

MARIA ELENA TOSTI

Introduzione	129
Analisi della sopravvivenza	129
• <i>La life-table</i> • <i>Il metodo di Kaplan-Meier</i> • <i>Altre considerazioni</i>	
Analisi basata sugli anni-persona	135
• <i>Vantaggi dell'utilizzo degli anni-persona</i> • <i>Svantaggi dell'utilizzo degli anni-persona</i>	

Introduzione

In questo capitolo verranno presentati due diversi approcci per analizzare le informazioni provenienti da uno studio longitudinale.* Gli stessi metodi possono essere applicati sia a disegni osservazionali, come lo studio osservazionale di coorte, sia a disegni sperimentali, come i trial clinici randomizzati; i trial, infatti, al di là della randomizzazione, sono basati sul follow-up dei soggetti arruolati e rientrano quindi pienamente negli studi di coorte.

Nella conduzione di uno studio che prevede il follow-up, soprattutto se osservazionale, generalmente non tutti i soggetti vengono arruolati nello stesso momento, quindi il primo passo da compiere è quello di riportare, in maniera fittizia, il punto in cui ogni soggetto è stato arruolato a un punto "zero" comune, che coincida con l'inizio dello studio (figura 10.1).

Una volta stabilito per tutti i pazienti inclusi nello studio un momento di arruolamento "fittizio" comune, il passo successivo, all'interno dell'analisi di uno studio longitudinale, è rappresentato dalla stima del rischio di evento. Il concetto di rischio può essere espresso attraverso una "probabilità di evento" o un "tasso di evento". Questi due diversi approcci per ottenere stime delle misure di rischio in esposti e non esposti portano a due ti-

pi di analisi diverse: a) analisi della sopravvivenza (rischio espresso come probabilità); b) utilizzo degli anni-persona come denominatore delle misure di rischio (rischio espresso come tasso di incidenza).

In entrambi i casi, il numeratore delle misure di frequenza dell'evento è costituito dal numero di eventi osservati nella popolazione in studio, durante il tempo di osservazione; quello che cambia è il metodo attraverso il quale si ricava il denominatore.

I successivi paragrafi mostreranno, nel dettaglio, i due diversi approcci all'analisi di uno studio di coorte.

Analisi della sopravvivenza

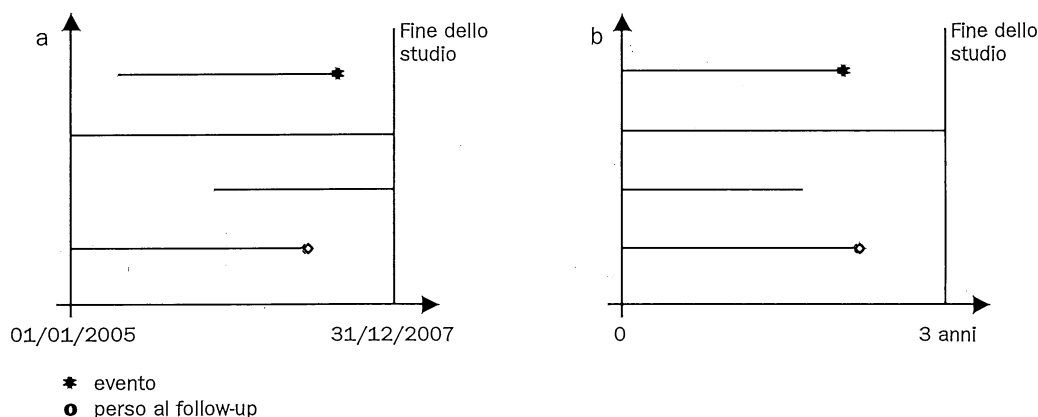
L'analisi della sopravvivenza prevede la valutazione della probabilità dell'evento in studio nel tempo. Le prime applicazioni di questo metodo furono fatte, negli anni Quaranta, su dati di mortalità; da qui il termine "analisi della sopravvivenza", che continua a essere utilizzato anche quando l'evento di interesse non sia la morte (ad esempio, guarigione, recidiva, ecc.).

Come evidenziato nell'introduzione, il primo passo da compiere è quello di riportare tutti i soggetti in studio a un punto "fittizio" di

* Studio che prevede il follow-up della popolazione arruolata.

