

La sostituzione dei non rispondenti nei sistemi di sorveglianza: problema o risorsa?

11 settembre 2014 - La sostituzione dei non rispondenti è una procedura utilizzata da Passi, che, se non impostata e gestita correttamente, può alterare il processo di rilevazione e l'accuratezza dei risultati della sorveglianza.

Per verificare se si siano verificati effetti negativi sul funzionamento e l'affidabilità del sistema, dovuti all'adozione di questa procedura, è stato realizzato uno studio sui dati Passi, i cui risultati sono stati pubblicati ad aprile 2014 sulla rivista internazionale *Annals of Epidemiology*. [1] Trattandosi di un argomento tecnicamente complesso, è opportuno fare qualche premessa per facilitarne la comprensione.

La "non risposta"

La *non risposta* è un fenomeno che interessa tutte le indagini svolte su un campione di popolazione (nel caso di Passi le informazioni vengono raccolte tramite interviste telefoniche). Il campione è costituito da persone selezionate in modo probabilistico da una lista, che idealmente dovrebbe comprendere tutte le unità della popolazione in studio. I "non rispondenti" sono le persone, che per qualunque motivo non vengono intervistate. [2]

Questo termine, pertanto, comprende non solo chi si rifiuta di rispondere, ma anche coloro che non sono rintracciabili, o che vengono esclusi perché non hanno le caratteristiche che li rendono eleggibili, cioè idonei ad essere intervistati: ad esempio, nel caso di Passi, un'età non compresa tra quelle previste dal protocollo, un'insufficiente padronanza della lingua, la non disponibilità di un recapito telefonico, ecc. [3]

La *non risposta* determina una riduzione del numero complessivo di unità da cui si raccolgono informazioni, e quindi una minor precisione delle stime degli indicatori.

I diversi tipi di non risposta ed i loro effetti

Se la *non risposta* coinvolge in modo del tutto casuale le diverse unità del campione (*Missing Completely At Random*), l'accuratezza delle stime rimane invariata [4]. Tuttavia questa evenienza si verifica raramente, perché di solito la non risposta interessa in modo differente i diversi gruppi che compongono il campione, e questi, a loro volta, sono spesso associati alle variabili oggetto di studio.

Per esempio, il peso corporeo varia in funzione della classe di età, della Regione di appartenenza, del sesso, ecc.

In pratica, se la non risposta è maggiore tra i giovani (che in media hanno un peso corporeo più basso degli anziani), il peso medio stimato della popolazione risulterà superiore al reale. In generale, l'alterazione delle stime è tanto più importante quanto maggiore è la quota di non rispondenti. [5]

Può capitare però che, all'interno delle singole categorie, la non risposta avvenga in modo casuale (*Missing At Random*). Riprendendo l'esempio precedente, tra i giovani ci possono essere meno rispondenti perché è più difficile reperire il loro recapito telefonico, ma è possibile che, per quanto riguarda il peso, non ci siano differenze tra rispondenti e non rispondenti. In questo caso, le stime dell'indicatore all'interno dei gruppi non vengono alterate dalla non risposta [4].

In molte indagini questi gruppi sono individuabili prima del campionamento, perché alcune importanti informazioni sono già disponibili nelle liste anagrafiche (sesso, età, residenza). Di conseguenza, in base a queste caratteristiche è possibile suddividere il campione in strati omogenei, di dimensioni proporzionali ai corrispondenti strati nella popolazione; ad esempio, in Passi il campionamento prevede sei strati per ciascuna Asl (sesso e tre classi di età).

La stratificazione è un accorgimento che consente di compensare l'alterazione della struttura del

campione dovuta alla non risposta, in modo da correggere la distorsione delle stime di popolazione. Gli interventi correttivi abitualmente adottati sono: la pesatura, il sovra-campionamento e la sostituzione, che possono essere adottati da soli o in associazione.

