

Soddisfatti e misurati. La qualità dei servizi alla prova della statistica

21 e 22 aprile 2022 - Palazzo della Gran Guardia - Piazza Bra, 1 - VERONA

## Monitorare e comprendere la pandemia nel nostro Paese: un bilancio da un punto di vista statistico e alcune proposte

Giorgio Alleva

Dipartimento di Metodi e modelli per l'economia, il territorio e la finanza (MEMOTEF)

Università Sapienza di Roma



# Sommario

- Centralità dei dati: il rito quotidiano della diffusione dei dati
- La mancanza di un quadro coerente
- I problemi principali e i dati necessari
- Il bilancio di un'occasione mancata
- Una variabilità nei dati non sempre decifrabile
- Il sistema di monitoraggio multi-fonte necessario
- Due proposte
- Conclusioni

# Centralità dei dati: il rito quotidiano della diffusione dei dati

Ancora oggi, a distanza di oltre due anni dall'insorgenza della pandemia, vengono resi disponibili giornalmente dati che riguardano:

- i **test processati**,
- i **contagiati** e il **tasso di positività al test**,
- lo stock dei **ricoverati** con il dettaglio per quelli in terapia intensiva e, per questa, gli ingressi nella giornata,
- i **morti**,
- l'**indice di trasmissione dei contagi  $R_t$**  con cadenza settimanale.

In generale i dati sono dettagliati per **regione**; in qualche caso, per provincia e grandi città.

Nelle tavole di sintesi dell'ISS sono riportati anche **età** e **genere**.

# La mancanza di un quadro di coerenza

Sulla spinta dell'emergenza di marzo-aprile 2020, i dati sono stati **raccolti senza un disegno orientato a comprendere** a fondo l'evoluzione dell'epidemia.

Non si è provveduto a raccogliere e integrare i dati necessari secondo un **disegno organico**, in un quadro di coerenza per contrastare l'epidemia.

- Si sarebbe dovuto capire che **non sarebbero stati sufficienti** a fornire un quadro informativo perspicuo.
- **Non è maturato un atteggiamento di “conoscere per decidere e operare”** (*knowledge for action*).
- Ci si sarebbe dovuto chiedere quali **dati** fossero **necessari** per definire le scelte di governo ai diversi livelli territoriali, per informare correttamente la popolazione e gli operatori economici e come si dovesse comporli in un quadro di coerenza generale.
- Invece, **la base informativa è rimasta sostanzialmente immutata** e ha mostrato la sua **fragilità** anche nelle ondate successive.

# I problemi principali/1

Quali sono stati i problemi principali, fin dall'inizio:

- la **manca**za di alcuni dati fondamentali;
- Il **quadro definitorio non nitido**, né per quanti preposti a raccogliere i dati né per la loro comprensione (i metadati pubblicati sono insufficienti).
- l'**assenza di collegamento fra i dati** disponibili (difficile comprendere le relazioni);
- la **scarsa iniziativa per l'approfondimento** di problemi e di opportunità emergenti;
- la **limitata produzione di informazione adeguata a sostenere i governi** - ai diversi livelli - per le decisioni da prendere; a convincere i cittadini ad assumere comportamenti adeguati e le imprese ad attivare misure di contrasto nei luoghi di lavoro;
- infine, la **scarsa "sensibilità statistica"** nella raccolta, nell'elaborazione e nella comunicazione dei dati.

# I problemi principali/2

*Dati slegati tra loro.* Non sono state prodotte stime delle probabilità di passaggio fra gli stati delle persone che si imbattono nella patologia: da asintomatico a eventualmente sintomatico, al trattamento domiciliare, al ricovero in terapia sub-intensiva e intensiva, alla guarigione o malauguratamente alla morte, informazioni fondamentali per consentire la lettura delle conseguenze sulla salute delle persone e una previsione migliore dell'impatto sul sistema sanitario.

*Assenza di un'informazione di carattere longitudinale.* Solamente da storie individuali si possono definire le diverse storie collettive che si osservano nel tempo.

# I dati necessari/1

Nella fase di insorgenza della pandemia, quali erano **le principali domande** che i cittadini si ponevano e alle quali occorreva dare risposta:

1. **qual è la probabilità di contagiarsi;**
  - se contagiati in modo lieve, di trovare risposta nella medicina territoriale;
  - se contagiati seriamente, di trovare accoglienza nelle strutture ospedaliere, se del caso per terapie sub-intensive e intensive;
  - se entrati in ospedale, di uscirne guariti;
2. **in quali luoghi e contesti c'è rischio maggiore di contagio** (anche sfruttando i big data);
3. infine, **sulla base di quali informazioni, vengono prese le decisioni** circa le misure restrittive da adottare.

# I dati necessari/4

## Oggi ci sono nuove domande:

1. gli effetti della campagna vaccinale sulla probabilità di contagio, tipologia di ricovero e decesso, e soprattutto
2. l'impatto di medio e lungo periodo del contagio sulla salute e la vita delle persone. E dunque, indirettamente, sulle strategie delle imprese e delle istituzioni.

Gli effetti della campagna vaccinale sui ricoveri e i decessi sono ben documentati dall'ISS, non altrettanto sui contagi.

Sull'impatto del Covid-19 attraverso l'impianto di uno studio longitudinale più avanti presento **una proposta** del nostro gruppo di ricerca.

# Centralità dei dati: il bilancio di un'occasione mancata/1

Il «rito quotidiano» doveva essere un appuntamento di grande chiarezza e autorevolezza, fondato sulla solidità dei processi di raccolta e integrazione dei dati, ben documentato, frutto di una collaborazione tra istituzioni con le diverse competenze necessarie (quotidiano?).

Occorreva implementare un **sistema di monitoraggio dinamico**, consapevoli dei limiti delle informazioni provenienti da quanto si manifestava nelle sole strutture ospedaliere, della necessità di miglioramento nei processi di raccolta dei dati a seconda delle criticità e delle domande via via emerse, della responsabilità di intervenire con rigore e autorevolezza nel dibattito sulle misure da adottare e su come fronteggiare l'emergenza (voce univoca?).

Non si sono sfruttati gli errori e l'impreparazione nella misura della dimensione e degli effetti della pandemia durante la prima ondata per fare meglio nella seconda e così via.

Non si è investito in un sistema informativo derivante dai dati delle autorità sanitarie integrato con quelli raccolti con indagini statistiche (campioni casuali).

# Centralità dei dati: il bilancio di un'occasione mancata/2

- La nostra tradizionale debolezza sulla *numeracy* ne è uscita rafforzata? La **fiducia nelle evidenze numeriche**, nelle previsioni da modello ne escono rafforzate? I dati e le informazioni statistiche si sono tradotti nella conoscenza necessaria per valutare e decidere?
- L'insufficienza e talvolta l'ambiguità dei dati hanno alimentato la **distanza delle posizioni** fra i tecnici e nel confronto politico e istituzionale.
- La **confusione** fra i cittadini e nei luoghi delle decisioni è stata grande.
- Non è stata **contrastata la comunicazione senza informazione**. Doveva risultare chiara la distinzione fra sprovvedutezza, genericità e competenza.
- La **trasparenza** non è stata sempre assunta come valore distintivo da parte delle istituzioni responsabili delle decisioni per il contrasto dell'epidemia.
- **Alle esortazioni non sono seguite sistematicamente le azioni**; agli errori, sempre possibili, deve seguire l'assunzione di responsabilità.

# Centralità dei dati: il bilancio di un'occasione mancata/3

- Non sono stati resi disponibili i **dati individuali anonimizzati alla comunità scientifica** (le resistenze nascondono un problema di qualità?).
- Non sono state superate le **gelosie istituzionali e disciplinari**.
- L'accademia si è esercitata sui **modelli** di tendenza e previsione della dinamica della diffusione del virus ma assai meno sulla **qualità dei dati** sottostanti.
- La **comunità statistica** non è riuscita ad avere un ruolo rilevante nei luoghi delle decisioni.

Al di là delle Regioni, i territori, i **Comuni** in primis, si sono proposti e sono stati coinvolti nei processi di raccolta, analisi e comunicazione dei dati sull'evoluzione della pandemia?

E ora, nella valutazione dell'impatto che ha avuto sulla vita delle persone, i Comuni possono contribuire, ad esempio, mostrandoci l'evoluzione della domanda e dell'offerta dei servizi resi ai cittadini?

# Una variabilità nei dati della diffusione della pandemia non sempre decifrabile

Molti fattori possono determinare la **variabilità osservata**.

