati positivi da Covid-19 e diminuzione o eccedenza dei decessi per tutte le cause nel primo quadrimestre del 2020 rispetto alla media per lo stesso periodo negli anni 2015-2019 per le province caratterizzate dalla maggiore eccedenza di decessi<sup>1</sup>**

Domini territoriali di stima	Decessi				Tassi di mortalità per Covid-19 (per 1.000.000 di residenti)	
	(1)	(2)	(3) = (2)-(1)	(4) = [(1)/(2)] x 100	(5) = [(1)/Pop] x 1.000.000	(6) = [(2)/Pop] x 1.000.000 (limite superiore)
	Certificati positivi da Covid-19	Diminuzione o eccedenza rispetto agli anni precedenti	Differenza	Certificati positivi da Covid-19 su eccedenza	Per i decessi certificati positivi da Covid-19	Per il totale stimato dei decessi da Covid-19
<i>Province (ordinate secondo l'eccedenza di decessi &gt; 600)</i>						
Bergamo	2.994	5.996	3.002	49,9	2.686	5.380
Milano	3.450	4.512	1.062	76,5	1.059	1.386
Brescia	2.466	4.208	1.742	58,6	1.948	3.324
Cremona	1.038	1.996	958	52,0	2.892	5.560
Pavia	1.047	1.504	457	69,6	1.918	2.755
Parma	704	1.317	613	53,5	1.559	2.916
Monza e della Brianza	751	1.249	498	60,1	859	1.429
Torino	962	1.190	228	80,8	426	527
Piacenza	884	1.129	245	78,3	3.079	3.932
Genova	663	930	267	71,3	788	1.106
Lodi	658	906	248	72,7	2.858	3.934
Lecco	413	826	413	50,0	1.225	2.451
Mantova	616	801	185	76,9	1.494	1.944
Alessandria	421	752	331	55,9	999	1.785
Como	485	675	190	71,8	809	1.127
Pesaro e Urbino	460	660	200	69,7	1.282	1.839
Reggio nell'Emilia	508	640	132	79,4	955	1.203

FONTE: NOSTRE STIME SU DATI ISTAT E PROTEZIONE CIVILE

<sup>1</sup> Le stime dei decessi a livello provinciale riguardano la totalità dei comuni.

<sup>21</sup> Se si procedesse alla standardizzazione si avvantaggerebbero nel confronto i domini territoriali con popolazione mediamente più anziana e viceversa.

Nord-Ovest da 1.119 a 1.782 e per la Lombardia da 1.366 a 2.334.

Al livello provinciale (tavola 9) Bergamo, Lecco, Brescia, Cremona, Parma e Alessandria registrano una mortalità molto più elevata rispetto a quella registrata ufficialmente per Covid-19. Il tasso di mortalità per Bergamo pari a 2.686 morti per 1.000.000 di abitanti passerebbe a 5.380 (+100,3%); aumenti analoghi risulterebbero a Lecco (+100,1%) e Cremona (+92,3%). Parma, Alessandria e Brescia registrerebbero pure aumenti assai consistenti.

Data la presenza dell'epidemia, non ci si aspetterebbe che il confronto della mortalità nei primi quattro mesi o nel bimestre marzo-aprile del 2020 con quella registrata mediamente negli stessi periodi del quinquennio 2015-2019 presentasse casi di diminuzione anziché eccedenza di decessi. Questo è certamente possibile in comuni di modeste dimensioni demografiche ma improbabile per grandi aree territoriali, come invece risulta distintamente per Centro, Sud e Isole nel confronto sia fra quadrimestri sia riferito al periodo acuto. Una possibile spiegazione, come si è già accennato, risiede nella circostanza che nei mesi di gennaio e febbraio del 2020 si è registrato un numero di decessi per tutte le cause estremamente basso rispetto agli stessi mesi dei cinque anni precedenti e tale andamento sarebbe prevedibilmente proseguito anche nei mesi successivi. Ad esempio, nel mese di gennaio 2020 i morti sono stati 61.743, un numero inferiore del 5,9% rispetto alla media dei cinque anni precedenti, pari a 65.611. Così pure risulta per il mese di febbraio con 55.548 morti nel 2020 (-2,7% rispetto alla media dei cinque anni precedenti, pari a 57.062). Per una stima migliore dell'eccedenza di mortalità nel periodo 1° marzo-30 aprile 2020 può quindi essere utile individuare, per le singole ripartizioni geografiche, un anno (o un gruppo di anni) meglio paragonabile rispetto alla media dei cinque anni precedenti<sup>22</sup>. L'ipotesi sottesa è che i regimi di mortalità nelle diverse aree territoriali del Paese siano differenti anche nello stesso anno. Sono state esplorate due diverse possibilità:

**Ipotesi A:** stima dell'eccedenza (o eventualmente diminuzione) di mortalità rispetto a quella normale, per ciascuna

ripartizione geografica, nei mesi di marzo e aprile 2020, mediante la differenza fra il numero di morti effettivo per tutte le cause e quello per gli stessi mesi di uno specifico anno (o gruppo di anni) individuato come paragonabile attraverso la comparazione delle serie mensili dei decessi fra gennaio 2014 e febbraio 2020;

**Ipotesi B:** si procede come per l'ipotesi A per individuare l'anno (o il gruppo di anni) paragonabile; l'eccedenza (o eventualmente la diminuzione) di mortalità viene stimata assumendo che fra febbraio e aprile 2020 si sia registrata la stessa dinamica temporale sperimentata nell'anno (o negli anni) paragonabile per ciascuna ripartizione geografica.

L'illustrazione delle ipotesi con qualche dettaglio metodologico è svolta nell'Appendice 2.

