

Morti COVID e numero oscuro: l'importanza della granularità

written by Paolo Milan | 21 Aprile 2020

Abstract

È ormai convinzione pressoché unanime che il numero di decessi COVID-19 positivi sia ampiamente sottostimato. A partire dai dati sulla mortalità generale pubblicati dal [SISMG](#) e dall'[ISTAT](#) diversi studi hanno cercato di individuare dei *moltiplicatori*, coefficienti calcolati come il rapporto tra il numero di decessi in eccesso rispetto alla mortalità attesa e il numero di decessi COVID-19 positivi ufficialmente censiti. Grazie a un [tweet](#) di Giorgio Gori sappiamo che a Bergamo dal 1 marzo al 12 aprile a fronte di 626 decessi in eccesso rispetto alla media storica risultavano appena 272 decessi ufficialmente COVID-19 positivi. Se per Bergamo, al 12 aprile, è possibile affermare che il moltiplicatore è 2.3, per la maggior parte dei comuni italiani non è possibile fare un calcolo esatto.

Ovviamente è sempre possibile tentare delle stime, ma uno dei problemi principali è la mancanza di dati ufficiali dotati di un livello di dettaglio significativo: anche conoscendo la mortalità in eccesso rispetto alla media storica di un determinato comune non è possibile calcolare con sicurezza alcun moltiplicatore proprio perché il dato sui decessi ufficialmente COVID-19 positivi è generalmente disponibile soltanto come aggregato regionale.

Tra i tanti lavori sull'eccesso di mortalità meritano particolare attenzione quelli dei fisici Daniele Del Re e Paolo Meridiani. Analizzando alcuni [dei loro risultati](#) è

possibile ad esempio ipotizzare che al 3 aprile i moltiplicatori di alcuni capoluoghi del Nord Italia (Milano 2.4, Torino 2.7, Genova 4.4) fossero significativamente inferiori rispetto a quelli di alcuni capoluoghi del Sud Italia (Messina 12.5, Bari 7.2, Palermo 6.8). Si tratta ovviamente di stime, per di più all'interno di intervalli di confidenza talvolta abbastanza ampi.

Scopo di questo articolo è evidenziare come la diffusa disponibilità di dati ufficiali meno aggregati rispetto a quelli della Protezione Civile possa contribuire a ridurre significativamente il margine di errore in cui ogni stima necessariamente incorre.

Introduzione

È ormai convinzione pressoché unanime che il numero di decessi ufficialmente positivi a COVID-19 sia ampiamente sottostimato. Il notevole incremento della mortalità generale registrato nelle zone più colpite dall'epidemia di COVID-19, rilevato inizialmente da inchieste indipendenti come [quella condotta dal Sindaco di Nembro \(BG\)](#), successivamente confermato in Italia dal [SISMG](#) ha recentemente trovato ulteriori riscontri grazie alla diffusione dei primi dati sulla mortalità generale rilasciati da [ISTAT](#).

Diversi studi, per diversi scopi, hanno cercato di individuare dei *moltiplicatori*, coefficienti calcolati ad esempio come il rapporto tra il numero di decessi in eccesso rispetto alla mortalità attesa e il numero di decessi COVID-19 positivi comunicati dalla Protezione Civile[1]. Uno dei problemi principali è la mancanza di dati ufficiali dotati di un livello di dettaglio significativo: senza conoscere ad esempio quanti sono i decessi COVID-19 positivi nel comune di Milano, giorno per giorno, o almeno settimana per settimana, risulta difficile stimare un ipotetico moltiplicatore per il comune di Milano. Giorgio Gori, sindaco di Bergamo, ha [comunicato](#) che

nella città da lui amministrata “dal 1/3 al 12/4 sono deceduti 795 residenti, 626 più della media – nello stesso periodo – dei dieci anni precedenti. L’aumento è del 370%. I decessi ufficialmente ricondotti a Covid-19 sono stati 272, gli unici sottoposti a tampone”. Per il comune di Bergamo calcolando il rapporto tra i decessi in eccesso (626) e i decessi COVID-19 positivi (272) si ottiene il moltiplicatore è 2.3. E nel resto d’Italia? Come ragionevolmente rilevato da Andrea Gentile, Gianluca Dotti e Riccardo Saporiti su [Wired](#), “la capacità del Sistema Sanitario di intercettare i malati di Covid-19 è cambiata nel tempo e non è omogenea nelle diverse aree geografiche” e quindi probabilmente il moltiplicatore non è né omogeneo sul territorio nazionale né costante nel tempo.

