



UNIVERSITA' DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
"M.FANNO"

DIPARTIMENTO DI SCIENZE STATISTICHE

CORSO DI LAUREA IN ECONOMIA

PROVA FINALE

**"STATISTICA E COMUNICAZIONE IN TEMPO DI PANDEMIA:
QUANDO I NUMERI FANNO LA DIFFERENZA"**

RELATORE:

CH.MO PROF. TOMMASO DI FONZO

LAUREANDA: ROSA PRENCIPE

MATRICOLA N. 1188384

ANNO ACCADEMICO 2020 – 2021

INDICE

Introduzione	4
Capitolo 1 - Il caso AstraZeneca	8
1.1 Le segnalazioni di eventi tromboembolici	8
1.2 Il nesso di causalità	10
1.3 Tanto panico per nulla?	11
1.4 Due vaccini a confronto: Pfizer e AstraZeneca	13
1.4.1 La situazione in UK	14
1.4.2 La situazione in Europa	15
Capitolo 2 – L’andamento dell’epidemia attraverso indicatori e grafici	19
2.1 Il tasso di positività	20
2.1.1 Il campione di convenienza	21
2.1.2 Il conteggio dei tamponi crea difformità nei risultati	24
2.1.3 Rapporto positivi/tamponi o positivi/casi testati?	28
2.2 La variazione percentuale giornaliera dei casi totali	29
Capitolo 3 – Il coprifuoco	33
3.1 Il coprifuoco è efficace contro il coronavirus?	33
3.2 Il pre-print e la diffusione nei media	33
3.3 I risultati di alcuni studi europei	34
Conclusioni	39
Bibliografia	40
Sitografia	40

Introduzione

Viviamo in un'epoca in cui raccogliere un'enorme quantità di dati non è più un problema. Grazie alle più recenti tecnologie, è possibile comprendere ed estrarre informazioni in possesso delle più svariate tipologie di organizzazioni. L'analisi dei dati, in forma aggregata o disaggregata, è uno strumento essenziale nelle mani dei decisori e consente loro di compiere scelte basate su relazioni e dipendenze che sussistono tra i vari dati, elaborando modelli e creando previsioni più o meno complesse e affidabili. Con l'arrivo della pandemia di Covid-19, la quantità di dati e di informazioni raccolte è stata come un'onda gigantesca che ha completamente stravolto le nostre vite, monopolizzando sia l'attenzione dei media che, da marzo 2020, hanno cercato in tutti i modi di offrire informazione e aggiornare il paese sull'evoluzione della pandemia, sia la quotidianità di ciascuno che si è trasformata in una corsa verso l'ultimo dato, dal quale pensavamo dipendesse e dipenda tuttora la nostra salute e quella dei nostri cari.

Il fenomeno pandemico si è diffuso molto rapidamente nel territorio italiano coinvolgendo numerose regioni e, fin dai primi tempi, si è presentata la necessità di raccolta di dati omogenei e uniformi per ottenere un quadro della situazione dal punto di vista nazionale. Non è stato facile ottenerli, anche per via della riforma del titolo V risalente al 2001 dell'articolo 117 della Costituzione, che prevede la competenza concorrente di Stato e Regioni in materia di tutela della salute.

Quali sono i numeri in mezzo ai quali per mesi abbiamo cercato di orientarci? I dati a disposizione sono forniti dal Ministero della Salute, vengono elaborati e gestiti dal dipartimento della Protezione Civile e resi pubblici tramite comunicati e bollettini giornalieri. Si tratta di dati aggregati con dettaglio nazionale, regionale, provinciale, comunale sul conteggio dei nuovi positivi, delle ospedalizzazioni, dei ricoveri in terapia intensiva, dei decessi, dei guariti, dei tamponi effettuati ecc. L'Istituto Superiore di Sanità (ISS) inoltre, dal 28 febbraio 2020, si occupa di coordinare un sistema di sorveglianza che integra a livello individuale dati epidemiologici forniti dalle Regioni e dalle Province Autonome e dal laboratorio nazionale di riferimento per SARS-CoV-2 dell'ISS. Sono i dati relativi alle caratteristiche personali, demografiche (sesso ed età) e cliniche (presenza di altre patologie tra i soggetti contagiati) di ogni singolo caso. Questi ultimi prendono il nome di dati disaggregati. È importante sottolineare la differenza tra le due tipologie di dati.

I dati aggregati o macro-dati sono dati risultanti da calcoli effettuati su micro-dati, i quali a loro volta sono dati elementari che si riferiscono alle singole unità di rilevazione. Si parla di dato aggregato quando l'unità di riferimento è un aggregato, ossia i dati vengono rilevati in un determinato ambito territoriale. La comunità scientifica ha purtroppo ancora a disposizione principalmente dati aggregati (Ventura, 2020). La mancanza di dati disaggregati non consente a ricercatori ed esperti di svolgere alcune attività importanti, tra le quali, valutare in modo quantitativo e rigoroso l'efficacia delle misure che vengono prese dal governo come mezzi di contrasto per la diffusione del virus.

Le statistiche presentate sono spesso espresse su base quotidiana e in numeri assoluti, vengono inoltre svolte delle analisi di confronto tra le regioni. Quello che conta è avere consapevolezza sulla natura del dato, soffermandosi sulle definizioni e sui concetti associati ad esso, in modo da vedere se è possibile compararlo direttamente con altri dati, e eventualmente con quali cautele (Origoni, 2020). I modelli statistici, tenendo a mente determinate accortezze sui dati dai quali sono tratti, sono quindi molto utili perché permettono di descrivere l'evoluzione nel tempo della pandemia; cercano di monitorare l'andamento del contagio tramite la stima di indici (stime puntuali e stime intervallari) e di modelli opportuni; permettono di prevedere almeno nel breve periodo l'andamento dei contagi; e infine simulano scenari di contagio (Ventura, 2020).

Purtroppo però ogni volta che si cerca di quantificare un fenomeno, a maggior ragione se si tratta di un fenomeno epidemiologico mondiale come questo, si ha accesso, anche se utile, ad una visione parziale di quello che succede realmente. È sempre rilevante tenere a mente, e questo un epidemiologo lo sa molto bene, che quando si registra anche solo un decesso per un virus sconosciuto, ci sono probabilmente centinaia se non migliaia di persone infette.

Da marzo 2020, il Covid-19 è diventato l'argomento principale su cui confrontarsi, dibattere e trarre, con l'aiuto degli esperti, delle conclusioni che permettessero di conoscerlo sempre più da vicino. I giornalisti e gli operatori della comunicazione hanno avuto un ruolo di fondamentale importanza in questo momento storico, hanno aiutato i cittadini ad avere una percezione più chiara di quello che stava accadendo. Un ruolo per nulla semplice, quello del mediatore intellettuale, che dovrebbe rendere facilmente intellegibile e in pochi secondi situazioni che invece l'esperto riesce a spiegare in molto più tempo e in maniera anche meno comprensibile alla massa (Università di Verona, 2020). I giornalisti hanno avuto a che fare con una mole di dati mai vista prima, hanno cercato di interpretarli minimizzando le distorsioni e cercando di cogliere la complessità che questi dati portano con sé ma senza ricorrere a semplificazioni. Ci sono stati operatori televisivi, giornalisti e reporter che hanno fatto numerosi servizi visitando le terapie intensive di alcuni dei maggiori ospedali italiani e, tramite essi, hanno cercato quindi di raccontare fatti e mostrare la realtà senza alcun tipo di pregiudizio e

filtro. Altri invece hanno comunicato in toni sensazionalistici, generato panico, talvolta esagerando, probabilmente però con lo scopo di comunicare agli italiani la gravità di quanto stava accadendo. Alcuni di loro, svolgendo il proprio lavoro, hanno contratto il virus, alcuni di loro sono morti, ma è importante riconoscere quello che ci hanno lasciato, un patrimonio sociale e informativo per l'opinione pubblica di grandissimo valore (Università di Verona, 2020).

La comunicazione è quindi stata e continua ad essere uno degli strumenti più importanti per fronteggiare l'attuale emergenza sanitaria. Si è puntato molto sulla comunicazione continua e si è cercato di renderla il più efficace possibile. Adottare una comunicazione eccessiva ma efficace significa mettere in moto tutta una serie di comportamenti collettivi di contrasto per la diffusione del virus, quali il mantenere la distanza di un metro quando si parla con una persona, evitare contatti diretti quando ci si saluta, igienizzarsi le mani, indossare sempre la mascherina, evitare luoghi affollati ecc.

Dall'altro lato però, comunicare i dati della pandemia ogni giorno ha il rischio di generare l'effetto "montagne russe": un giorno salgono i contagi e l'epidemia peggiora, un altro scendono e allora le cose migliorano (Sestili, 2020). Purtroppo le cose sono ben più complicate di così e spesso i media, nel comunicare i dati sul coronavirus, a causa di una eccessiva semplificazione non riescono a trasmettere tale complessità.

Per questo motivo, sono stati di ampio utilizzo i grafici, ma non solo, anche infografiche e dashboard come lab gedi digital del gruppo GEDI o lab 24 de Il Sole 24 Ore, tutti strumenti utili a far comprendere meglio i dati durante la pandemia. Strumenti che se non gestiti attentamente possono diventare un'arma a doppio taglio: infatti, se sono poco chiari o troppo complessi e senza una buona spiegazione, possono generare un effetto repulsivo verso coloro che hanno difficoltà ad interpretare una curva su un grafico cartesiano. È importante ricordarsi che ci si sta rivolgendo ad un pubblico ampio e di "non addetti ai lavori", e pertanto è bene non dare nulla per scontato.

Un giornalismo che ha avuto una forte accelerazione in questo periodo è stato il data journalism, una fusione tra dati, scienza e comunicazione, che unisce quindi gli strumenti della matematica e della statistica a quelli classici del giornalismo (Sestili, 2020). Ottenere risultati soddisfacenti in termini di resa fruibile, chiara e comprensibile dei dati comunicati dagli epidemiologi, richiede delle competenze specifiche soprattutto in campo statistico, e non tutti coloro che lavorano in questo settore possiedono o possiedono ma non ad un livello sufficiente.

Vista la natura del fenomeno che ci ha travolto a marzo dello scorso anno, era prevedibile una generale impreparazione, tutti ne hanno sofferto e quindi sono stati commessi molti errori, alcuni evitabili e altri inevitabili, non solo a livello di statistiche ma spesso anche notizie e informazioni che sono state trasmesse.

Il libro di Darrell Huff “Mentire con le statistiche” mi ha dato un enorme spunto per l’argomento che ho intenzione di approfondire. Il libro si presenta come un manuale che offre, tramite esempi ed illustrazioni in parte didattiche e in parte ironiche, degli strumenti per capire bene le statistiche. A differenza di quanto si può intuire dal titolo, il libro non offre metodi per ingannare la gente con medie e grafici, è un libro al servizio del cittadino, a cui suggerisce di dare sempre una seconda occhiata quando gli capita di leggere in un articolo di giornale o in un libro un dato statistico, per capire se esiste un legame sospetto tra il dato statistico presentato e le finalità di chi lo ha prodotto. Si parla di inganno nell’uso di dati statistici quando sono presenti dati che supportano una certa verità quando invece la verità è un’altra o quando ci sono tante altre interpretazioni legittime che possono essere date. Il punto è capire il perché si danno queste verità assolute o si comunicano in malo modo determinati concetti. È difficile individuare i motivi precisi: a volte chi utilizza il dato statistico ignora le proprietà e i limiti delle conclusioni, si limita a spiegare o a commentare la cosa in maniera superficiale senza curarsi delle conseguenze; altre volte ancora chi scrive o utilizza il dato statistico non ha come finalità quella di informare ma di “fare sensazione” creando così falsi allarmi o fornendo interpretazioni poco corrette (Huff, 2009).

Il libro raccoglie esempi interessanti di errori che sono stati commessi utilizzando la statistica e raccomanda al lettore, che si confronta spesso con dati statistici e grafici, di non dare per scontato che quello che viene scritto è vero e attendibile.

Con questo lavoro, ho intenzione di mettere in pratica quanto appena detto, non fermandomi alla prima occhiata nella lettura delle notizie, ma portando alcuni esempi di informazione e comunicazione riguardanti accadimenti e dati statistici tratti dall’emergenza sanitaria che stiamo vivendo, per valutarne la qualità attraverso gli strumenti della teoria statistica.

Il lavoro è organizzato in tre capitoli:

Il *primo capitolo* tratterà il caso AstraZeneca, si concentrerà sulla diffusione di alcuni titoli allarmistici in seguito alle segnalazioni di insorgenza di eventi tromboembolici avvenuti in soggetti ai quali era stato somministrato il vaccino AstraZeneca. Si valuterà, utilizzando i dati a disposizione forniti dall’EMA e dell’AIFA, se i dubbi sorti sulla sicurezza di questo vaccino siano in qualche modo giustificabili. Nel *secondo capitolo*, si metteranno in evidenza alcuni indici e grafici che sono stati ricorrenti nel racconto dell’evoluzione della pandemia e si evidenzieranno i motivi per cui non sempre sono stati chiari e precisi. Nel *terzo capitolo* infine, si proverà, citando degli studi effettuati, ad analizzare l’efficacia della misura di coprifuoco come mezzo di contrasto per la diffusione del virus. Seguiranno le *conclusioni*.