FIGURA 10.1

Esempio ipotetico di quattro soggetti arruolati nell'arco di tempo di tre anni, 1° gennaio 2005-31 dicembre 2007.



Nella figura 10.1a è rappresentato l'arruolamento e il follow-up di una ipotetica coorte di quattro soggetti; ogni linea parte da un punto rappresentato dalla data in cui è avvenuto l'arruolamento; l'asse delle ascisse raffigura, quindi, il tempo di calendario.

Nella figura 10.1b, il follow-up dei soggetti appartenenti alla coorte è rappresentato su un'altra scala temporale che indica la lunghezza del follow-up (e non più le date reali di arruolamento e uscita dallo studio); in questo modo il follow-up di ogni soggetto inizia in corrispondenza di un'origine comune.

arruolamento che sia comune all'intera coorte. Il passo successivo consiste nello studio dell'andamento della probabilità di evento in base al tempo trascorso dall'inizio del follow-up. Questa valutazione dell'andamento delle probabilità avviene sia attraverso tabelle sia attraverso grafici nei quali vengono rappresentate le curve che descrivono l'andamento della probabilità: le curve di sopravvivenza.

Esistono due metodi per la costruzione delle curve di sopravvivenza: il metodo delle "life-table"¹ e il metodo di Kaplan-Meier.²

LA LIFE-TABLE

Il metodo delle life-table (o metodo classico) è stato il primo a essere impiegato per la costruzione delle curve di sopravvivenza. Le procedure che contraddistinguono questo metodo sono le seguenti: l'intero arco dello studio viene suddiviso in intervalli e, per ogni intervallo, viene calcolata la probabilità dell'evento studiato attraverso il rapporto tra il numero degli eventi osservati in quell'intervallo e i soggetti a rischio di sviluppare l'evento. Il gruppo di soggetti a rischio è costituito dai soggetti presenti all'inizio dell'inter-

vallo a cui va sottratta metà dei censored (soggetti con follow-up troncato, vedi capitolo 7) che si sono avuti nell'intervallo in esame. Si ipotizza, infatti, che ogni soggetto uscito dall'osservazione in quell'intervallo di tempo sia rimasto nello studio in media metà del tempo che costituisce l'intervallo. Per ogni intervallo, la probabilità di non evento viene calcolata come $1 - \text{probabilità di evento}$. La probabilità cumulativa di non evento, ossia la sopravvivenza cumulativa, viene calcolata attraverso il prodotto della probabilità di sopravvivenza di quell'intervallo per la probabilità di sopravvivenza cumulativa relativa all'intervallo precedente. Il procedimento di calcolo è esplicitato attraverso l'esempio mostrato nella tabella 10.1.

La sopravvivenza cumulativa ottenuta attraverso una life-table viene rappresentata su un grafico, la *curva di sopravvivenza*, che si ottiene unendo i punti relativi alla probabilità cumulativa calcolata nei vari intervalli.

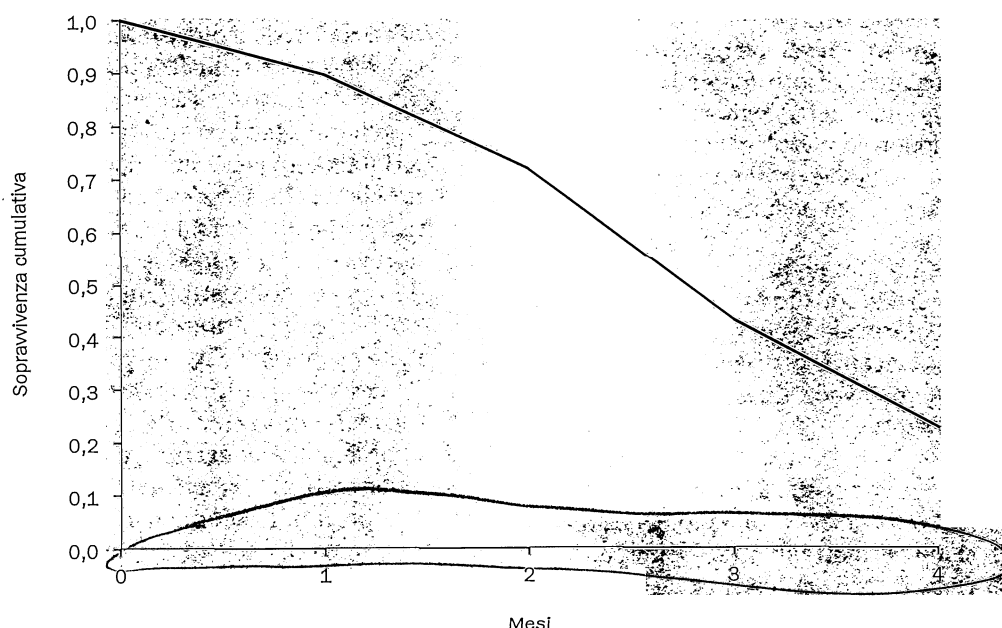
La figura 10.2 contiene la rappresentazione grafica della curva di sopravvivenza, calcolata con il metodo delle life-table, basata sui dati della tabella 10.1.