- Specificità della **sindrome e mutazioni** del virus.
- Caratteristiche della **popolazione**, specialmente composizione per **età**.
- **Stato dei sistemi sanitari**.
- Tipologia e intensità delle **misure di contenimento** introdotte.
- Stato più o meno avanzato dei **sistemi di raccolta** delle informazioni.
- Disomogeneità nelle **definizioni adottate**.
- Da ultimo nel corso del tempo, l'andamento della **campagna vaccinale** e l'adesione da parte delle persone rispetto al ciclo proposto.

# I test

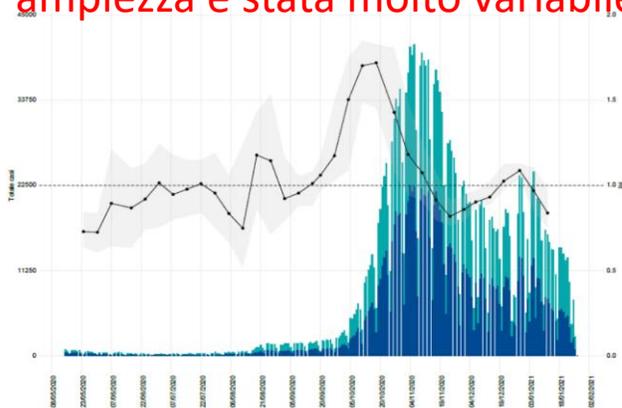
- Fin **dall'inizio** si è trattato di **test molecolari** attraverso tampone oro-naso-faringeo; da **metà gennaio 2021** si sono aggiunti i **test rapidi**, da **qualche mese anche i self test**. L'inclusione dei test rapidi e l'esclusione dei self test alterano la serie storica sono **motivo di confusione** nella lettura delle dinamiche in atto.
- Il **numero di tamponi processati per 100.000 abitanti** è risultato molto **variabile nel tempo e fra territori**.
- A marzo-aprile 2020, i test erano riservati essenzialmente ai sintomatici; attualmente, **si distribuiscono fra sintomatici, contatti, screening**. Il **mix** è **variabile** nel tempo e fra regioni.

# I contagi

- Il **numero dei contagi** accertati è **strettamente collegato al numero di test** processati (non solo alla diffusione dell'infezione).
- I **test** effettuati e processati **sui contatti e per screening** sono stati via via più numerosi nel corso del tempo e **hanno fatto emergere gli asintomatici e i pauci-sintomatici**.
- I **“guariti clinicamente”** sono riconducibili presumibilmente a **re-testing**.
- I casi di **re-testing** e anche quelli emersi da **screening andrebbero scorporati** ai fini del calcolo del tasso di positività, sia dal numeratore (test positivi) sia dal denominatore (test eseguiti).

## L'indice di trasmissione del contagio $R_t$

- **Presentato** dai media **come l'elemento di preoccupazione o conforto fondamentale** e soprattutto come l'elemento di prova della bontà degli interventi operati nelle settimane precedenti o come indizio di quelli da assumere successivamente.
- Calcolato **sui soli sintomatici dei quali si conosca la data di comparsa dei sintomi**.
- Comunicato come stima puntuale e **trattato come deterministico**, mentre sconta un errore statistico abbastanza ampio.
- Nel corso del tempo sono stati forniti un **intervallo di credibilità** al 95% e poi un **range** pure al 95% la cui **ampiezza è stata molto variabile**, indizio preoccupante per l'**affidabilità della stima**.



Stima dell'indice  $R_t$  medio a 14 giorni e numero di casi per data di prelievo/diagnosi distinti per presenza o assenza della data di inizio dei sintomi

# Ricoveri

- I dati sui ricoveri **sono meno ambigui**.
- Rappresentano il **riferimento principale** per valutare la pressione sul sistema sanitario e quindi calibrare le restrizioni da adottare per contrastare il diffondersi dell'epidemia.
- Durante la campagna vaccinale sarebbe stato importante dare maggiore risalto, in modo appropriato, agli stock, ai flussi e ai tassi di ricovero **secondo la fase del ciclo vaccinale** in corso da parte dei ricoverati (1°, 2° e buster), separatamente per reparti Covid, terapia sub-intensiva e terapia intensiva.

# Decessi

- Per i **morti** registrati 'da Covid-19' si dispone di **informazioni** di sintesi **abbastanza dettagliate**: luogo, genere ed età.
- Il **numero di morti per Covid-19** è stato **sottostimato** in una prima fase e forse **sovrastimato** nella fase successiva.
- Il **bilancio finale** dei morti per l'epidemia dovrà **tenere conto** in qualche misura anche dell'**eccesso di mortalità** generale.
- Si possono calcolare **tassi di letalità** (morti per 100 contagiati) e **tassi di mortalità** (morti per 100.000 abitanti).
- Per il primo tasso, si dispone dei **solli contagiati accertati**, con una presenza (assoluta e relativa) di asintomatici e *pauci*-sintomatici variabile nel tempo e territorialmente.
- La **letalità** è **correlata positivamente** con **l'età dei contagiati**, e **soprattutto con la presenza di altre patologie**. Dati e studi sulla co-morbilità dei decessi registrati da Covid sono importanti.

# Il sistema di monitoraggio necessario/1

## 1. Integrare informazioni sanitarie, raccolte in modo rigoroso, con altre campionarie.

- Campione ripetuto nel corso del tempo per **stimare la prevalenza** e seguirne l'evoluzione.
- Campione panel per stimare i **cambiamenti di stato** dei contagiati.
- Campione **panel di scuole e alunni** per stimare il tasso di contagio nel sistema educativo e seguirne l'evoluzione.
- Campione **panel di medici di base** per stimare il numero di pazienti con quadri sintomatici riconducibili a Covid-19, il loro tasso di positività al test e il decorso successivo.
- Monitoraggio costante della situazione nelle **residenze sanitarie assistite**.

## 2. Individuare, attraverso sistemi di tracciamento, dei **luoghi e dei contesti nei quali la trasmissione del SARS-CoV-2 è più frequente**.

## 3. Utilizzare le **tecnologie dell'informazione** e i **big data** per stimare gli effetti delle restrizioni adottate

## 4. Copertura informativa del **piano vaccinale**, non limitata al solo numero dei vaccinati ma alla lettura dei contagi e delle loro conseguenze sul sistema sanitario.

# Il sistema di monitoraggio necessario/2

- Valutazione delle **conseguenze di Covid-19 sulle condizioni di salute delle persone** dopo la guarigione.
- L'**Invalsi** dovrebbe condurre un progetto specifico per la valutazione dell'effetto differenziale della didattica a distanza rispetto a quella in presenza sul rendimento scolastico degli alunni.
- Le **università**, in modo coordinato, dovrebbero misurare gli effetti sulle carriere e sul rendimento agli esami della didattica a distanza e in modalità mista.
- L'**Istat** potrebbe aggiungere stabilmente alcune domande nei questionari di rilevazione periodica sulle imprese e sulle forze di lavoro (relativamente a presenza e intensità dello *smart working*; l'adozione delle norme igieniche essenziali; il rispetto delle misure di contenimento via via introdotte da parte delle persone).
- Il sistema dovrebbe essere in grado di **attrezzarsi rapidamente** per produrre stime affidabili **su questioni rilevanti che emergessero** via via.

# Due proposte

Marzo 2020

① Indagine campionaria ripetuta nel tempo per stimare **il tasso di contagio** riferito alla popolazione (passaggio da campione 'di convenienza' a campione probabilistico).

Marzo 2022

② Impianto di una indagine campionaria longitudinale sugli **effetti della pandemia** sulla vita delle persone, nel medio e lungo periodo.

Le idee progettuali sono state elaborate da un gruppo di ricerca che ha portato avanti, dall'inizio del 2020, studi relativi all'impianto di sistemi integrati di raccolta di dati sulla diffusione della pandemia Covid-19.

Oltre al sottoscritto ne fanno parte, Alberto Zuliani, Piero Falorsi, Giuseppe Arbia, Vincenzo Nardelli. Sulla seconda proposta anche Alessandra De Rose e Elena Ambrosetti della Sapienza.

La proposta ② che oggi vi porto è nuova e mi fa piacere presentarla e valutarla con voi.

## ① Proposta di campionamento (Alleva et al., 2022, JOS)

Con un gruppo di ricercatori ho avanzato a più riprese di realizzare un'indagine campionaria ripetuta nel tempo per fare inferenza sul tasso di contagio riferito alla popolazione.

La nostra proposta era quella di un campione rappresentativo della popolazione italiana, eventualmente a partire da una certa età, sul quale effettuare il test diagnostico per valutare la prevalenza, e al quale proporre un questionario per raccogliere elementi utili a interpretare le dinamiche di contagiosità.

Avanzata nel **marzo 2020**, in piena prima ondata dell'epidemia, quando si conosceva ancora molto poco di Covid-19.

