

Appendice

La regressione logistica

Si prefigge di stimare o prevedere l'associazione tra una variabile di risposta con due livelli (dicotomica) e variabili indipendenti di diversa natura. La logistica consente pertanto di individuare i fattori che sono ipotizzati influenzare la probabilità del verificarsi di un dato evento o attributo, controllando per gli effetti di altre variabili in esame. Questa procedura di aggiustamento viene effettuata quando si ipotizza che la relazione tra il verificarsi dell'evento e il determinante studiato sia contemporaneamente legata ad un'altra variabile, la cui presenza distorce l'effetto dell'agente e può produrre sia una accentuazione che una riduzione dell'effetto del fattore di rischio studiato.

Dai dati PASSI 2006 si è osservato, ad esempio, che l'adesione al "5 a day" relativo al consumo giornaliero di frutta e verdura non risulta associato con alcun fattore (classi di età, sesso, istruzione e livello di attività fisica) se analizzati singolarmente ma, quando vengono studiati contemporaneamente in un modello logistico, si individua un'associazione positiva statisticamente significativa nelle donne rispetto agli uomini, per la classe 50-69 anni rispetto a quella 18-24 anni e per chi svolge una regolare attività fisica rispetto ai sedentari.

La misura di associazione tra il verificarsi del fenomeno studiato e la variabile considerata è data dall'*odds ratio* (OR), cioè dal rapporto tra la probabilità del verificarsi dell'evento tra gli esposti al fattore di rischio e la probabilità del verificarsi dell'evento tra i non esposti.

Monitoraggio

Per la valutazione della qualità del sistema di sorveglianza si utilizzano alcuni indicatori di processo forniti dal monitoraggio, ricavati dal sito internet di servizio della sorveglianza PASSI (www.passidati.it):

- Tasso di risposta
- Tasso di sostituzione
- Tasso di rifiuto
- Tasso di non reperibilità
- Tasso di eleggibilità "e"
- Distribuzione percentuale dei motivi di non eleggibilità
- Modalità di reperimento del numero telefonico
- Distribuzione delle interviste per orari/giorni

La tabella seguente mostra i valori dei principali indicatori del monitoraggio nella regione Emilia-Romagna, nelle singole AUSL e nel pool PASSI complessivo:

	Tasso di risposta	Tasso di sostituzione	Tasso di rifiuto	Tasso di non reperibilità	Tasso di eleggibilità
AUSL di Piacenza	84,4	13,8	11,3	2,5	95,2
AUSL di Parma	85,3	14,7	9,7	5	97,8
AUSL di Reggio Emilia	86,2	8,9	7,1	1,8	97,1
AUSL di Modena	84,4	15,6	13,1	2,5	94,8
AUSL di Bologna	87,4	11,6	5	6,6	97,3
AUSL di Imola	88,1	11,9	7,5	4,5	98,5
AUSL di Ferrara	88,1	12,6	10,4	2,2	93,3
AUSL di Ravenna	88,4	9,6	7	2,6	99,6
AUSL di Forlì	90,4	6,6	4,2	2,3	98,1
AUSL di Cesena	91,1	11,9	7,6	4,3	93
AUSL di Rimini	93,4	13,6	9,2	4,4	92,8
Regione Emilia-Romagna	87,6	12,4	8,8	3,6	95,7
Pool ASL	85	15	11	4,1	95,3

Per meglio comprendere il significato dei dati sopra indicati, si riportano alcune definizioni importanti e le descrizioni degli indicatori utilizzati:

- **Popolazione indagata:** persone residenti nell'AUSL, di età 18-69 anni, registrate nell'anagrafe sanitaria degli assistiti, presenti nel mese di indagine, che abbiano la disponibilità di un recapito telefonico e siano capaci di sostenere una conversazione in Italiano (o in altra lingua ufficiale della Regione/PA).
- **Eleggibilità:** si considerano eleggibili tutti gli individui campionati di età compresa tra 18 e 69 anni, residenti nel comune di riferimento per l'AUSL, in grado di sostenere un'intervista telefonica.
- **Non eleggibilità:** le persone non-eleggibili sono coloro che sono state campionate e quindi inserite nel diario dell'intervistatore, ma che successivamente sono state escluse dal campione per i motivi previsti dal protocollo, cioè residente altrove, senza telefono, istituzionalizzato, deceduto, non conoscenza della lingua italiana, grave disabilità, età minore di 18 oppure maggiore di 69 anni.
- **Non reperibilità:** si considerano non reperibili le persone di cui si ha il numero telefonico, ma per le quali non è stato possibile il contatto nonostante i 6 e più tentativi previsti dal protocollo (in orari e giorni della settimana diversi).
- **Rifiuto:** è prevista la possibilità che una persona eleggibile campionata non sia disponibile a collaborare rispondendo all'intervista, per cui deve essere registrata come un rifiuto e sostituita.
- **Senza telefono rintracciabile:** le persone che non sono in possesso di un recapito telefonico o di cui non è stato possibile rintracciare il numero di telefono seguendo tutte le procedure indicate dal protocollo.
- **Sostituzione:** coloro i quali rifiutano l'intervista o sono non reperibili devono essere sostituiti da un individuo campionato appartenente allo stesso strato (per sesso e classe di età).

• **Tasso di risposta**

Questo indicatore, che fa parte degli indicatori standard internazionali, misura la proporzione di persone intervistate su tutte le persone eleggibili (intervistati e non).

Il tasso grezzo di risposta sulla popolazione indagata è così espresso:

$$RRI = \frac{n^{\circ} \text{ interviste}}{n^{\circ} \text{ interviste} + \text{rifiuti} + \text{non reperibili}} * 100$$

Si tratta di un indicatore fondamentale, anche se generico, per valutare l'adesione all'indagine.

• **Tasso di sostituzione**

Questo indicatore, che fa parte degli indicatori standard internazionali, misura la proporzione di persone eleggibili sostituite per rifiuto o non reperibilità sul totale delle persone eleggibili, così indicato:

$$\frac{\text{non reperibili} + \text{rifiuti}}{n^{\circ} \text{ interviste} + \text{rifiuti} + \text{non reperibili}} * 100$$

Pur avendo i sostituiti lo stesso sesso e la stessa classe di età dei titolari, un numero troppo elevato di sostituzioni potrebbe ugualmente alterare la rappresentatività del campione.

• **Tasso di rifiuto**

Questo indicatore, che fa parte degli indicatori standard internazionali, misura la proporzione di persone che hanno rifiutato l'intervista su tutte le persone eleggibili:

$$REFI = \frac{\text{rifiuti}}{n^{\circ} \text{ interviste} + \text{rifiuti} + \text{non reperibili}} * 100$$

Nel caso in cui il tasso dovesse risultare troppo alto viene raccomandato di verificare:

- la percentuale di lettere ricevute (se è bassa si può tentare di ridurre i rifiuti attraverso una maggiore diffusione delle lettere);
- il grado di coinvolgimento del medico di medicina generale (se risulta poco interpellato si potrebbe tentare di ridurre i rifiuti coinvolgendo maggiormente il medico per convincere la persona a partecipare).

Nel monitoraggio si verifica inoltre che i rifiuti siano distribuiti uniformemente e che non ci siano realtà, quali regioni, AUSL, intervistatori, con livelli particolarmente elevati di rifiuti o grosse differenze tra uomini e donne e/o tra classi di età diverse.

• **Tasso di non reperibilità**

Questo indicatore misura la proporzione di persone eleggibili che non sono state raggiunte telefonicamente su tutte le persone eleggibili:

$$\frac{\text{non reperibili}}{\text{n}^\circ \text{int erviste} + \text{rifiuti} + \text{non reperibili}} * 100$$

Nel caso in cui il tasso risulti troppo alto viene raccomandato di verificare che i non reperibili si distribuiscano uniformemente e che non ci siano realtà (regioni, AUSL, intervistatori) con livelli particolarmente elevati di non reperibili o con grosse differenza tra uomini e donne e/o tra classi di età diverse.