Tempo (intervalli)	N. a rischio all'inizio dell'intervallo	N. di eventi	Censored	Denominatore corretto per i C	Probabilità di evento	Probabilità di sopravvivenza	Sopravvivenza cumulativa
i	O_i	d_i	C_i	$O'_i = O_i - \frac{1}{2} C_i$	$q_i = d_i/O'_i$	$p_i = 1 - q_i$	S_i
1	100	10	8	96	0,10	0,90	0,90
2	82	15	12	76	0,20	0,80	0,72
3	55	20	10	50	0,40	0,60	0,43
4	25	9	12	19	0,47	0,53	0,23

TABELLA 10.1

Tabella di sopravvivenza costruita in base al metodo delle life-table.

Questo studio fittizio si svolge nell'arco di quattro mesi; per procedere con l'analisi, si suddivide l'intero periodo in quattro intervalli all'interno dei quali si "svolgono" i calcoli. All'inizio dello studio vi sono 100 persone in osservazione. Durante il primo intervallo di tempo si verificano 10 eventi (d_i) e 8 censored (C_i); il denominatore corretto (O'_i) diventa $100 - 8/2 = 96$ e la probabilità di evento (q_i) sarà data da $10/96 = 0,10$, mentre il suo complemento a 1 (probabilità di sopravvivenza, p_i) sarà $1 - 0,10 = 0,90$. La probabilità cumulativa di sopravvivenza (S_i) per il primo intervallo sarà così uguale a 0,90. Per i successivi intervalli si fanno le stesse operazioni tranne per il fatto che la probabilità cumulativa di sopravvivenza si ottiene attraverso il prodotto della sopravvivenza cumulativa dell'intervallo precedente moltiplicata per la sopravvivenza specifica di quell'intervallo ($S_2 = 0,90 \times 0,80 = 0,72$ per il secondo intervallo, mentre la sopravvivenza cumulativa del terzo intervallo sarà data da $0,72 \times 0,60 = 0,43$). A proposito delle due colonne relative alla probabilità di sopravvivenza vanno fatte alcune precisazioni: la probabilità p_i si chiama anche "probabilità condizionale di sopravvivenza" e rappresenta la sopravvivenza specifica dell'intervallo (è applicabile solo ai soggetti che erano in osservazione all'inizio di quell'intervallo); per quanto riguarda la probabilità cumulativa di sopravvivenza (S_i), essa invece stima la proporzione di soggetti che sopravvivranno alla fine dell'intervallo, tra tutti i soggetti che sono stati reclutati per lo studio.

**FIGURA 10.2**

Curva di sopravvivenza "disegnata" con il metodo delle life-table.

sistemata

La curva di sopravvivenza relativa a una "life-table" si ottiene collegando tra loro i punti che esprimono la probabilità di sopravvivenza cumulativa dei diversi intervalli. Il punto di partenza sarà quello relativo al tempo 0 e a una probabilità cumulativa di sopravvivenza = 1,0 poiché all'inizio dello studio nessuno dei soggetti ha ancora sperimentato l'evento.