Capitolo 1

Il caso AstraZeneca

1.1 Le segnalazioni di eventi tromboembolici

Uno dei casi più eclatanti e che merita di essere analizzato come primo esempio di mancata corretta comunicazione di un fenomeno con all'interno una buona dose di statistica, riguarda quanto avvenuto con il vaccino AstraZeneca. Raccontiamo, senza peraltro scendere nei particolari, l'accaduto. AstraZeneca, insieme ai vaccini Pfizer-Biontech e Moderna, è stato uno tra i primi vaccini anti-Covid ad essere somministrato e a dare avvio alla campagna vaccinale in Italia. Ripercorrendo la cronologia, a partire dal 27 dicembre 2020, è stato somministrato per la prima volta Pfizer, a seguire Moderna dal 14 gennaio e infine AstraZeneca il 30 gennaio; date a parte, i problemi iniziano a sorgere solo in un secondo momento, nel mese di marzo (Ministero della Salute, 2021).

Domenica 14 marzo 2021, a seguito della morte di un docente di 57 anni di Biella vaccinato con AstraZeneca, la regione Piemonte decide di sospendere la somministrazione delle fiale appartenenti al lotto da cui è stata presa la dose del professore (Voltattorni, 2021). La regione ha preso una decisione autonomamente, senza consultarsi prima con l'AIFA (Agenzia Italiana del Farmaco). Sempre nella prima decade di marzo, si registrano in Sicilia e Campania decessi legati a trombosi e coaguli di sangue in persone che si erano da poco vaccinate con AstraZeneca. È il caso di 3 deceduti in Sicilia, due militari e un poliziotto, nonostante le dosi provenissero da un lotto diverso rispetto a quello piemontese. Un altro caso di morte sospetta è stata quella di un uomo in Campania e sempre in Campania, più precisamente a Napoli, una donna di 54 anni in stato di coma farmacologico (Anon., 2021). Seguendo la scia piemontese viene sospeso l'utilizzo di due lotti nelle due regioni Sicilia e Campania e iniziano le indagini per cercare di capire la causa di queste morti.

A fronte di queste decisioni, sia il Ministro della Salute Roberto Speranza sia il presidente dell'AIFA Giorgio Palù cercano di tranquillizzare la popolazione sulla sicurezza del vaccino dicendo che presto l'EMA (European Medicines Agency) emetterà un nuovo rapporto sulle reazioni avverse provocate dai vaccini. Lunedì 15 marzo, arriva però la decisione da parte dell'AIFA di sospendere la somministrazione di AstraZeneca in tutto il territorio, si tratta di una "sospensione temporanea e precauzionale", decisione presa anche dalla Germania e successivamente da Francia, Spagna e Portogallo, unendosi così ai paesi europei che lo avevano

da poco già sospeso: Danimarca, Irlanda, Norvegia, Bulgaria, Islanda e Olanda, tutti in via cautelativa. Il tweet qui riportato è stato scritto dall'AIFA il 14 marzo, non sono passate nemmeno 24 ore, non sono state condotte indagini, completate verifiche, nulla che apparentemente giustifichi la decisione della sospensione.



Probabilmente è stata una decisione presa nel tentativo di fermare la psicosi, di rassicurare la popolazione spaventata dalle morti sospette, con il rischio di aver fatto però più danni. L'unico buon motivo potrebbe essere quello che nei giorni precedenti al 15 marzo, a causa dei decessi sospetti registrati, ci sono stati diversi casi di rifiuto del vaccino, tanto che il generale Figliuolo (commissario straordinario per l'attuazione e il coordinamento delle misure per il contenimento

dell'emergenza COVID-19) ha disposto la creazione delle liste di riserva per evitare che dosi di vaccino non somministrate andassero sprecate. Si può quindi vedere la decisione presa come un modo per sospendere momentaneamente il giudizio su quanto è accaduto, al fine di attuare le dovute verifiche necessarie per fornire un'informazione rigorosa e scientifica alla popolazione.

Cerchiamo di capire cosa è stato a causare una paura tale da dover rassicurare i cittadini, sia coloro che avevano ricevuto la dose di AstraZeneca nei giorni precedenti alla sospensione, e sia coloro che probabilmente, in seguito ad una risposta da parte dell'EMA, avrebbero ricevuto quel vaccino. È innegabile che ci sia stato un problema dal punto di vista della comunicazione; le testate giornalistiche avrebbero dovuto riportare la decisione presa dall'AIFA senza violare altre semplici regole del giornalismo, quali, ad esempio: evitare titoli sensazionalistici, i quali causano infatti la nascita di timori o di speranze infondate, creando incertezza e paura su un argomento delicato, ma soprattutto si sarebbe dovuto evitare di diffondere notizie non verificate da fonti qualificate. La situazione creatasi in quei giorni appare anche a distanza di mesi caotica e fumosa, oltre a ciò si stava sviluppando la consapevolezza che non sarebbe stato facile ristabilire il clima che si era diffuso qualora la pronuncia da parte dell'EMA fosse stata positiva, ovvero si fosse dimostrata assenza di nesso di causalità tra la somministrazione del vaccino e gli eventi tromboembolici avversi.

Procediamo con ordine, analizziamo i toni utilizzati per comunicare le notizie di morti sospette.

Il vaccino anti-Covid, poi l'attacco cardiaco Sandro Tognatti, 57 anni, non soffriva di altre malattie

La Stampa 15 marzo 2021

Militare morto dopo vaccino, dieci indagati per omicidio

Stesso lotto di AstraZeneca di cui Aifa ha chiesto il sequestro

ANSA 11 marzo 2021

AstraZeneca, gravissima donna di 54 anni vaccinata a Napoli: la dose da un lotto sequestrato

NAPOLI > CRONACA
Lunedì 15 Marzo 2021

Il Mattino 15 marzo 2021

AstraZeneca, donna danese muore dopo il vaccino. L'agenzia della sanità: «Coaguli del sangue insoliti»

Il messaggero 11 marzo 2021

Quello che traspare da questi titoli è l'esistenza di un rapporto causa-effetto tra la somministrazione del vaccino AstraZeneca e la morte sospetta dei soggetti, non è scritto esplicitamente ma presentare un titolo nel quale vengono messe vicine le parole “morto” e “dopo il vaccino” fa quantomeno venire il dubbio che l'evento tromboembolico, ovvero l'arresto cardiaco, siano stati causati dal vaccino. Diverso invece è un titolo come questo:

Non ce l'ha fatta il carabiniere colpito da trombosi cerebrale: sarà disposta l'autopsia per far luce sulle cause

Di redazione - 17 Marzo 2021

Mantova uno 17 marzo 2021

Questo titolo mette molto meno allarme rispetto ai precedenti, si evidenzia l'accaduto ma non si traggono conclusioni affrettate e ancora tutte da dimostrare. L'AIFA a fronte degli avvenimenti dichiara che nulla ancora dimostra la causalità tra vaccino ed eventi tromboembolici.

1.2 Il nesso di causalità

Cos'è il nesso di causalità? Come si arriva ad una conclusione che il nesso di causalità esiste? Il problema è riassumibile con questa asserzione: “se B segue ad A, vuol dire che A è la causa di B”, questa non è la definizione del nesso di causalità, è però quello che spesso si vuole far

credere (Huff, 2009). Ci possono essere molte spiegazioni ragionevoli da prendere in considerazione. Iniziamo con la differenza tra correlazione e nesso di causalità: la correlazione è una misura statistica che esprime la relazione lineare tra due variabili ed è molto usata per descrivere semplici relazioni senza parlare di causa ed effetto. Un nesso di causalità tra due eventi, invece, può essere definito come la relazione tra un evento (la causa) ed un secondo evento (l'effetto) dove il secondo evento viene interpretato come conseguenza del primo (JMP Statistics Knowledge Portal).

La correlazione quindi non è necessariamente indice di causalità. Una forte correlazione potrebbe essere conseguenza di un nesso di causalità ma potrebbe anche trattarsi di una caratteristica per la quale questa chiave interpretativa non vale. Si può riuscire a mettere insieme un gruppo di numeri per dimostrare qualcosa di improbabile ma non è detto che lo stesso accada se si prova di nuovo con un altro gruppo di numeri; oppure potrebbe esserci una terza variabile, un terzo evento C e gli eventi A e B sono conseguenze di C (JMP Statistics Knowledge Portal). Per poter stabilire l'esistenza di una causa diretta ed unica, è bene eseguire una valutazione più complessa su quello che accade intorno ai fenomeni e non basarsi solo su poche osservazioni. Entrando nel merito, l'epidemiologia indaga e cerca di trovare risposte su associazioni tra possibili cause e malattie, mentre è la statistica che ci aiuta a capire se le associazioni sono sostenute dall'evidenza empirica, oppure no.

Ed è quello a cui, con l'aiuto dei dati e con il parere degli esperti, cercheremo di rispondere.

1.3 Tanto panico per nulla?

Andiamo con ordine, partendo dalle segnalazioni di eventi tromboembolici. L'Agenzia Italiana del Farmaco (AIFA) nel suo terzo rapporto, che comprende il periodo che va dal 27/12/2020 al 26/03/2021, descrive le segnalazioni di reazioni che sono state osservate dopo la somministrazione del vaccino. Come detto finora e come ribadito da AIFA, non è detto che queste reazioni siano state causate dal vaccino. Ad occuparsi di indagare sul significato e sulle cause di queste reazioni è la farmacovigilanza. Distinguiamo tre tipi di fenomeni:

- l'**evento avverso** è un qualsiasi episodio sfavorevole che si verifica dopo la somministrazione di un farmaco o di un vaccino, non è necessariamente causato dall'assunzione del farmaco o dall'iniezione del vaccino;
- l'**effetto indesiderato** è un effetto non intenzionale connesso alle proprietà del farmaco o del vaccino, che non è necessariamente nocivo ed è stato osservato in un certo numero di persone;
- la **reazione avversa** è invece una risposta nociva e non intenzionale ad un farmaco o ad un vaccino per la quale è possibile stabilire una relazione causale con il farmaco o con

la vaccinazione stessa. Bisogna valutare se è possibile risalire alla causa legata al prodotto e non c'entra nulla il fatto che l'evento si sia verificato a breve distanza dall'assunzione del farmaco o del vaccino (AIFA, 2021).

Nel nostro caso, le segnalazioni di eventi trombotici sono eventi avversi, fino a che non si riuscirà a dimostrare che esiste una relazione causale. Il rapporto sulla sorveglianza dei vaccini Covid-19 dedica una sezione all'insorgenza di eventi tromboembolici causati dal vaccino AstraZeneca viste le segnalazioni e visti i blocchi, prima dei singoli lotti e poi dell'intera somministrazione all'interno di alcuni paesi. In data 12/03/21 il PRAC (Comitato per la Valutazione del Rischio in Farmacovigilanza) ha iniziato la valutazione scientifica dei casi segnalati. In base a quanto inserito fino all'11/03 da EudraVigilance (banca dati dell'agenzia europea dei medicinali che si occupa di raccogliere segnalazioni di sospette reazioni indesiderate a farmaci) gli eventi trombotici e tromboembolici di qualunque natura sono stati 269 (258 dei quali gravi), di cui 224 dal Regno Unito, 30 dai paesi dell'UE/EEA (1 dall'Italia) e 15 da paesi non europei (AIFA, 2021). È bene ricordare che, alla data del 16/03, l'Agenzia regolatoria del Regno Unito riportava circa 11 milioni di dosi somministrate; a differenza dei paesi UE, dove EMA riportava, alla data del 10/03, circa 5 milioni di dosi somministrate, sempre considerando solo AstraZeneca.

Proseguendo, quello che riporta il rapporto dell'AIFA è che il numero dei casi sospetti osservati in seguito alla vaccinazione con AstraZeneca era inferiore rispetto al numero dei casi che si sarebbero osservati nella popolazione in condizioni normali, ovvero senza la somministrazione del vaccino. Questo è importante per capire se è presente una potenziale associazione statistica tra il vaccino e l'evento avverso. Se infatti il numero di reazioni sospette osservate dopo la somministrazione del farmaco è inferiore rispetto a quello che si otterrebbe indipendentemente da esso, allora l'associazione tra evento e medicinale è "improbabile e verosimilmente accidentale". Al contrario, se l'assunzione del medicinale comporta un numero di casi sospetti più elevato rispetto al numero che si otterrebbe senza la somministrazione, allora l'associazione è "potenzialmente causale". Rimane sempre da sottolineare il fatto che questo tipo di analisi ci dà indicazioni sulla forza statistica dell'associazione tra i due eventi, l'effettiva esistenza di un nesso di causalità può essere dimostrata solo in seguito ad una valutazione clinica dei singoli casi. Anche il professor Paolo Gresele, presidente della Società Italiana per lo Studio dell'Emostasi e della Trombosi, commenta il caso AstraZeneca dicendo che in Italia, si verificano "circa 65mila casi di tromboembolismo polmonare all'anno. Non c'è incremento". Confrontando i 65mila casi all'anno con i 30 casi su 5 milioni di dosi, e annualizzando quest'ultimo dato, abbiamo tra i 71 e i 117 casi di eventi tromboembolici su 100 mila abitanti nella popolazione generale, per contro 0,6 casi di eventi tromboembolici su 100 mila abitanti

che potrebbero verificarsi in seguito al vaccino (Renda, 2021). “Sono incidenze ben più basse rispetto di quelle attese nella popolazione generale” ribadisce il professor Gresele. Stessa situazione confermata anche in Gran Bretagna, dove sono state fatte molte più dosi rispetto a quelle somministrate in Italia, circa 11 milioni, e l’incidenza di trombi sanguigni è stata ancora più bassa, dello 0,0002% (Renda, 2021).