• **Tasso di eleggibilità "e"**

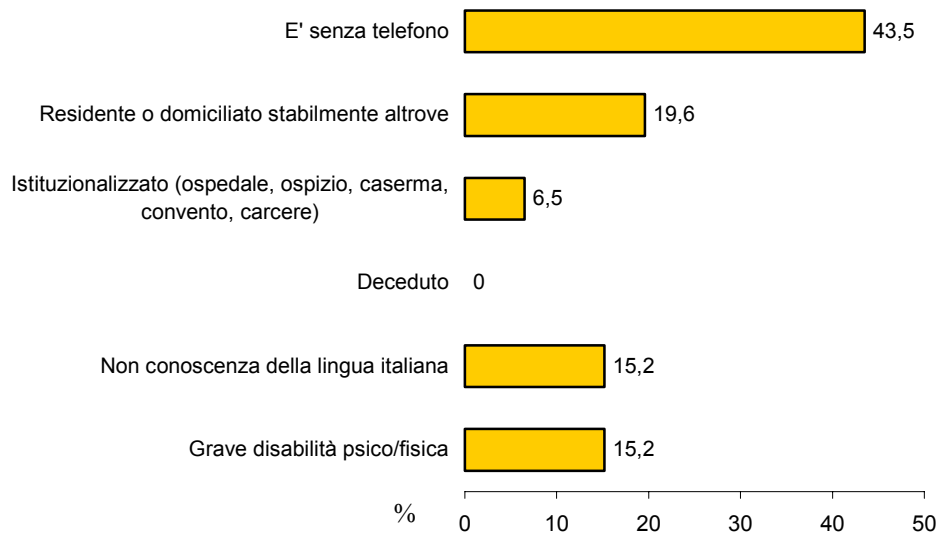
Questo indicatore misura la proporzione di persone eleggibili contattate sul totale delle persone di cui si ha un'informazione certa circa la condizione di eleggibilità. Per stimare gli eleggibili tra le persone che risultano "non reperibili" o "senza telefono rintracciabile" si moltiplica per "e" il numero di individui classificati in queste due categorie. Queste stime sono importanti per il calcolo di indicatori più complessi:

$$e = \frac{\text{n}^\circ \text{int erviste} + \text{rifiuti}}{\text{n}^\circ \text{int.} + \text{rif.} + \text{residenti altrove} + \text{istituz.} + \text{deceduti} + \text{no italiano} + \text{disabili} + \text{fuori range età}} * 100$$

• **Distribuzione percentuale dei motivi di non eleggibilità**

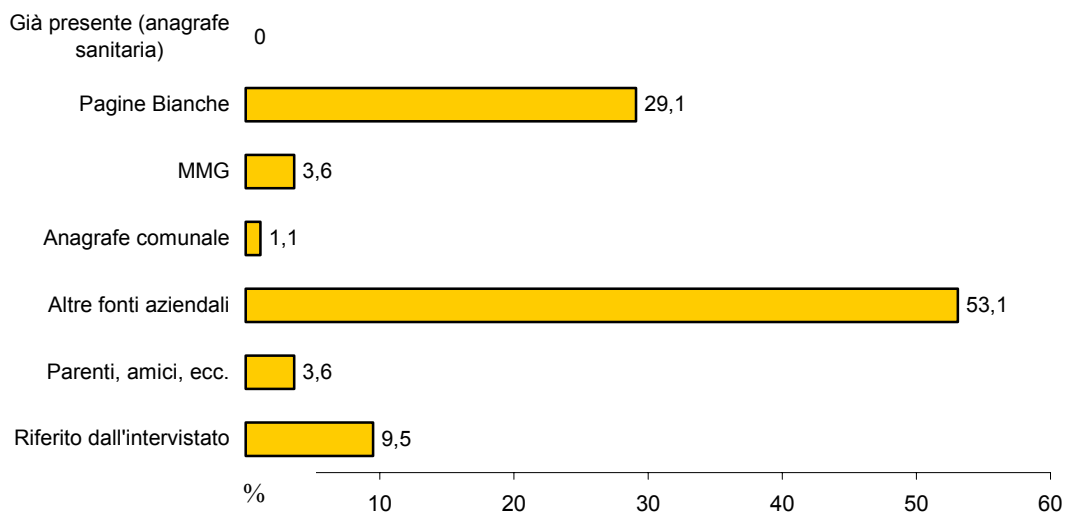
E' la distribuzione percentuale dei motivi che hanno portato all'esclusione dal campione di persone inizialmente campionate. In questo caso il rimpiazzo del non eleggibile non viene considerato una sostituzione vera e propria.