1. La *pesatura* consiste nell'uso di coefficienti (*pesi*), che tengono in considerazione le dimensioni degli strati nel campione intervistato rispetto a quelli della popolazione: in pratica, se il numero di interviste in uno strato si riduce, il rispettivo peso aumenta, consentendo di aggiustare le stime. [6]
2. Con il *sovra-campionamento* si pianifica, prima della fase di raccolta dei dati, un aumento delle dimensioni di ciascuno strato in ciascuna area geografica, proporzionato alla non risposta attesa in quello specifico strato; in tal modo, alla fine, la numerosità complessiva e la composizione del campione intervistato si avvicinano a quanto previsto nel disegno dello studio. [7]
3. La *sostituzione*, detta "sul campo" perché avviene nella fase operativa dello studio, prevede che ogni unità del campione originario (*titolare*), che non si riesce ad intervistare, sia rimpiazzata da un'altra che ne prende il posto (*sostituto*).

A questo riguardo è importante specificare che il disegno dello studio può prevedere diverse modalità per selezionare e contattare i sostituti. In particolare, nelle indagini svolte tramite interviste faccia a faccia, più impegnative dal punto di vista organizzativo ed esecutivo, per motivi di praticità si ricorre talvolta a sostituti non selezionati con meccanismi probabilistici. Ad esempio, se l'unità campionaria è la famiglia, in caso di non risposta può venire sostituita da un'altra famiglia con caratteristiche simili [8]; oppure il non rispondente può essere sostituito da un vicino di casa, che appartenga, per sesso ed età, allo stesso strato del titolare. [9]

Diverso è il caso, se i sostituti sono selezionati e gestiti con gli stessi criteri e procedure dei titolari, come avviene in Passi, il cui protocollo prevede l'estrazione casuale dei sostituti dalla stessa lista e negli stessi strati dei titolari, e la predisposizione di elenchi di riserva, che individuano in modo univoco, per ciascun titolare, gli eventuali sostituti da intervistare, e in quale ordine. A questi elenchi gli intervistatori attingono secondo necessità, seguendo criteri e passaggi ben codificati, sotto la supervisione di un coordinatore, e registrando in un diario le procedure eseguite e il loro esito. [10]

Tornando a esaminare i possibili tipi di *non risposta*, può accadere che essa coinvolga in maggior misura alcune categorie, diverse e indipendenti dagli strati in cui è stato diviso il campione (*Not Missing At Random*). [4]

Se queste categorie sono associate alla variabile indagata, i meccanismi di aggiustamento basati sulla stratificazione diventano inefficaci.

Assumendo, ad esempio, che la mancanza di un recapito telefonico sia più frequente nelle persone con un basso livello di istruzione, e che questa variabile sia associata con un elevato peso corporeo, ci saranno meno persone obese nel campione intervistato, per cui la stima del peso medio si abbasserà, sia nei singoli strati sia nell'intera popolazione.

Essendo stata modificata la struttura non solo dell'intero campione ma anche degli strati che lo compongono, le procedure correttive sopra nominate (*pesatura*, *sovra-campionamento*, *sostituzione*) non sono in grado di compensare la conseguente distorsione delle stime. [4, 6]

Da qui deriva l'importanza di adottare tutte le misure idonee per ridurre la non risposta e intervistare il maggior numero possibile di titolari in ciascuno strato (adeguato disegno dello studio, esecuzione rigorosa delle procedure, verifiche della qualità della rilevazione, ecc.).

Il bias di sostituzione

La sostituzione offre dei vantaggi, in quanto consente di mantenere immutate le dimensioni del campione nonché la sua composizione demografica e geografica, in modo che anche i sottogruppi di piccole dimensioni conservino una numerosità sufficiente per analisi dedicate.

