Coronavirus: facciamo un’indagine nazionale su un campione rappresentativo

written by Luca Ricolfi | 20 Marzo 2020

L’idea è venuta tempo fa a un mio amico viticoltore, che ha l’azienda in Friuli (si chiama Nicola Manferrari, le sue vigne sono a Borgo del Tiglio): perché non cerchiamo di capire quanti sono i contagiati dal coronavirus con un’indagine statistica?

Se si sottoponesse a tampone e (possibilmente) ad esame del sangue un campione rappresentativo della popolazione italiana, potremmo rapidamente sapere alcune cose fondamentali, che al momento non sappiamo ancora:

- quanti sono i contagiati totali (sintomatici, asintomatici, guariti)
- qual è la percentuale di asintomatici
- qual è la distribuzione per genere, età ed altre caratteristiche
- qual è, approssimativamente, il tasso di letalità.

Tutte informazioni preziose per contrastare la diffusione del virus.

Allora (appena tre settimane fa) l’idea mi parve impraticabile perché le stime che circolavano sul numero di contagiati erano troppo basse. I casi ufficiali erano circa 1000, e si pensava ancora che i casi totali potessero essere il triplo, magari anche il quadruplo o il quintuplo, ma certo non più di 10 mila. Diciamo che un numero verosimile poteva essere non lontano da 6000, il che significa 1 italiano su 10000.

Di qui la mia perplessità. Se il numero di contagiati è sconosciuto ma ancora molto basso (diciamo 1 caso su 10 mila,

giusto per fissare le idee), per poter stimare accuratamente il numero effettivo di contagiati nella popolazione complessiva occorrerebbe un campione enorme: con un campione di 10 mila casi potremmo anche non intercettarne nessuno (o averne appena 1, o 2), con un campione di 100 mila casi potremmo ragionevolmente attenderci di intercettarne una decina (con un errore atteso di circa 3 casi). Per poter sperare di intercettare anche solo un centinaio di casi occorrerebbe un campione di 1 milione di casi, che è decisamente troppo grande (anche perché sarebbe opportuno sottoporre a test tutti i soggetti nei medesimi giorni).

Ora però le cose sono cambiate, per quattro motivi.

1. Il numero di casi è molto più alto di tre settimane fa (al 20 marzo, i 1000 casi di inizio mese sono diventati circa 50 volte tanti).
2. Sono sempre più numerosi gli esperti che congetturano che i casi effettivi possano essere molto più numerosi (anche 10 volte tanti, se non di più) rispetto ai casi ufficiali.
3. La regola enunciata dal prof. Andrea Crisanti (lo studioso che ha guidato l'indagine di popolazione sul comune di Vo'), secondo cui l'ordine di grandezza del numero di asintomatici è 10 volte il numero di ospedalizzati, conduce a valutare in 200 mila il numero totale di casi positivi.
4. Una stima indipendente, condotta dalla Fondazione Hume con un approccio del tutto diverso, conduce a valutazioni del medesimo ordine di grandezza.

E allora il calcolo è presto fatto. Se riuscissimo a organizzare una rilevazione per fine mese, anche un campione relativamente piccolo (per esempio 50 mila casi: più o meno i tamponi che oggi si fanno in 3 giorni) potrebbe fornirci le informazioni che cerchiamo.

I 45 mila casi ufficiali attuali (20 marzo), infatti, tra la

fine del mese e i primi di aprile saranno diventati circa 150 mila, cui andrebbero aggiunti tutti gli asintomatici, che adottando la “regola Crisanti” (asintomatici = ospedalizzati x 10) potrebbero essere un po’ più di 400 mila.

Questo significa che, già nei primi giorni di aprile, potremmo avere a che fare con una popolazione di 5-600 mila contagiati su 60 milioni (l’1% della popolazione). Un campione di 20 mila casi ne intercetterebbe circa 200, un campione di 50 mila casi ne intercetterebbe circa 500.

Certo, resterebbe il problema di stratificare il campione per zona geografica, perché il numero di positivi per abitante è molto alto nel centro-nord, e molto più basso nel centro-sud. Ma è un problema risolvibile sovracampionando il centro-sud e sottocampionado il centro-nord, in modo da avere pressappoco il medesimo numero di casi in entrambi i territori.

Il vero problema è di estrarre un campione davvero rappresentativo, cosa che si può fare usando le liste elettorali, l’anagrafe tributaria o altra fonte amministrativa capace di raggiungere la quasi totalità della popolazione.

Insomma, il contagio è andato così avanti che l’idea di ripetere con un campione rappresentativo ciò che a Vo’ è stato fatto con la quasi totalità della popolazione sta diventando praticabile.

Non sappiamo quanti siano i contagiati, ma conoscerne l’ordine di grandezza (1 italiano su 100 ai primi di aprile) è più che sufficiente per pianificare un’indagine campionaria, che potrà finalmente fornire una risposta affidabile alle nostre domande iniziali: quanti sono effettivamente i contagiati, qual è il peso degli asintomatici, qual è la composizione per età, qual è il tasso di letalità.

Tutte cose che al momento non sappiamo, e che sarebbe molto meglio sapere. Come ha detto il prof. Crisanti nell’ultima intervista a “Italia Oggi” a proposito dell’idea di

un’indagine statistica nazionale: “potrebbe essere molto utile dopo il picco quando si tratterà di capire se è il caso di sollevare le misure restrittive”.

Il picco è previsto per martedì 24 marzo. Credo che – se c’è la volontà politica di farla – l’indagine potrebbe partire lunedì 30 marzo e terminare lunedì 6 aprile.

[20 marzo 2020]