I risultati ottenuti sono compendati nella tavola 10 nella quale vengono riportati, per confronto, il tasso di mortalità certificato da Covid-19 e la stima ottenuta nella tavola 8. Assumendo ancora l'eccedenza dei decessi come limite superiore del numero dei morti per Covid-19, esso per l'intero Paese risulterebbe pari a circa 46.000 per il bimestre marzo-aprile 2020, il 65,2% in più di quello certificato e il 34,0% in più rispetto all'eccedenza stimata mediante il confronto con la media dei cinque anni precedenti (2015-2019). Il tasso di mortalità (numero di morti per 1.000.000 di residenti), pure per il periodo 1° marzo-30 aprile 2020, risulterebbe a livello nazionale circa 1,7 volte più elevato di quello calcolato considerando le sole morti certificate per Covid-19. La nuova stima di eccedenza penalizza ulteriormente il Nord-Ovest dove il limite superiore del tasso di mortalità si attesterebbe intorno a 2.040 morti per 1.000.000 di abitanti (2,7 volte più elevato rispetto a quello per l'intera Italia). Il tasso per il Nord-Est risulterebbe pari a poco più di 820 per 1.000.000, lievemente più elevato di quello medio italiano. Nel Centro e nel Sud tasso minimo (per i decessi certificati positivi da Covid-19) e tasso massimo di mortalità (calcolato sull'eccedenza di decessi rispetto alla mortalità normale) risultano molto bassi. Per quanto riguarda le Isole, occorre considerare che il numero di morti è molto inferiore rispetto alle altre ripartizioni territoriali e che Sicilia e Sardegna risultano alquanto diverse

<sup>22</sup> La media quinquennale è stata adottata, comunque con buone ragioni, dall'Istat (Istituto superiore di sanità e Istat, 2020).

Tavola 10

**Decessi certificati positivi da Covid-19 ed eccedenza dei decessi per tutte le cause nel periodo 1° marzo-30 aprile 2020 secondo differenti ipotesi**

Ripartizioni geografiche	Decessi Certificati positivi da Covid-19	Eccedenza dei decessi per tutte le cause (limite superiore dei decessi da Covid-19)		Tassi di mortalità da Covid-19 (per 1.000.000 di residenti)								
	(1)	(2)	(3)	(4)	(5) = [(1)/Pop]* 1.000.000		(6) = [(2)/Pop]* 1.000.000a,b		(7) = [(3)/Pop]* 1.000.000b		(8) = [(4)/Pop]* 1.000.000b	
	Mortalità certificata	Confronto con media 2015-2019 <sup>1</sup>	Ipotesi A	Ipotesi B	Certificato	Ipotesi di tabella 8	Ipotesi A	Ipotesi B	stima	N.I. (Italia = 100)	stima	N.I. (Italia = 100)
					Stima	N.I. (Italia = 100)	Stima	N.I. (Italia = 100)	Stima	N.I. (Italia = 100)	Stima	N.I. (Italia = 100)
Nord-Ovest	18.017	28.688	32.945	32.829	1.119	242,2	1.782	313,2	2.047	268,3	2.040	262,9
Nord-Est	6.039	7.679	9.476	9.701	518	112,1	659	115,8	813	106,6	833	107,3
Centro	2.280	-386	1.891	1.919	190	41,1	-32	-	157	20,6	159	20,5
Sud	1.165	-725	1.654	1.512	83	18,0	-52	-	120	15,7	110	14,2
Isole	345	-927	-94	681	52	11,3	-140	-	-	-	102	13,1
Italia	27.846	34.330	45.872	46.642	462	100,0	569	100,0	763	100,0	776	100,0

FONTE: NOSTRE STIME SU DATI ISTAT, ISTITUTO SUPERIORE DI SANITÀ E PROTEZIONE CIVILE

<sup>1</sup> Dati dalla colonna 2 della tavola 8.

<sup>2</sup> Le tre ipotesi assumono che tutta l'eccedenza di mortalità sia da attribuire all'epidemia, cosicché i tassi di mortalità rappresentano quindi un limite superiore.

fra loro. Non è apparso opportuno procedere a una stima separata, data la natura esplorativa dell'esercizio.

Qualche dato comparativo internazionale è stato reso disponibile dall'*Economist* (*The Economist*, 2020; Tozer, 2020): nella città di New York tra metà marzo e il 4 aprile ci sono stati 5.099 morti in eccesso rispetto agli anni precedenti ed è stato stimato che il 98% fosse da attribuire al coronavirus; in Francia il 78%, cioè 8.886 dell'eccedenza pari a 11.417. In Svezia, dove l'eccedenza è molto minore, la percentuale è 71%. Tutte le evidenze confermano un numero sommerso di decessi per Covid-19 molto consistente. Sarà comunque difficile fare un bilancio preciso a livello mondiale e anche italiano poiché i dati presentano margini di imprecisione relativamente elevati, soprattutto con riferimento alle prime fasi dell'epidemia.

Una riflessione a parte meritano la diffusione del contagio e la mortalità nelle residenze sanitarie assistite, Rsa, dove, in base ai dati resi noti, c'è stata una strage<sup>23</sup>. Una strage annunciata: il virus colpisce gli anziani; li colpisce in modo più violento; colpisce i più fragili, già con compromissioni

di salute importanti; si propaga rapidamente in ambienti comunitari. Gli ospiti delle residenze erano quindi vittime predestinate. L'epidemia ha fatto emergere l'inadeguatezza generale delle Rsa alla quale si sono aggiunti la sottovalutazione iniziale della gravità del Covid-19, l'impreparazione a fronteggiare in quegli ambienti situazioni di emergenza diffusa, i ritardi nelle decisioni e talvolta gli errori, la carenza di dispositivi di protezione per gli operatori e per gli ospiti. Poiché la domanda di ricovero nelle residenze tenderà a crescere nei prossimi anni per il progressivo invecchiamento della popolazione, occorre procedere rapidamente a un loro adeguamento strutturale per il rispetto e la riconoscenza che si deve alla generazione dei genitori e a quelle precedenti ancora in vita.