Tuttavia, il campione dei sostituti intervistati corre il rischio di risultare differente dal campione dei titolari intervistati. Ciò può capitare se – come sopra accennato – il meccanismo del rimpiazzo non è pianificato secondo criteri probabilistici, oppure se le procedure previste per la sostituzione, pur essendo corrette, non vengono seguite rigorosamente: ad esempio, l'intervistatore può essere indotto, per guadagnare tempo e risparmiare risorse, a utilizzare come rimpiazzo le persone più facilmente accessibili, riducendo il numero dei tentativi di contatto previsti dal protocollo in caso di non reperibilità. [11]

Le persone più facilmente accessibili possono avere caratteristiche differenti da quelle meno agevolmente rintracciabili; in tal caso, all'alterazione delle stime dovuta alla non risposta si aggiungono gli effetti legati all'imperfetto processo di sostituzione (*bias di sostituzione*). Se invece si riescono a evitare questi fenomeni, il *bias* di non risposta non viene aggravato. Per questo motivo si ritiene che, in linea teorica, non vi siano obiezioni all'uso della sostituzione e che l'utilità del suo impiego debba essere valutata empiricamente, caso per caso. [12]

Queste premesse servono a evidenziare che i *bias* di *non risposta* e di *sostituzione* sono concettualmente e praticamente distinti:

- per valutare il *bias di non risposta* sono necessari confronti tra il gruppo dei rispondenti e quello dei non rispondenti. In altre parole, bisogna effettuare un'ulteriore indagine *ad hoc*, cercando di ottenere informazioni anche sulle persone che non sono state intervistate, perché più difficilmente raggiungibili o meno disponibili all'intervista
- viceversa, per misurare gli effetti del *bias di sostituzione* è sufficiente utilizzare informazioni che il sistema di sorveglianza già raccoglie: si confrontano le stime ottenute nel gruppo dei sostituti con quelle dei titolari, che fungono da riferimento, in quanto non affette dal *bias di sostituzione*. Proprio questa valutazione è l'obiettivo dello studio qui presentato. [1]

Lo studio della sostituzione in Passi

Lo studio è stato eseguito sui dati Passi relativi al 2011: le persone selezionate nel campione originario (titolari) erano 37.252; le persone non rispondenti per vari motivi (rifiuto, indisponibilità del numero telefonico, non eleggibilità) sono state rimpiazzate da sostituti, sino a ottenere un numero di intervistati (37.162) praticamente uguale al numero originariamente programmato. La composizione demografica del campione finale (titolari intervistati + sostituti intervistati) era uguale a quella del campione dei titolari inizialmente selezionati.

Il tasso grezzo di non risposta è risultato pari al 25% nei titolari e al 32% nei sostituti; la differenza tra i due gruppi è attesa, in quanto tra i sostituti sono più rappresentati gli strati demografici e le aree geografiche con tasso di non risposta più elevato (una spiegazione più esauriente di questo concetto è esposta nell'articolo pubblicato su *Annals of Epidemiology* [1]).

Va precisato che, ai fini di questa analisi, sono utilizzati i tassi grezzi, comprendenti anche le unità non eleggibili, che vengono anch'esse sostituite. Per questo motivo, il tasso *grezzo* di non risposta sull'intero campione (titolari + sostituti), pari al 27%, è inferiore al tasso di non risposta *corretto* (17%), idoneo a valutare la qualità del processo di rilevazione.

A completamento dello studio, sono stati calcolati i valori di 11 variabili importanti indagate da Passi (caratteristiche socio-economiche, fattori di rischio comportamentali, indicatori dello stato di salute), sia nei sostituti sia nei titolari: sono stati osservati scostamenti tra le rispettive stime molto modesti in termini assoluti (0,6% al massimo per gli indicatori di tipo sanitario), in ogni caso senza differenze statisticamente significative.

In conclusione, anche ammettendo che le stime, sia dei titolari sia dei sostituti, siano affette da *bias di non risposta*, le analisi effettuate non evidenziano differenze tra i due gruppi e consentono di escludere che in Passi il *bias di sostituzione* abbia aggravato l'eventuale *bias di non risposta*.

Questo è un elemento cruciale, perché giustifica la fusione dei due gruppi (titolari e sostituti) in un solo *dataset* più ampio, su cui possono essere calcolati gli indicatori ed effettuate le analisi. In questo modo vengono mantenute immutate le dimensioni e la struttura del campione, rispetto a quanto originariamente pianificato: un vantaggio non indifferente per una sorveglianza con un disegno stratificato ed ampiamente articolato sul territorio, come Passi.