Per costruire una life-table, evitando una distorsione dei risultati della stessa, è necessario che siano verificate alcune condizioni (assunzioni richieste per l'applicazione del procedimento illustrato in precedenza):

1. *I persi hanno la stessa probabilità di sopravvivenza dei soggetti rimasti in studio.* Questa assunzione è fondamentale per uno studio che prevede un follow-up e non solo per la costruzione delle curve di sopravvivenza; il fatto che questa assunzione sia vera è la base affinché i risultati ottenuti attraverso uno studio di coorte non siano distorti.
2. *Assenza di trend temporale.* Questa assunzione si riferisce alla non introduzione, nel corso dell'arruolamento dei soggetti e del successivo follow-up, di nuove procedure diagnostiche o terapeutiche della malattia in studio, che potrebbero influire sia sulla selezione dei partecipanti sia sull'esito. Il linfoma di Hodgkin e la sindrome da immunodeficienza acquisita (AIDS) sono due esempi molto chiari per illustrare questo punto. Entrambe queste malattie hanno cambiato il loro decorso grazie ai trattamenti introdotti negli anni, per cui non è ragionevole mettere insieme, ad esempio in uno studio di prognosi, pazienti reclutati nel corso di un lungo arco di tempo.

Queste due prime assunzioni sono indipendenti dal metodo di analisi. Le altre assunzioni necessarie per la conduzione di uno studio di sopravvivenza analizzata attraverso le life-table sono, invece, dipendenti dal metodo di analisi scelto.

3. *Gli eventi si equidistribuiscono all'interno di ogni intervallo.* Questa assunzione permette di unire i punti così come viene fatto nella costruzione delle curve di sopravvivenza ottenute attraverso le life-table. I punti sul grafico, che rappresentano le probabilità cumulative a fine intervallo, si uniscono attraverso una retta proprio per indicare che la sopravvivenza, all-

l'interno dell'intervallo, diminuisce in maniera graduale, ossia ipotizzando che gli eventi accadano con cadenza regolare.

Per comprendere il significato di questa assunzione si può ricorrere a un esempio. Si supponga di dover condurre uno studio sulla mortalità infantile; per condurre un'analisi di sopravvivenza (utilizzando il metodo delle life-table) sarà necessario dividere l'intero tempo dello studio (ad esempio, un anno) in intervalli all'interno dei quali si andrà poi a stimare la sopravvivenza. Nel primo anno di vita la mortalità sarà certamente disomogenea, cioè "alta" nel primo giorno, un po' più bassa durante la prima settimana e ancora decrescente fino alla fine del primo anno di vita (termine dello studio). Se si decidesse di dividere l'anno dello studio in intervalli di un mese ciascuno, nel primo mese si stimebbe una sopravvivenza "media" che non rappresenta bene né la probabilità di sopravvivenza del primo giorno di vita né quella degli ultimi giorni del mese; di qui la necessità di ricorrere a intervalli più stretti, all'interno dei quali la probabilità di sopravvivenza possa assumersi omogenea. In generale, per la scelta dell'ampiezza degli intervalli è necessario conoscere bene il problema che si sta studiando. Va sottolineato che gli intervalli non devono avere necessariamente la stessa ampiezza, ma devono essere costruiti in base all'assunzione di equidistribuzione degli eventi all'interno del singolo intervallo.

4. *Le perdite si equidistribuiscono all'interno di ogni intervallo.* Nel calcolo delle probabilità di evento, al denominatore viene sottratta la metà dei censored, assumendo una loro equidistribuzione, cioè che, in media, ogni individuo, poi "perso" durante l'intervallo, sia rimasto in osservazione per metà del periodo.

Nel caso in cui non si sia sicuri riguardo al soddisfacimento di queste ultime due assunzioni, si possono utilizzare degli intervalli più picco-

li, all'interno dei quali è più facile che queste condizioni siano rispettate. Ovviamente, per lavorare con intervalli più stretti è necessario avere a disposizione dati più dettagliati.

IL METODO DI KAPLAN-MEIER

L'altro metodo per ottenere le curve di sopravvivenza è quello di Kaplan e Meier.² La differenza principale rispetto alle life-table è che le probabilità calcolate non sono relative a un intervallo ma sono istantanee. In particolare, viene calcolata una probabilità di evento per ogni istante nel quale si è verificato almeno un evento. Questo secondo metodo è, attualmente, quello maggiormente utilizzato quando vi sia la possibilità di conoscere il momento esatto dell'occorrenza dell'evento, cosa che in genere è nota negli studi di prognosi.