La proposta era articolata: oltre a stimare la prevalenza puntuale, era previsto di seguire longitudinalmente parte del campione per valutare i passaggi di stato. In particolare mediante un approccio misto a partire da:

- un campione di contagiati e di loro contatti (Gruppo A);
- un campione di residenti (Gruppo B), rappresentativo della popolazione, e delle persone che hanno avuto contatti con tra loro sono risultati infetti.

Le stime ottenibili soddisfacevano le caratteristiche di non distorsione e presentavano un'efficienza di gran lunga superiore a quella ottenibile mediante un campionamento casuale semplice.

# Proposta di campionamento ①

- **Gruppo A** è il sottogruppo costituito dagli individui per i quali è stato **accertato uno stato di infezione** (potrebbero essere ricoverati in ospedale o in quarantena obbligatoria) e da tutte le persone che hanno avuto contatti con loro nei giorni precedenti.

Conosciute attraverso i dati raccolti dalle autorità sanitarie pubbliche.

Pertanto, concentrare gli investimenti in risorse sull'osservazione dei contatti di questo gruppo massimizza il numero di persone infette osservate nel campione.

- **Gruppo B** contiene persone sane, persone infette ma non consapevoli di esserlo, coloro le cui infezioni sono ancora in una fase di incubazione e tutte le persone che hanno avuto contatti con loro.

È necessario osservare il Gruppo B per produrre stime affidabili per l'intera popolazione, e questo è obbligatorio per stimare correttamente il tasso di persone infette e il tasso di mortalità

# Proposta di campionamento ①

Era previsto che la rilevazione fosse **ripetuta ogni due settimane** in modo da seguire l'evolvere dell'epidemia nel corso del tempo, modulando opportunamente i due campioni e il panel.

**Allora non se ne fece nulla.**

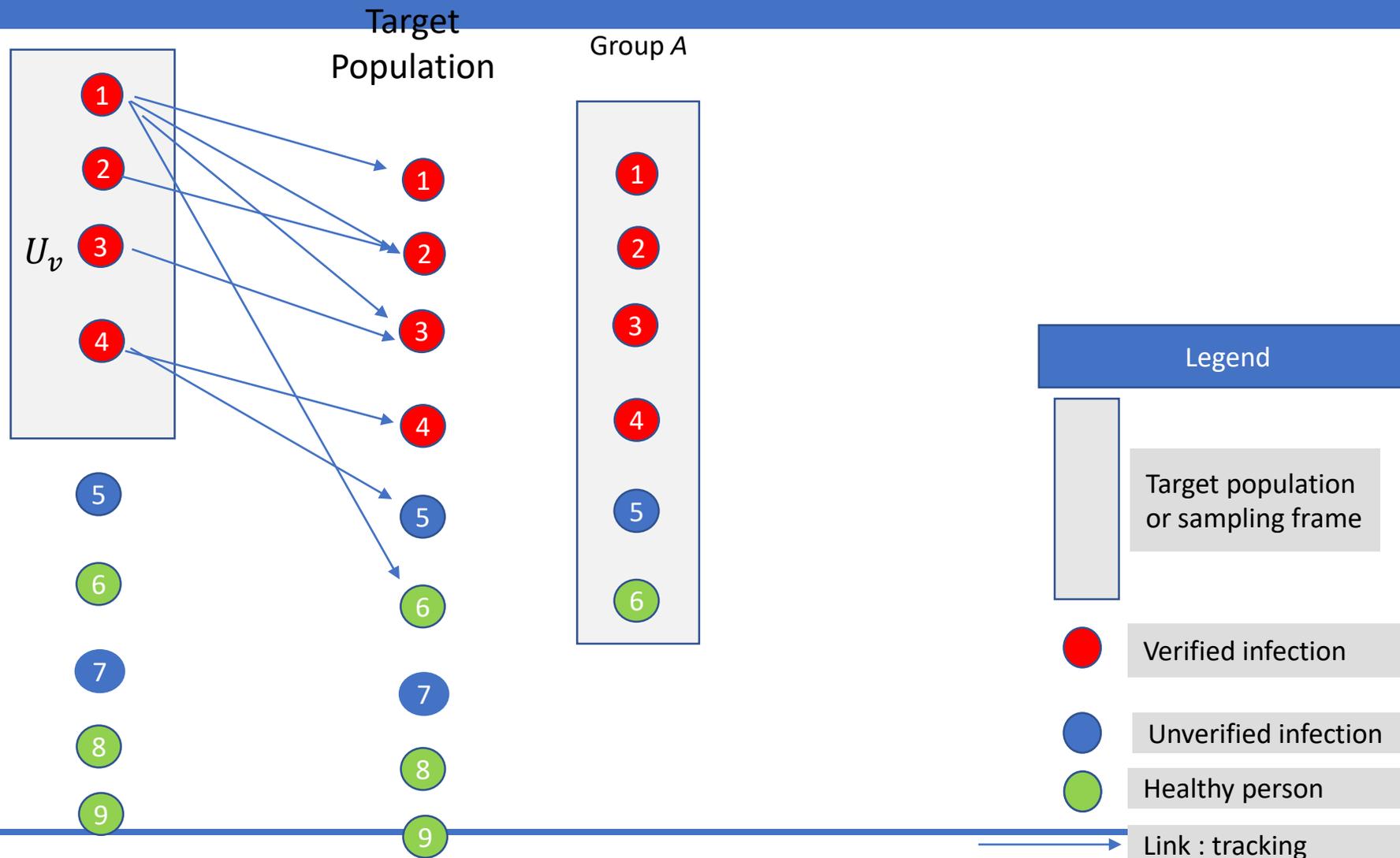
Una giustificazione potrebbe essere rinvenuta nella situazione di emergenza sanitaria che si stava vivendo, che suggeriva di destinare i tamponi ai soli sintomatici.

Nel corso della seconda ondata si processavano circa 200.000 tamponi ogni giorno e destinarne il 5% una volta ogni due settimane alla stima della prevalenza sarebbe stato del tutto fattibile. A maggior ragione nelle ondate successive.

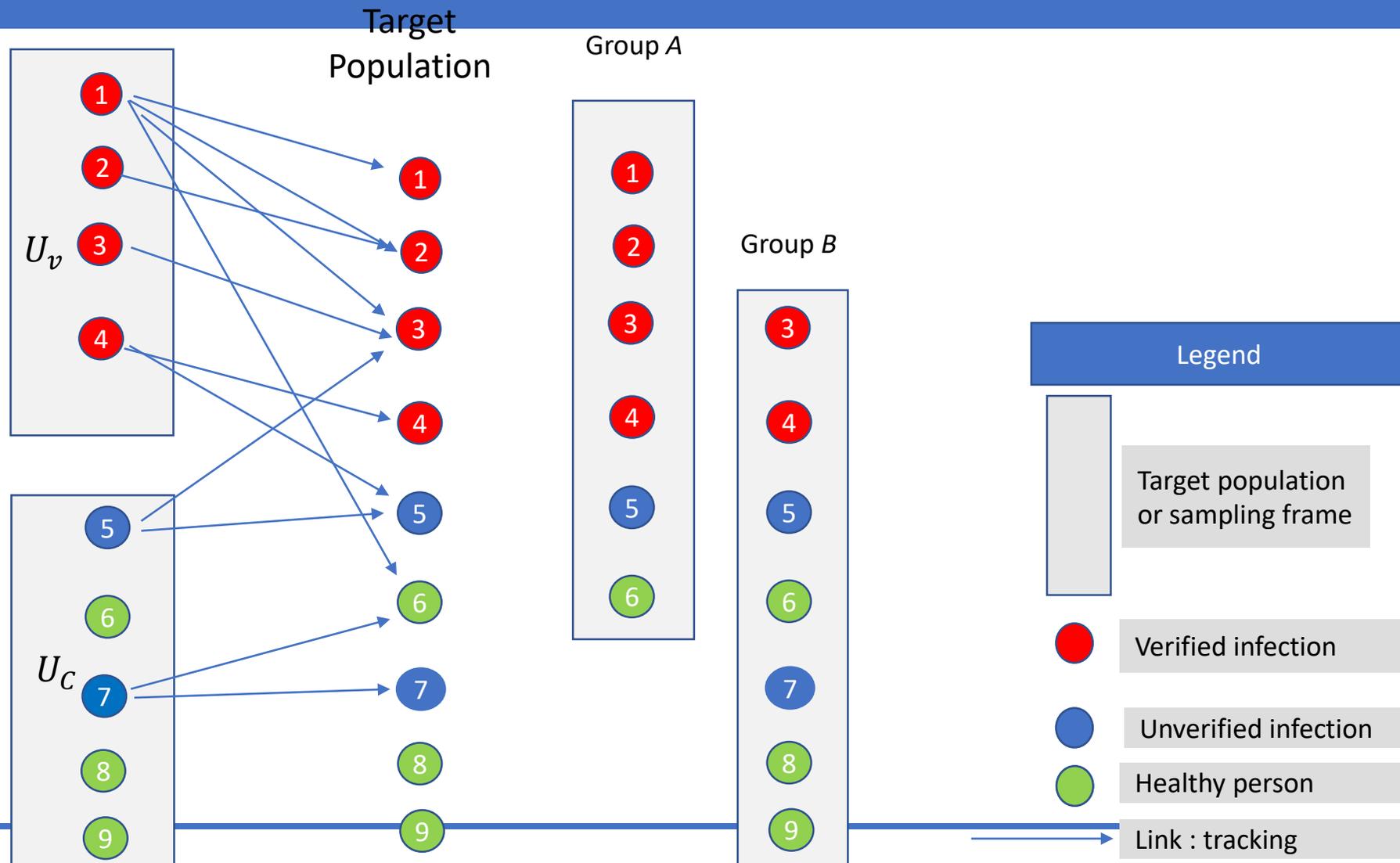
Complessivamente valuto **un'occasione perduta** da parte della statistica e della statistica ufficiale nel riuscire a contribuire alla definizione di un quadro evolutivo affidabile della pandemia.