L’idea dei fisici Daniele Del Re e Paolo Meridiani è partire dal dato ufficiale dei decessi a livello regionale per stimare il numero di decessi ufficiali a livello provinciale, e a partire da quest’ultimo stimare il numero di decessi ufficialmente COVID-19 positivi a livello comunale o per la porzione di territorio interessata. Analizzando alcuni [dei loro risultati](#) è possibile calcolare i moltiplicatori come rapporto tra decessi COVID-19 stimati (calcolati a partire dalla mortalità in eccesso rilevata dal SISMG) e decessi COVID-19 ufficiali (ricavati a partire dai dati della Protezione Civile) e ipotizzare che al 3 aprile i moltiplicatori di alcuni capoluoghi del Nord Italia (Milano 2.4, Torino 2.7, Genova 4.4) fossero significativamente inferiori rispetto a quelli di alcuni capoluoghi del Sud Italia (ad esempio Messina 12.5, Bari 7.2, Palermo 6.8). Si tratta ovviamente di stime, per di più all’interno di intervalli di confidenza talvolta abbastanza ampi.

In questo documento viene preliminarmente illustrata una procedura per stimare il numero di decessi ufficialmente COVID-19 positivi a livello comunale che riprende l’idea generale proposta e collaudata da Del Re e Meridiani. Successivamente si evidenzia come la diffusa disponibilità di

dati ufficiali meno aggregati rispetto a quelli della Protezione Civile possa contribuire a ridurre significativamente il margine di errore in cui ogni stima necessariamente incorre.

L'importanza della granularità

Per capire l'importanza della granularità dei dati – qui intesa come livello di dettaglio delle informazioni di pubblico dominio sull'esatto numero di decessi di positivi al COVID-19 – può essere conveniente partire dal pionieristico lavoro condotto da Sergio Cima su [Scienza in Rete](#): confrontare giorno per giorno l'eccesso di mortalità registrato dal SISMG con il numero di decessi COVID-19 positivi registrato nella città di Brescia. L'esperimento di Cima è stato possibile soltanto perché l'ATS di Brescia dirama quotidianamente un bollettino con il dato dei decessi giornalieri e questo bollettino viene ripreso da alcune testate locali, tra cui [Brescia Today](#).

Nella tabella seguente sono riportati il numero di decessi rilevati dal [SISMG](#) (dati ricavati dal quarto rapporto pubblicato il 17 aprile), la mortalità attesa (SISMG baseline), l'eccesso di mortalità rilevato (SISMG excess) e il numero di decessi positivi stimato a partire dai bollettini ATS Brescia. Dal momento che ATS Brescia non comunica l'età dei deceduti positivi, per consentire il confronto è necessario stimare il numero di decessi positivi appartenenti alla popolazione 65+ anni (Decessi 65+ positivi), calcolato per semplicità come meglio spiegato più avanti come il 90% del numero di decessi ufficialmente positivi. Viene infine mostrato il moltiplicatore calcolato come il rapporto tra l'eccesso di mortalità rilevato dal SISMG e il numero di decessi 65+ positivi stimato.

Settimana	SISMG decessi rilevati	SISMG baseline	SISMG excess (rilevati - baseline)	Decessi 65+ positivi (stima dai dati ATS BS)	Moltiplicatore
07/03 - 13/03	86	33	53	29	1.8
14/03 - 20/03	131	36	95	60	1.6
21/03 - 27/03	147	28	119	68	1.8
28/03 - 03/04	150	32	118	59	2
Totali	514	129	385	216	1.8

Purtroppo per la maggior parte dei comuni italiani il dato dei morti COVID-19 positivi non è disponibile. Una possibile soluzione è quindi stimare il numero di decessi a partire dai pochi dati resi quotidianamente disponibili per tutto il territorio nazionale dalla Protezione Civile: il numero di decessi a livello regionale e il numero totale di casi positivi a livello provinciale. Nei prossimi paragrafi viene presentata una procedura per la stima dei decessi comunali COVID-19 positivi, che ricalca il metodo adottato da [Del Re e Meridiani](#). Sostanzialmente alla stima finale si arriva per gradi: prima si passa dal regionale al provinciale e quindi dal provinciale al comunale. Per confrontare le stime ottenute con i dati di mortalità rilevati dal SISMG si rende necessario un terzo passaggio, dal comunale al comunale ristretto alla popolazione 65+ anni.