Prendendo come riferimento un determinato istante, la probabilità di evento si ottiene attraverso il rapporto tra il numero degli

eventi che si sono verificati e i soggetti in osservazione. Dato che la probabilità è istantanea, in un determinato momento, un soggetto o è nello studio o è già uscito dallo studio, quindi non vi è il problema di dover calcolare un denominatore corretto per le perdite, come nel caso della life-table classica. Una volta calcolata la probabilità di evento, la sopravvivenza cumulativa si ottiene in maniera analoga a quanto visto a proposito delle life-table. La tabella 10.2 mostra un esempio di calcolo di una tabella di sopravvivenza secondo il metodo di Kaplan-Meier.

La rappresentazione grafica della sopravvivenza cumulativa stimata con il metodo di Kaplan-Meier (figura 10.3) è diversa dalla rappresentazione vista nella figura 10.2, a causa della diversa modalità di calcolo delle probabilità di evento.

Per quanto riguarda le assunzioni necessarie, non lavorando con intervalli di tempo

Tempo (giorni)	Soggetti a rischio	N. di eventi	Censored	Probabilità di evento	Probabilità di sopravvivenza	Sopravvivenza cumulativa
i	n_i	d_i	$C_{(i) - (i + 1)}$	$q_i = d_i / n_i$	$p_i = 1 - q_i$	S_i
0	20	0	0	$0/20 = 0,000$	1,000	1,000
15°	20	1	3	$1/20 = 0,050$	0,950	0,950
46°	16	2	0	$2/16 = 0,125$	0,875	0,831
83°	14	2	2	$2/14 = 0,143$	0,857	0,713
102°	10	1	3	$1/10 = 0,100$	0,900	0,641
155°	6	2	0	$2/6 = 0,333$	0,667	0,428
209°	4	1	3	$1/4 = 0,250$	0,750	0,321