**Il tasso di letalità.** La sottostima dei contagi e dei morti per coronavirus, peraltro in proporzione differente, impedisce di stimare correttamente il tasso di letalità (numero di morti/numero di contagiati conclamati e asintomatici da Covid-19). Disponendo di dati completi, sarebbe utile det-

<sup>23</sup> L'evidenza non è soltanto italiana. L'Organizzazione Mondiale della Sanità, a fine aprile, ha riferito che la metà dei decessi per coronavirus avvenuti in Europa si sarebbe verificata in residenze per anziani.

tagliare il tasso per i due gruppi e per altre collettività di interesse, fra le quali certamente i ricoverati nelle residenze sanitarie assistite e, in generale, nelle residenze per anziani. I dati attualmente disponibili non consentono stime accurate. Anche per questo motivo i tassi di letalità sono estremamente variabili paese da paese, territorio da territorio, situazione da situazione.

John Ioannidis, professore presso la School of Medicine dell'Università di Stanford, ha stimato il rischio di morte per coronavirus in alcune situazioni sulla base dei dati di 11 paesi europei, Canada e 12 stati degli Usa. Ecco qualche risultato: 1,7 morti per milione di ultrasessantacinquenni in Germania e 79 per milione a New York; per gli ultraottantenni uno ogni 6.000 in Germania e uno ogni 420 in Spagna. Le persone di età inferiore a 65 anni avrebbero un rischio di morte inferiore da 36 a 84 volte rispetto a quelle più anziane. Sulla Diamond Princess, il tasso di contagio è stato del 19% e il tasso di letalità dell'1% (Ioannidis, 2020). L'incertezza rimane molto forte.

#### 4. Raccogliere e rendere disponibile informazione affidabile

Un'indagine campionaria rigorosa, svolta al momento dell'insorgere dell'epidemia nel paese, avrebbe messo a disposizione dati affidabili sui quali basare le decisioni e utili per convincere la popolazione e le imprese ad adottare comportamenti restrittivi. L'abbiamo proposta insieme ad altri tre statistici (Alleve et al., 2020 a) con l'aspettativa che il Ministero della salute, la Protezione civile, l'Istituto superiore di sanità e l'Istituto nazionale di statistica potessero farla propria e procedere<sup>24</sup>.

Il suggerimento era quello di riferirsi a due campioni: il primo estratto a partire dalla popolazione dei contagiati conclamati, ricoverati presso strutture ospedaliere o in quarantena presso la propria abitazione, riferito ai loro contatti nei quattordici giorni precedenti l'accertamento del contagio; il secondo relativo alla restante popolazione, un panel da seguire nel corso del tempo a cadenze temporali fissate. Per individuare il primo campione si sarebbe potuto avvalersi

degli elenchi dei contagiati che si fossero riferiti a presidi sanitari all'interno di finestre temporali fissate; in mancanza di elenchi, si sarebbe potuto utilizzare un campione di strutture ospedaliere (unità di primo stadio) e quindi selezionare i campioni di contagiati (unità di secondo stadio). Tramite questi, ripercorrendo le loro reti sociali<sup>25</sup>, si sarebbe arrivati ai contatti nei quattordici giorni precedenti fra i quali estrarre i campioni di interesse (unità di terzo stadio). Su questi ultimi campioni sarebbero stati effettuati i test diagnostici; i test si sarebbero potuti ripetere più volte nel corso del tempo. Il secondo campione, il panel, sarebbe stato selezionato mediante un pre-screening riferito a quattro categorie di soggetti. Un primo gruppo composto da persone ad alto rischio, in prima approssimazione con età superiore a 65 anni, diviso in due sottogruppi: persone con attività intensa e contatti frequenti nel corso della giornata e persone con vita più ritirata e pochi contatti interpersonali. Un secondo gruppo composto da persone a basso rischio, con età fino a 65 anni, distinto nei medesimi due sottogruppi. Il pre-screening avrebbe consentito l'utilizzazione di tassi variabili di campionamento per i quattro sottogruppi, opportuna tenuto conto della comunque limitata presenza di contagiati nella popolazione generale, e avrebbe consentito di migliorare l'efficienza delle stime. Per il panel era prevista, nel corso del periodo di osservazione, una quota di ricambi, necessari per sostituire le persone nel frattempo contagiate o che si fossero rifiutate di proseguire. Anche sul secondo campione (popolazione altra) si sarebbero effettuati opportuni esami diagnostici (virologici e/o sierologici, da decidere in relazione alle finalità perseguite); quelli virologici ripetuti a distanza di quattordici giorni uno dall'altro per il periodo di tempo ritenuto necessario in base all'evoluzione dell'epidemia<sup>26</sup>. Le stime statistiche degli aggregati di interesse sarebbero state prodotte con il concorso di ambedue i campioni<sup>27</sup>. La proposta delinea un protocollo di osservazione in grado di fornire stime statisticamente affidabili su variabili rilevanti dell'epidemia, uno strumento dinamico di monitoraggio, con la possibilità di accentuare l'osservazione sul primo campione nel caso si avessero segnali del riaccutizzarsi del contagio o dell'arrivo di una seconda ondata e sul secondo

<sup>24</sup> La proposta è rintracciabile in italiano nel sito [https://web.uniroma1.it/memoteff/sites/default/files/Proposta\\_0.pdf](https://web.uniroma1.it/memoteff/sites/default/files/Proposta_0.pdf) e in inglese nel sito <http://arxiv.org/abs/2004.0606>. A essa ha dato risonanza un articolo di Dario Di Vico sul *Corriere della Sera* (Di Vico, 2020). Molti consensi sono stati manifestati agli autori (Eurostat, Instituto nacional de estadística, Commissione europea, World Bank, ricercatori di università e centri di ricerca italiani e stranieri, dell'Istat e della Banca d'Italia).