E' un indicatore che serve per verificare la qualità e l'aggiornamento dell'anagrafe da cui è stato fatto il campionamento (deceduti, cambi di residenza); la proporzione di persone che risulta "senza telefono rintracciabile", cioè che non sono in possesso di un recapito telefonico o di cui non è stato possibile rintracciare il numero di telefono seguendo tutte le procedure indicate dal protocollo e la presenza di altri motivi di esclusione.



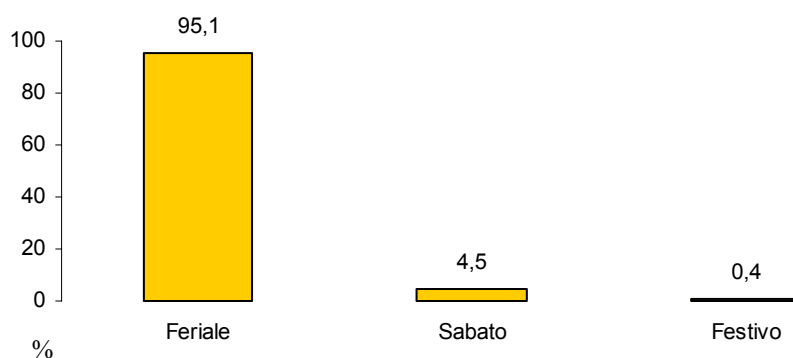
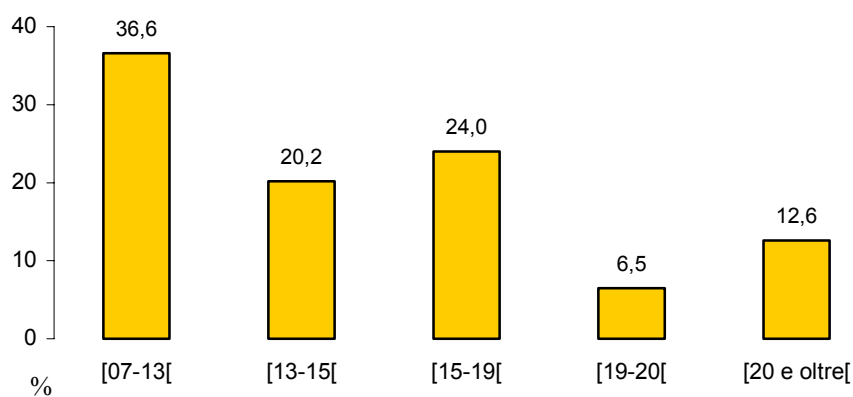
• Modalità di reperimento del numero telefonico

Questo indicatore riflette il lavoro associato al reperimento del numero di telefono. Assume una particolare importanza in caso di un'alta percentuale di "senza telefono rintracciabile". L'indicatore può variare molto da realtà a realtà per cui eventuali azioni correttive vanno contestualizzate alla situazione locale.



- Distribuzione delle interviste per orari/giorni**

La distribuzione oraria e settimanale delle interviste serve soprattutto a stimare la proporzione di interviste svolte in ore e/o giorni presumibilmente da considerare "fuori orario di lavoro" dell'intervistatore.



Utilizzo della pesatura

Il sistema di sorveglianza nasce soprattutto per fornire informazioni sulle condizioni di salute e gli stili di vita della popolazione a livello delle AUSL, quindi è stato effettuato un campionamento a rappresentatività aziendale. Il tipo di campionamento scelto per la sorveglianza PASSI è stratificato proporzionale per sesso e classi di età (18-34, 35-49, 50-69) direttamente effettuato sulle liste delle anagrafi sanitarie delle AUSL.