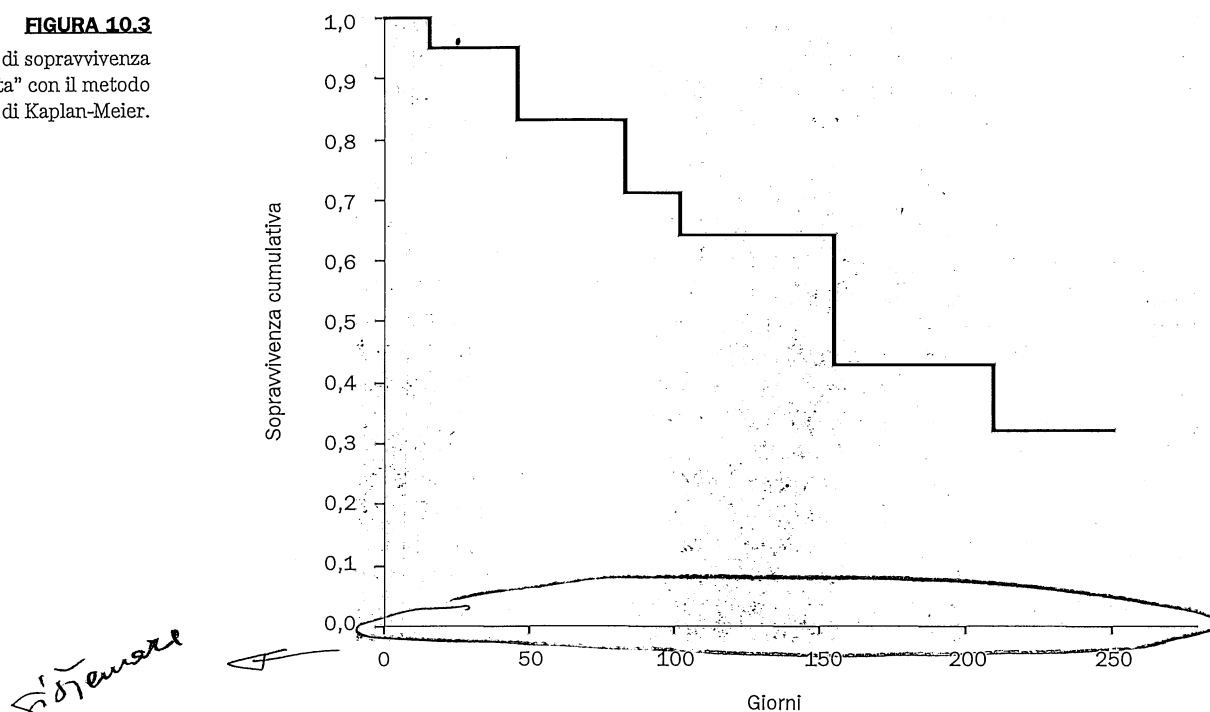
TABELLA 10.2

Tabella di sopravvivenza costruita in base al metodo di Kaplan-Meier.

La costruzione di una tabella di sopravvivenza fatta con il metodo di Kaplan-Meier è analoga a quanto visto con la life-table. La differenza sostanziale è che, col metodo di Kaplan-Meier, non esistono gli intervalli ma ogni probabilità stimata è "istantanea" (nella pratica come "istante", generalmente, si intende un giorno). Il punto di partenza è rappresentato dal tempo 0, in cui viene arruolata la coorte e non si hanno eventi. Nell'esempio fittizio di questa tabella il primo evento si verifica al 15° giorno, ed è proprio al 15° giorno che vengono calcolate le probabilità; le formule per il calcolo sono le stesse della life-table, tranne per il fatto che non c'è bisogno di ricorrere a un denominatore "aggiustato" per le perdite. Nel caso in cui vi sia concomitanza tra una perdita e un evento, si assume che l'evento sia avvenuto un attimo *prima della perdita*. Per quanto riguarda la colonna dei censored, va fatta una precisazione: questa colonna rappresenta i soggetti che sono usciti dal follow-up nel periodo che va dal momento a cui fa riferimento la riga dove si collocano i persi fino all'istante in cui avviene l'evento successivo e si ha quindi un nuovo calcolo della probabilità (riga successiva); per quanto detto a proposito del metodo di Kaplan-Meier questa colonna non è funzionale al calcolo delle probabilità (non esiste infatti un denominatore "corretto" per le perdite), come avviene nelle life-table, ma è semplicemente un "promemoria" per ricordarsi che questi soggetti non sono più presenti nel follow-up nel momento in cui si va a calcolare la successiva probabilità di sopravvivenza.

FIGURA 10.3

Curva di sopravvivenza
"disegnata" con il metodo
di Kaplan-Meier.



La curva avrà una caratteristica forma a gradini, in cui ogni gradino corrisponde al punto in cui si è verificato almeno un evento, ossia nel punto in cui si calcola la probabilità istantanea. Nell'intervallo di tempo compreso tra due eventi consecutivi si sa che non si sono verificati nuovi eventi e quindi la probabilità di sopravvivenza non si modifica. Da questo deriva la rappresentazione della sopravvivenza attraverso una linea orizzontale in tutti i periodi compresi tra due eventi consecutivi.

ma con probabilità istantanee, non sono richieste quelle sull'equidistribuzione degli eventi e delle perdite; mentre rimangono invariate le assunzioni che riguardano il destino dei soggetti con un follow-up troncato e l'assenza di trend temporale.

ALTRE CONSIDERAZIONI

In uno studio di coorte generalmente si confrontano gruppi di soggetti: esposti e non esposti (studio analitico osservazionale), pazienti con o senza un determinato fattore prognostico (studio di prognosi), terapia sperimentale e terapia di controllo (trial clinico). In questi casi, le curve di sopravvivenza che si rappresentano sono una per ogni gruppo e, oltre a un primo esame visivo, è necessario l'utilizzo di un test statistico per verificare se la sopravvivenza nei diversi gruppi è uguale o

diversa. Esistono numerosi test che possono essere utilizzati a questo scopo; il più frequente è il log-rank test, che mette a confronto gli eventi osservati nei diversi gruppi con quelli attesi sotto l'ipotesi nulla di mancanza di differenza tra i gruppi a confronto.⁸