<sup>25</sup> Le reti sociali da considerare si distinguono usualmente in reti primarie (parenti, amici, vicini di casa, ecc.), reti secondarie (colleghi di lavoro, commessi di negozi alimentari, farmacisti etc.), reti formali (medici, forze dell'ordine, ope-

ratori sanitari etc.), reti informali (comunità, gruppi di aggregazione, ecc.) (Ascani, 2020; Yang et al., 2016).

<sup>26</sup> Non siamo i soli a ritenere importante una procedura campionaria rigorosa che consenta di fare inferenza sull'intera popolazione. Lo hanno affermato, fra gli altri, Enrico Bucci, biologo dell'Università di Philadelphia e Roberto Burioni insieme ai medici riuniti sul sito Medical Facts.

<sup>27</sup> Esse godono di proprietà piuttosto importanti: la non distorsione, e l'efficienza (Alleve et al., 2020 b).

campione, il panel, per valutare gli andamenti in particolare nel corso della fase 2 e di quelle successive. Ai soggetti dei diversi campioni interessati dall'indagine (contagiati conclamati e loro contatti e componenti del panel nell'occasione di ingresso) sarebbe stato somministrato un questionario volto a raccogliere informazioni individuali e famigliari utili per la comprensione della diffusione dell'epidemia e delle misure da adottare nella sua fase ascendente o discendente)<sup>28</sup>. La rilevazione campionaria rappresenta un tassello essenziale di un sistema informativo integrato su Covid-19 nel quale far convergere, inoltre, le informazioni sanitarie, importanti per valutare la tenuta del sistema, e quelle tecnologiche (flusso di dati riferiti alla presenza del contagio nelle persone e al tracciamento dei loro contatti negli  $x$  giorni precedenti, acquisibile attraverso i telefoni cellulari, con la necessaria tutela della riservatezza)<sup>29</sup>. A esse dovrebbero aggiungersi quelle, pure molto importanti, raccolte a livello di micro area di contagio (le così dette «zone rosse»)<sup>30</sup>. Tutti questi dati, opportunamente raccordati, possono fornire un quadro informativo coerente di sorveglianza sull'epidemia sul quale fondare le decisioni delle autorità competenti e motivare i comportamenti dei cittadini e delle imprese.

L'approccio campionario è l'unico che consente di sfruttare un consolidato paradigma inferenziale per avanzare affermazioni sulle caratteristiche della popolazione, affiancandovi gradi prefissati di fiducia. D'altra parte, l'indagine campionaria può avere due distinte finalità principali. La prima è di natura conoscitiva a sostegno di decisioni, come per la nostra proposta. In questo caso, il disegno di campionamento e la dimensione del campione sono definiti prefissando la misura dell'accuratezza delle stime dei parametri d'interesse (ad esempio, il tasso di contagio, di letalità o di immunizzazione). La seconda finalità è di natura più strategica, intesa a verificare che il tasso di trasmissione del contagio  $R$  raggiunga o si mantenga entro un certo limite in un determinato arco di tempo. In questo secondo caso, il disegno di campionamento, oltre che considerare la popolazione generale, ne deve prevedere una stratificazione opportuna secondo categorie con rischio differente. Suggerimenti in questa direzione sono rintracciabili nei contributi di David Vines e di

Paul Romer per un campionamento massivo negli Stati Uniti. Sul tema si è acceso un vivace dibattito fra i sostenitori di ciascuna delle due possibili alternative (Cleevely et al., 2020; Romer e Shah, 2020).

La nostra proposta ha contribuito a smuovere le acque. La consapevolezza che la statistica debba essere coinvolta nell'azione di sorveglianza sull'epidemia si è fatta strada. A metà aprile, l'Istat<sup>31</sup> è stato indicato per predisporre un piano di campionamento importante, di 150.000 unità, alle quali chiedere alcune informazioni e sulle quali effettuare l'analisi sierologica<sup>32</sup>. L'ampiezza e la complessità organizzativa della rilevazione renderanno disponibili dati verificati ed elaborati non prima di tre-quattro mesi. Disporremo allora di una fotografia molto distante nel tempo.

Abbiamo avanzato in diverse occasioni qualche suggerimento migliorativo riguardo alla rilevazione. Si sarebbe potuto scegliere una parte delle unità campionarie previste, all'incirca 15-20.000, rappresentative della popolazione residente a livello nazionale e costruire un campione «anticipatorio». Si sarebbe potuto elaborare rapidamente i dati rilevati su di esse, offrendo così una prima rappresentazione affidabile della situazione relativa ai contagi avvenuti e risolti positivamente (se si fosse effettuato il solo test sierologico). Si sarebbe potuto utilizzare il campione anticipatorio come panel da seguire nel corso del tempo con scansione opportuna. Sulle unità del panel si sarebbe potuto effettuare oltre all'esame sierologico anche quello per tampone in modo da accertare l'eventuale sopravvenuto contagio successivo. Questo sarebbe risultato particolarmente importante se si fosse riattivata la virulenza o fosse arrivata una seconda ondata dell'epidemia. Si sarebbe potuto fare, secondo noi, un servizio maggiormente utile al Paese, senza un aggravio elevato dei costi. Parecchi dei miglioramenti suggeriti, d'altronde di buon senso, sono presenti nel «Protocollo metodologico per un'indagine di siero-prevalenza sul Sars-CoV-2 condotta dal Ministero della Salute e dall'Istat» (Ministero della Salute e Istat, 2020).

A nostro modo di vedere, si sarebbe dovuto imporre l'obbligo di risposta, giustificato dall'esigenza di tutela della salute pubblica. Si sarebbe dovuto fare un'azione di comunicazione

<sup>28</sup> Tra le informazioni individuali, oltre al genere e all'età, rivestono importanza il titolo di studio, la condizione professionale, il tipo e la frequenza degli spostamenti, il luogo di residenza e quello eventuale di lavoro o studio; tra le variabili di carattere famigliare, il tipo e la composizione del nucleo, il tipo e la frequenza degli spostamenti.