Di fronte alla grande incertezza che ha generato il Covid-19 la statistica, che di incertezza se ne nutre per costituzione doveva avere un ruolo più importante. Un'occasione perduta considerato il mandato e l'indipendenza degli Istituti, le competenze e professionalità che possiede, la fiducia di cui gode da parte degli utenti.

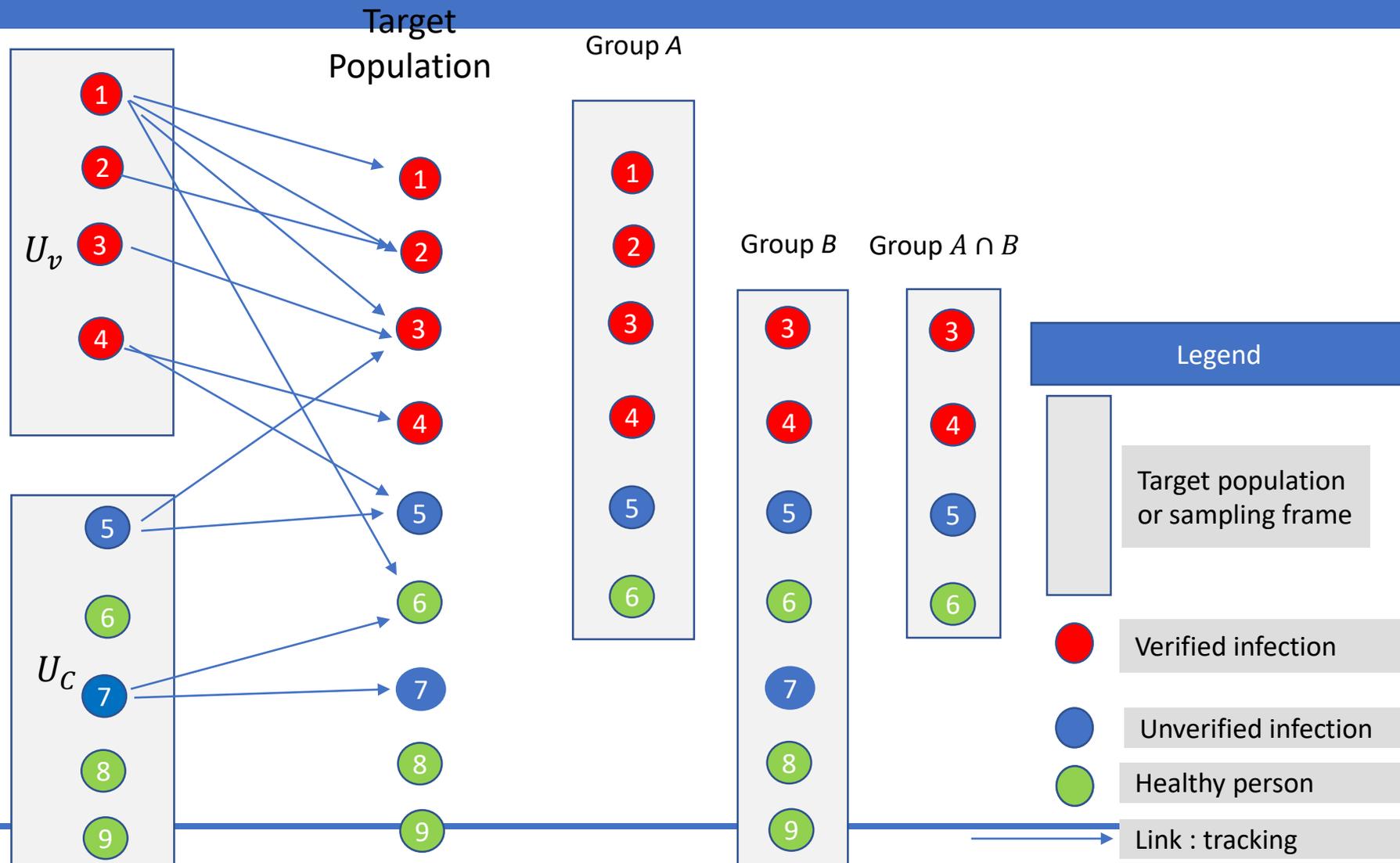
# Metodologia di campionamento ①



# Metodologia di campionamento ①



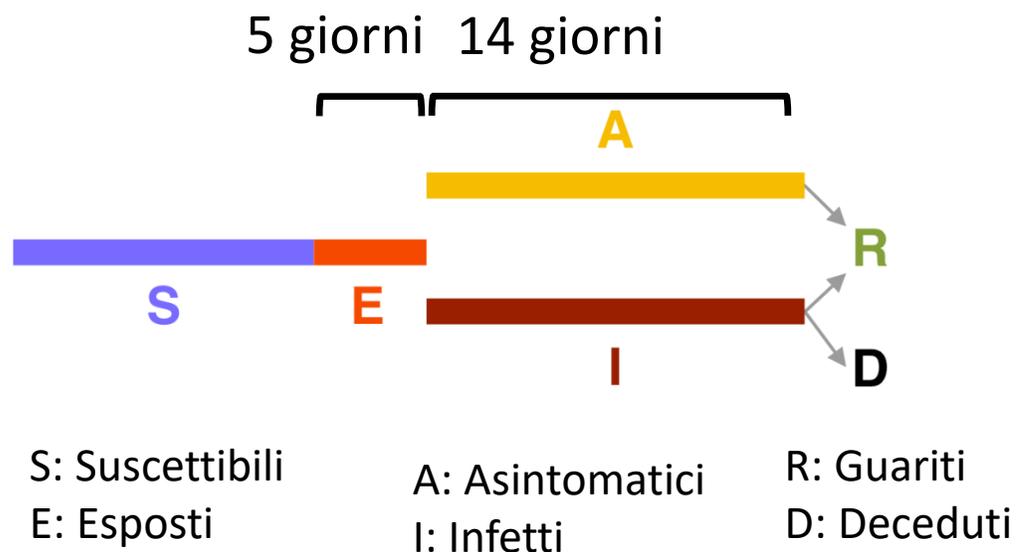
# Metodologia di campionamento ①



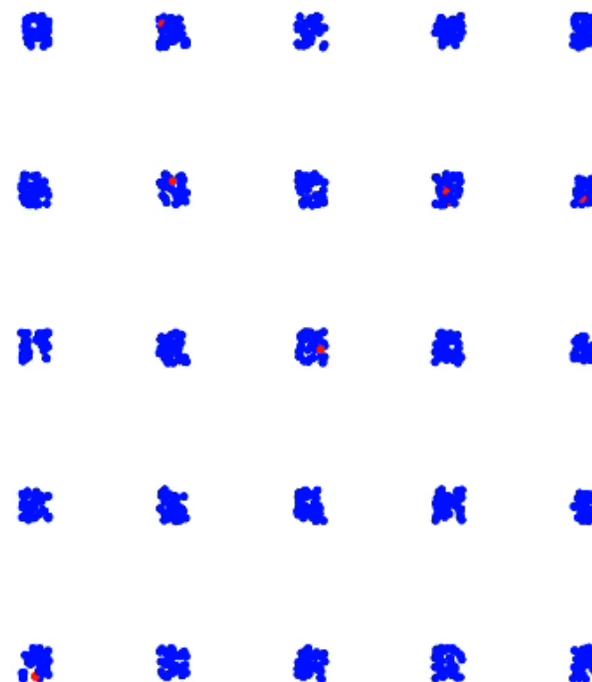
# Simulazione ①

Considerata l'impossibilità di ottenere **dati individuali anonimizzati** per finalità scientifiche, i metodi di campionamento sono testati su un **dataset generato** in cui vengono tracciati 20.000 individui.