A sostegno di AstraZeneca, anche Silvio Garattini, fondatore e presidente dell’Istituto di ricerche farmacologiche Mario Negri Irccs di Milano dice che: “le patologie che normalmente avvengono continueranno ad avvenire, non c’è dubbio che avremo tanti di questi episodi. In Italia abbiamo circa 2mila decessi al giorno per varie cause. È probabile che qualcuno muoia 3, 5 o 10 giorni dopo aver fatto il vaccino, ma magari sarebbe morto in ogni caso”. Il 18 marzo 2021 arriva il responso da parte dell’EMA, la quale dichiara che il vaccino AstraZeneca anti-Covid è “sicuro ed efficace”, i benefici del vaccino nel prevenire la malattia da Covid-19 superano di molto i rischi e “non può essere associato ad un aumento di eventi tromboembolici”. Gli unici casi rarissimi per i quali non è possibile escludere in maniera definitiva un’associazione causale tra i due eventi riguardano i casi rappresentati da trombosi dei vasi che drenano il sangue dal cervello (18 casi) e trombosi di più vasi sanguigni (7 casi). Afferma dunque che verranno sicuramente effettuate nuove indagini e si valuterà se aggiornare il foglio illustrativo del vaccino AstraZeneca. Rimane comunque un vaccino efficace e la campagna vaccinale deve proseguire. Tanto panico per nulla? Una comunicazione fatta in toni allarmistici è stata in grado di suggestionare fortemente l’opinione pubblica e di spingere la politica a reazioni inconsulte e dannose, montando così un caso sanitario in assenza di un effettivo caso sanitario. A fronte di un’incidenza così bassa di insorgenza di eventi tromboembolici, gli effetti di una sospensione anche solo cautelativa di un vaccino possono essere rilevanti.

1.4 Due vaccini a confronto: Pfizer e AstraZeneca

Come annunciato all’inizio del capitolo, la campagna vaccinale era da poco iniziata e AstraZeneca non è stato l’unico vaccino ad essere stato somministrato, insieme a lui Pfizer-BioNTech e Moderna, i due vaccini statunitensi a base di mRNA. Potrebbe quindi essere interessante provare a fare un confronto tra i casi sospetti segnalati in seguito a Pfizer e in seguito ad AstraZeneca, prima nel Regno Unito e poi a livello europeo con i dati forniti dall’EMA.

Innanzitutto, evidenziamo i soggetti che fino al mese di marzo 2021 erano stati vaccinati. Il piano strategico nazionale italiano per la vaccinazione anti Covid-19 adottato il 2 gennaio, teneva conto di alcuni fattori che vale la pena citare: presenza di uno scarso numero di dosi, pochi dati disponibili per poter estendere a qualunque categoria di persone la vaccinazione e il

perseguimento dello scopo fondamentale iniziale della campagna vaccinale di ridurre il tasso di mortalità e di morbilità (Ministero della Salute, AIFA et al., 2021). Tenuto conto di questi elementi, le autorità competenti hanno preso delle decisioni e segmentato la campagna vaccinale in varie fasi: la prima fase riguardava personale sanitario e sociosanitario, personale ed ospiti dei presidi residenziali per anziani, anziani over 80, persone dai 70 ai 79 anni e a seguire dai 60 ai 69 anni e la popolazione con almeno una comorbosità cronica. Successivamente, con l'arrivo di un numero maggiore di dosi, anche insegnanti e personale scolastico, forze dell'ordine, personale delle carceri e dei luoghi di comunità (Ministero della Salute, AIFA et al., 2021). Gruppi target che nel corso delle settimane potevano essere soggetti a modifiche in base all'evoluzione delle conoscenze e delle informazioni disponibili a proposito dei vaccini in circolazione.

Pfizer, vaccino a base di mRNA, insieme a Moderna, era un vaccino consigliato al personale dedicato a fronteggiare l'emergenza sanitaria e ai soggetti con più alto rischio di sviluppare una malattia grave, quindi i soggetti estremamente vulnerabili sia dal punto di vista dell'età che delle condizioni patologiche. AstraZeneca invece poteva essere utilizzato nella fascia di età compresa tra i 18 e i 65 anni ad eccezione dei soggetti estremamente vulnerabili, quindi sì ai soggetti che possono aumentare il rischio di sviluppare forme severe di COVID-19 ma senza quella connotazione di gravità riportata per le persone definite estremamente vulnerabili (Ministero della Salute, 2021). Poco tempo dopo (8 marzo) un'altra circolare emessa dal Ministero della Salute estende il vaccino AstraZeneca anche ai soggetti di età superiore ai 65 anni. Tale indicazione non è da intendersi applicabile ai soggetti identificati come estremamente vulnerabili in ragione di condizioni di immunodeficienza, primitiva o secondaria a trattamenti farmacologici o per patologia concomitante che aumenti considerevolmente il rischio di sviluppare forme fatali di COVID-19 (Ministero della Salute, 2021).

1.4.1 La situazione in UK

Quindi, a marzo 2021, anche per questioni di numerosità di dosi disponibili, Pfizer veniva somministrato molto di più rispetto ad AstraZeneca, in Italia e anche in altri paesi europei. Così non è stato in Gran Bretagna, dove sono state somministrate un numero molto alto e simile di dosi tra i vaccini AstraZeneca e Pfizer, circa 10 milioni ciascuno. Tramite l'utilizzo di un report tratto dal Yellow Card Scheme, il sistema britannico di raccolta informazioni sulle reazioni avverse sospette ai farmaci (ADRs), possiamo trarre il numero degli eventi di embolia polmonare sorti in seguito ad AstraZeneca e in seguito a Pfizer al 9 marzo 2021:

Name: COVID-19 vaccine AstraZeneca analysis print		
Report Run Date: 09-Mar-2021	Data Lock Date: 08-Mar-2021 19:00:03	
Reaction Name	Total	Fatal
Respiratory disorders Respiratory disorders cont'd		
<i>Pulmonary thrombotic and embolic conditions</i>	13	
Pulmonary embolism		1

Name: COVID-19 mRNA Pfizer- BioNTech vaccine analysis print		
Report Run Date: 09-Mar-2021	Data Lock Date: 08-Mar-2021 19:00:03	
Reaction Name	Total	Fatal
Respiratory disorders Respiratory disorders cont'd		
<i>Pulmonary thrombotic and embolic conditions</i>	15	
Pulmonary embolism		1

Screenshot dal tweet di Martin Moder

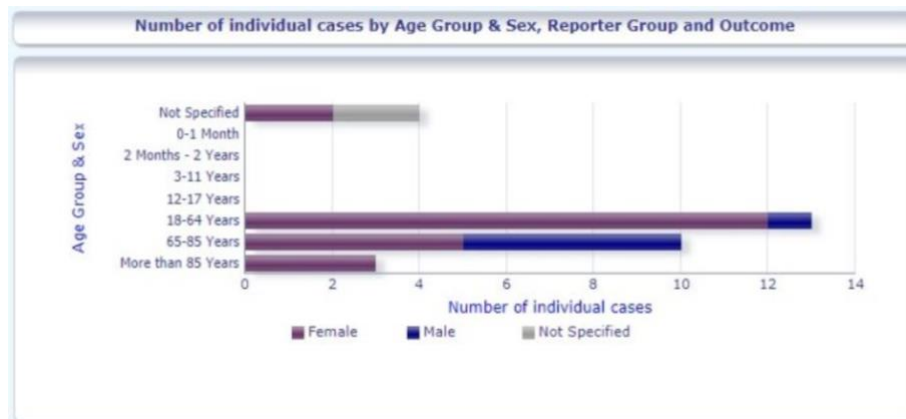
Come possiamo notare, il numero di casi di embolismo polmonare è stato molto simile, addirittura un numero maggiore, di un paio di eventi, nel caso di Pfizer; eppure non è mai stata messa in discussione la sicurezza del prodotto come invece è avvenuto per AstraZeneca.

1.4.2 La situazione in Europa

Dalle segnalazioni di reazioni avverse riportate dall'EMA al 13 marzo 2021, sono stati riportati **30** casi di embolia polmonare su circa **7 milioni** di dosi somministrate del vaccino AstraZeneca (Puente, 2021):



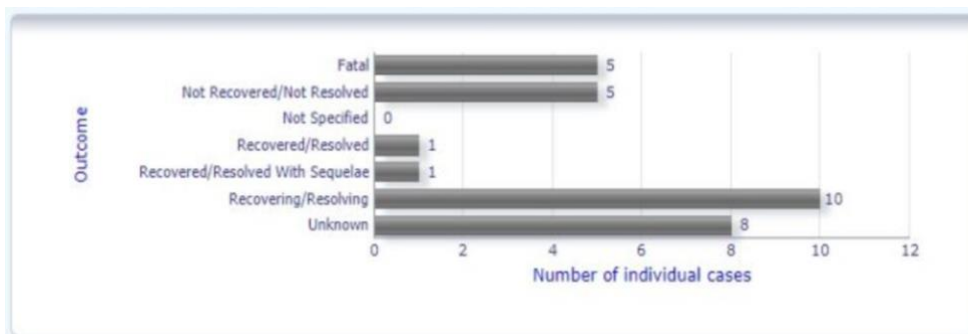
I soggetti riportati hanno le seguenti caratteristiche demografiche:



Come possiamo notare, la maggior parte delle reazioni avverse in seguito ad AstraZeneca si sono registrate nella fascia d'età 18-64, sia a causa dell'ampiezza della fascia considerata e sia perché alcuni studi clinici hanno dimostrato che la popolazione anziana risulta essere meno

incline allo sviluppo di reazioni avverse. Un'altra osservazione importante può essere condotta analizzando il sesso dei soggetti segnalatori: su 30 reazioni avverse segnalate, 22 provengono da donne, di cui la maggior parte nella fascia d'età 18-64. Questa differenza, secondo l'AIFA, sembra dipendere dalla diversa esposizione ai vaccini (soprattutto donne), dalla maggiore sensibilità delle donne alle segnalazioni e dalla diversa frequenza di reazioni avverse tra i due sessi.

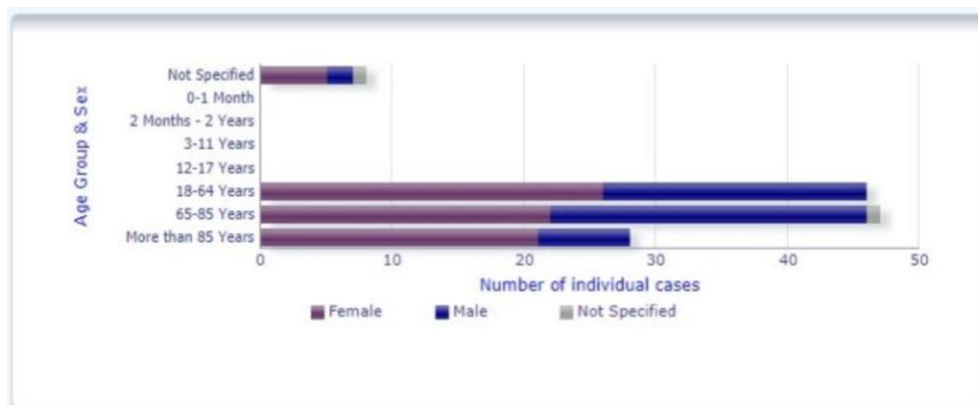
E questi sono stati gli outcome delle segnalazioni:



Per quanto riguarda invece i casi di embolia polmonare verificatisi in seguito al vaccino Pfizer, se ne sono registrati **129** a fronte di **36 milioni di somministrazioni** (Puente, 2021):



Con le seguenti caratteristiche demografiche:

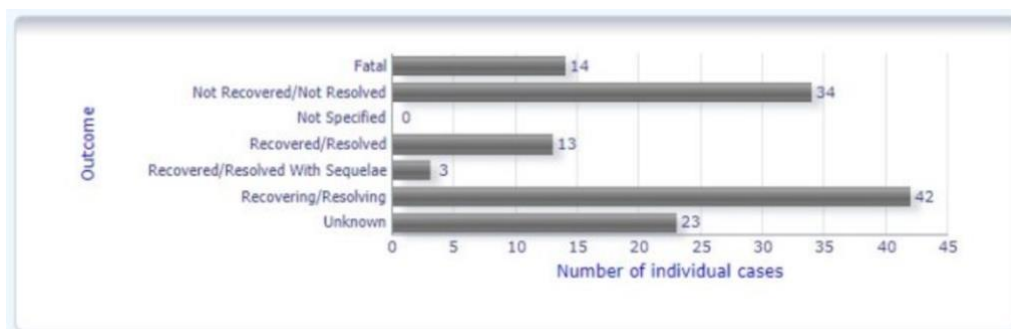


La situazione è in linea con quanto registrato per AstraZeneca. Sono state somministrate molte più dosi, di conseguenza il numero delle reazioni avverse è maggiore in tutte le fasce d'età

considerate. Anche per la fascia d'età > 85 si registrano numerose segnalazioni, Pfizer infatti è stato un vaccino, nei primi tempi, somministrato ad individui più fragili o esposti a situazioni di maggiore rischio.

I dati confermano che il numero di segnalazioni provengono principalmente da donne, delle 129 segnalazioni, 74 sono state effettuate da donne, distribuite similmente tra le varie fasce d'età.

E questi sono stati i risultati:



Dati tratti dalla banca dati europea delle segnalazioni di sospette reazioni avverse ai farmaci

Proviamo a rapportare il numero di segnalazioni registrate sul numero di dosi somministrate per entrambi i vaccini.

AstraZeneca: 30 segnalazioni/7 milioni di dosi = 0,0000043

Pfizer: 129 segnalazioni/36 milioni di dosi = 0,0000036

Apparentemente, Pfizer presenta un rapporto di poco inferiore rispetto ad AstraZeneca, vediamo se questa differenza risulta essere significativa.