Il log-rank test permette di confrontare le curve di sopravvivenza tenendo conto solo di un fattore alla volta, oppure stratificando attraverso la combinazione di un numero di fattori comunque limitato, affinché abbia poi senso il confronto. Invece, in presenza di diversi fattori in grado di influenzare la prognosi, è necessario ricorrere a un'analisi multivariata, che sia in grado di tener conto contemporaneamente di ciascuna variabile prognostica per indagare il suo effetto sulla prognosi, indipendente dalle altre. Il metodo statistico che si utilizza con i dati di sopravvivenza è la

Regressione di Cox o Modello dei Rischi Proporzionali. Si rimanda a testi specialistici per l'approfondimento di questi concetti.^{3 4}

Qualunque sia il metodo di calcolo, attraverso l'analisi della sopravvivenza quello che si stima è l'andamento, nel tempo, della probabilità di sopravvivenza, ossia della probabilità di "non evento". Si è detto nel paragrafo introduttivo che l'analisi della sopravvivenza è uno dei due metodi per il calcolo della frequenza dell'evento di interesse negli studi longitudinali: attraverso il complemento a 1 della sopravvivenza cumulativa si ottiene la probabilità cumulativa di evento.

Analisi basata sugli anni-persona

Il concetto di "tempo-persona" è già stato introdotto nel capitolo 2; in questo paragrafo verrà approfondito il suo utilizzo, evidenziandone vantaggi e limiti. Poiché l'unità di misura del tempo più ricorrente è l'anno, in questo paragrafo si utilizzerà a volte la denominazione "anni-persona" al posto di "tempo-persona".

Come si è detto, i soggetti arruolati in uno studio di coorte non vengono seguiti tutti per un periodo di tempo di uguale lunghezza: alcuni verranno arruolati in un momento successivo rispetto all'inizio dello studio, mentre altri termineranno l'osservazione prima del termine dello stesso. In questa situazione, il calcolo dei tassi di incidenza può essere fatto utilizzando due diversi approcci per ottenere il denominatore delle misure di frequenza:

1. si può ricorrere all'utilizzo del numero di persone in osservazione, ottenendo così un tasso che viene definito "tasso di incidenza cumulativo";
2. si può tener conto del tempo esatto in cui ogni soggetto è stato in osservazione: in questo modo il denominatore del tasso sarà costituito, non più dal numero di persone, ma dalla somma dei tempi di osservazione degli individui arruolati (ossia dalla somma dei tempi in cui, complessivamente, i soggetti arruolati sono stati "a ri-

schio" di sviluppare l'evento). Il tasso che si ottiene è definito "tasso medio di incidenza" o "densità di incidenza".

All'interno degli studi epidemiologici analitici il metodo corretto per il calcolo di un tasso di incidenza è il secondo, ossia l'utilizzo degli anni-persona come denominatore.

Un esempio di calcolo del tempo-persona è il seguente: un soggetto arruolato in uno studio il 01/01/2007 e seguito fino al 31/12/2008 contribuirà al denominatore del tasso per 2 anni-persona. La tabella 10.3 mostra un esempio di calcolo del tempo-persona e del conseguente tasso medio di incidenza.

Il tempo-persona può essere espresso in diverse unità di tempo (anni, mesi, giorni). La scelta dipende solo dalla malattia che si sta studiando e dalla frequenza (incidenza) con cui questa si presenta nella popolazione in studio: ad esempio, in caso di epidemie si utilizzeranno i mesi o anche i giorni, in caso di malattie rare si utilizzeranno gli anni. Come già detto, l'unità di tempo più frequentemente utilizzata per esprimere un tasso è l'anno: si parla infatti generalmente di tassi per anni-persona.

Così come accade per l'analisi della sopravvivenza, anche nel caso dell'analisi attraverso gli anni-persona è necessario che siano verificate alcune assunzioni:

1. *I persi hanno la stessa probabilità di sopravvivenza dei soggetti che sono rimasti in studio.*
2. *Assenza di trend temporale.*
3. *Il rischio è uniforme nel periodo considerato.* Questa assunzione è richiesta affinché il tasso medio, che si viene a calcolare con questo metodo, sia rappresentativo dell'intero periodo.
4. *Le perdite si equidistribuiscono all'interno del periodo considerato.* Questa assunzione è necessaria solamente nel caso in cui, non conoscendo il momento esatto di uscita dei soggetti dallo studio, il denominatore del tasso di incidenza si stimi attraverso la popolazione media del periodo.