<sup>29</sup> L'opportunità tecnologica è stata sfruttata con successo nella Corea del Sud.

<sup>30</sup> È il caso, ad esempio, della ricerca condotta sulla popolazione di Vo' Euganeo da parte dell'Università di Padova e dell'Imperial College di Londra, coordinata da Andrea Crisanti, infettivologo presso l'Università di Padova, il quale ha molto

influenzato le misure di contrasto all'epidemia nella Regione Veneto. Uno dei risultati ottenuti è che la maggior parte dei contagi in quel comune è avvenuta in ambiente famigliare attraverso persone asintomatiche; e il 43,2% dei casi confermati dai tamponi era relativo ad asintomatici.

<sup>31</sup> Ricordiamo, incidentalmente, che il Presidente dell'Istat non è stato indicato fra i 70 componenti della Task force dati che il decreto Cura Italia ha previsto di istituire presso il Ministro dell'Innovazione. In generale gli Istituti nazionali di statistica europei non hanno assunto tempestivamente il ruolo che sarebbe stato necessario – e naturale – per produrre un'informazione rigorosa sull'epidemia. Eurostat ha dato segno della sua presenza sul tema soltanto il 10 aprile

molto più incisiva e convincente. Per avvicinare e motivare le persone facenti parte del campione a prendere parte alla rilevazione si sarebbe potuta usare la rete dei medici di famiglia piuttosto che la Croce rossa, benemerita in parecchie emergenze del Paese, ma relativamente poco organizzata per un'impresa come questa, molto impegnativa sul piano dimensionale ed estremamente complessa. A rilevazione ormai iniziata, l'auspicio è che possa essere garantita la necessaria qualità dei dati raccolti e che essi possano essere valorizzati rendendoli completamente disponibili senza limitazioni alla comunità scientifica.

## 5. Conclusioni

Abbiamo espresso un punto di vista da statistici, quali siamo ambedue, consapevoli che per produrre una informazione corretta su Covid-19 sia necessaria la collabora-

zione di molte competenze: virologi, epidemiologi, responsabili sanitari, modellisti fisici e matematici, economisti, valutatori di interventi pubblici, oltre che di statistici. Occorre una reale intenzione interdisciplinare. È necessaria anche una forte collaborazione inter-istituzionale al centro e sul territorio. Purtroppo, la presenza di decisioni che appaiono in qualche caso improvvisate o non sostenute dal conforto scientifico, la scarsa trasparenza delle proposte avanzate dalle molte e spesso pletoriche task force, l'incapacità di predisporre soluzioni selettive riguardo sia alle misure restrittive<sup>33</sup> sia, in molti casi, ai beneficiari delle provvidenze apprestate, la fiducia immotivata che i problemi possano essere risolti attraverso la sola decretazione, la tradizionale allergia per gli aspetti organizzativi delle soluzioni indicate, l'insofferenza dei comparti produttivi e – c'è da aspettarselo – della popolazione, lasciano molta incertezza.

### Appendice 1

#### Stima dell'eccedenza di mortalità rispetto al livello normale nel periodo 1° marzo-30 aprile 2020

Il problema che si pone è quello di stimare il numero totale di decessi per Covid-19 nel periodo 1° marzo-30 aprile 2020, compresa la parte che non emerge dalle rilevazioni condotte dalla Protezione civile attraverso le comunicazioni giornaliere regionali. Queste danno conto con sufficiente precisione dei decessi avvenuti fra i ricoverati presso le strutture ospedaliere ma non dei morti nelle proprie abitazioni o nelle residenze per anziani ai quali non sia stato occasionalmente effettuato il tampone.

La differenza fra il numero di decessi per tutte le cause nel periodo 1° marzo-30 aprile 2020 con quello che si può considerare normale, cioè in assenza di Covid-19, consente di stimare il limite superiore della mortalità riconducibile, in qualche misura, all'epidemia.

È quindi necessario valutare il livello di mortalità normale nei mesi di marzo e aprile 2020. In proposito, occorre tenere conto che il 2020 è iniziato con un numero di decessi decisamente più basso rispetto agli anni precedenti. Nel mese di gennaio i morti sono stati 61.743, un numero inferiore del 5,9% rispetto alla media dei sei anni 2014-2019, ai quali ci siamo riferiti, che risulta pari a 65.611. Così pure è avvenuto per il mese di febbraio con 55.548 morti (-2,7% rispetto alla media dei sei anni precedenti, pari a 57.062). Per procedere alla stima della mortalità normale occorre quindi individuare un anno (o un gruppo di anni) paragonabile al 2020. Si propongono due ipotesi diverse, A e B. Per ambedue, si parte dai dati mensili di mortalità relativi alle cinque ripartizioni territoriali (Nord-Ovest, Nord-Est, Centro, Sud, Isole) per gli anni 2014-2020. Si considerano i decessi relativi ai periodi «settembre di un dato anno-febbraio dell'anno successivo» nel presupposto che durante essi si esprima una mortalità con caratteristiche differenti per le diverse ripartizioni geografiche. Non si va oltre febbraio in quanto a marzo del 2020 è esplosa la mortalità per Covid-19. Inoltre, poiché non si dispone ancora della stima ufficiale dei morti nel mese di dicembre 2019, non si tiene conto di tale mese nella comparazione fra le serie mensili di periodo.

Per selezionare l'anno o eventualmente il gruppo di anni meglio paragonabili con il 2020 è stato definito un indice di dissomiglianza della generica serie mensile di periodo rispetto a quella relativa a settembre 2019-febbraio 2020 (con esclusione, come si è già detto, del mese di dicembre).