Assumiamo che entrambe le popolazioni, che indichiamo con Π_a e Π_p (rispettivamente AstraZeneca e Pfizer), abbiano distribuzione Bernoulliana. Vogliamo comparare il tasso di segnalazione di embolia polmonare dei soggetti a cui è stato somministrato AstraZeneca con il tasso di segnalazione di embolia polmonare dei soggetti a cui è stato somministrato Pfizer. È importante precisare però che gli eventi che stiamo confrontando sono eventi rarissimi (0,0000043 e 0,0000036 sono infatti dei tassi di segnalazione molto bassi) e indipendenti, che si sono verificati casualmente in un determinato periodo di tempo. Per questo motivo, sarebbe più adatto approssimare le due popolazioni (AstraZeneca e Pfizer) con la distribuzione di Poisson, e non con quella Bernoulliana. Ricorreremo, nonostante ciò, ad una semplificazione per poter trarre evidenza statistica, assumendo che le due popolazioni abbiano distribuzione Bernoulliana.

Il test d'ipotesi è il seguente:

$H_0: \Pi_a = \Pi_p$ ipotesi nulla stesso tasso di segnalazione

H1: $\Pi_a > \Pi_p$ ipotesi alternativa, tasso di segnalazione AstraZeneca maggiore di quello di Pfizer

I due campioni casuali di persone a cui è stato somministrato il vaccino hanno numerosità:

$n_a = 7$ milioni; $n_p = 36$ milioni

Quando le numerosità campionarie sono sufficientemente elevate, possiamo verificare le ipotesi attraverso l'uso della statistica test:

$$Z = \frac{\bar{x}_a - \bar{x}_p}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_a} + \frac{1}{n_p}\right)}} \quad \text{dove } \bar{p} \text{ è lo stimatore congiunto di } \Pi = \frac{n_a \bar{x}_a + n_p \bar{x}_p}{n_a + n_p}$$

$$\bar{x}_a = \frac{30}{7.000.000}, \quad \bar{x}_p = \frac{129}{36.000.000}, \quad \bar{p} = \frac{30+129}{36.000.000+7.000.000}.$$

$$Z = 0,8843$$

Controllo sulla tavola statistica e l'area sottesa alla curva della normale standardizzata è 0,8117.

$1 - 0,8117 = 0,1883$ è il p value

Fissiamo alfa, il livello di significatività, a 0,05.

Confronto $0,1883 > 0,05$, quindi p value $> 0,05$

Accetto H_0 e concludo che non c'è evidenza statistica per poter dire che le reazioni avverse di embolia polmonare di AstraZeneca sono maggiori di quelle di Pfizer.

La differenza tra i due tassi di segnalazione calcolati in precedenza è dovuta alla variabilità campionaria.

In conclusione, molteplici sono i fattori che portano a pensare che non ci sia una sostanziale differenza tra i tassi di segnalazione degli eventi avversi dei due vaccini. È stato cruciale quindi il ruolo della comunicazione che ha trasmesso un messaggio di fatto sbagliato. Non sappiamo con certezza il motivo per cui i canali mediatici abbiano alimentato questa narrazione.

Come già successo in altre occasioni, la situazione è stata presentata in maniera superficiale e senza preoccuparsi delle conseguenze che purtroppo ci sono state. Nei mesi successivi infatti, molte sono state le regioni, soprattutto al Sud, a registrare numerose dosi inutilizzate del vaccino AstraZeneca. In Sicilia, ad esempio, in data 8 maggio 2021, si è registrato un costante ed elevato sotto utilizzo delle dosi disponibili, che ha costretto il governatore Musumeci a ritarare il piano aprendo agli over 50 sperando di poter ampliare la platea di vaccinati. La regione ha, in data 5 maggio, il 49,3% di dosi consegnate ferme nei frigoriferi, notizia shock a tal punto da attirare l'attenzione di una regione come la Lombardia dove solo l'1% della popolazione rifiuta AstraZeneca e che richiede che vengano ridistribuite le dosi inutilizzate e consegnate alle "regioni più virtuose" (Savelli, 2021).

Capitolo 2

L'andamento della pandemia attraverso indicatori e grafici

La diffusione del Covid-19 in Italia, ha fatto emergere fin da subito l'esigenza di raccontare in maniera chiara e ordinata ai cittadini quanto stava accadendo. Cittadini che, in alcuni comuni del Nord Italia, a partire dal 22 febbraio, si sono visti privati delle usuali e scontate libertà di circolazione ed isolati all'interno dei propri comuni con il solo permesso di uscire per motivi di necessità ed emergenza. Molte famiglie con figli al Nord per motivi di lavoro o di studio e genitori che volevano capire quando avrebbero potuto riabbracciarli piuttosto che albergatori in zone di montagna che hanno smesso in piena stagione di accogliere nuove prenotazioni o hanno dovuto annullare quelle già fatte.

È nata così la pagina di approfondimento lab24 in una sezione specifica de Il Sole 24 Ore, con lo scopo iniziale di localizzare all'interno di una mappa questi 11 comuni isolati. La situazione era in continua evoluzione, ben presto (2 marzo 2020) tutte le regioni italiane avevano almeno un caso di infezione e il 9 marzo 2020 venne esteso il divieto di spostamento a tutte le regioni con sospensione di molte attività, tra cui quelle sportive, manifestazioni ed eventi, scuole, università ecc. L'Italia, dopo la Cina era diventato il nuovo epicentro di coronavirus, quindi molte regioni e paesi imposero misure di quarantena o divieti d'ingresso per cittadini o turisti provenienti dall'Italia.

La pagina lab24 piano piano si allargò, non bastava più sapere dove erano localizzati i contagi ma serviva monitorarne la crescita. E così la pagina diventa una pagina visual piena di tabelle, grafici e infografici interamente dedicati alla diffusione del virus in Italia, ampliando l'offerta di dati e informazioni relativi ai nuovi contagi, ai decessi, ai ricoveri in terapia intensiva ecc. e, facendo confronti anche con i paesi europei ed extraeuropei nei quali il virus si era maggiormente diffuso.

Lab24 riceve un forte successo e ben presto diventa la pagina de Il Sole 24 Ore più visitata da parte degli utenti, viene definita dal coordinatore come un progetto partecipato, riceve centinaia di mail e accoglie suggerimenti provenienti da lettori, ma anche da parte di analisti finanziari, designer o fisici, che suggeriscono delle modalità attraverso le quali rendere più evidente un dato o un numero.

La cosa di cui ci si accorge subito è la complessità della situazione che si vuole provare a spiegare attraverso dati, numeri e grafici. Gli ideatori della pagina sono infatti consapevoli delle numerose imprecisioni a cui vanno incontro, alcune inevitabili, ad esempio il fatto che la pagina venga aggiornata quotidianamente alle 18 e non tutte le regioni, soprattutto all'inizio, rispettavano le tempistiche di “consegna” dei dati al Ministero della Salute.

Le informazioni che ricevono i giornalisti sono informazioni parziali, è impossibile avere una visione chiara e completa di quanto accade, nemmeno gli epidemiologi ce l'hanno. Il punto è dare informazioni non troppo semplicistiche o ottimistiche ma ricche di chiarimenti.

In questo capitolo, la mia intenzione è quella di portare alla luce alcune informazioni riportate su lab24 che possono essere fuorvianti o mal interpretate da lettori disattenti e poco inclini alla verifica se comunicate senza le dovute precisazioni.

2.1 Il tasso di positività

Il tasso di positività è calcolato come il rapporto tra il numero dei nuovi test positivi e il totale dei nuovi test effettuati, è una percentuale che indica quanti, tra i test effettuati giornalmente, sono risultati positivi rispetto al totale dei soggetti testati (Capozzoli, 2020). Viene utilizzato per calcolare l'incidenza del virus, ovvero per sapere in una determinata quantità di tempo quanti sono i nuovi contagiati. Su lab24 ogni giorno viene tenuto il conto dei nuovi tamponi effettuati e in base al numero dei tamponi positivi viene fatto un confronto con il giorno precedente.

Ma ecco quanto riportato da alcune testate giornalistiche:

ITALIA

Il tasso di positività cresce al 14,1%. Aumentano i contagi

Gli attualmente affetti da Covid oltre quota 500mila. Le persone che si sono vaccinate sono quasi 36mila, soprattutto personale sanitario

Il Manifesto, 2 gennaio 2021

“Il tasso di positività cresce al 14,1%. Aumentano i contagi” cosa ci suggerisce? Che rispetto al giorno o ai giorni precedenti in cui il tasso di positività era inferiore, il 1° gennaio 2021 si è registrato un aumento del tasso di positività.

Alla prima lettura, senza porsi troppi interrogativi, la prima cosa da pensare è “stiamo peggiorando, cresce di nuovo il tasso di positività!” oppure “le misure di lockdown che sono state messe in atto nel periodo di Natale non hanno funzionato” e via dicendo.

Ma è realmente così? È corretto arrivare a conclusioni così puntuali solo perché in una settimana il tasso di positività oscillerà tra il 14% e il 17%?

giorno	diff_guariti	diff_tamponi	diff_casi_testati	diff_totale_casi	rapp_totale_casi_tamponi
--------	--------------	--------------	-------------------	------------------	--------------------------

2020-12-29	17044	128740	45702	11224	8,718347056082026
2020-12-30	19960	169045	63748	16202	9,584430181312669
2020-12-31	17421	186004	76798	23477	12,621771574804843
2021-01-01	16877	157524	70010	22211	14,100073639572383
2021-01-02	9166	67174	30451	11831	17,61246911007235
2021-01-03	14746	102974	44310	14245	13,833589061316449
2021-01-04	16206	77993	35417	10800	13,847396561229854

Dati scaricati da lab24

La risposta è che non è possibile semplificare così tanto senza tener conto di alcune precisazioni, vediamo quali.

2.1.1 il campione di convenienza

Il tasso di positività è il rapporto positivi/tamponi, il quale dipende dal numero di tamponi processati quotidianamente. Il tampone è un test diagnostico per accertare l'infezione da SARS-CoV-2, viene eseguito con un lungo cotton fioc con cui si raccolgono il muco e i liquidi della gola. Il campione viene poi analizzato in laboratorio attraverso il metodo RT-PCR, che serve per amplificare e quantificare il codice genetico all'interno del quale cercare sezioni tipiche dell'RNA del coronavirus.

L'Organizzazione Mondiale della Sanità (OMS) raccomanda che vengano sottoposte al test "il maggior numero di persone possibili", il coronavirus può essere trasmesso infatti non solo da persone sintomatiche ma anche dai paucisintomatici e dagli asintomatici, i più difficili da scovare e i più pericolosi per la diffusione del contagio. Lo scopo è quello di trovare il numero più alto di persone contagiate, isolarle e cercare di fermare la catena del contagio.

La domanda, sorta ai primi scoppi della pandemia una volta perfezionato il test negli istituti di ricerca di tutto il mondo, è stata a quali persone sottoporre il test.

Ebbene, ci sono delle categorie precise: residenti nelle Rsa e strutture per lungodegenti, coloro che assistono persone anziane o fragili in ambienti privati, chi ha avuto contatti con un positivo, chi ha sviluppato sintomi, chi dopo aver contratto il covid fa il/i tampone/i di controllo, chi desidera viaggiare all'estero ma anche in determinate regioni italiane, chi per motivi di lavoro, chi si trova in un contesto in cui è opportuno verificare periodicamente la negatività (atleti, giornalisti, sanitari) (Bocci, 2020). Per fortuna il tampone oggi è molto più fruibile rispetto a prima, ha trovato anche rapida diffusione all'interno delle farmacie, e chiunque abbia intenzione di controllare la propria positività e non rientra nelle categorie elencate, può farlo, ma non gratuitamente.

Se ognuno di noi fosse obbligato a sottoporsi ad un test diagnostico settimanalmente, la situazione potrebbe essere maggiormente sotto controllo rispetto a com'è stata e com'è tuttora.

Vediamo la realizzabilità dal punto di vista statistico, la popolazione in questione, prendo come riferimento la popolazione italiana (60,36 milioni di abitanti), è una popolazione finita e potrebbe essere studiata in modo esaustivo enumerando e osservando tutte le unità statistiche che la compongono. Un'indagine condotta in questo modo prende il nome di indagine totale o censimento e consiste nell'osservare il valore assunto dal carattere d'interesse X in ciascuna delle N unità della popolazione (Borra, Di Ciaccio, 2014 pg.244). Purtroppo, o per fortuna, questa modalità per eseguire l'indagine presenta molti limiti, di tempo e di costi, nel caso di una pandemia la variabile tempo ha un impatto non di poco conto, il tampone conferma la presenza o meno del virus per quel determinato istante e non per i giorni successivi, considerando che la popolazione dovrebbe essere sottoposta al tampone interamente, non solo ci vorrebbero giorni ma sarebbe anche inutile perché i primi tamponati potrebbero essere dopo qualche giorno positivi.

Per queste ragioni, è preferibile osservare solo una porzione della popolazione obiettivo, il cosiddetto campione.

Il punto cruciale per definire il campione è dato dalla regola di selezione, dalla procedura con la quale le unità vengono estratte dalla popolazione, il cosiddetto campionamento.

Qual è la regola di selezione del campione? Generalmente di tipo probabilistico, l'estrazione del campione avviene secondo una specifica distribuzione di probabilità.

Se gli individui 1 e 2 sono selezionati a caso, il valore di x_1 non ha contenuto informativo per x_2 (dove X è il carattere d'interesse), allora x_1 e x_2 sono distribuiti indipendentemente.

Se x_1 e x_2 provengono dalla stessa distribuzione, allora x_1 , x_2 sono identicamente distribuiti.

Solo sotto queste condizioni si ha il campionamento casuale, e si dice che x_1 e x_2 sono identicamente e indipendentemente distribuiti (Pozzolo, 2020).

Esiste anche il campionamento di tipo non probabilistico, tra le tipologie troviamo il campionamento di convenienza, il quale non offre a tutte le unità della popolazione la stessa possibilità di entrare a far parte del campione (Bottarelli, 2020).

La scelta tra i due tipi di campionamento dipende da fattori diversi: le caratteristiche della popolazione, il carattere di interesse che si vuole esaminare, il tempo e il budget che si hanno a disposizione. Statisticamente parlando, è preferibile il campionamento di tipo probabilistico perché nella maggior parte dei casi risulta essere quello più rappresentativo.