Passo 1 – Dal regionale al provinciale

Ogni giorno, e quindi ogni settimana, il numero di decessi COVID-19 positivi provinciali è ovviamente una parte di quelli regionali, e ovviamente il numero di decessi di un giorno sono una parte dei casi positivi rilevati nei giorni precedenti. In altri termini i decessi di oggi sono funzione dei casi attivi dei giorni precedenti, ammettendo di trascurare due categorie: le persone decedute il giorno stesso in cui vengono dichiarate positive e i deceduti a cui viene effettuato un tampone post mortem. Purtroppo il numero di casi attivi è un'informazione disponibile soltanto su base regionale, e comunque è risultata essere priva della dovuta precisione e attendibilità, come evidenziato da [un'eccellente inchiesta](#) di

GIMBE e Youtrend.

Per stimare il numero di decessi provinciali si estrapolano dunque dai bollettini quotidiani diramati dalla [Protezione Civile](#):

- il dato del totale decessi a livello regionale
- il dato del totale casi positivi a livello regionale
- il dato del totale casi positivi a livello provinciale per la provincia di interesse [\[2\]](#)

Su base settimanale si calcola quindi il rapporto tra i nuovi casi positivi registrati nella provincia di interesse e i nuovi casi positivi registrati in tutta la regione e tale rapporto viene utilizzato per stimare il numero di decessi della settimana successiva. Il rapporto tra nuovi casi positivi provinciali e nuovi casi positivi regionali della settimana X viene moltiplicato per il numero di decessi regionali della settimana X+1 per ottenere una stima del numero di decessi provinciali nella settimana X+1.

Assumere che i decessi di una certa settimana siano proporzionali ai nuovi casi positivi registrati nella settimana precedente è chiaramente una forte semplificazione ed è sicuramente possibile studiare criteri più raffinati e precisi. Si segnala ad esempio un interessante e approfondito studio di [Francesco Furno](#) che per altri scopi stima come trascorrono “cinque giorni in media tra il momento in cui viene diagnosticato il caso (e quindi venga conteggiato nei casi totali) e il momento in cui viene registrato il decesso”, evidenziando inoltre come questo dato presenti “significative differenze regionali”. In generale – parafrasando [Gentile, Dotti e Saporiti](#) – dal momento che la capacità del Sistema Sanitario di *salvare la vita* ai malati di Covid-19 è cambiata nel tempo e non è omogenea nelle diverse aree geografiche, potrebbe essere complicato individuare criteri precisi e validi continuamente nel tempo e ovunque nello spazio per stimare il numero di decessi provinciale a partire dal dato regionale.

Passo 2 – Dal provinciale al comunale

In assenza di dati migliori, una volta stimato il numero di decessi a livello provinciale è possibile attribuire ad ogni comune un numero di decessi proporzionale al numero di abitanti. Questa scelta è chiaramente un'approssimazione: così come passando dal livello regionale a quello provinciale è meglio tener conto del numero di casi positivi piuttosto che guardare al numero di abitanti, un discorso analogo potrebbe valere per il passaggio dal livello provinciale a quello comunale.

Effettuare il passaggio dalla provincia al comune utilizzando come criterio il numero di residenti potrebbe causare errori significativi in alcuni comuni. Si segnala tra tutti il caso di Roma: in base ai dati resi disponibili dal [DEP-LAZIO](#) a fronte del 66% dei residenti nella capitale rispetto al totale dei residenti della Provincia, risultano soltanto il 47% dei casi positivi provinciali (1732 positivi nel comune di Roma a fronte di 3665 positivi nella provincia di Roma, dato aggiornato al 15 aprile). Senza considerare altri fattori, sarebbe quindi lecito aspettarsi che il rapporto tra il numero di decessi nel Comune di Roma rispetto al numero di decessi in Provincia di Roma sia più vicino al 47% piuttosto che al 66%.