Questi risultati, inoltre, inducono a ritenere che le procedure di sostituzione previste dal protocollo sono state eseguite in modo sufficientemente corretto ed omogeneo, e indicano che il sistema di sorveglianza, per quanto riguarda questo aspetto del suo funzionamento, ha operato in modo soddisfacente.

Più in generale, lo studio fornisce un contributo alla discussione su questo metodo, poco noto e poco studiato [13], di cui non di rado vengono sottolineati gli aspetti problematici, sicuramente presenti, ma non insuperabili, come evidenzia l'esperienza di Passi.

Bibliografia

1. Baldissera S, Ferrante G, Quarchioni E, Minardi V, et al. Field substitution of nonresponders can maintain sample size and structure without altering survey estimates—the experience of the Italian behavioral risk factors surveillance system (PASSI). *Annals of Epidemiology* 24 (2014) 241-245
2. Stoop IAL. The Hunt for the Last Respondent, <http://www.scp.nl/dsresource?objectid=22020&type=org>; 2005
3. Baldissera S, Campostrini S, Binkin N, Minardi V, Minelli G, Ferrante G, et al. Features and initial assessment of the Italian Behavioral Risk Factor Surveillance System (PASSI), 2007-2008. *Prev Chronic Dis* 2011; 8(1):A24.
4. Cobben F. Nonresponse in Sample Surveys. *Methods for Analysis and Adjustment*. p. 10-12; 118-120 (2009) *Statistics Netherlands - The Hague* [<http://www.cbs.nl/NR/rdonlyres/2C300D9D-C65D-4B44-B7F3-377BB6CEA066/0/2009x11cobben.pdf>]
5. Philippens M, Billiet J. Monitoring and Evaluating Non-response Issues and Fieldwork Efforts in the European Social Survey. *Proceedings of the European Conference on Quality and Methodology in Official Statistics*, [http://www.europeansocialsurvey.org/index.php?option=com_docman&task=cat_view&gid=55&Itemid=80; 2004]
6. Chapman DW. Substitution for missing units. *Proceedings of the Survey Research Methods Section, American Statistical Association, 1982* [http://www.amstat.org/sections/srms/Proceedings/papers/1985_048.pdf]
7. Pickery J, Carton A. Oversampling in relation to differential regional response rates. *Survey Research Methods* 2008; 2(2): 83e92.
8. Yansaneh IS. Overview of sample design issues for household surveys in developing and transition countries, in: Department of Economic and Social Affairs. *Household Sample Surveys in Developing and Transition Countries*, p.30. (2005). United Nations Publication - New York [http://unstats.un.org/unsd/publication/seriesf/Seriesf_96e.pdf]
9. Gallus S, Muttarak R, Martínez-Sánchez JM, Zuccaro P, Colombo P, La Vecchia C. Smoking prevalence and smoking attributable mortality in Italy, 2010. *Prev Med*. 2011 Jun; 52(6):434-8.
10. Gruppo Tecnico di Coordinamento del Progetto di sperimentazione del "Sistema di Sorveglianza PASSI". *Sistema di Sorveglianza PASSI. Progressi delle Aziende Sanitarie per la Salute in Italia..* Roma: Istituto Superiore di Sanità; 2007. (Rapporti ISTISAN 07/30). [<http://www.iss.it/binary/publ/cont/07-30.1195128446.pdf>]
11. Demarest S, Gisle L, Van der Heyden J. Playing hard to get: field substitutions in health surveys. *Int J Public Health* 2007; 52(3):188e9.
12. Eurostat. Variance estimation methods in the European Union. *Monographs of official statistics. Official Publications of the European Communities*, p. 54-56 [http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CR-02-001/EN/KS-CR-02-001-EN.PDF]
13. Vehovar V. Field substitution and unit nonresponse. *Journal of Official Statistics* 1999; 15(2): 335e50.