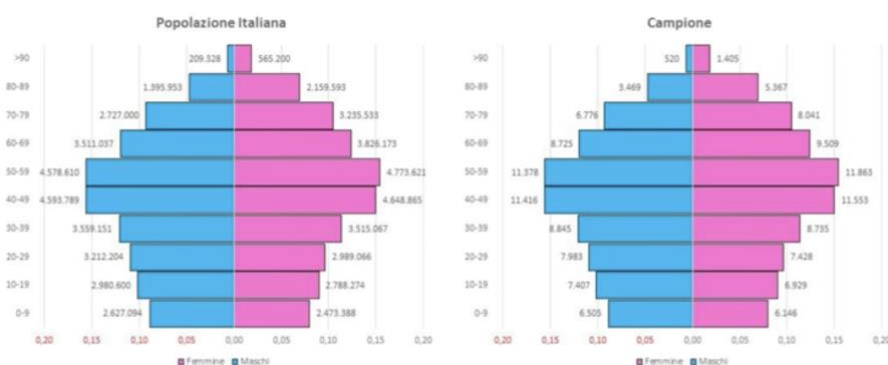
Vediamo cosa si intende per campione rappresentativo attraverso un esempio.

Il Ministero della Salute insieme all'Istat e alla Croce Rossa Italiana, in data 25 maggio 2020, hanno avviato un'indagine sierologica per l'infezione del virus SARS-CoV-2 per capire il numero di persone che avevano sviluppato anticorpi al nuovo coronavirus, anche in assenza di sintomi. Il campione a cui è stato rivolto lo studio era formato da 150.000 persone residenti in

duemila comuni, distribuite per sei fasce d'età, sesso e attività lavorativa. Il campione era stato costruito in modo tale da essere rappresentativo. 150.000 persone su 60 milioni sembrerebbe non essere un numero tale da rappresentare la popolazione italiana, è importante però sottolineare che non sempre i campioni grandi sono rappresentativi e che tutto dipende da come vengono selezionati i soggetti. L'importante è tener conto delle caratteristiche che il campione deve avere per rispecchiare la popolazione, in questo caso, i fattori correlati con la probabilità di ammalarsi: il genere, l'età ma anche prendendo persone di diverse professioni. Ad esempio, affinché il campione sia rappresentativo dal punto di vista del genere, è necessario rispettare il rapporto tra maschi e femmine nella popolazione italiana. Se i maschi sono il 48,7%, la stessa percentuale deve essere mantenuta nel campione, ovvero 73.050. E lo stesso ragionamento vale per la caratteristica "fascia d'età". Purtroppo l'indagine non è stata così efficace come si pensava o come poteva essere perché ha risposto solo il 40% del campione, a questo si è aggiunto un problema di selezione, coloro che hanno dato la disponibilità a partecipare, probabilmente sono soggetti molto simili tra loro, magari persone che hanno avuto i sintomi e sono curiose di sapere se hanno sviluppato gli anticorpi, piuttosto che persone che non lavorano e che hanno risposto ad una chiamata da un numero sconosciuto rispondendo al questionario (Paknazar, 2020). È davvero molto difficile ottenere un campione rappresentativo.

Si dice che un campione è rappresentativo quando la distribuzione delle variabili correlate con il fenomeno che si vuole studiare è uguale a quella della popolazione.

Esempio di campione rappresentativo:



Esempio di uno non rappresentativo:



tratti dal video Youtube "#9: le parole dell'epidemia: campione rappresentativo"

Il tampone, o test diagnostico, per accertare la presenza del coronavirus nei soggetti, viene sottoposto ad un sottoinsieme della popolazione italiana, quindi ad un campione.

Prendendo in considerazione quanto appena detto, emerge che il campionamento utilizzato per selezionare le persone a cui fare il tampone non è casuale in quanto non di tipo probabilistico, x_1 e x_2 sono indipendenti tra di loro ma non identicamente distribuiti, e di conseguenza scarsamente rappresentativo. Il campionamento utilizzato è un campionamento di convenienza, il quale, per definizione, è un campione dove la selezione avviene in maniera non casuale e fornisce un campione bias, non rappresentativo. È importante sottolineare però che spesso è possibile utilizzare una procedura in grado di rendere il campione rappresentativo: la riponderazione. Viene modificato artificialmente in fase di elaborazione dei dati per renderlo più simile alla distribuzione nella popolazione (Scienze Statistiche Unipd).

2.1.2 Il conteggio dei tamponi crea difformità nei risultati

La seconda considerazione riguarda invece la tipologia di tamponi che viene conteggiata nel calcolo del tasso di positività. Esistono tre tipi di test per rilevare l'infezione da coronavirus: i tamponi molecolari, gli unici validi per la diagnosi, i tamponi rapidi antigenici e i test sierologici. I tamponi molecolari sono la procedura di riferimento a livello internazionale proprio per l'altissimo livello di specificità e sensibilità. Sono efficaci anche nelle fasi primissime dell'infezione. I test antigenici, come i test molecolari, sono dei test di determinazione diretta, si va a determinare direttamente la presenza del virus nei campioni biologici. Essi mostrano la stessa sensibilità e specificità di quelli molecolari solo però nelle fasi immediatamente successive all'infezione. Non vengono utilizzati per la diagnosi ma sono molto utili per indagini di screening. I test sierologici invece non sono test diretti, determinano indirettamente la presenza del virus andando a determinare la quantità di anticorpi sviluppata. È un test la cui sensibilità e specificità è molto bassa, non possono essere assolutamente usati

per la diagnosi, né per dimostrare che c'è un'infezione in atto, ci dice solo che il soggetto è entrato in contatto con il virus (Messaggero Veneto, 2020).

Escludiamo subito il test sierologico, in quanto non è mai stato preso in considerazione per il calcolo del tasso di positività. I test molecolari e i test antigenici sono invece quelli più utilizzati, generalmente il tampone molecolare viene eseguito su un soggetto che chiaramente presenta i sintomi del Covid-19 mentre si procede con un test antigenico per i pazienti che presentano sintomatologia non chiaramente Covid-19 oppure per coloro che hanno bisogno di un risultato rapido o un tampone di controllo, e se esso presenta esito positivo allora si prenota subito un test molecolare al fine di ottenere una conferma diagnostica.

Capiamo bene che è possibile anzi doveroso eseguire un tampone molecolare a seguito di un esito positivo del tampone antigenico, un unico soggetto quindi può essere sottoposto a due tamponi in una sola giornata. Come vengono conteggiati i diversi tamponi?

Recuperando i dati forniti da lab24 sui tamponi giornalieri che vengono processati, possiamo notare subito come la distinzione tra tamponi molecolari e tamponi antigenici venga fatta solo a partire dal 15 gennaio 2021, data nella quale i tamponi antigenici vengono inclusi per il calcolo del tasso di positività. Fino al 15 gennaio venivano contati solo i molecolari.

Per essere ancora più precisi, il 15 gennaio non è stata una data rispettata da parte di tutte le regioni. 8 regioni quali Sicilia, Liguria, Marche, Abruzzo, Umbria, Basilicata, Molise, Veneto e Valle d'Aosta non si sono affrettate a segnalare i loro positivi tracciati con il tampone antigenico. Situazione identica fino al 26 gennaio tranne per il Veneto che finalmente ha inserito anche i test rapidi.

Il problema di una situazione del genere sta nella difformità di risultati che si ottengono.

Il tasso di positività che viene trasmesso attraverso i più disparati mezzi di comunicazione, tra cui quello del 14,1% citato prima, rappresenta il rapporto positivi/tamponi complessivo nazionale. Avere delle differenze nel conteggio dei tamponi tra regione e regione ha delle ripercussioni non di poco conto sul tasso di positività.

Proviamo a calcolare i tassi di positività per alcune regioni in data 15 gennaio 2021:

Regioni	Casi identificati da test rapido	Nuovi positivi giornalieri	Numero tamponi molecolari giornalieri	Numero tamponi rapidi giornalieri	Totale tamponi per il calcolo del tasso	Tasso di positività
Lombardia	27	2.205	26.535	7.412	33.947	6,50
Veneto	0	1.079	19.121	38.721	57.842	1,87

Piemonte	133	871	5.858	12.188	18.046	4,83
Campania	73	1.150	12.576	1.353	13.929	8,26
Emilia Romagna	0	1.794	15.667	6.713	22.380	8,01
Lazio	65	1.394	12.132	16.110	28.242	4,94

Bollettino aggiornamento casi Covid-19 Quotidiano Nazionale 15 gennaio 2021

La regione Veneto conferma quanto detto, non inserisce il numero dei casi identificati grazie al tampone rapido mentre l'Emilia Romagna, che dovrebbe registrare tale numero, riporta 0 casi identificati da test rapido.

Vediamo cosa cambia nel giro di un giorno prendendo sempre le stesse regioni:

Regioni	Casi identificati da test rapido	Nuovi positivi giornalieri	Numero tamponi molecolari giornalieri	Numero tamponi rapidi giornalieri	Totale tamponi per il calcolo del tasso	Tasso di positività
Lombardia	45	2.134	19.304	16.013	35.317	6,04
Veneto	0	1.929	36.262	56.213	36.262	5,32
Piemonte	237	1.056	24.375	27.228	24.375	4,33
Campania	74	1.132	13.247	1.364	14.611	7,74
Emilia Romagna	0	1.674	8.459	12.353	20.812	8,04
Lazio	120	1.282	29.848	33.328	29.848	4,30

Bollettino aggiornamento casi Covid-19 Il Messaggero 16 gennaio 2021

I dati in grassetto indicano la presenza di incongruenze. Nel caso della regione Veneto, in data 16 gennaio non venivano ancora registrati i casi tracciati grazie ai tamponi rapidi; per il Lazio il tasso di positività viene invece calcolato contando solo i tamponi molecolari processati quel giorno; infine Emilia Romagna, degli oltre ottomila tamponi rapidi nemmeno un caso positivo registrato grazie ad essi.

Un'altra osservazione sempre riguardante la regione Veneto e mettendo a confronto le due tabelle è l'estrema variabilità del tasso di positività, passa dall'1,87% a 5,32%.

Inserire all'interno del conteggio dei tamponi anche i test rapidi provoca un aumento notevole del denominatore del rapporto positivi/tamponi con conseguente diminuzione del tasso di positività. Infatti, in data 15 gennaio 2021, sono stati processati 273.506 tamponi, compresi

anche i test rapidi antigenici. La differenza tra il 14 e il 15 gennaio è notevole, contando solo i tamponi molecolari (156.647), il tasso di positività è del 10,3% (14 gennaio); contando anche i test rapidi il tasso scende al 5,9% (15 gennaio) (Fassari, 2021).

È importante essere cauti nella comunicazione di questi dati, è bene ricordare anche che i test rapidi vengono sottoposti anche a soggetti non a rischio, è un test di facile accesso.

Non solo, ci sono problemi anche per quanto riguarda la variabilità del tasso di positività calcolato prendendo in considerazione solo i tamponi antigenici tra regione e regione.

Nelle regioni che indicano il dato, in data 27 gennaio 2021:

Regioni	Casi identificati da test rapido	Nuovi positivi giornalieri	Numero tamponi rapidi giornalieri	Tasso di positività tamponi rapidi
Sardegna	2	251	6.877	0,03
Friuli-Venezia Giulia	1.961	584	28.361	6,9
Veneto	1.498	2.385	298.000	0,5
Lombardia	1.029	2.293	99.495	1,03
Emilia Romagna	48	923	83.930	0,05
Puglia	300	1.233	14.959	2
Lazio	868	1.338	197.820	0,44

Bollettino aggiornamento casi Covid-19 Quotidiano Sanità 27 gennaio 2021

È evidente che ci stati dei problemi nel raccogliere i dati, la regione Veneto ad esempio, presenta un tasso di positività bassissimo, ha però incluso nel numero dei tamponi rapidi tutti i tamponi processati dal 15 gennaio al 27 gennaio 2021.

Il Friuli-Venezia Giulia con 584 nuovi positivi quando solo i positivi tracciati con il test rapido sono stati 1.961.

Così come lo strano numero di casi identificati con il test rapido della Sardegna sia stato 2 a fronte di 6.877 tamponi rapidi.

In questa tabella qui abbiamo considerato alcuni tassi di positività per mostrare l'estrema variabilità, variabilità causata anche dalle diverse tipologie di test antigenico esistenti ed utilizzate. Esistono infatti varie generazioni di test antigenico con diverse qualità, l'ultima generazione presenta la qualità più elevata ma le regioni non precisano la tipologia usata.

La percentuale dei positivi/tamponi è inoltre uno degli indicatori secondo i quali le regioni vengono collocate in zona rossa, arancione, gialla o bianca, sono legittimi dei dubbi sulla validità dei dati raccolti soprattutto in quel periodo vista la mancanza di dati da parte di numerose regioni.

Non è possibile dire quanto, sia l'ingresso dei test rapidi nel conteggio e sia la mancanza di ingresso dei test rapidi da parte di alcune regioni, influiscano negativamente ai fini del calcolo del tasso di positività nazionale che ogni giorno viene però messo in risalto dai mezzi di comunicazione.

Coronavirus, i dati di oggi: 18.020 nuovi casi con 121mila tamponi, altri 414 morti. Il tasso di positività risale al 14,9%

Il Fatto Quotidiano 7 gennaio 2021

Il problema è scrivere la parola “risale” o “cresce” quasi a far percepire che la situazione è peggiorata. È un dato troppo puntuale rispetto a tutta la complessità che sta dietro.