L'indice utilizzato per il generico periodo  $i$ -esimo assume la forma seguente:

$${}^iW = \frac{2 \sum u | {}^i m_u - {}^t m_u |}{{}^i M + {}^t M} \quad (u = \text{set., ott., nov., gen., feb.}; {}^i W \geq 0)$$

dove:

${}^i m_u$  e  ${}^t m_u$  sono il numero dei decessi nel mese  $u$ -esimo, rispettivamente del periodo  $i$ -esimo e del periodo settembre 2019-febbraio 2020;

${}^i M$  e  ${}^t M$  sono le medie aritmetiche dei decessi, rispettivamente del periodo  $i$ -esimo e del periodo settembre 2019-febbraio 2020.

Il periodo che si sceglie come paragonabile è quello che presenta il valore minimo di  ${}^i W$ .

attraverso una circolare. È forse superfluo sottolineare quanto l'omogeneità delle classificazioni sia necessaria per i confronti internazionali e quanto i confronti internazionali siano indispensabili per comprendere l'andamento dell'epidemia e per fronteggiarla. Per una breve rassegna sui modi in cui l'esigenza di informazione è stata corrisposta in altri paesi si veda l'Appendice 2.

**32** Si è perso quindi molto tempo. La formalizzazione dell'incarico è avvenuta più tardi con il decreto legge n. 30 del 10 maggio 2020. La rilevazione sul campo è iniziata ufficialmente lo scorso 25 maggio con i primi contatti telefonici alle unità campionarie da parte dei call center della Croce rossa italiana. Una valutazione dell'indagine è stata richiesta a uno degli autori dalla Commissione

Igiene e sanità del Senato della Repubblica in occasione della conversione in legge del decreto-legge 10 maggio 2020, n.30 recante misure urgenti in materia di studi epidemiologici e statistiche sul Sars-CoV-2 (Alleva, 2020).

**33** Ci si può chiedere se sarebbe stato lo stesso se l'epidemia si fosse manifestata con virulenza nel Mezzogiorno anziché nel Nord del Paese.

I valori minimi sono stati rinvenuti nel periodo settembre 2015-febbraio 2016 per il Nord-Ovest; nei due periodi settembre 2015-febbraio 2016 e settembre 2018-febbraio 2019 per il Nord-Est; nei due periodi settembre 2015-febbraio 2016 e settembre 2017-febbraio 2018 per il Centro; nel periodo settembre 2017-febbraio 2018 per il Sud e nei due periodi settembre 2017-febbraio 2018 e settembre 2018-febbraio 2019 per le Isole. Il numero di decessi per i mesi di marzo e aprile della serie scelta come paragonabile viene assunto come livello normale per marzo e aprile 2020 nelle singole ripartizioni geografiche. Quando i periodi di riferimento considerati paragonabili sono due si assume la loro media aritmetica.

Le stime dell'eccedenza (o della diminuzione) ottenute come differenza fra il numero effettivo di decessi per tutte le cause nel periodo 1° marzo-30 aprile 2020 e il numero normale di decessi nello stesso periodo (ipotesi A) sono scritte nel prospetto seguente. I totali per l'Italia sono ottenuti per somma.

Aree territoriali	Nord-Ovest	Nord-Est	Centro	Sud	Isole	Italia
Numero effettivo di decessi nel periodo 1° marzo-30 aprile 2020	61.447	30.106	24.071	25.217	12.396	153.237
Stima del numero normale di decessi nel periodo 1° marzo-30 aprile 2020	28.502	20.630	22.180	23.563	12.490	107.365
Eccedenza o diminuzione (-)	32.945	9.476	1.891	1.654	-94	45.872

Per procedere alle stime secondo l'ipotesi B vengono calcolati i tassi di variazione del numero di decessi fra febbraio e marzo e fra marzo e aprile per i periodi selezionati come paragonabili per le cinque ripartizioni territoriali. Quando i periodi di riferimento sono due il numero di decessi viene mediato. Il tasso di variazione fra febbraio e marzo è applicato al numero di decessi di febbraio 2020, ottenendo la stima di mortalità normale per marzo 2020; su questo valore è quindi applicato il tasso di variazione fra marzo e aprile per ottenere la stima di mortalità normale per aprile. Le stime così ottenute sono scritte nel prospetto seguente insieme al numero effettivo di decessi per tutte le cause nel periodo 1° marzo-30 aprile 2020 e all'eccedenza (o diminuzione). I totali per l'Italia sono ottenuti per somma.

Aree territoriali	Nord-Ovest	Nord-Est	Centro	Sud	Isole	Italia
Numero effettivo di decessi nel periodo 1° marzo-30 aprile 2020	61.447	30.106	24.071	25.217	12.396	153.237
Stima del numero normale di decessi nel periodo 1° marzo-30 aprile 2020	28.618	20.405	22.152	23.705	11.715	106.595
Eccedenza o diminuzione (-)	32.829	9.701	1.919	1.512	681	46.642

Le due ipotesi conducono a stime della mortalità normale per il periodo 1° marzo-30 aprile 2020 per l'Italia molto vicine tra di loro.

Le divergenze territoriali sono molto limitate, salvo che per le Isole, dove la differenza è di 775 unità. Su di essa può influire sia il relativamente limitato numero di decessi in quell'area, sia la possibile diversificazione dei regimi di mortalità nelle due isole.

## Appendice 2

### La raccolta dei dati sull'epidemia Covid-19; una rassegna delle strategie ed esperienze nazionali attuate per la stima delle variabili di interesse sull'intera popolazione

La rassegna degli approcci adottati in diversi paesi per la raccolta di dati sull'epidemia da Sars-CoV-2 fino ad aprile 2020 consente di individuare quattro differenti modi di operare.

La prima strategia consiste in una campagna massiva di test, a prescindere dalla presenza di sintomi, condotta senza seguire un disegno campionario con lo scopo essenziale di identificare le persone contagiate. Essa è stata adottata da Corea del Sud, Hong Kong, Emirati arabi uniti, Australia, Islanda e anche dalla Germania; in Italia, dalla Regione Veneto.

