

Giorgio Alleva  
Sapienza Università di Roma

Roma, 27 maggio 2020

### **Premessa**

In questo contributo fornirò una valutazione del decreto-legge 10 maggio 2020, n. 30, recante misure urgenti in materia di studi epidemiologici e statistiche sul SARS-COV-2, sperando possa essere utile alla discussione nella 12<sup>a</sup> Commissione permanente *Igiene e sanità* in questa fase della sua conversione in legge.

Dopo una introduzione sull'importanza di conoscere i veri numeri dell'epidemia attraverso l'impianto di indagini statistiche condotte da parte di soggetti competenti su mandato delle istituzioni responsabili della salute pubblica, entrerà nel merito del provvedimento, indicando alcuni aspetti critici e suggerendo possibili miglioramenti della norma che possano accrescere l'utilità dell'indagine a sostegno delle decisioni pubbliche e dell'approfondimento della comunità scientifica.

A fronte della soddisfazione per l'avvio di operazioni che avevo sollecitato fin dal mese di marzo con una proposta rivolta alle istituzioni preposte ed elaborata insieme ad alcuni colleghi statistici (tra i quali il prof. Alberto Zuliani, come il sottoscritto già Presidente dell'Istat) debbo rilevare l'avvio tardivo rispetto all'esigenza conoscitiva e l'adozione di un disegno e protocollo di indagine che ne può limitare fortemente l'utilità.

### **1. L'importanza di conoscere e comprendere l'evoluzione dell'epidemia**

Dati statistici affidabili sono essenziali per almeno quattro finalità:

- a) stimare le modalità di propagazione dell'epidemia;
- b) stimare il numero effettivo dei contagiati, sintomatici, paucisintomatici e asintomatici, dei guariti, dei decessi e degli immuni in modo da poter calcolare correttamente i tassi di contagio, per le diverse categorie e in totale, i tassi di guarigione, letalità e immunizzazione e altri parametri indicativi della diffusione dell'epidemia (ad esempio,  $R_0$  al momento della completa suscettibilità della popolazione e successivamente  $R_t$ ); seguire la loro evoluzione nel tempo e fare previsioni, costruendo scenari anche differenti e fornendo la relativa probabilità di accadimento;

- c) valutare gli effetti delle misure via via introdotte in modo da poter orientarne nel corso del tempo inasprimento, attenuazione e rimozione;
- d) 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; 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. In ogni caso, 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. 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 con appropriate misure di accuratezza.

Quanto alla seconda finalità, il numero di contagi così come viene valutato attualmente<sup>1</sup> è certamente sottostimato; mancano quasi completamente gli asintomatici.<sup>2</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 ad essa un grado probabilistico di affidabilità.

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 paucisintomatici 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 risulterebbe 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 dove non vengono effettuati molti test e comunque che non stimino la numerosità della popolazione infetta sulla base di indagini campionarie.

Un'indagine a campione 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

---

<sup>1</sup> I conteggi presentano comunque qualche lacuna nel sistema di raccolta e contabilizzazione, testimoniata da fluttuazioni delle serie regionali dovute a registrazioni dei dati in giorni diversi da quelli ai quali si riferiscono, dando adito a letture di inattese accelerazioni o di agognati rallentamenti.

<sup>2</sup> Sfuggono anche le forme lievi e, purtroppo, anche quelle gravi e spesso letali che non sono riuscite ad arrivare ai pronto soccorso e alle strutture specializzate. Una delle maggiori carenze portate allo scoperto dall'epidemia è relativa alla medicina territoriale.

popolazione e le imprese ad adottare comportamenti restrittivi. L'avevo proposta insieme ad altri quattro statistici (Alleva *et al.*, 2020 a, b, c) 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.