2.1.3 Rapporto positivi/tamponi o positivi/casi testati?

La terza e ultima considerazione da non dimenticare è l'esistenza di diversi tipi di tassi di positività. Finora abbiamo detto che il tasso di positività è il rapporto tra il numero dei positivi e il numero dei tamponi processati in un stesso arco di tempo.

Individuiamo due tipi di tasso di positività: il tasso di positività dato dal rapporto positivi/numero totale dei tamponi e il tasso di positività dato dal rapporto positivi/totale casi testati. Il primo include qualsiasi tipo di tampone processato quel giorno, sia se si tratta di un tampone rapido a cui è seguito un tampone molecolare, sia se si tratta di un tampone di controllo in seguito ad un periodo di quarantena causa infezione. Il secondo invece riguarda il numero di positivi sul totale dei primi test eseguiti, si tiene conto delle persone che non sono state mai tracciate quel giorno. Quest'ultimo ha fatto ingresso tra i dati non solo di lab24 ma in generale di coloro che raccontano l'andamento dell'epidemia, solo in data 19 aprile 2020. La differenza tra i due tassi è notevole, prendendo ad esempio il giorno 30 maggio 2021, il rapporto positivi/tamponi totali è 1,79% mentre quello rapporto positivi/casi testati 5,38% (lab24, 2021).

Dovrebbe spaventarci di più il secondo tasso in quanto indica un valore più realistico.

Eppure, abbiamo mai letto in un titolo di un articolo di giornale questa precisazione? Si tratta del tasso di positività sul numero dei tamponi totali o sul numero dei casi testati?

Prendiamo gli articoli che ho già riportato nel capitolo e individuiamo se quanto scritto si rifa al primo o al secondo tipo di tasso di positività.

14,1% del 1 gennaio 2021: positivi/numero totale tamponi

14,9% del 7 gennaio 2021: positivi/numero totale tamponi

Continua infatti ad essere maggiormente utilizzato il tasso di positività basato sul rapporto positivi/tamponi effettuati soprattutto perché permette un confronto con gli altri paesi che continuano ad utilizzare il primo tasso di positività e non quello sui casi testati.

2.2 La variazione percentuale giornaliera dei casi totali

Nella pagina di approfondimento lab24 è presente un altro dato molto interessante che riguarda come varia la crescita in termini percentuali dei casi totali giorno per giorno, ovvero la variazione percentuale giornaliera. Viene proposto attraverso un grafico sia in scala lineare che in scala logaritmica. Il grafico in scala lineare comprende un intervallo di tempo che copre l'intero arco di durata della pandemia quindi dal 25 febbraio ad oggi, con aggiornamenti quotidiani:

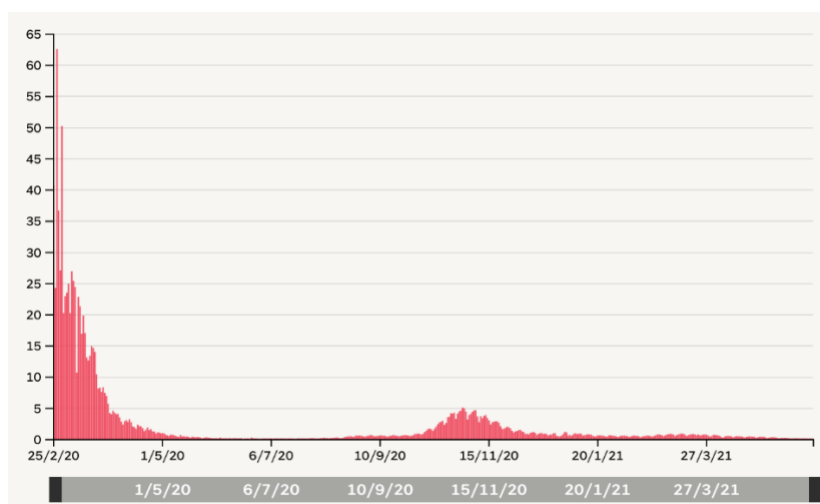


Grafico tratto da lab24

Sull'asse orizzontale troviamo la linea del tempo mentre sull'asse verticale il valore della variazione % giornaliera.

La variazione percentuale, come ogni variazione percentuale che si voglia calcolare, è data dal rapporto:

$$\text{variazione \%} = \frac{\text{valore finale} - \text{valore iniziale}}{\text{valore iniziale}}$$

Calando la formula nel contesto che ci interessa:

$$\text{variazione \%} = \frac{\text{casi totali finali} - \text{casi totali iniziali}}{\text{casi totali iniziali}}$$

Prendendo quindi due giorni differenti e consecutivi, 29 e 30 maggio 2021 (lab24, 2021):

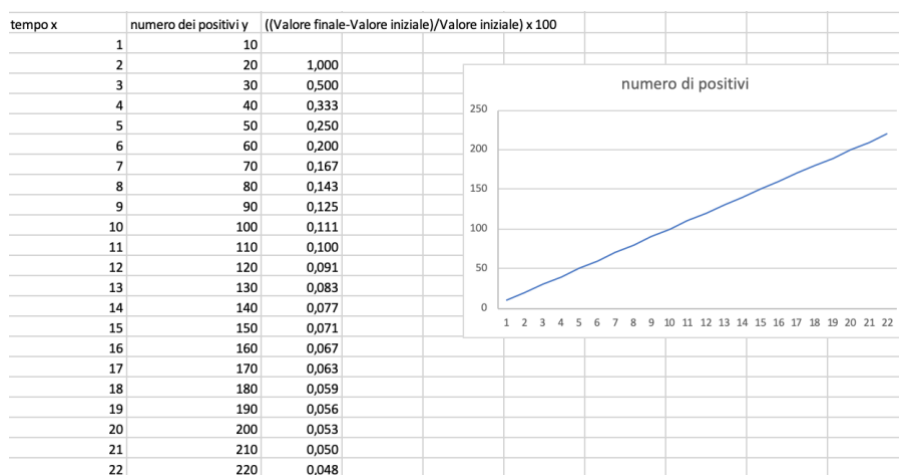
$$\text{variazione \%} = \frac{\text{casi totali 30 maggio} - \text{casi totali 29 maggio}}{\text{casi totali 29 maggio}} = \frac{4.216.003 - 4.213.055}{4.213.055} = 0,07\%$$

È importante notare numeratore e denominatore della formula appena analizzata, il numeratore è dato dalla differenza tra i casi totali del giorno e quelli del giorno precedente, ci dice quindi il numero dei nuovi positivi; il denominatore invece è il numero complessivo dei contagiati da inizio della pandemia, anche quello con il passare dei giorni aumenta sempre di più, o almeno fino a che aumenteranno i contagi. Ricordiamo sempre che più alto diventa il denominatore e più basso sarà il valore del rapporto.

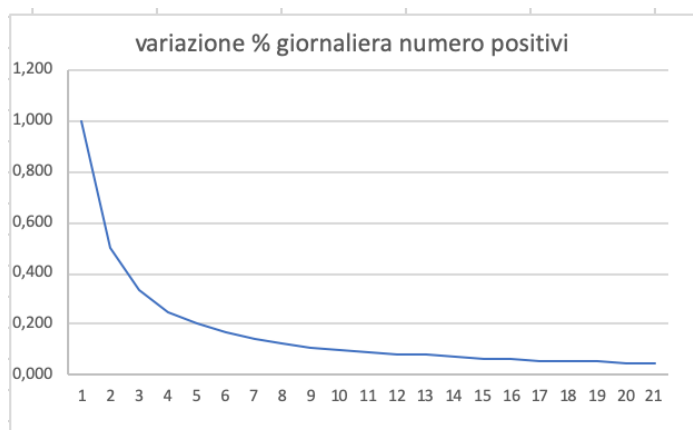
Procediamo con ordine. Cosa mostra il grafico rappresentato qui sopra? Si può notare come all'inizio, nei mesi di febbraio, marzo e aprile, si sia registrata una variazione percentuale elevatissima, mentre nei mesi di ottobre e novembre 2020 durante la seconda ondata, nonostante l'elevato numero di contagi, la variazione percentuale è stata piuttosto bassa, niente di paragonabile alla prima ondata. Le conclusioni che quindi possiamo trarre sono a dir poco ottimistiche, ad una prima occhiata il grafico ci fa ben sperare.

Ma è un grafico utile? È giusto quanto comunicato alla prima occhiata?

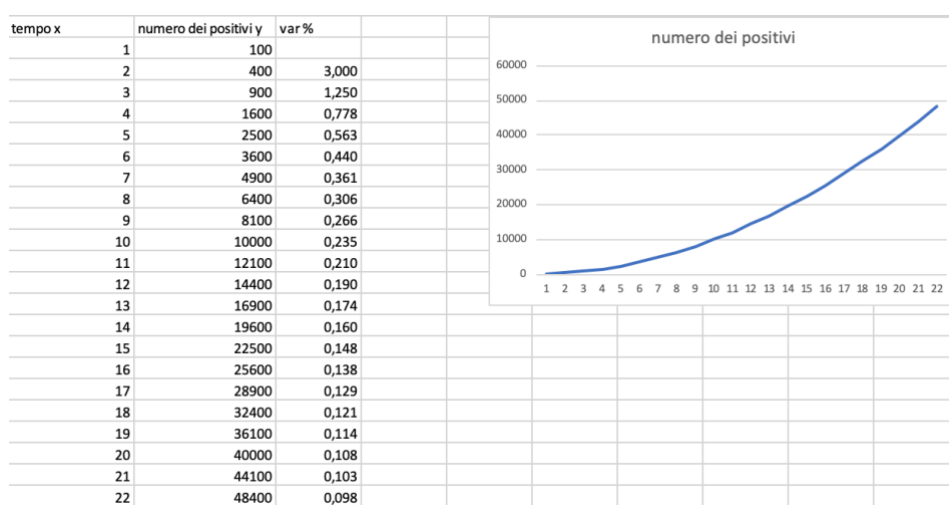
Facciamo un paio di esempi. Supponiamo una **crescita lineare** del numero dei positivi, in particolare ogni giorno i positivi aumentano di 10 unità. Riproduciamo il grafico con l'aiuto di excel per un periodo piuttosto breve.



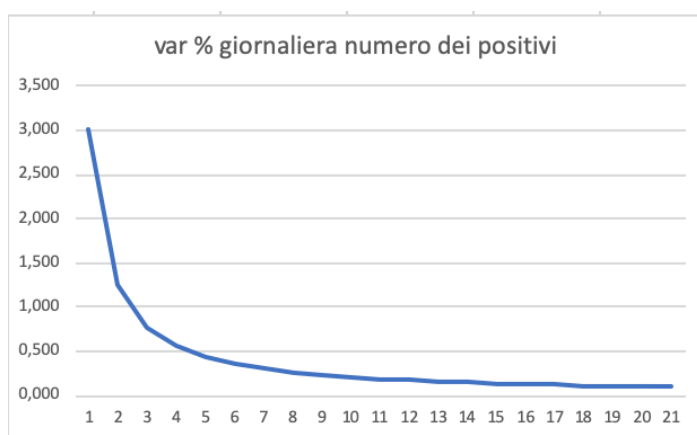
La variazione percentuale:



Stesso discorso ma proviamo addirittura con una **crescita quadratica** del numero dei nuovi positivi giornalieri:



La variazione percentuale:



Con un aumento lineare del numero, la variazione nel corso del tempo tende a zero, provando con un aumento quadratico, quindi un aumento spropositato e distante dalla realtà, il grafico della variazione percentuale giornaliera mostra un andamento che nel tempo tende sempre verso lo zero.

Sono dei grafici inadeguati, non descrivono il fenomeno e non danno indicazioni precise e sensate a coloro che li leggono. Sono informazioni imprecise, è naturale che quel numero

tenderà sempre a 0 perché ad aumentare non è solo il numeratore, quindi solo il numero dei nuovi positivi, ma anche il denominatore che è un dato cumulato e che aumenta quotidianamente.

Capitolo 3

Il coprifuoco

3.1 Il coprifuoco è efficace contro il coronavirus?

L'esempio conclusivo che ho intenzione di trattare si concentra su una delle misure di salute pubblica adottate negli ultimi mesi che mirano a prevenire e/o controllare la trasmissione della SARS-CoV-2 nella comunità: il coprifuoco. Fa parte dei cosiddetti interventi non farmaceutici (NPI in inglese, ovvero non-pharmaceutical interventions) introdotti in 30 paesi europei che, combinati con misure farmaceutiche come i vaccini, possono rispondere in maniera consistente alla pandemia di Covid-19. Essendo il coprifuoco una misura in vigore da molti mesi, ormai da novembre 2020, in molti, sia comuni cittadini che politici che esperti in epidemiologia piuttosto che in statistica, si sono chiesti l'utilità e l'effettiva efficacia in termini di riduzione dei contagi o delle ospedalizzazioni o dei ricoveri in terapia intensiva. Insomma, in che modo il coprifuoco è stata una misura così utile da essere tenuta per così tanti mesi? Si è accusato il governo di mantenere in vigore misure prive di fondamento scientifico. È davvero così? Una risposta a questa domanda la troviamo grazie ad alcuni studi che hanno cercato di misurare gli effetti di numerose misure non farmaceutiche adottate dai paesi, tra le quali il coprifuoco, in termini di riduzione dell'indice Rt e del numero medio dei nuovi positivi registrati.

L'indice Rt, noto come indice di trasmissibilità o indice di riproduzione netto di un virus, è un parametro che misura la potenziale trasmissibilità di una malattia infettiva. È il numero medio di infezioni prodotte da ciascun individuo infetto in seguito all'applicazione di specifici interventi atti a contenere la diffusione del virus. È quindi un modo per calcolare l'efficacia di tali misure e risulta essere l'indice più affidabile e più utilizzato nonostante tenda a sottostimare la reale trasmissione del virus, in quanto tiene conto solo delle persone sintomatiche (Anon., 2021).