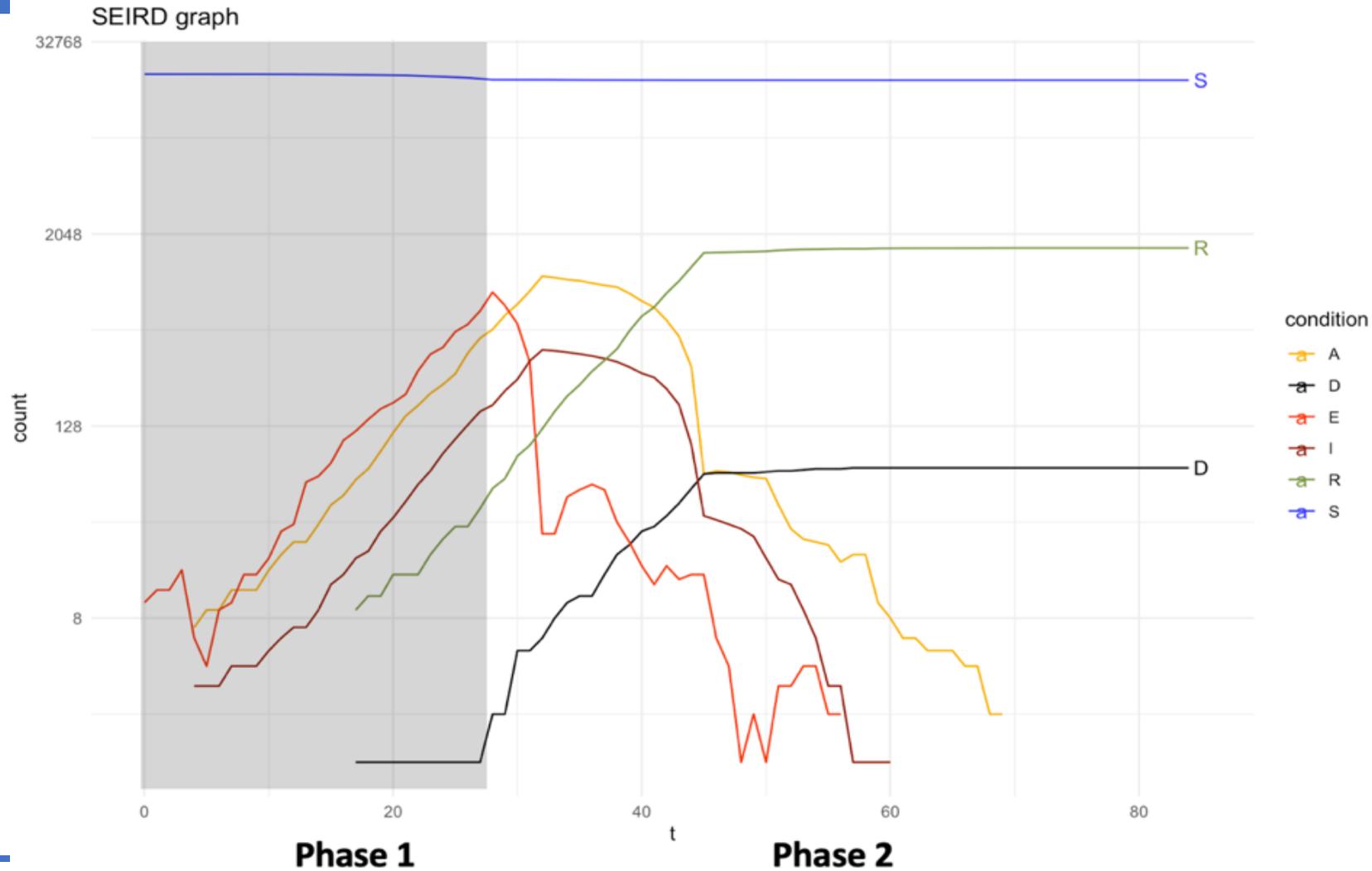
## Simulazione epidemia



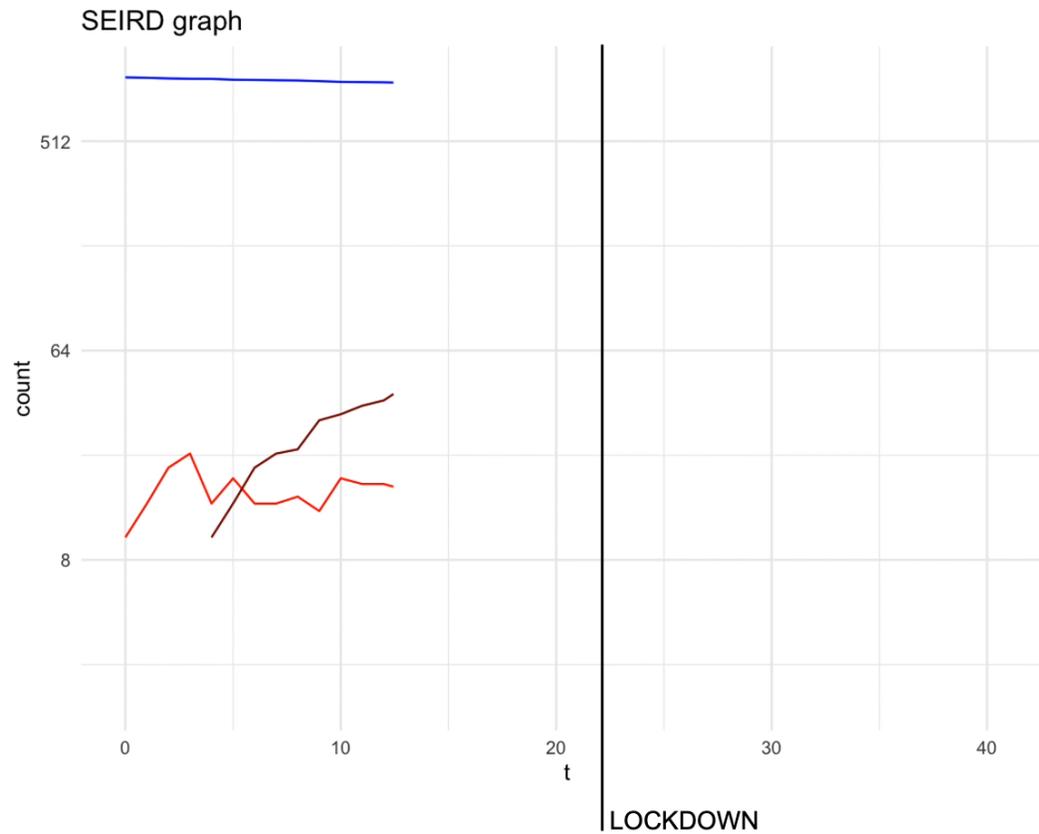
## Simulazione mobilità ed interazioni sociali



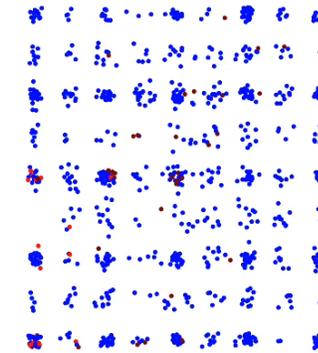
# Simulazione ①



# Simulazione ①

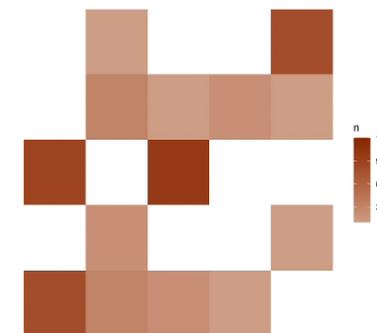


## MOBILITY



## OUTBREAKS

Day: 14



## Risultati della simulazione: Distorsione ed Efficienza ①

<b>Giorno</b>	<b>Percentuale individui infetti nella popolazione</b>	<b>Distorsione relativa</b>	<b>Numerosità campione tradizionale equivalente a 1000 testati con tracciamento</b>
<b>15</b>	0.58%	0.0019	72
<b>25</b>	3.94%	0.0000	161
<b>35</b>	6.54%	0.0045	190

# Cosa si poteva fare? L'esempio dell'Office for National Statistics (ONS)

- **Coronavirus (COVID-19) Infection Survey, UK**
- **Sample survey from 18 April 2020, results every two weeks.**
- Estimates for England, Wales, Northern Ireland and Scotland, and for the main cities.
- This survey is being delivered in partnership with University of Oxford, University of Manchester, UK Health Security Agency and Wellcome Trust.
- This study is jointly led by the Office for National Statistics (ONS) and the Department for Health and Social Care (DHSC) working with the University of Oxford and Lighthouse laboratory to collect and test samples.