Nella proposta si suggeriva 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 14 giorni precedenti l'accertamento del contagio; il secondo relativo alla restante popolazione, un panel da seguire a cadenze temporali fissate nel corso del tempo. Su questi campioni - contatti e panel - sarebbero stati raccolti i dati relativi alle caratteristiche di interesse ed effettuati i test diagnostici; i test sarebbero stati ripetuti più volte nel corso del tempo. Il panel, in particolare, 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 di 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. Per tutti i componenti del primo campione (contatti) e del secondo campione (popolazione altra) si sarebbero effettuati opportuni esami diagnostici (virologici e/o sierologici, da decidere in relazione alle finalità perseguite) ripetuti a distanza di quattordici giorni uno dall'altro (o un mese) per il periodo di tempo ritenuto necessario per il monitoraggio dell'evoluzione dell'epidemia. Le stime statistiche degli aggregati di interesse sarebbero state prodotte con il concorso di ambedue i campioni. Esse godono di proprietà piuttosto importanti: la non distorsione e l'efficienza (Alleva *et al.*, 2020 b).<sup>3</sup> La proposta delinea, quindi, 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 riacutizzarsi del contagio o dell'arrivo di una seconda ondata e sul secondo campione, il panel, per valutare gli andamenti, in particolare nel corso della fase 2 e di quelle successive. È importante sottolineare che ai diversi campioni interessati dall'indagine (i contagiati conclamati e i loro contatti e i componenti del panel nella occasione di ingresso) sarebbe stato somministrato anche un questionario volto a raccogliere informazioni individuali e familiari utili per la comprensione della diffusione dell'epidemia e delle misure da adottare nella sua fase ascendente o discendente)<sup>4</sup>.

---

<sup>3</sup> Non siamo stati 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>4</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 familiare, il tipo e la composizione del nucleo, il tipo e la frequenza degli spostamenti.

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 è stato chiamato a predisporre un piano di campionamento importante, di 150.000 unità, alle quali chiedere alcune informazioni e sulle quali effettuare il prelievo per l'analisi sierologica.

Con uno sviluppo ulteriore della nostra proposta abbiamo avanzato qualche suggerimento migliorativo che ricordo in questa audizione. Si sarebbe potuto scegliere una parte delle unità campionarie previste, diciamo 15-20.000, in modo da garantire la rappresentatività della popolazione residente a livello nazionale e costituire un campione "anticipatorio". Si sarebbe potuto elaborare rapidamente i dati rilevati su di esse, offrendo una prima rappresentazione affidabile della situazione relativa ai contagi avvenuti e risolti positivamente (se si effettuasse - come è previsto - il solo test sierologico). Si sarebbe potuto proseguire con le altre 130.000 unità campionarie. Il campione "anticipatorio" avrebbe potuto costituire un panel da seguire nel corso del tempo con scansione opportuna, in prima approssimazione ogni due settimane.<sup>5</sup> Sulle unità del panel si sarebbe potuto effettuare oltre all'esame sierologico anche quello per tampone in modo da accertare l'eventuale sopravvenuto contagio. Questo sarebbe risultato particolarmente importante se, nel corso dei prossimi mesi, si fosse riattivata la virulenza o fosse arrivata una seconda ondata dell'epidemia. Il panel si sarebbe dovuto alimentare, nel corso del periodo previsto per l'osservazione di sorveglianza, con una quota di ricambi, necessari per sostituire le persone nel frattempo contagiate o che si fossero rifiutate di sottoporsi ulteriormente ai test. Si sarebbe potuto attingere alle unità del campione coinvolte successivamente (le altre 130.000), sempre attraverso una procedura probabilistica idonea a mantenere l'affidabilità statistica delle stime. Si sarebbe fatto in questo modo un servizio maggiormente utile al Paese, senza un aggravio elevato dei costi. Parecchi di questi suggerimenti sembrano essere stati recepiti a una prima lettura del "Protocollo metodologico per un'indagine di siero-prevalenza sul SARS-CoV-2 condotta dal Ministero della salute e dall'ISTAT" reso disponibile il 25 maggio, data di avvio dell'indagine sul campo.

## **2. Questioni di carattere generale: obiettivi dell'indagine e coerenza delle modalità di conduzione**

Le finalità dell'indagine, come osserva la relazione illustrativa del disegno di legge di conversione, appaiono molto estese, andando dalla stima dello stato immunitario della popolazione, alla stima delle infezioni paucisintomatiche e asintomatiche, alla misura più accurata della letalità e delle dinamiche di trasmissione del *virus*.