3.2 Il pre-print e la diffusione nei media

La maggior parte degli studi che prenderò in considerazione o che citerò, ai tempi in cui sono stati pubblicati e quando hanno ricevuto una certa attenzione mediatica, soprattutto da parte dei giornali, sono scritti in formato pre-print, ovvero in formato pre-stampato. È una modalità di pubblicazione che, soprattutto nell'ultimo periodo, ha trovato una forte diffusione vista la numerosità di centri di ricerca e centri universitari che l'hanno adottata. Lo stato di emergenza

provocato dallo scoppio della pandemia di Covid-19 ha avuto ripercussioni su vari piani, sicuramente sul piano sociale, economico e psicologico, ma anche su un aspetto quale la produzione, circolazione e validazione dei dati scientifici. È tutto nato da una forte spinta che è avvenuta in questi mesi nella ricerca del virus e che ha cambiato in maniera notevole anche le modalità attraverso le quali la ricerca biomedica viene condotta e comunicata. La modalità di pubblicazione prevede di fornire risultati prima del controllo a cura di esperti del settore che, secondo abitudine, sottopongono al referaggio i risultati pubblicati (Lovisolò, 2021). Il primo studio britannico a cui mi rifarò per una parte delle considerazioni è stato pubblicato su un archivio online di nome MedRxiv, il principale archivio di ricerca biomedica, nato di recente nel 2019 e gestito dai Cold Spring Harbors Laboratories, dal British Medical Journal e dall'Università di Yale.

Lo studio a cui faremo riferimento vede come autori, docenti e ricercatori dei più disparati dipartimenti dell'Università di Oxford, dell'Università di Bristol, dell'Imperial College di Londra, sono nomi che hanno un certo rilievo e sui quali l'affidabilità a primo impatto sembrerebbe essere indiscussa, eppure è importante sottolineare, e lo sottolinea anche l'archivio stesso, che quanto viene pubblicato in formato pre-print non fornisce dati definitivi, e i risultati non possono essere usati per pratica clinica e per diffusione nei media.

Invece, per quanto riguarda la diffusione nei media, è proprio quello che è accaduto.

Questo quanto riportato da alcuni titoli di giornale:

CRONACA 29/04/2021 10:50 CEST

"Il coprifuoco riduce l'Rt del 13%": i risultati di uno studio europeo

I risultati di una ricerca sulle "misure non farmacologiche" condotta in 7 paesi da alcune delle maggiori università europee

HuffPost



il Corriere della Sera, 29 aprile 2021

LO STUDIO DI OXFORD

Il coprifuoco riduce l'Rt del 13%, come le mascherine: lo studio europeo

I migliori gruppi di ricerca europei hanno calcolato gli effetti delle chiusure sulla riduzione dei casi. Quella dei negozi è la più «utile» e anche la limitazione degli incontri

di Margherita De Bac

Huffington Post, 29 aprile 2021

3.3 I risultati di alcuni studi europei

Iniziamo con lo studio europeo a cui si riferiscono le testate giornalistiche citate, studio dal titolo "understanding the effectiveness of government interventions in Europe's second wave

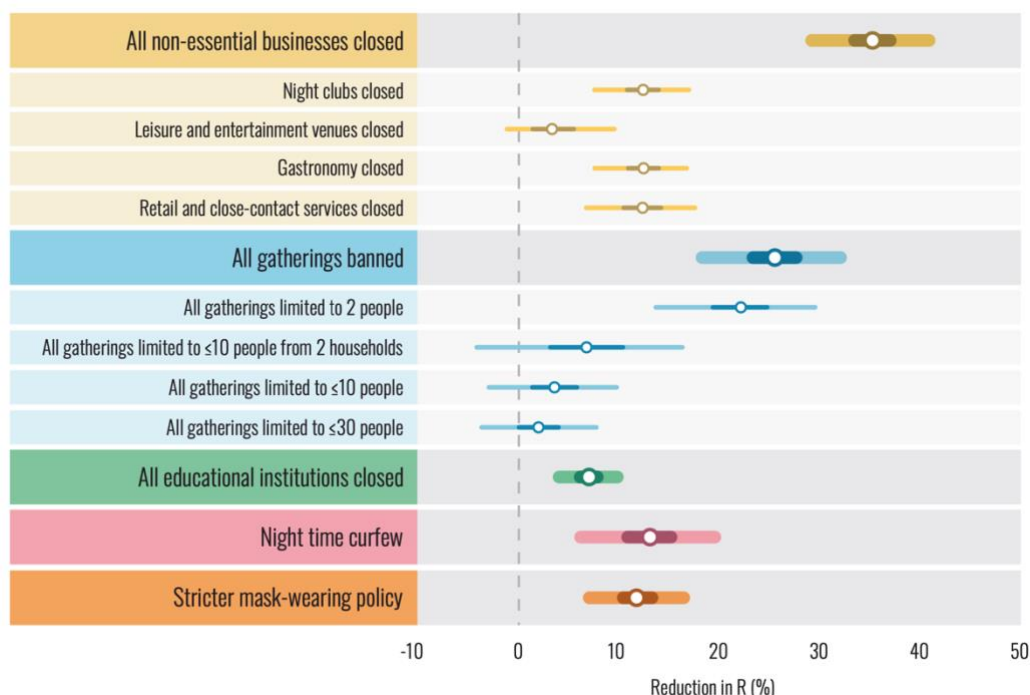
of COVID-19”, ovvero “capire l’efficacia degli interventi messi in campo dai governi durante la seconda ondata di COVID-19”.

Lo studio prende in esame 7 paesi europei e analizza 114 regioni diverse all’interno dei paesi Austria, Repubblica Ceca, Inghilterra, Germania, Italia, Paesi Bassi e Svizzera. Il periodo nel quale si svolge l’analisi è durante la seconda ondata di Covid-19, in particolare dal 1° agosto 2020 al 9 gennaio 2021 (senza considerare il problema dell’insorgenza di varianti del virus sorte in UK prima della data 9 gennaio). Le NPI che sono state prese in considerazione non sono misure esclusivamente adottate durante la seconda ondata, anzi sono state e rimangono uno degli strumenti principali, se non lo strumento principale fino all’inizio della campagna vaccinale, per controllare la diffusione del virus nel breve periodo. Perché allora è stata considerata la seconda ondata e non la prima? Non solo perché le misure sono state implementate e aggiornate nel corso del tempo, ma anche per via di differenze significative che hanno caratterizzato i due periodi di tempo. Le caratteristiche della prima ondata sono state imprevedibilità, incertezza, impreparazione generale, si è dovuta gestire un’emergenza in pochissimo tempo sperimentando delle misure che non si sapeva se sarebbero state efficaci fino all’utilizzo. È stata la fase iniziale, non si sapeva se la mascherina andasse indossata anche in luoghi chiusi, piuttosto che il distanziamento sociale, se tenere la distanza di un metro l’uno dall’altro potesse bastare e moltissime altre domande. La seconda ondata invece è stata qualcosa di previsto da parte degli esperti e gli individui erano già pronti ad affrontarla, prendendo come riferimento quanto vissuto nei mesi precedenti con un minimo spiraglio di libertà durante l’estate. I modelli di contatto erano cambiati già nella prima ondata e determinati NPI venivano già adottati, evitando luoghi affollati, lavorando da casa, frequentando le lezioni da casa vista la chiusura di scuole e università così come le organizzazioni, che si erano piano piano mobilitate per affrontare la situazione. Inoltre gli NPI adottati nella prima ondata sono stati provvedimenti messi in campo subito e tutti insieme, era quindi molto difficile scovare gli effetti delle singole misure adottate. Nella seconda ondata invece, molte misure, ad esempio la chiusura di scuole, centri commerciali, musei, cinema, teatri e centri di attività sportiva è avvenuta in momenti diversi, si è adottato un approccio più scaglionato nel tempo e questo sicuramente ha favorito l’analisi delle singole misure. Complessivamente quindi, si è previsto che i dati della seconda ondata fornissero un quadro più dettagliato, solido e contemporaneo in grado di misurare l’impatto delle misure specifiche.

Sono stati utilizzati dei modelli statistici per stimare l’efficacia degli NPI confrontando le osservazioni epidemiologiche con e senza gli NPI, quindi evidenziando il numero di giorni nei quali le misure di contrasto erano attive.

Quali sono gli NPI inseriti nella ricerca? Ne sono 19 e tramite i modelli statistici è stata stimata la misura dell’impatto dei singoli interventi, in alcuni casi raggruppandoli (ad esempio sono stati messi insieme i vari raduni pubblici e privati), in termini di riduzioni percentuali, sull’indice di trasmissibilità Rt.

Ecco quanto riportato nello studio, concentriamoci soprattutto sugli effetti degli ultimi due NPI, l’utilizzo delle mascherine e il coprifuoco notturno.



tratto dallo studio britannico “understanding the effectiveness of government interventions in Europe’s second wave of COVID-19”

L’utilizzo della mascherina e del coprifuoco notturno hanno avuto “effetti moderati ma statisticamente significativi”, rispettivamente, una riduzione del 12% dell’indice Rt con intervallo di confidenza al 95%: 7-17% e una riduzione del **13%** dell’indice Rt con intervallo di confidenza: 6-20%. Un effetto quindi, quello del coprifuoco, superiore alla misura “chiusura delle scuole” per esempio (Sharma, M., Minderman, S., et al., 2021).

Uno studio francese pubblicato su Eurosurveillance dedica l’attenzione all’impatto della misura di coprifuoco nel gennaio 2021 in Francia per contrastare la diffusione della nuova variante del virus, inizialmente localizzata in Inghilterra ma poi diffusa velocemente in vari paesi europei. Il coprifuoco, misura già adottata dal governo francese, venne inasprito e anticipato alle ore 18 in alcune regioni della Francia ma ben presto esteso a tutta la nazione. Quello che emerge dallo studio è che il coprifuoco e gli altri NPI adottati nei mesi appena precedenti a gennaio 2021, hanno aiutato a ridurre l’indice di trasmissibilità ma non sono riusciti ad evitare un aumento del numero di casi positivi e di ospedalizzazioni provocate dalla nuova variante del virus. Il

coprifuoco, secondo lo studio, è una misura utile ma non sufficiente (Di Domenico, L., Sabbatini, C., 2021).

Uno studio sempre francese ha invece analizzato il numero medio di casi positivi al Covid-19 in un periodo di tempo di 7 giorni per vedere come evolveva questo numero con le misure di contrasto in azione, tra cui il coprifuoco. Quello che è emerso è che l'impatto positivo c'è stato, ma dato che le varie misure sono state adottate tutte insieme (coprifuoco, lockdown, chiusura di attività commerciali), è stato impossibile separare gli effetti causati da una singola misura (Valdano, E., Lee, J. et al., 2021).

Un altro studio analizza gli impatti di vari interventi, farmaceutici e non, come mezzi di contrasto al Covid-19 nello stato del Lussemburgo nel mese di ottobre 2020. Per quanto riguarda la misura coprifuoco, quanto emerge è che l'impatto del coprifuoco dalle 11 di sera alle 6 di mattina è relativamente piccolo se confrontato con l'impatto che ha una misura restrittiva come il lockdown. Il problema di questo modello, e viene specificato dallo studio stesso, è che non si considera come il coprifuoco influenzi in realtà il comportamento degli individui ad inizio serata, si considera che gli individui intorno a quelle ore si muovano solo per dirigersi verso le proprie abitazioni a meno che non si trovino in ospedale. Gli individui invece potrebbero adottare comportamenti diversi dal solito, proprio in vista del coprifuoco, ad esempio andare a fare la spesa prima (Thompson, J., Wattam, S., 2021).

Tra questi studi appena citati esistono elementi in comune che è bene presentare: tutti coprono un arco temporale piuttosto breve, quasi tutti comprendono il periodo di autunno-fine inverno; sono studi che precedono l'inizio della campagna vaccinale; vengono utilizzati dei modelli rigidi, quasi astratti, per ammissione degli stessi studiosi, che non prendono in considerazione la variabilità comportamentale degli individui; infine, l'aspetto più importante, è quasi impossibile isolare i dati dalle altre misure che sono state adottate, la curva dei contagi diminuisce grazie al lockdown, alla chiusura delle scuole, all'utilizzo delle mascherine, forse anche grazie al coprifuoco ecc.

Gli studi appena visti indicano che il coprifuoco può probabilmente aiutare a ridurre la trasmissione del virus, gli studi più ottimistici dicono che esso riduce l'indice R_t del 13% ma sono presenti molti cavilli, i modelli sono incompleti, inevitabilmente incompleti.

I modelli servono per costruire un sistema che si comporti e racconti una situazione il più possibile vicina alla realtà ma è impossibile simulare i comportamenti individuali.

Determinare quindi l'impatto della singola misura di coprifuoco è molto impegnativo. Ci sono dati ma sono frammentari. E perché allora continua a rimanere in vigore? Probabilmente per motivi precauzionali, è una misura che soprattutto nei mesi di ottobre, novembre in piena seconda ondata era utile a scoraggiare le uscite, di sicuro non è risolutiva nel frenare la crescita del contagio.

A fronte di queste conclusioni, e riprendendo i titoli citati prima nel periodo del 19/29 aprile, periodo di pubblicazione della maggior parte di questi studi, è bene ribadire che una comunicazione di quel tipo è fuorviante. Tenendo conto di tutti questi elementi, non è proprio vero che il coprifuoco riduca l'indice R_t del 13%. È quindi bene ancora una volta prestare attenzione nella lettura delle informazioni, controllare le fonti e non farsi convincere subito dell'efficacia di una misura.