Un sistema informativo dinamico, che ha sfruttato il mandato istituzionale e la competenza dell'Istituto nazionale di statistica, e la storica indagine longitudinale [Understanding Society](#).

**Explore by topic**

Comparisons

Infections

Hospitals

Deaths

Vaccines

Antibodies

Well-being

Lifestyle

## Infections decreased

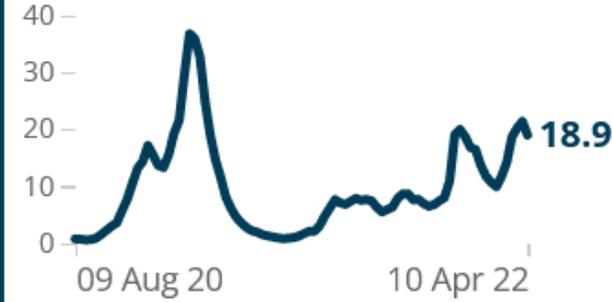
Percentage testing positive for COVID-19, England



Office for National Statistics

## Hospital admissions decreased

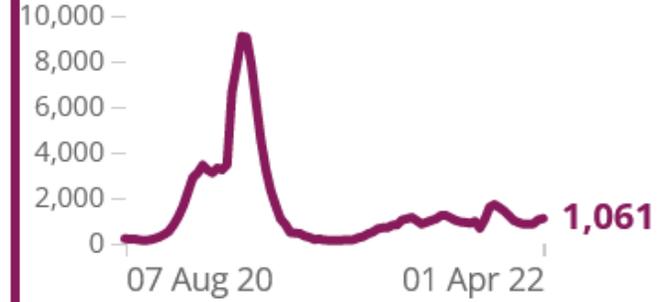
Hospital admissions involving COVID-19 per 100,000 people, England



UK Health Security Agency

## Deaths continued to increase

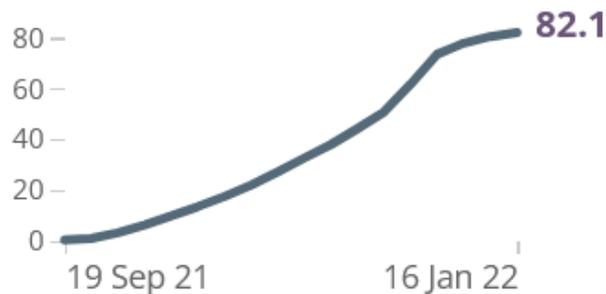
Deaths registered by week involving COVID-19, UK



Office for National Statistics

## Around 8 in 10 adults have had 3 vaccine doses

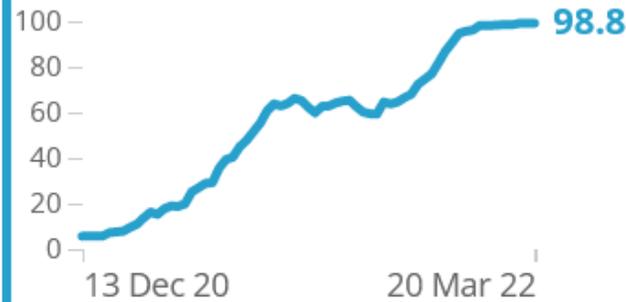
Percentage who have received three doses of a COVID-19 vaccine, England



Office for National Statistics

## Over 9 in 10 adults had antibodies

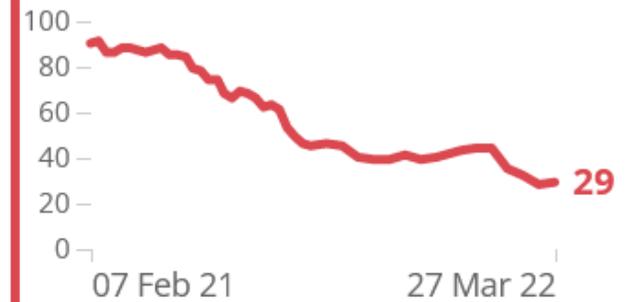
Percentage testing positive for antibodies to COVID-19, England