"L'esigenza di condurre con urgenza l'indagine nasce dalla necessità di disporre di dati affidabili e completi sullo *stato immunitario della popolazione*, nonché in ragione della consapevolezza delle difficoltà di valutare la frazione di *infezioni paucisintomatiche o asintomatiche* che richiedono assistenza medica". E ancora. "L'indagine di sieroprevalenza risulta necessaria per guidare decisioni strategiche di carattere socio sanitario. Non solo. Utilizzando i dati raccolti mediante l'indagine di sieroprevalenza, potranno essere compiuti studi che consentiranno anche di

---

<sup>5</sup> Gli epidemiologi potrebbero suggerire cadenze meglio calibrate.

effettuare *misurazioni più accurate della letalità*, migliorare gli sforzi di modellizzazione e aiutare a meglio definire le *dinamiche di trasmissione del virus*".

Tuttavia, a fronte di questi diversi obiettivi, nel titolo dell'articolo 1, l'indagine è definita come di siero-prevalenza sul SARS-COV-2, e al comma 1 è indicato un unico obiettivo conoscitivo: studi epidemiologici e statistiche affidabili e complete sullo *stato immunitario della popolazione*.

La scelta di condurre l'indagine esclusivamente attraverso un esame sierologico (art. 1, comma 2,)), e non anche con il tampone naso-faringeo, non consente di conoscere e monitorare con riferimento a date certe il livello dell'epidemia in corso (lo stato infettivo della popolazione); soprattutto non permette di attivare il tracciamento dei contatti di persone risultate contagiate (paucisintomatici e asintomatici) e la conseguente messa in sicurezza di quanti abbiano avuto con loro rapporti. D'altra parte, come noto, l'esito dell'esame sierologico, seppur nell'incertezza dovuta ai casi di falsi-positivi e di falsi-negativi (dei quali il decreto in esame non stabilisce valori soglia per i dispositivi da utilizzare), consente di conoscere la popolazione che ha avuto contatti pregressi con il *virus* (senza tuttavia poterne stabilire una data e dunque fornire una rappresentazione dell'infezione nel passato) ma non di assegnare un patentino di immunità, non essendoci sufficienti basi scientifiche per farlo.

Ciò riduce fortemente la rilevanza dell'indagine, certamente con riferimento alle finalità espresse nella relazione di presentazione del DDL.

Nel Protocollo metodologico per l'indagine di siero-prevalenza, pubblicato il 25 maggio sul sito web dell'Istat (cfr. bibliografia), è previsto che "In ogni caso, ciascun soggetto trovato positivo ad un test sierologico dovrà essere preso in carico e andranno eseguiti ulteriori accertamenti attraverso un tampone naso-faringeo per stabilire lo stato di eventuale contagiosità". Si tratta di una importante specificazione del follow-up dell'indagine che si ritiene tuttavia dovesse essere prevista nel DDL, anche per garantirne la comunicazione alla popolazione che farà parte del campione. Rimane la critica che l'identificazione e selezione dei contatti di quanti risulteranno contagiati dagli accertamenti con i tamponi naso-faringei non siano stati considerati elementi di un unico disegno d'indagine di carattere statistico e dinamico, sullo stato e l'evoluzione dell'epidemia in corso. Inoltre, condizionare la somministrazione del tampone all'esito dell'esame sierologico non rappresenta la modalità corretta se oltre allo stato immunitario si fosse inteso valutare la dimensione della popolazione infetta al di fuori di quella sintomatica.

Non ho dubbio che il campione della popolazione da sottoporre all'esame sierologico sia stato costruito rigorosamente in modo da fornire stime affidabili; la Direzione metodologica dell'Istat ha competenze di alto livello. Tuttavia, la sua ampiezza e la complessità organizzativa dell'intera operazione renderanno disponibili dati verificati ed elaborati non prima di tre-quattro mesi dalla statuizione normativa ed è una previsione ottimistica. Disporremo in quel momento di una fotografia molto distante nel tempo. Informazioni poco tempestive su fenomeni che evolvono con velocità e intensità diverse nel Paese assumono una rilevanza ridotta nel supportare decisioni e piani operativi a livello nazionale, regionale e locale.