Passo 3 – Dal comunale alla popolazione 65+ anni

Partendo dalle informazioni disponibili nel “Report sulle caratteristiche dei pazienti deceduti positivi a COVID-19 in Italia” curato e aggiornato bisettimanalmente dall'[ISS](#), si assume che il 90% dei decessi COVID-19 positivi abbia più di 65 anni. Si tratta di un'approssimazione semplificativa, anche perché questa percentuale potrebbe essere variabile nel tempo e disomogenea sul territorio nazionale.

Esempio di verifica della stima per il comune di Brescia

Nella seguente tabella si riportano i dati raccolti per la Regione Lombardia e la Provincia di Brescia che consentono di stimare il numero di decessi COVID-19 positivi comunali secondo la procedura precedentemente illustrata. Viene anche mostrato un confronto con i dati ufficiali forniti da ATS Brescia e calcolato l'errore come differenza tra la stima e i dati ufficiali.

Settimana	Decessi positivi Regione	Nuovi casi Prov.	Decessi positivi Provincia (stima)	Decessi positivi comune		Errore comunale (Stima - ATS)
		Nuovi casi Reg.		stima*	dati ATS BS	
29/02 - 06/03	-	8.1%	-	-	-	-
07/03 - 13/03	755	22.2%	61	10	32	-22
14/03 - 20/03	1659	23.1%	369	58	67	-9
21/03 - 27/03	2853	17.6%	659	103	75	28
28/03 - 03/04	2909	16.7%	512	80	65	15
04/04 - 10/04	1927	15.9%	322	50	50	0
Totali				301	289	12

*Pur essendo disponibile a partire dal 5 marzo il dato dei casi positivi per la città di Brescia, per mostrare i risultati ottenuti con la procedura classica - cioè utilizzando i soli dati della Protezione Civile - il numero di decessi positivi comunali è stato calcolato in base al criterio dell'anagrafe residenti, e cioè come il 15.7% dei decessi provinciali.

Nella tabella seguente si propone un confronto tra il moltiplicatore calcolato a partire dai dati ufficiali di ATS Brescia e quello calcolato a partire dai dati stimati, indicati come stima classica.

Settimana	SISMG excess	Decessi positivi 65+ (stima dai dati ATS BS)	Moltiplicatore "esatto"	Decessi positivi 65+ (stima classica)	Moltiplicatore "stimato"
07/03 - 13/03	53	29	1.8	9	5.9
14/03 - 20/03	95	60	1.6	52	1.8
21/03 - 27/03	119	68	1.8	93	1.3
28/03 - 03/04	118	59	2	72	1.6
Totali	385	216	1.8	226	1.7

Sebbene per la città di Brescia la stima classica possa ritenersi almeno complessivamente una buona approssimazione del dato ufficiale, per altri comuni l'approssimazione potrebbe risultare significativamente lontana