Office for National Statistics

## Social distancing remained low

Percentage always/often maintaining social distancing

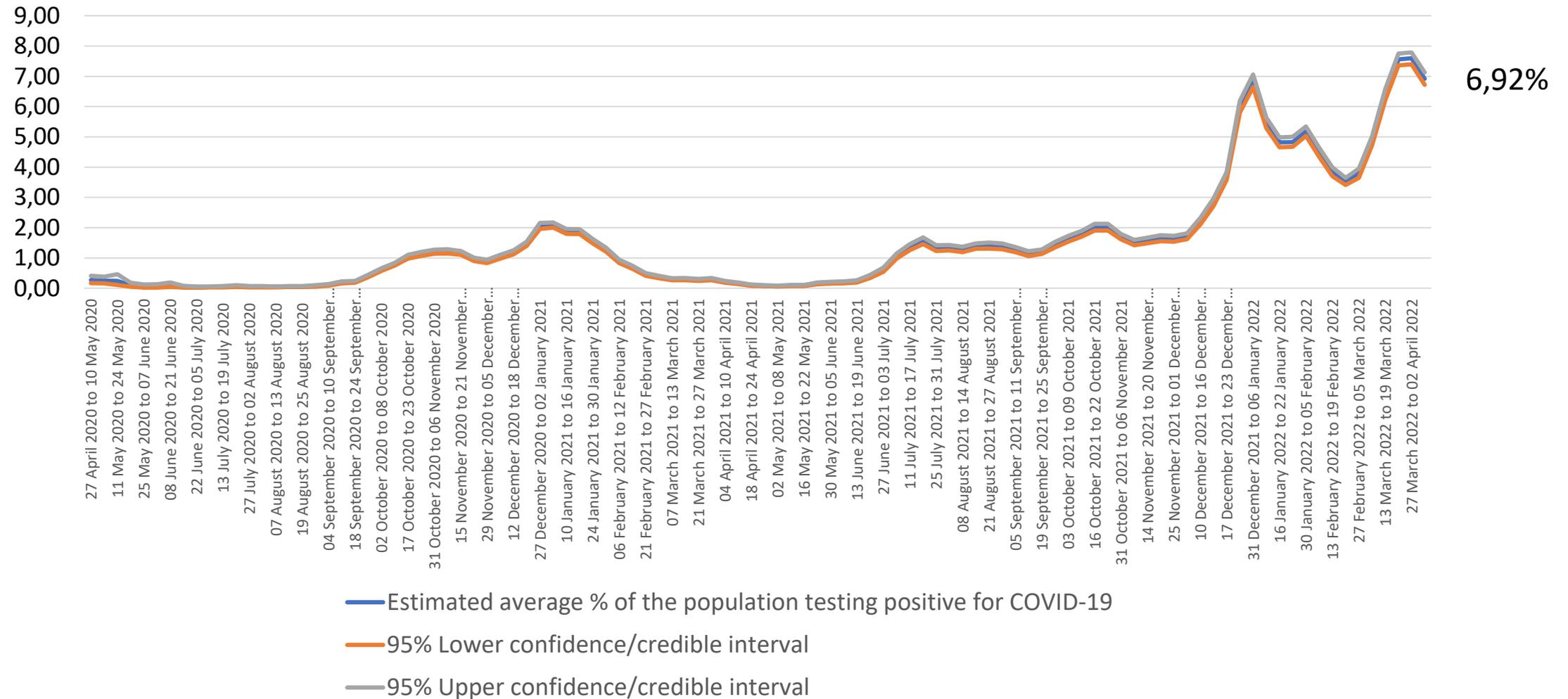


Office for National Statistics

## Number of people in sample



## Official reported estimates of the percentage of the population testing positive for COVID-19, UK countries



# Coronavirus (COVID-19) latest insights

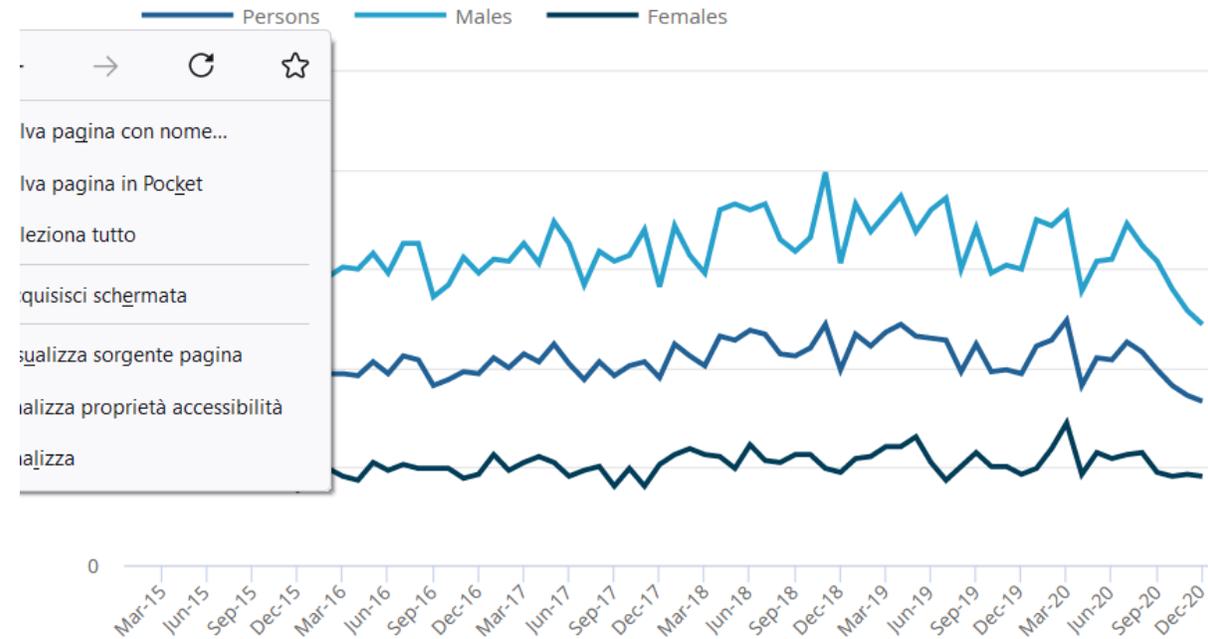
## 14 April 2022

### Latest COVID-19 headlines

- [Infections remain high in Wales but fall in other UK countries](#)
- [No evidence suicides increased during the COVID-19 pandemic](#)
- [Retail and recreation visits remain below pre-COVID-19 levels](#)

### Suicide rates in 2020 remained fairly consistent with previous years

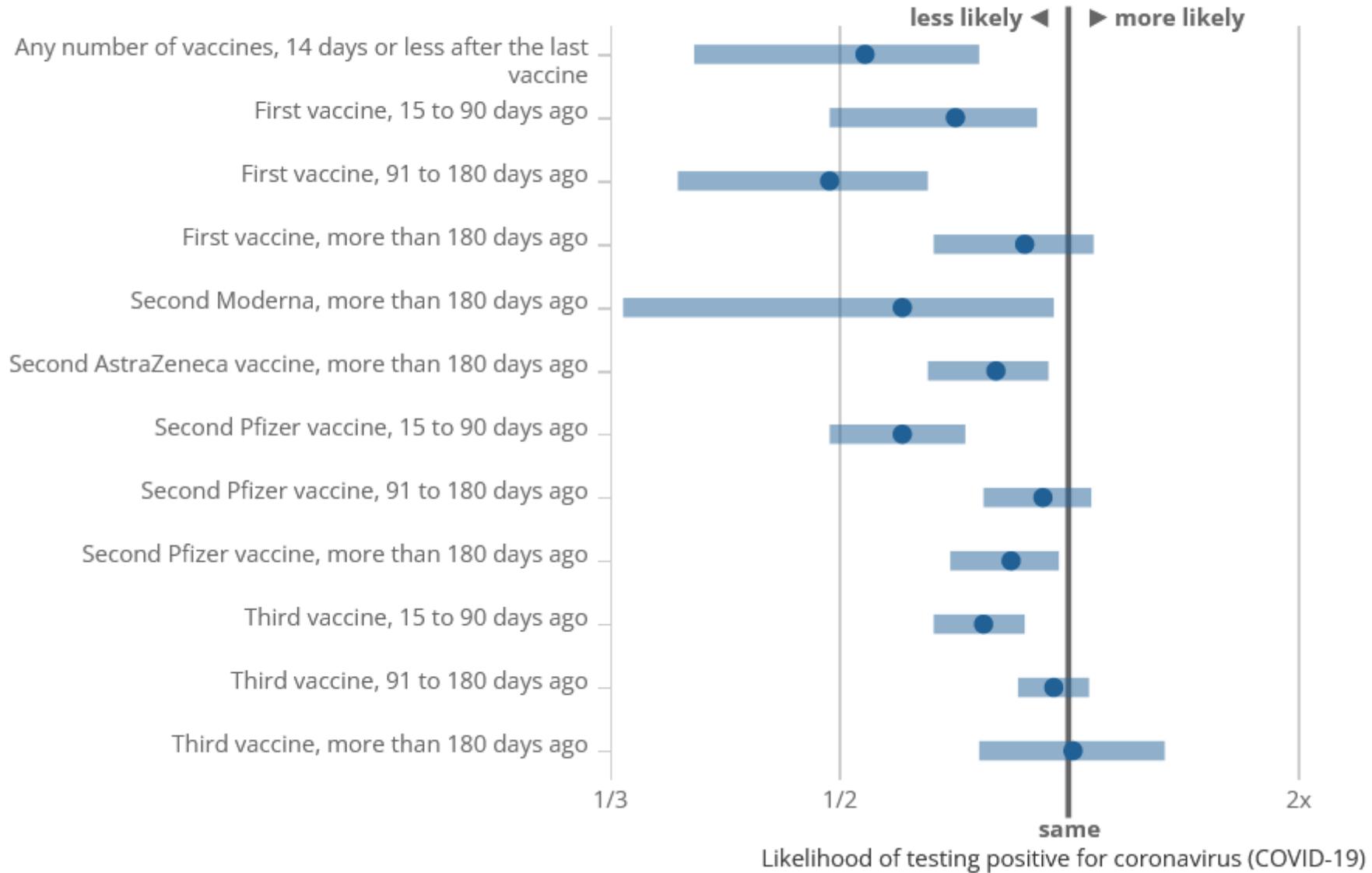
Monthly age-standardised suicide rates by sex, England and Wales, deaths occurring between January 2015 and December 2020



Source: Office for National Statistics – Deaths from suicide that occurred in England and Wales

● Estimates with 95% confidence intervals

**Vaccination status (Reference group: Not vaccinated)**



# Impact of coronavirus (COVID-19) on lifestyle

---

- [Travel](#)
- [Impact on society](#)
- [Household spending and disposable income](#)
- [Preventative measures](#)
- [Homeworking](#)
- [Impact on education](#)
- [Impact on disabled people](#)
- [Further information](#)

## ② Proposta di per uno studio longitudinale sugli effetti del Covid-19 (Allewa et al., in corso)

- A distanza di due anni dall'insorgenza della pandemia, **manca in Italia uno studio longitudinale** che consenta di valutare in modo sistematico gli effetti di Covid-19 sulla vita delle persone, elemento cruciale per comprenderne l'impatto sulla società e sostenere le scelte ai diversi livelli di governo e ambiti di responsabilità.
- Le **esperienze in corso in altri Paesi** hanno consentito e consentono di dare risposta a numerose questioni che altrimenti rimarrebbero sospese.
- E' stata delineata una proposta di indagine longitudinale riferita a Covid-19, tratteggiando lo schema di disegno campionario e fornendo una stima delle risorse finanziarie e organizzative necessarie.
- E' fondamentale prevedere la regia da parte delle autorità sanitarie nazionali e il coinvolgimento attivo delle Regioni e dell'Istat.

# Gli studi longitudinali/1

Gli studi longitudinali offrono informazioni importanti su aspetti della società e sono condotti attualmente da diverse istituzioni e centri di ricerca in numerosi paesi.

Sono progettati per **seguire i cambiamenti che avvengono nel corso del tempo negli individui e nelle famiglie.**

Permettono di seguire i partecipanti fornendo una prospettiva sulla loro vita e di esplorare i cambiamenti e allo stesso tempo la stabilità dei loro comportamenti.

I **tipi di studi longitudinali** sono:

- Indagini panel;
- Indagini per coorte;
- Studi retrospettivi.

Alcune delle indagini longitudinali in atto sono state utilizzate per rilevare gli **effetti della pandemia Covid-19.**

## Gli studi longitudinali/2

L'indagine più significativa **a livello europeo** è la **SHARE** (*Survey of Health, Ageing and Retirement in Europe*), un'infrastruttura di ricerca per lo studio degli effetti delle politiche sanitarie, sociali, economiche e ambientali sul percorso di vita dei cittadini europei.

È realizzata dal 2004 in 28 paesi europei e Israele ed è coordinata presso il *Max Planck Institute for Social Law and Social Policy*. Finora sono state realizzate otto waves d'indagine tra il 2004 e il 2020.

Il campione italiano è di circa 2000 individui di 50 anni e più.