Come ho accennato in precedenza, un campione anticipatorio di tipo panel si sarebbe potuto programmare per monitorare nel corso del tempo i parametri d'interesse dell'epidemia,

realizzando un sistema di sorveglianza statisticamente rigoroso, anziché prevedere una fotografia statica, disponibile molto tempo dopo essere stata scattata. D'altra parte, il monitoraggio da realizzare attraverso la somministrazione del tampone, ad esempio ogni quindici giorni, è molto rilevante con riferimento allo stato dell'epidemia in corso (la popolazione contagiata). Importante, ma in misura minore, è la conoscenza dell'evoluzione della popolazione che ha avuto contatti pregressi con il virus realizzata attraverso l'esame sierologico.

Al comma 3 dell'articolo 1 è prevista anche una estensione longitudinale del campione; tuttavia il DDL non ne specifica il ruolo, a mio avviso cruciale, che dovrebbe assumere all'interno dell'indagine: ruolo di campione anticipatorio, ruolo di sorveglianza di specifici parametri d'interesse sull'epidemia, ruolo di informazione da integrare con quella predisposta dal sistema di sorveglianza dell'Istituto superiore della sanità e del Ministero della salute.

Facendo uscire dalle nebbie quanto indicato nel citato comma, il già citato Protocollo metodologico per l'indagine di siero-prevalenza chiarisce che è stato previsto un campione anticipatorio di circa 20.000 unità, che costituisce un sottoinsieme dei circa 195.000 individui selezionati e mantiene le caratteristiche del campione totale e, quindi, potrà essere utilizzato per diffondere tempestivamente primi risultati e anche per auspicabili ritorni successivi di approfondimento. Il disegno dello studio di siero-prevalenza viene pertanto definito come *cross sectional*, rappresentativo della popolazione italiana, con valutazione longitudinale di un sottogruppo di popolazione.

### **3. Insufficiente specificazione di alcuni aspetti fondamentali per la garanzia della qualità dei risultati dell'indagine**

Considerato che l'indagine statistica prevede esami diagnostici individuali, l'interazione tra gli errori possibili nella valutazione dei loro esiti, la disponibilità delle persone del campione nel rendersi disponibili alla loro somministrazione e l'organizzazione sul campo dell'indagine rappresentano elementi fondamentali per garantire la necessaria qualità dei risultati. Come anche la loro valorizzazione attraverso l'accesso ai dati da parte della comunità scientifica.

Un primo importante aspetto di qualità è rappresentato dall'*affidabilità dei risultati dei test sierologici*, importante non soltanto per la scelta dello specifico prodotto diagnostico da adottare, ma soprattutto per le decisioni che si possono prendere a valle dell'esito dell'esame: l'isolamento (se ancora necessario) e il tracciamento dei contatti nel caso di esito positivo; la continuazione o ripresa delle normali attività nel caso di esito negativo. Nel DDL non si fa riferimento alla dimensione massima 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 stabilire quale aspetto fra i due tenere sotto controllo in misura maggiore. È più importante un'elevata sensibilità del test in modo da minimizzare il numero di falsi-negativi? Oppure un'elevata specificità, in modo da minimizzare quello dei falsi-positivi? 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. A testimonianza della rilevanza di questo tema, nella progettazione dell'indagine campionaria che l'Istat ha avviato in questi giorni, nel bando di gara per la scelta del

test sierologico è stato indicato che il kit avrebbe dovuto assicurare una sensibilità almeno pari al 90% e una specificità almeno pari al 95%. È la scelta giusta? Chi l'ha fatta? Sarebbe forse stato opportuno che la norma indicasse esplicitamente uno o più soggetti responsabili della decisione (Ministero della salute, Istat, Istituto superiore di sanità, Comitato tecnico scientifico).

Un secondo aspetto fondamentale per assicurare la qualità dei risultati è la copertura del campione programmato, e dunque la *elevata partecipazione all'indagine delle persone estratte nel campione*. Seppure insieme alla lista campionaria di base si estrae usualmente un maggior numero di unità (le c.d. unità di riserva), una quota di rifiuti elevata, e il conseguente ricorso esteso alle altre unità, introduce un'autoselezione dei partecipanti nel campione che può compromettere la rappresentatività della popolazione d'interesse. Nel Protocollo metodologico dell'indagine è indicato che a fronte dell'obiettivo di copertura di 150.000 unità è stata selezionata una lista di 195.000, implicitamente prevedendo un tasso di partecipazione non inferiore al 76,9% (e dunque un tasso di rifiuto inferiore ad un quarto). Considerata la natura dell'indagine, che prevede un esame sierologico e la risposta a un questionario sulla propria salute, *la non obbligatorietà di risposta prevista dal DDL rappresenta un forte rischio sulla qualità dei risultati e il successo dell'indagine*. Sarebbe stato opportuno, come per altre indagini importanti, prevedere l'obbligo di risposta (come ad esempio per i censimenti generali della popolazione per aggiornare le anagrafi). La capacità dell'Associazione Croce Rossa Italiana nel convincere le persone facenti parte del campione nel corso del contatto telefonico previsto al comma 5 dell'art. 1 diviene un elemento strategico, unitamente a una efficace campagna promozionale dell'indagine presso la popolazione.

