

modelli previsionali

Quell'incertezza dei dati scientifici

Il 30 aprile è finalmente comparso sul sito Epicentro un primo bollettino su Covid19 stilato dai ricercatori di Istituto Superiore di Sanità e Fondazione Kessler, seguito il 7 maggio da un secondo. È un segno di fiducia verso la società civile e la comunità scientifica, in controtendenza rispetto alla mancata divulgazione dell'irrealistico e controverso modello previsionale di ignota paternità (quello dei 150mila ricoverati in terapia intensiva l'8 giugno, per intenderci), invalidato altrove da stimati colleghi e che ha ispirato la fase 2 del Governo. Le regole della scienza sono semplici: i report scientifici devono essere precisi, chiari e basarsi su dati messi a disposizione della comunità scientifica. Altrimenti, la fiducia nelle conclusioni tratte e nella possibilità di utilizzarle si riduce. Purtroppo, va sottolineato come anche questi due rapporti siano costellati da imprecisioni e avventate conclusioni.

Nel primo rapporto, da una figura all'altra appaiono valori di R_t descritti «su dati al 27 aprile», o riferiti graficamente al 27 aprile, diversissimi fra loro: si passa da 0,84 a 0,22 in Molise, da 0,52 a 0,19 in Lombardia, da 0,63 a 0,3 nel Lazio, da 0,51 a 0,18 in Veneto. Il secondo rapporto chiarisce la questione. Il 30 aprile, il valore più alto è il «vero» R_t : è calcolato il 27 aprile ma si riferisce al 12 aprile; il 7 maggio, il valore R_t si riferisce al 19 aprile. Gli altri valori sono stimati. Questo è un problema piuttosto serio: i valori pubblicati e i loro limiti vanno infatti sempre descritti con accuratezza. Inoltre, un modello matematico offre una fotografia epidemiologica così vecchia solo se la raccolta dati è inefficace. In particolare, una conclusione cruciale non sembra valida alla luce dei dati presentati e delle caratteristiche della malattia. Si sostiene che il lockdown abbia «avuto un impatto nell'invertire l'andamento delle infezioni» perché «il picco dei casi per data di inizio sintomi» è «stato raggiunto qualche giorno dopo l'adozione delle misure» di lockdown. In realtà questo picco è tra il 12 e il 13 marzo, due giorni dopo l'inizio del lockdown. Con un intervallo seriale di 5-7 giorni ed altri due giorni dai sintomi all'autoisolamento, il picco sarebbe dovuto essere attorno al 18 marzo affinché queste conclusioni potessero esser valide. Tutto ciò denota un cortocircuito tra scienziati, che dovrebbero essere agnostici rispetto alle misure adottate, e politici.

Sorgono quindi molti dubbi. Perché vengono divulgati e usati rapporti con R_t “reali” vecchi di 18 giorni? Se così non fosse, quali sono i report realmente usati, e perché non vengono sottoposti al vaglio della comunità scientifica in tempo reale, depositandoli insieme a tutti i dati in un repository pubblico? Perché non si usa il valore R_t stimato ieri, o una settimana fa? Nel secondo rapporto si stima che già dal 7 maggio il valore R_t sia già 0 in ben 13 regioni. Perché questi dati previsionali apparentemente molto solidi non sono stati comunicati prima agli scienziati e poi alla popolazione intera? Perché nel rapporto previsionale sulla fase 2, redatto il 27 aprile, si è usato per simulare lo scenario apocalittico dei mesi estivi un R_t del 12 aprile, ben 21 giorni prima dell'inizio della fase 2?

I modelli matematici hanno limiti noti, ma l'innegabile vantaggio di fornire previsioni, non fotografie del passato. Altrimenti, diventano foglie di fico per ingiustificate restrizioni delle libertà costituzionali. Non «decidono gli scienziati», come spesso abbiamo sentito in questi giorni: decide il governo, purtroppo -alla luce di quanto abbiamo rilevato-, sulla base di traballanti report non vagliati dalla comunità scientifica.

* Università di Padova

** Danish Cancer Center

*** Emory University

© RIPRODUZIONE RISERVATA

Luca Scorrano*

Francesco Cecconi**

Guido Silvestri***