**A livello nazionale** in miglior esempio di indagine longitudinale sul Covid-19 è quello del **Regno Unito**, nell'ambito della storica indagine longitudinale **Understanding Society** condotta *dall'Office for National Statistics* (ONS). Sono state realizzate numerose ondate di indagini specifiche su Covid-19, a partire dal mese di aprile 2020.

Valuta l'impatto della pandemia sul benessere degli individui e delle famiglie.

Pure nel Regno Unito sono state attuate modifiche e aggiornamenti ai questionari di altre indagini panel esistenti per monitorare gli effetti della pandemia sulla popolazione.

## Gli studi longitudinali/3

L'indagine panel francese **ELIPSS** (Étude Longitudinale par Internet Pour les Sciences Sociales), progettata e gestita da Sciences Po e CNRS. Rilevazione mensile partita nel 2012, si è giovata di un finanziamento dell'Agence Nationale de la Recherche (ANR). I partecipanti sono dotati di un tablet e di un abbonamento a internet mobile rispondendo da remoto, via web, in occasione delle diverse waves.

E' realizzata ad hoc per la comunità scientifica in modo da colmare la carenza di mezzi per la realizzazione di indagini longitudinali sulle tematiche proprie alle scienze umane e sociali.

Altri paesi europei hanno proposto questionari ad hoc nell'ambito di indagini longitudinali esistenti per studiare gli effetti di Covid-19 sulla salute fisica e mentale dei cittadini e valutarne l'impatto sociale ed economico.

In Italia si segnala l'indagine longitudinale [Italian Lives](#) (ITA.LI) realizzata dal Dipartimento di Sociologia e ricerca sociale dell'Università di Milano-Bicocca nel corso del 2020.

Un campione di 950 intervistati ha risposto, tra aprile e agosto 2020, a un breve questionario inteso a individuare i cambiamenti intervenuti per i principali aspetti di qualità di vita a causa di Covid-19.

# Gli studi longitudinali/4

Per la statistica ufficiale sarebbe di estremo interesse realizzare uno studio campionario longitudinale sulla popolazione residente in Italia per valutare gli effetti dell'attuale pandemia.

Si colmerebbe una lacuna; si disporrebbe di un quadro evolutivo analitico, indispensabile per interpretare i risultati di *focus* specifici.

## La proposta ②

La proposta è di suddividere la popolazione in tre sottopopolazioni:

A: **ricoverati finora** in strutture ospedaliere per Covid-19 e guariti o ancora in terapia ospedaliera;

B: **contagiati conosciuti finora**, non ospedalizzati e guariti, o ancora in terapia domiciliare;

C: **restante popolazione** (compresi i contagiati non conosciuti perché asintomatici o *pauci*-sintomatici, non diagnosticati o volutamente occultati).

Potrebbero verificarsi transiti e sovrapposizioni fra le tre sottopopolazioni l'inizio della rilevazione.

Per la sottopopolazione A la soglia di età sarebbe indicata tenendo conto del consenso dei genitori dei rispondenti. L'idea sarebbe di estenderla alle persone con 10 anni o più attingendo le risposte dai soggetti adulti della famiglia, tipicamente la coppia genitoriale.

Nella proposta è suggerita una possibile estensione dello studio ai nuclei famigliari.

## La proposta ②

Per la **sottopopolazione A** degli ospedalizzati sarebbe assai conveniente poter effettuare il campionamento a partire da **elenchi in possesso delle autorità sanitarie**.

Collettività di questo primo gruppo sono già oggetto di analisi longitudinali locali da parte delle strutture ospedaliere e delle autorità sanitarie. Sarebbe interessante procedere a una ricognizione dell'esistente e all'integrazione delle informazioni eventualmente disponibili.

Anche per la **sottopopolazione B** dei contagiati sarebbe importante poter accedere a elenchi disponibili (presumibilmente **elenchi della medicina territoriale**).

La costruzione del panel relativo alla terza **sottopopolazione C** non pone problemi particolari.

La periodicità delle ondate andrà definita tenendo conto dell'evoluzione della pandemia.

Sono state formulate ipotesi operative per poter procedere alla stima dei costi.

## La dimensione campionaria ②

Si ritiene idonea una numerosità di 1.000 unità per ciascuna delle tre sottopopolazioni per ogni realtà geografica di interesse (ad esempio una regione, una città metropolitana).

Sarebbero prodotte stime accettabili (con  $CV < 15$ ) nel caso di un'incidenza del carattere investigato nella popolazione superiore al 5% (tabella 1).

*Errori di campionamento (CV) secondo l'incidenza del carattere sulla popolazione e la numerosità campionaria*

Valore della proporzione espressa in termini percentuali	Numerosità campionaria			
	1.000	5.000	10.000	20.000
30%	4,83	2,16	1,53	1,08
20%	6,32	2,83	2,00	1,41
10%	9,49	4,24	3,00	2,12
5%	13,78	6,16	4,36	3,08
1%	31,46	14,07	9,95	7,04

Il suggerimento per un'indagine nazionale che punti a produrre stime regionali è di operare su campioni di circa 1.000 unità per ognuna delle tre sottopopolazioni di interesse per ogni regione e di puntare, quindi, a un campione di 20.000 unità per ciascuna sottopopolazione, e dunque di complessive 60.000 unità.

## La dimensione campionaria ②

Per le sottopopolazioni A e B, **se non fosse possibile accedere a elenchi** in possesso delle autorità sanitarie, delle strutture ospedaliere e della medicina territoriale, si potrebbe procedere a un **campionamento in due fasi**, con screening nella prima fase.

Per raggiungere le numerosità di 1.000 unità per ciascuna delle due sottopopolazioni si dovrebbe coinvolgere nella prima fase campioni delle numerosità indicate nella tabella sottostante, al netto delle mancate risposte. Il dispendio di risorse per effettuare lo screening sarebbe molto alto e, **nel primo caso, praticamente insostenibile.**

Numerosità campionarie da coinvolgere nella fase di screening per raggiungere 1.000 unità campionarie

Sottopopolazioni -->	Contagiati ospedalizzati	Contagiati non ospedalizzati
Numero di interviste nella fase di screening per raggiungere 1.000 unità campionarie d'interesse	121.484	12.283

# Le variazioni nel corso del tempo da osservare attraverso i campioni delle tre sottopopolazioni ②

- salute percepita e salute oggettiva;
- prevenzione e cura;
- impatto sulla salute mentale di Covid-19;
- comportamenti assunti rispetto alla pandemia;
- comportamenti a rischio;
- stili di consumo;
- relazioni all'interno o all'esterno della famiglia;
- aiuti forniti e ricevuti;
- utilizzo e soddisfazione dei servizi pubblici e privati, con specifica attenzione a quelli sanitari;
- situazione di benessere oggettivo e percepito;
- condizione occupazionale;
- situazione abitativa,
- altre.

L'indagine qui proposta dovrebbe essere promossa dal Ministero della salute, dall'Istituto superiore di sanità e dall'Istat, con l'auspicabile collaborazione delle regioni e dei comuni, e potrebbe giovare di un finanziamento del PNRR.

# Conclusioni / 1

- Di fronte a un virus sconosciuto **l'incertezza è ineludibile ma può essere ridotta** e si può alimentare la consapevolezza sociale ponendosi le **domande giuste** e cercando di dare loro **risposte adeguate**.
- La **produzione di dati** è **centrale**. È necessario che essi siano raccolti ed elaborati all'interno di un disegno organico nel quale siano dominanti i valori dell'**interdisciplinarietà** e della **cooperazione**.
- Occorre promuovere l'integrazione fruttuosa fra i dati della **sorveglianza sanitaria** e quelli derivanti da **rilevazioni**, totali o campionarie.
- **Statistica e statistica ufficiale dovrebbero essere coinvolte maggiormente**. È importante che **l'Istat** diventi **protagonista**. L'esempio dell'ONS e della partnership con autorità sanitaria e università mostra la strada da intraprendere quanto prima-
- Uno specifico finanziamento all'interno del *Recovery Plan* potrebbe essere riservato all'allestimento di un **quadro informativo dinamico su Covid-19 e i suoi effetti** su diversi aspetti della vita delle persone. Occorre attrezzarsi per il futuro.
- L'impianto anche in Italia di una **indagine longitudinale importante** sarebbe l'investimento coerente.

# Conclusioni/2

- Deve essere **contrastata la comunicazione senza informazione**. Deve risultare chiara la distinzione fra sprovvedutezza, genericità e competenza.
- La **trasparenza** deve essere assunta come valore distintivo da parte delle istituzioni responsabili delle decisioni per il contrasto dell'epidemia.
- **Alle esortazioni devono seguire le azioni**; agli errori, sempre possibili, deve seguire l'assunzione di responsabilità.
- **Il cammino è ancora lungo.**

Grazie per  
l'attenzione