Conclusioni

La statistica ha svolto un ruolo di vitale importanza nel corso della pandemia da Covid-19, essenzialmente perché è stata usata per cercare di spiegarci attraverso i numeri quello che stava accadendo. Gli epidemiologi, gli statistici, gli economisti, gli esperti che, grazie all'uso di indicatori e grafici, hanno in questi mesi offerto spiegazioni, a noi cittadini ma soprattutto a coloro che dovevano prendere delle decisioni che avrebbero avuto ripercussioni, non solo a livello nazionale ma anche internazionale, vista la portata dell'evento che ci ha travolto.

Provare a limitare il più possibile i danni economici tenendo chiuse le attività oppure salvaguardare la popolazione in tema di salute? Chiudere le scuole per prevenire la diffusione del virus o meglio lasciare i bambini a casa? Quanto aiuta la mascherina nel prevenire il contagio? Sono tutte domande alle quali si è cercato di rispondere ma non sempre in maniera ragionata e coerente, anzi. Inevitabilmente le risposte erano caratterizzate da incertezza, per via della scarsa disponibilità di dati disaggregati, per via dello scarso coordinamento tra enti locali e governo centrale, per via delle diverse linee di pensiero di alcuni epidemiologi, ma soprattutto per via delle scarse conoscenze che si avevano a riguardo del virus. Ognuno, con il materiale a disposizione, ha però cercato di fare il meglio che poteva.

I tre esempi di incorretta comunicazione di dati statistici che ho presentato in questo lavoro rappresentano solo dei casi, ma sono, a mio parere, molto utili per una serie di motivi.

Innanzitutto, ci hanno permesso di comprendere l'importanza di affidarsi alle informazioni provenienti solo da fonti qualificate, e di non dare per scontato che gli indici siano sempre i portatori di verità. È indispensabile che chi pubblica i valori degli indici ne chiarisca il significato, offra istruzioni e consigli le necessarie cautele per interpretarli correttamente.

Inoltre, abbiamo avuto modo di vedere gli effetti di una comunicazione di dati fatta in maniera superficiale e poco attenta alle conseguenze. L'esempio è il vaccino AstraZeneca: nei mesi successivi a marzo 2021, molte sono state le regioni a registrare dosi inutilizzate del vaccino.

A conclusione del lavoro svolto, mi sento di affermare che la statistica è sì, una scienza che permette grande manipolazione, perché riesce a convincere facilmente, soprattutto coloro che sono poco inclini alla verifica o che sono disattenti, a interpretare la realtà non sempre in maniera oggettiva, ma è fondamentale e illuminante perché ci consente di esplorare, di dare risposte alle nostre curiosità e di comprendere più a fondo quanto accade intorno a noi. È utile comprenderla perché a guadagnarci siamo noi cittadini.

Bibliografia

- AIFA, periodo dal 27/12/2020 al 26/03/2021. Terzo rapporto sulla sorveglianza dei vaccini Covid-19.
- Borra, Di Ciaccio, 2014. Statistica: metodologie per le scienze economiche e sociali. McGraw-Hill Education.
- ECDC, 24 Settembre 2020. Guidelines for the implementation of non- pharmaceutical interventions against COVID-19.
- Huff, D., 2009. Mentire con le statistiche.
- Ministero della Salute, 22 Febbraio 2021. Aggiornamento sulle modalità d'uso del vaccino COVID-19 VACCINE ASTRAZENECA.
- Ministero della Salute, 8 Marzo 2021. Utilizzo del vaccino COVID-19 VACCINE ASTRAZENECA nei soggetti di età superiore ai 65 anni.
- Ministero della Salute, AIFA et al., 10 Marzo 2021. Raccomandazioni ad interim sui gruppi target della vaccinazione anti-SARS-CoV-2/COVID-19.
- Savelli, F., 8 Marzo 2021. AstraZeneca troppe dosi inutilizzate. Moratti: datele alle regioni virtuose. *Il Corriere della Sera*, 8.
- Voltattorni, C., 16 Marzo 2021. L'Italia ferma AstraZeneca. "Sospensione temporanea". *Il Corriere della Sera*, 2.

Sitografia

- Anon., 15 Gennaio 2021. Covid, bollettino Italia del 15 gennaio. I numeri sul Coronavirus nelle regioni. Quotidiano nazionale. Disponibile su: <https://www.quotidiano.net/cronaca/covid-oggi-italia-numeri-1.5916765> [Ultimo accesso: 28/05/2021].
- Anon., 15 Marzo 2021. AstraZeneca, gravissima donna di 54 anni vaccinata a Napoli: la dose da un lotto sequestrato. *Il Mattino*. Disponibile su: https://www.ilmattino.it/napoli/cronaca/astrazeneca_napoli_donna_gravissima_54_an_ni_sonia_battaglia-5832861.html [Ultimo accesso: 04/06/2021].
- Anon., 16 Gennaio 2021. Aggiornamento casi Covid-19. *Il Messaggero*. Disponibile su: https://www.ilmessaggero.it/uploads/ckfile/202101/covid_bollettino_16_gennaio_2021_16172246.pdf [Ultimo accesso: 31/05/2021].

- Anon., 27 Gennaio 2021. Covid. Il bollettino: oggi 15.204 nuovi casi e 467 decessi. Da inizio pandemia superati i 2,5 milioni di casi. Quotidiano sanità. Disponibile su: http://www.quotidianosanita.it/studi-e-analisi/articolo.php?articolo_id=91923 [Ultimo accesso: 31/05/2021].
- Anon., 7 Maggio 2021. Indice Rt: cos'è, come si calcola e qual è la differenza con l'R0. Sky tg24. Disponibile su: <https://tg24.sky.it/salute-e-benessere/approfondimenti/indice-rt> [Ultimo accesso: 03/06/2021].
- Bocci, M., 3 Aprile 2020. Coronavirus, il ministero chiarisce: ecco chi deve fare il tampone. *La Repubblica*. Disponibile su: https://www.repubblica.it/salute/medicina-e-ricerca/2020/04/03/news/coronavirus_il_ministero_chiarisce_ecco_chi_deve_fare_il_tampone-253063119/ [Ultimo accesso: 31/05/2021].
- Bottarelli, E., 26 Ottobre 2020. Campionamento con metodo non probabilistico. Quaderno di epidemiologia veterinaria. Disponibile su: http://www.quadernodiepidemiologia.it/epi/campion/non_pro.htm [Ultimo accesso: 29/05/2021].
- Capozzoli, G., 31 Agosto 2020. Roma. Uno studio sul tasso di positività di Covid-19 in Italia. *L'Espresso*. Disponibile su: http://fino-a-prova-contraria.blogautore.espresso.repubblica.it/2020/08/31/roma-uno-studio-sul-tasso-di-positivita-di-covid-19-in-italia/?fbclid=IwAR3KmXHt3RVYZvUgrRNO58kWIJ9_u4j0l-MIq5Bqsou96-i2-B6wHQcRXc4 [Ultimo accesso: 27/05/2021].
- Di Domenico, L., Sabbatini, C., et al., 15 Aprile 2021. Impact of January 2021 curfew measures on SARS-CoV-2 B.1.1.7 circulation in France. *Eurosurveillance*. Disponibile su: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.15.2100272> [Ultimo accesso: 04/06/2021].
- EMA, 11 Marzo 2021. COVID-19 Vaccine AstraZeneca: PRAC investigating cases of thromboembolic events-vaccine's benefits currently still outweigh risks. Disponibile su: <https://www.ema.europa.eu/en/news/covid-19-vaccine-astrazeneca-prac-investigating-cases-thromboembolic-events-vaccines-benefits> [Ultimo accesso: 02/06/2021].
- Fassari, L., 27 Gennaio 2021. Tamponi rapidi: grande difformità nei risultati e otto regioni non forniscono i dati dei referti. Quotidiano sanità. Disponibile su: http://www.quotidianosanita.it/studi-e-analisi/articolo.php?articolo_id=91901 [Ultimo accesso: 25/05/2021].

- Jarry, J., 23 Aprile 2021. Do curfews work? McGill. Disponibile su: <https://www.mcgill.ca/oss/article/covid-19-general-science/do-curfews-work> [Ultimo accesso: 03/06/2021].
- JMP Statistics Knowledge Portal. Correlazione o causalità? Disponibile su: https://www.jmp.com/it_it/statistics-knowledge-portal/what-is-correlation.html [Ultimo accesso: 30/05/2021].
- Lovisolo, D., 16 Febbraio 2021. Come Covid-19 cambia i meccanismi di pubblicazione e validazione dei dati. Scienza in rete. Disponibile su: <https://www.scienzainrete.it/articolo/come-covid-19-cambia-meccanismi-di-pubblicazione-e-di-validazione-dei-dati/davide-lovisolo> [Ultimo accesso: 04/06/2021].
- Messaggero Veneto, 4 Dicembre 2020. Tamponi molecolari, antigenici e sierologici: i test anti-Covid spiegati dall'esperta. Disponibile su: <https://www.youtube.com/watch?v=36LuYHQ59b8> [Ultimo accesso: 31/05/2021].
- Ministero della Salute, ultimo aggiornamento 12 giugno 2021. Piano vaccini anti Covid-19. Disponibile su: <https://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?lingua=italiano&id=5452&area=nuovoCoronavirus&menu=vuoto> [Ultimo accesso: 30/05/2021].
- Origoni, P., Novembre 2020. Statistica e pandemia. Statistiche e società. Disponibile su: https://www3.ti.ch/DFE/DR/USTAT/allegati/articolo/2701dss_2020-2_0.pdf [Ultimo accesso: 27/05/2021].
- Pagliani, S., 8 Maggio 2021. La farmacovigilanza sui vaccini è essenziale per la nostra salute. Scienza in rete. Disponibile su: <https://www.scienzainrete.it/articolo/farmacovigilanza-sui-vaccini-essenziale-nostra-salute/simonetta-pagliani/2021-05-08> [Ultimo accesso: 30/05/2021].
- Paknazar, B., 14 Luglio 2020. Covid-19: bassa adesione all'indagine statistica ISTAT sulla sieroprevalenza. Un'occasione persa. Il Bo Live Unipd. Disponibile su: <https://ilbolive.unipd.it/index.php/it/news/covid19-bassa-adesione-allindagine-istat-sulla> [Ultimo accesso: 26/05/2021].
- Puente, D., 16 Marzo 2021. La strage che non c'è del vaccino AstraZeneca (e Pfizer). Basta paure infondate sui vaccini anti Covid-19! OPEN. Disponibile su: <https://www.open.online/2021/03/16/covid-19-nessuna-strage-vaccino-astrazeneca-pfizer-basta-paure-infondate/> [Ultimo accesso: 03/06/2021].
- Renda, S., 15 Marzo 2021. "In Italia circa 65 mila casi di tromboembolismo polmonare all'anno. Non c'è incremento." Huffington Post. Disponibile su:

- https://www.huffingtonpost.it/entry/in-italia-circa-65mila-casi-di-tromboembolismo-polmonare-allanno-non-ce-incremento_it_604f448ec5b65bed87dc64b1 [Ultimo accesso: 04/06/2021].
- Ruffino, L., 19 Aprile 2021. Il coprifuoco serve contro il coronavirus? YouTrend. Disponibile su: <https://www.youtrend.it/2021/04/19/il-coprifuoco-serve-contro-il-coronavirus/> [Ultimo accesso: 24/06/2021].
 - Scienze Statistiche Unipd, 19 marzo 2020. #8 Le parole dell'epidemia: stima dei contagi. Disponibile su: <https://www.youtube.com/watch?v=zSuoi-Bxyy4> [Ultimo accesso: 25/05/2021].
 - Scienze Statistiche Unipd, 22 Dicembre 2020. "Parola agli statistici. Affrontare il Covid-19 con dati e competenze". Disponibile su: <https://www.youtube.com/watch?v=vAQb56zdB9w&t=4418s> [Ultimo accesso: 24/05/2021].
 - Scienze Statistiche Unipd, 25 maggio 2020. #9 Le parole dell'epidemia: campione rappresentativo. Disponibile su: <https://www.youtube.com/watch?v=zSuoi-Bxyy4> [Ultimo accesso: 25/05/2021].
 - Sestili, G., 16 Dicembre 2020. Comunicare i dati della pandemia di Covid-19. Data journalism by Giorgio Sestili. Disponibile su: <https://www.giorgiosestili.it/comunicare-i-dati-della-pandemia-di-covid-19/> [Ultimo accesso: 04/06/2021].
 - Sharma, M., Minderman, S., et al., 26 marzo 2021. Understanding the effectiveness of government interventions in Europe's second wave of COVID-19. MedRxiv. Disponibile su: <https://www.medrxiv.org/content/10.1101/2021.03.25.21254330v1.full.pdf> [Ultimo accesso: 03/06/2021].
 - Thompson, J., Wattam, S., 26 Marzo 2021. Estimating the impact of interventions against COVID-19: from lockdown to vaccination. MedRxiv. Disponibile su: <https://www.medrxiv.org/content/10.1101/2021.03.21.21254049v1.full.pdf> [Ultimo accesso: 03/06/2021].
 - Università di Verona, 26 Ottobre 2020. Comunicazione e informazione statistica nell'epoca del Covid-19 | diffusioni. Disponibile su: <https://www.youtube.com/watch?v=tm7eUv3YJVQ> [Ultimo accesso: 26/05/2021].
 - Valdano, E., Lee, J. et al., 14 Aprile 2021. Highlighting socio-economic constraints on mobility reductions during COVID-19 restrictions in France can inform effective and equitable pandemic response. Journal of Travel Medicine. Disponibile su:

<https://academic.oup.com/jtm/article/28/4/taab045/6225375> [Ultimo accesso: 05/06/2021].

- Ventura, L., 15 Dicembre 2020. Covid-19: La statistica racconta la pandemia. Università degli studi di Padova. Disponibile su: https://www.stat.unipd.it/sites/dipartimenti.it/files/ombragalileo_LV.pdf [Ultimo accesso: 24/05/2021].