Al fine di un confronto tra i dati di ciascuna realtà locale a quella complessiva della regione di appartenenza, è importante ottenere delle stime anche a livello regionale, aggregando i dati delle singole AUSL, così come a livello complessivo è stato fatto per l'intero "pool PASSI".

L'analisi dei dati a livello regionale richiede perciò meccanismi complessi di controllo e pesatura dei dati. La pesatura migliora l'affidabilità delle stime, soprattutto per le variabili con forte eterogeneità interaziendale, e la procedura di pesatura influenza l'ampiezza degli intervalli di confidenza (garantisce la correttezza delle stime pur accettando, di solito, una minor precisione).

La modalità di pesatura è dipendente dal tipo di campionamento stabilito.

In considerazione del fatto che la sorveglianza PASSI ha utilizzato un campionamento stratificato proporzionale, i pesi sono strato-dipendenti, quindi calcolati uno per ogni singolo strato, perciò ogni AUSL avrà sei valori di peso.

La variabile peso rappresenta quanto il singolo strato di AUSL "pesa" sul campione aggregato di regione. Ad ogni intervista andrà associato il peso relativo allo strato di appartenenza dell'individuo intervistato.

Sono stati calcolati due diversi pesi, uno che riporta i dati all'universo di riferimento ("*Peso1*") e un altro che invece mantiene la numerosità campionaria ("*Peso2*").

Il "*Peso1*" è dato dal rapporto tra la proporzione di popolazione ${}_i P_k$ (prendendo come riferimento quella ISTAT al 31/12/2006) dello strato k-esimo della AUSL i-esima rispetto alla regione di appartenenza e la proporzione delle interviste effettivamente svolte ${}_i \hat{p}_k$ in quel dato periodo dello strato k-esimo dell'AUSL i-esima rispetto a quelle svolte nell'intera regione, formalmente:

$${}_i \text{Peso1}_k = \frac{{}_i P_k}{{}_i \hat{p}_k}$$

$${}_i \hat{P}_k = \frac{\text{pop strato}_k \text{ ASL}_i}{\text{pop strato}_k \text{ Reg}}$$

e

$${}_i \hat{p}_k = \frac{\text{numero int strato}_k \text{ ASL}_i}{\text{numero int strato}_k \text{ Reg}}$$

Il "*Peso2*" è l'inverso della frazione campionaria, dato dal rapporto tra la popolazione ISTAT della i-esima AUSL dello strato k e il numero di interviste della i-esima AUSL dello strato k, formalmente:

$${}_i \text{Peso2}_k = \frac{\text{pop strato}_k \text{ ASL}_i}{\text{numero int strato}_k \text{ ASL}_i}$$

Il "*Peso1*" è stato impiegato nella regressione logistica multivariata mentre il "*Peso2*" è stato utilizzato nell'analisi univariata.

Per quelle sezioni del rapporto PASSI 2007 in cui il target di popolazione analizzato è relativo ad età differenti da quelle sopra citate, quali screening con Pap-test (25-64 anni), vaccinazione influenzale (18-64 anni) e carta del rischio cardiovascolare (40-69 anni), sono stati ricalcolati appositamente entrambi i pesi per le rispettive classi.

Anche per i dati delle singole Aziende Unità Sanitarie Locali sono stati utilizzati i meccanismi complessi di controllo e pesatura, allo scopo di tener conto dei piccoli disallineamenti tra il campione effettivo e quello atteso legati ad arrotondamenti nell'assegnazione delle interviste ai sei strati per sesso e classi di età. Inoltre la pesatura a livello di AUSL è imposta per ottenere stime corrette per l'AUSL di Bologna ed Imola che vengono considerate assieme.

Come il dato regionale deriva da una sintesi pesata delle varie AUSL appartenenti alla Regione, così i valori per l'intero Pool PASSI 2007 sono il risultato di un'aggregazione di tutte le Aziende Sanitarie Locali partecipanti alla sorveglianza PASSI (che hanno raggiunto un livello minimo di rappresentatività), utilizzando le stesse procedure impiegate a livello regionale ed aziendale.