La seconda strategia prevede test diagnostici su un campione probabilistico orientati alla stima delle variabili di interesse con prefissati livelli di precisione e includendo le persone asintomatiche. In questa direzione si muove il progetto dell'Helmholtz Center for Research on Infections in Germany per la ricerca di anticorpi specifici del Covid-19 su 100.000 persone (Hackenbroch, 2020)<sup>34</sup>. Un altro caso è quello della Romania riferito a un campione casuale di 10.500 persone di Bucarest testate sotto la supervisione del Matei Bals Institute of Infectious Diseases in Bucharest (Romania-insider.com, 2020). Infine, si è operato così per due localizzazioni in Canberra, dove il testing è stato condotto dall'Australian Capital Territory (Abc, 2020). Recentemente è stata assegnata all'Istituto nazionale di statistica italiano, Istat, la definizione di un campione di 150.000 unità sulle quali effettuare il test sierologico.

<sup>34</sup> La proposta per la Germania ha caratteristiche e finalità analoghe a quelle descritte in Alleve et al., 2020a.

La terza strategia consiste in una rilevazione via web estesa, con partecipazione su base volontaria. Essa è stata sperimentata in Israele su circa 60.000 persone le quali hanno fornito informazioni relative a età, genere, indirizzo, stato di salute generale, sintomi particolari legati a Covid-19 ed eventuale stato di isolamento. Lo scopo principale è stato quello di confrontare i risultati dei test e le situazioni di contagio per persone con differenti caratteristiche socio-economiche. L'impianto è stato realizzato dall'Istituto Weizmann (Rossmann et al., 2020). Infine, un'ulteriore possibile strategia è quella di utilizzare impianto e formati di rilevazioni preesistenti, opportunamente integrati per raccogliere informazioni sull'epidemia. È questa la proposta del Centre for European Policy Studies, presentata da Daniel Gros (2020) per l'utilizzazione dell'ampio campione panel europeo denominato «Eu statistics on income and living conditions» (Eu-Silc). Più specificamente, Dewatripont et al. (2020) hanno suggerito di impiegare il panel Eu-Silc per due sondaggi: il primo per verificare la presenza del contagio; il secondo per verificare l'immunizzazione. In direzione analoga, l'Organizzazione internazionale del lavoro ha sollecitato gli Istituti nazionali di statistica a introdurre nelle rilevazioni statistiche sul lavoro un modulo destinato a valutare l'impatto del Covid-19, rimuovendo eventualmente sezioni con minore priorità (Ilo, 2020). Alcuni paesi hanno adottato più di un approccio contemporaneamente e hanno proceduto a una forte integrazione di dati provenienti da fonti diverse.

## Bibliografia

- Abc** (2020), *Random Coronavirus Testing to Begin in Canberra Next Week at Drive-Through Centre and Clinic*, Abc net, 03-04-2020. <https://www.abc.net.au/news/2020-04-03/random-coronavirus-testing-begins-in-canberra/12119364>.
- Adam D.** (2020), «Special report: The Simulations Driving the World's Response to Covid-10», in *Nature*, 2 aprile.
- Alleva G., Arbia G., Falorsi P.D., Pellegrini G., Zuliani A.** (2020 a), *Indagine a campione per una stima affidabile dei parametri fondamentali della epidemia da Sars-CoV-2*, [https://web.uniroma1.it/memotef/sites/default/files/Proposta\\_0.pdf](https://web.uniroma1.it/memotef/sites/default/files/Proposta_0.pdf).
- Alleva G., Arbia G., Falorsi P.D., Zuliani A.** (2020 b), *A Sample Approach to the Estimation of the Critical Parameters of the Sars-Cov-2 Epidemics: An Operational Design with a Focus on the Italian Health System*, <http://arxiv.org/abs/2004.0606>.
- Alleva G., Zuliani A.** (2020), *Numeri affidabili per uscire dalle nebbie dell'epidemia*, *lavoce.info*, 30 marzo, <https://www.lavoce.info/archives/64813/numeri-affidabili-per-uscire-dalle-nebbie-dellepidemia/>.
- Alleva G.** (2020), *Contributo per la 12ª Commissione permanente Igiene e sanità del Senato della Repubblica*, Roma, 27 maggio, [https://www.senato.it/application/xmanager/projects/leg18/attachments/documento\\_evento\\_procedura\\_commissione/files/000/135/501/giorgio\\_alleva.pdf](https://www.senato.it/application/xmanager/projects/leg18/attachments/documento_evento_procedura_commissione/files/000/135/501/giorgio_alleva.pdf).
- Arbia G.** (2020), *Fase 2, quando inizierà? Ecco perché i numeri non bastano a rispondere*, *Il sussidiario.net*, 13 aprile, <https://www.ilsussidiario.net/news/fase-2-quando-iniziera-ecco-perche-i-numeri-non-bastano-a-rispondere/2008955/>.
- Ascani P.** (2020), *Technical Note on the Methods of the Data Collection Phase, Annex to the Proposal: A Sample Design For Reliable Estimates Of The Sars-Cov-2 Epidemic's Parameters. Calling for a Protocol Using Panel Data*, <https://web.uniroma1.it/memotef/sites/default/files/TechNote.pdf>.
- Bucci E., Leuzzi L., Marinari E., Parisi G., Ricci Tersenghi F.** (2020), *Verso una stima di morti dirette e indirette per Covid*, *Scienze in rete*, 24 aprile.
- Burioni R.** (2020), «Convivere con Covid-19: proposta scientifica per riaprire l'Italia, gestendo in modo sicuro la transizione da pandemia a endemia», in *Medical Facts*, 14 aprile.
- Cleevly M., Susskind D., Vines D., Vines L., Wills S.** (2020), *A Workable Strategy for Covid-19 Testing: Stratified Periodic Testing Rather Than Universal Random Testing*, *Covid Economics Vetted and Real-Time Papers*, Issue 8, 22 April, The Centre for Economic Policy Research (Cepr).
- Crisanti A. et al.** (2020), *Suppression of Covid-19 Outbreak in the Municipality of Vo', Italy*, *MedRxiv*, <https://www.medrxiv.org/content/10.1101/2020.04.17.20053157v1.full.pdf>.
- D'Alessio G., De Bonis R.** (2020), «Covid: una caporetto statistica», in *Doppio Zero*, 6 maggio, <https://www.doppiozero.com/materiali/covid-una-caporetto-statistica>.
- Dewatripont M., Goldman M., Muraille E., Platteau J.P.** (2020), *Rapidly Identifying Workers Who Are Immune to Covid-19 and Virus-Free Is a Priority for Restarting the Economy*, *VoxEU.org*, 23 March.
- Di Vico D.** (2020), «E se la statistica aiutasse i virologi?», in *Corriere della Sera*, 26 marzo, [https://www.corriere.it/opinioni/20\\_marzo\\_26/se-statistica-aiutasse-virologi-0bc1f966-6f95-11ea-b81d-2856ba22fce7.shtml](https://www.corriere.it/opinioni/20_marzo_26/se-statistica-aiutasse-virologi-0bc1f966-6f95-11ea-b81d-2856ba22fce7.shtml).
- Fuggetta M.** (2020), *Testing for the Base Rate*, [massimofuggetta.com/2020/04/28/testing-for-the-base-rate/](https://www.massimofuggetta.com/2020/04/28/testing-for-the-base-rate/).
- Grasso A.** (2020), «Cresce la "quota d'ascolto" dell'informazione e i tg fanno record di share», in *Corriere della Sera*, 4 aprile.
- Gros D.** (2020), *Creating an Eu 'Corona Panel': Standardised European Sample Tests to Uncover the True Spread of the Coronavirus*, *VoxEU.org*, 28 March.
- Hackenbroch V.** (2020), «Große Antikörperstudie soll Immunität der Deutschen gegen Covid-19 feststellen», in *Spiegel*, 26 marzo, [https://www.spiegel.de/wissenschaft/medizin/coronavirus-grosse-antikoerper-studie-soll-immunitaet-der-deutschen-feststellen-a-c8c64a33-5c0f-4630-bd7348c17c1bad23?d=1585300132&sara\\_ecid=soci\\_upd\\_wbMbjhOSvVilJc8RPU89NcCvtlFcl](https://www.spiegel.de/wissenschaft/medizin/coronavirus-grosse-antikoerper-studie-soll-immunitaet-der-deutschen-feststellen-a-c8c64a33-5c0f-4630-bd7348c17c1bad23?d=1585300132&sara_ecid=soci_upd_wbMbjhOSvVilJc8RPU89NcCvtlFcl).
- International Labour Organization** (2020), *Covid-19 Impact on the Collection of Labour Market Statistics*, <https://ilostat.ilo.org/topics/covid-19/covid-19-impact-on-labour-market-statistics/>.
- Ioannidis J.P.A.** (2020), *Population-Level Covid-19 Mortality Risk for Non-Elderly Individuals Overall and for Non-Elderly Individuals Without Underlying Diseases in Pandemic Epicentres*, <https://www.medrxiv.org/content/10.1101/2020.04.05.20054361v2>.
- Istat** (2020), *Decessi anni 2015-2020*, <https://www.istat.it/it/archivio/240401>.