Un terzo elemento decisivo per il successo dell'indagine riguarda le operazioni sul campo e i ruoli assunti dai soggetti coinvolti. È positivo aver previsto al comma 5 dell'art. 1 di comunicare con modalità sicure ai medici di medicina generale e ai pediatri di libera scelta i nominativi dei loro assistiti rientranti nei campioni, in modo da favorire la partecipazione all'indagine. Anche positivo appare il coinvolgimento dell'Associazione Croce Rossa Italiana per il primo contatto telefonico con la popolazione facente parte del campione, per spiegare le finalità dell'indagine, comunicare le garanzie sul trattamento dei dati personali previste dal Regolamento UE 2016/679 e per fissare l'appuntamento per il prelievo. Nel Protocollo metodologico dell'indagine è indicato che a tal fine viene organizzato un servizio di *call center* con personale CRI adeguatamente formato. Tuttavia, appare dubbio che la stessa Associazione abbia la competenza e l'esperienza necessaria per la somministrazione di un questionario per via telefonica su temi sensibili come quelli sulla salute. L'attività di formazione risulta quindi cruciale ed è auspicabile che possa essere fatta in modo efficace nei tempi purtroppo brevi che è possibile prevedere. *Il ricorso a società specializzate sarebbe stata una maggiore garanzia per raccogliere informazioni in modo completo e accurato*.

Sottolineando di nuovo la ridotta rilevanza di una costosa fotografia statica sulla sieroprevalenza nella popolazione rispetto all'impiego di un sistema di sorveglianza anche sullo stato infettivo, in grado di innescare, in tempo reale, conseguenti provvedimenti sanitari, alcuni elementi critici emergono sul piano della valorizzazione dei risultati attraverso la loro condivisione con la comunità scientifica.

Compatibilmente con il rispetto della normativa sulla privacy, nelle previsioni del DDL si notano *alcune asimmetrie circa l'accesso ai dati prodotti attraverso l'indagine*.

In particolare nell'articolo 1:

- al comma 6 è indicato che il titolare del trattamento dei dati raccolti nella banca biologica è il Ministero della salute e "l'accesso ai dati da parte di altri soggetti, per le suddette finalità di ricerca, è consentito esclusivamente nell'ambito di progetti di ricerca congiunti con il medesimo Ministero". Si tratta di una *limitazione severa per la ricerca e burocratizza il processo di valorizzazione dei dati rilevati con l'indagine*. Si suggerisce di eliminare "esclusivamente nell'ambito di progetti di ricerca congiunti con il medesimo Ministero".
- al comma 7, si indica che i dati raccolti nell'ambito dell'indagine, privi di identificativi diretti, possono essere comunicati, per finalità scientifiche, ai soggetti del SISTAN (art 5-ter dlgs n. 33/2013). Nel medesimo è indicato che l'Istituto superiore di sanità può trattare i dati raccolti nell'ambito dell'indagine per finalità di ricerca scientifica. Considerato che l'Istituto è un soggetto facente parte del SISTAN tale previsione si ritiene costituire *un'asimmetria nel panorama degli enti di ricerca appartenenti al SISTAN e appare superflua*. Nello stesso comma 7, la garanzia della tutela della privacy attraverso la trasmissione di dati 'privi di identificativi diretti' non appare conforme con il principio della non identificabilità. Per evitare controversie interpretative si suggerisce di sostituire con 'privi di ogni riferimento che ne permetta il collegamento diretto con gli interessati' come peraltro indicato al comma 8 del DDL a proposito dell'accesso ai dati in forma individuale da parte delle regioni e province autonome;
- al comma 8 è indicato che "per quanto riguarda i dati relativi agli assistiti delle altre Regioni e province autonome, l'accesso, da parte degli enti territoriali in esame, è consentito in forma anonima e aggregata, *a soli fini comparativi*". *Tale finalità non è specificata in modo sufficientemente chiaro per consentire l'accesso a dati sensibili, seppur previsto solamente in forma anonima e aggregata*.