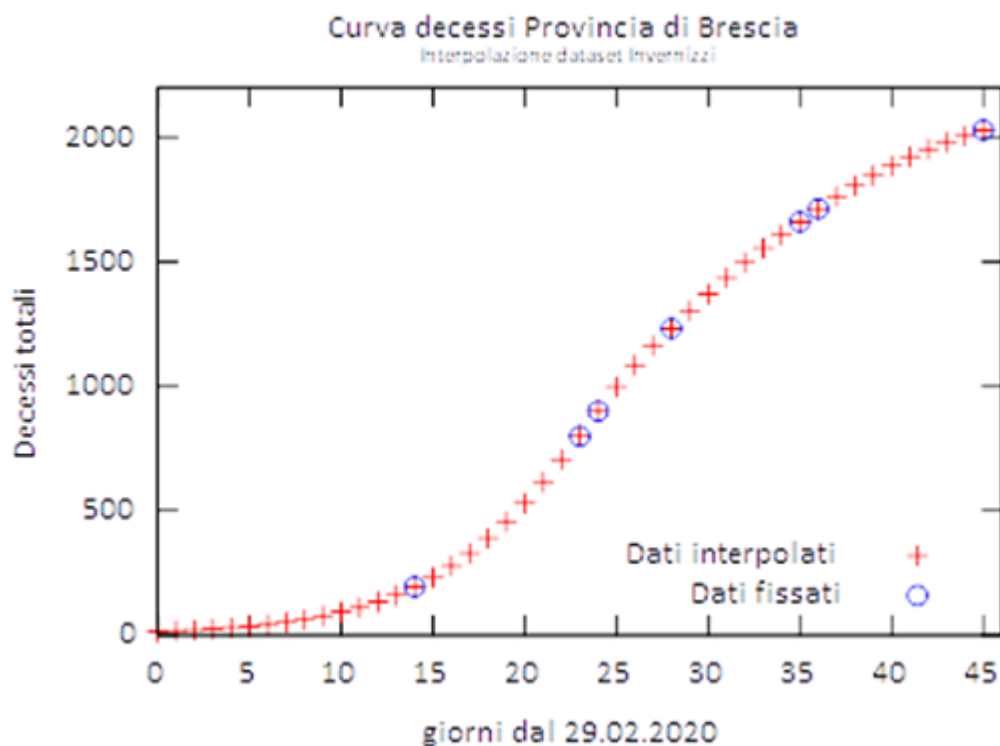
dai dati reali. E sebbene per la maggior parte dei comuni non sia disponibile un'informazione dettagliata e precisa come nel caso di Brescia, va segnalato che esiste una diffusa disponibilità di dati oltre a quelli della Protezione Civile che rendono possibile migliorare la stima finale. Nei paragrafi seguenti vengono mostrati due esempi a tal proposito.

Migliorare la stima: Lombardia

Per la Lombardia è disponibile un [dataset](#) che contiene informazioni sul numero di decessi provinciali, curato e aggiornato grazie al prezioso lavoro di documentazione e analisi dal giornalista dell'Eco di Bergamo [Isaia Invernizzi](#). Si ritiene in generale che la disponibilità di tali dati possa contribuire a migliorare significativamente la stima finale, e che la metodologia di seguito descritta possa ragionevolmente essere generalizzata a tutto il territorio lombardo.

Verifichiamo quindi come e quanto l'utilizzo del "dataset Invernizzi" consenta di migliorare la stima finale, sempre utilizzando come parametro di confronto il comune di Brescia per il quale si hanno a disposizione informazioni precise sul numero di decessi COVID-19 positivi.

Il dataset contiene al momento informazioni sul numero di decessi provinciali per sei giornate comprese tra il 14 marzo e il 14 aprile. A partire da questi punti fissati e attraverso una serie di interpolazioni è possibile costruire una curva che descriva l'andamento del numero totale di decessi provinciali negli intervalli di tempo compresi tra i giorni per cui è noto il numero totale di decessi. Il grafico seguente mostra una stima del numero di decessi totali di COVID-19 positivi per la provincia di Brescia ottenuta a partire dai punti fissi contenuti nel dataset.



Si precisa che tale curva ha un valore puramente descrittivo e non predittivo: se può essere utile per stimare il numero di decessi avvenuto tra il 14 marzo e il 14 aprile, non avrebbe molto senso utilizzarla per stimare il numero di decessi delle settimane a venire. Nella speranza che questo dataset continui ad essere aggiornato come lo è stato finora, si ritiene che la stima del numero di decessi provinciali COVID-19 positivi attraverso questi dati sia da preferire ad altre possibili tecniche di stima.

Nella tabella seguente la stima dei decessi settimanali della Provincia di Brescia calcolata utilizzando la procedura classica viene confrontata con quella ottenuta partendo dal dataset Invernizzi.

Settimana	Decessi Provincia di Brescia		Errore provinciale (Stima classica - stima dataset Inv.)
	(stima classica)	(stima da dataset Invernizzi)	
07/03 - 13/03	61	119	-58
14/03 - 20/03	369	369	0
21/03 - 27/03	659	632	27
28/03 - 03/04	512	450	62
04/04 - 10/04	322	320	2
Totali	1923	1890	33

Per concludere si propone un confronto sulla stima del moltiplicatore, confrontando il moltiplicatore "esatto", calcolato a partire dai dati di ATS Brescia e il moltiplicatore stimato, calcolato a partire dai dati presenti nel dataset di Invernizzi.