**Istituto Superiore di Sanità** (2020), *Bollettino-Sorveglianza-Integrata-Covid-19*, Aggiornamento nazionale, 7 maggio, [https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-Covid-19\\_9-giugno-2020.pdf](https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-Covid-19_9-giugno-2020.pdf).

**Istituto Superiore di Sanità-Istat** (2020), *Impatto dell'epidemia covid-19 sulla mortalità totale della popolazione residente primo quadrimestre 2020*, 4 giugno, [https://www.epicentro.iss.it/coronavirus/pdf/Rapp\\_Istat\\_Iss\\_3Giugno.pdf](https://www.epicentro.iss.it/coronavirus/pdf/Rapp_Istat_Iss_3Giugno.pdf).

**Ministero della salute-Istat** (2020), *Protocollo metodologico per un'indagine di siero-prevalenza sul Sars-Cov-2 condotta dal Ministero della salute e dall'Istat*, 25 maggio.

**Our World in Data**, *Programme on Global Development*, University of Oxford, <https://ourworldindata.org/coronavirus>.

**Protezione civile**, *Dati giornalieri sull'epidemia, dati delle regioni e province autonome*, <https://github.com/pcm-dpc/Covid-19/tree/master/dati-regioni>.

**Romania-insider.co** (2020), *Coronavirus in Romania: Over 10,000 Bucharest*

*Residents Will Be Tested for Covid-19 as Part of a Study*, 3 aprile, <https://www.romania-insider.com/coronavirus-romania-bucharest-testing-streinu-cerchel>.

**Romer P., Shah R.** (2020), «Testing Is Our Way Out», in *The Wall Street Journal*, April 2, <https://www.wsj.com/articles/testing-is-our-way-out-11585869705>.

**Rosman H., Keshet A., Shilo S., Gavrieli A., Bauman T., Cohen O., Balicer R., Geiger B., Dor Y., Segal E.** (2020), *A Framework for Identifying Regional Outbreak and Spread of Covid-19 From One-Minute Population-Wide Surveys*, <https://doi.org/10.1101/2020.03.19.20038844>.

**The Economist** (2020), *Why a Study Showing That Covid-19 Is Everywhere Is Good News*, 11 aprile.

**Tozer J.** (2020), «Measuring the True Toll of the Pandemic. How The Economist Measures Excess Deaths Caused by Covid-19», in *The Economist*, 24 aprile, <https://medium.economist.com/measuring-the-true-toll-of-the-pandemic-fa7e003b3ff4>.

**Yang S., Keller F. B., Zheng L.** (2016), *Social Network Analysis: Methods and Examples*, Sage Publications, London.