#### **4. Indeterminatezza e limitata utilità di alcuni punti del provvedimento.**

Il comma 3 dell'articolo 1 stabilisce in modo confuso le modalità di selezione dei campione e non specifica quali siano i domini di stima pianificati. Con riferimento alla formazione dei campioni si suggerisce di sostituire "tramite i propri registri statistici individui, unità economiche, luoghi e tematico del lavoro" con "tramite i propri registri statistici"; i quattro registri statistici di base (individui, unità economiche, luoghi e attività), costituiscono infatti il sistema integrato dei registri dell'ISTAT; nello stesso comma si specificano le variabili di stratificazione, delle quali si tiene conto per l'estrazione del campione, ma non i domini di stima, ossia i sottoinsiemi di popolazione per i quali ci si impegna a fornire stime con margini di errore prefissati; qualora le variabili di stratificazione non coincidessero con i domini di stima pianificati (classi di età, genere settore di attività economica a livello regionale) questo andrebbe specificato. Come già sottolineato nel paragrafo 2, il riferimento, solamente per inciso, a campioni longitudinali non è accompagnato dalla specificazione del ruolo di tale campione nella strategia dell'indagine.

Il comma 14 autorizza l'ISTAT a conferire fino ad un massimo di 10 incarichi di lavoro autonomo anche di collaborazione coordinata e continuativa, della durata di sei mesi, per una spesa complessiva di 385.000 euro. Si rileva che per un'attività di così breve durata un reclutamento ad hoc non consente di individuare soggetti effettivamente qualificati; d'altronde l'Istat già dispone di molte risorse umane,



con adeguata varietà di specializzazioni e competenze. Per i tempi necessari al reclutamento del personale negli enti pubblici, queste persone sarebbero prevedibilmente inquadrati al termine delle operazioni d'indagine.

Il comma 15 reca alcune autorizzazioni di spesa. La valorizzazione di alcune attività riportata nella relazione tecnica non sembra appropriata. La previsione di una spesa di 280.000 euro per l'attività svolta dalla Croce Rossa Italiana per la "chiamata, identificazione, somministrazione del questionario, appuntamento, raccolta e trasferimento dei campioni dai singoli laboratori alla banca biologica centrale dei campioni" appare sottostimata (per coinvolgere 150.000 persone fra le 195.000 definite per il campionamento). come anche la spesa di 1.500.000 euro per l'acquisto dei dispositivi idonei alla somministrazione delle analisi sierologiche.

## 5. Conclusioni

L'indagine oggetto del provvedimento legislativo in discussione è stata avviata sul campo lo scorso 25 maggio e parte delle osservazioni critiche, suggerimenti e raccomandazioni per migliorarne il disegno e accrescerne la rilevanza per i decisori pubblici e la comunità scientifica che ho descritto sono oramai tardive.

D'altra parte, credo sia importante che la Commissione, oltre a condividere la soddisfazione per l'avvio dell'iniziativa, abbia consapevolezza dei suoi limiti e dell'opportunità perduta di attivare (ben prima) una strategia più ambiziosa per la gestione più consapevole e documentata dell'epidemia.

Resta invece possibile migliorare il testo della legge, in primo luogo favorendo la valorizzazione dei risultati da parte della comunità scientifica.

## Riferimenti bibliografici

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. e Zuliani A. (2020 c), "Numeri affidabili per uscire dalle nebbie dell'epidemia", *lavoce.info*, 30 marzo 2020. <https://www.lavoce.info/archives/64813/numeri-affidabili-per-uscire-dalle-nebbie-dellepidemia/>.

Ministero della salute e Istat (2020), Protocollo metodologico per un'indagine di siero-prevalenza sul SARS-CoV-2 condotta dal Ministero della salute e dall'ISTAT. <https://www.istat.it/it/files//2020/05/Protocollo-approvato.pdf>