Settimana	SISMG excess	Decessi positivi 65+ stima dai dati ATS BS	Moltiplicatore "esatto"	Decessi positivi 65+ (stima dai decessi dataset Invernizzi)*	Moltiplicatore stimato (Inv)
07/03 - 13/03	53	29	1.8	17	3.1
14/03 - 20/03	95	60	1.6	52	1.8
21/03 - 27/03	119	68	1.8	89	1.3
28/03 - 03/04	118	59	2	63	1.9
Totali	385	216	1.8	221	1.7

*Pur essendo disponibile a partire dal 5 marzo il dato dei casi positivi per la città di Brescia per mostrare il contributo dei dati del dataset Invernizzi il numero di decessi positivi comunali a partire dal quale calcolare i decessi positivi della classe 65+ anni è stato calcolato in base al criterio dell'anagrafe residenti, e cioè come il 15.7% dei decessi provinciali.

Migliorare la stima: Piemonte, Puglia, Emilia Romagna

Alcune Regioni (tra le altre [Piemonte](#), [Puglia](#), [Emilia Romagna](#)) forniscono quotidianamente il dato relativo al numero di decessi a livello provinciale. Si propone quindi per la Provincia di Torino un confronto tra la stima classica e il dato ufficiale.

Settimana	Decessi Regione Piemonte	Nuovi casi Prov.	Decessi positivi Provincia		Errore Provinciale
		Nuovi casi Reg.	(stima)	(dati Regione Piemonte)	
29/02 - 06/03	-	17,40%	-	-	-
07/03 - 13/03	42	38.9%	7	12	-5
14/03 - 20/03	163	47.7%	63	37	26
21/03 - 27/03	360	49,70%	172	144	28
28/03 - 03/04	474	50.8%	236	223	13
04/04 - 10/04	489	46.9%	249	225	24
		Totali	727	641	86

Si evidenzia come la stima risulti complessivamente superiore ai dati ufficiali. Questo può essere legato al fatto che in Provincia di Torino l'età media dei casi positivi è inferiore rispetto alla media regionale [3]: a parità di casi positivi in Provincia di Torino si registra un numero inferiore di decessi rispetto alle altre province della regione.

Si ritiene che la disponibilità di informazioni ufficiali sul numero di decessi provinciali consenta di migliorare significativamente la stima finale, evitando le approssimazioni dovute al passaggio dal livello regionale a quello provinciale.

A proposito dell'importanza del livello di dettaglio dei dati va segnalato che proprio mentre veniva scritto questo articolo l'Associazione OnData "liberava" un [database](#) sui tamponi effettuati dalla Regione Lombardia. Una miniera di dati finalmente a disposizione di tutti.

Note

[1] A scanso di equivoci si ritiene che non si possa che far corrispondere automaticamente il numero di morti in eccesso rispetto alla media storica con il numero "reale" di decessi COVID-19 positivi. Da un lato come evidenziato tra gli altri

da Carra e Satolli su Scienza in Rete “è probabile infatti che anche altre malattie siano esitate in morte prima del tempo probabilmente anche a causa dell’intasamento dei pronto soccorso”. Dall’altro una parte dei decessi clinicamente correlabili al COVID-19 rientrano nella mortalità attesa. Se il Governatore della Liguria Toti probabilmente sbaglia quando afferma “escludo che ci siano morti di COVID-19 non censiti in Liguria”, ha almeno parzialmente ragione quando spiega che esistano “morti censiti come COVID-19 [che] probabilmente sarebbero morti ugualmente data la loro età”.

[\[2\]](#) Si segnala che quotidianamente una parte di casi positivi è privo di attribuzione provinciale (nei bollettini della Protezione Civile è indicato alla voce “Altro/In fase di verifica”): a voler essere precisi tale dato, soprattutto dove e quando appare significativo, andrebbe tenuto in considerazione.

[\[3\]](#) In base ai dati elaborati dal SEREMI (aggiornamento al 16 aprile) risulta ad esempio che in Provincia di Torino i contagiati con più di 40 anni siano l’82.5%, mentre ad Alessandria siano il 90.2%.