TABELLA 10.3

Esempio del calcolo del "tempo-persona" in uno studio di coorte iniziato il 01/01/2007 e terminato il 31/12/2008.

N. soggetto	Arruolamento	Termine follow-up	Tempo di osservazione	Evento
1	01/01/2007	31/12/2008	24 mesi	no
2	01/01/2007	31/08/2008	20 mesi	no
3	01/02/2007	30/09/2007	8 mesi	sì
4	01/04/2007	31/12/2008	21 mesi	no
5	01/05/2007	30/09/2008	17 mesi	no
6	01/09/2007	31/12/2007	4 mesi	no
7	01/09/2007	31/12/2008	16 mesi	sì
8	01/01/2008	31/10/2008	10 mesi	no
9	01/05/2008	31/05/2008	1 mese	no
10	01/05/2008	31/12/2008	8 mesi	sì
Totale			129 mesi-persona o 10,75 anni-persona	

Nella tabella è indicato il periodo di follow-up di 10 soggetti arruolati all'interno di uno studio di coorte; i soggetti evidenziati in neretto sono quelli che hanno sperimentato l'evento in studio (ad esempio, insorgenza di malattia cardiovascolare). Complessivamente i 10 soggetti sono stati seguiti per 129 mesi (o 10,75 anni). Avendo osservato 3 eventi, il tasso di incidenza, in questo studio, sarà:

$$\text{Incid.} = \frac{3}{129 \text{ mesi}} = 2,3 \times 100 \text{ mesi-persona}$$

$$\text{Incid.} = \frac{3}{10,75 \text{ anni}} = 27,9 \times 100 \text{ anni-persona}$$

Le assunzioni viste finora sono analoghe a quelle necessarie per il calcolo della sopravvivenza attraverso le life-table. Un'ulteriore assunzione è specifica del metodo di analisi attraverso gli anni-persona:

5. *Il rischio deve aumentare proporzionalmente all'aumentare della lunghezza dell'osservazione.* Questa assunzione verrà approfondita nel paragrafo sugli svantaggi legati all'utilizzo di questo metodo.

VANTAGGI DELL'UTILIZZO DEGLI ANNI-PERSONA

Il principale vantaggio dell'utilizzo degli anni-persona è che, nel caso siano vere le assunzioni 3 e 4 (equidistribuzione degli eventi ed equidistribuzione delle perdite), il denominatore delle misure di incidenza calcolate corrisponde alla popolazione a metà del periodo di osservazione. Questo fa sì che i tassi ottenuti con questo metodo siano direttamente confrontabili con le statistiche correnti (vedi nota a p. 46). L'importanza di ciò sta nel fatto

che è possibile costruire uno studio epidemiologico analitico reclutando solo una popolazione di esposti, in quanto il confronto può essere fatto utilizzando, come popolazione di riferimento, i dati provenienti dalle statistiche correnti: ad esempio, la mortalità in un determinato gruppo a rischio può essere confrontata con la mortalità dell'intera popolazione, descritta dalle statistiche correnti.

Un altro vantaggio dell'utilizzo degli anni-persona è che esso consente di tener conto dei cambiamenti di stato, rispetto all'esposizione, durante il follow-up. Un esempio è rappresentato da uno studio sul rischio di infarto del miocardio nelle donne in post-menopausa: in questo caso l'esposizione è costituita dalla post-menopausa, mentre i periodi in cui la donna non era ancora in menopausa sono considerati la "non esposizione"; avendo utilizzato le stesse donne sia come esposte che come non esposte, i tassi nei due gruppi, pre- o post-menopausa, possono essere confrontati tra loro, tenendo però presente che i due gruppi possono avere diverse distribuzioni

per età (può essere quindi necessario procedere a un aggiustamento per età; vedi capitolo 11).

Un altro esempio di questo passaggio tra la categoria degli esposti e quella dei non esposti è rappresentato dagli studi di coorte sugli utilizzatori di specifici farmaci, i quali, in momenti diversi del follow-up, possono passare dalla categoria degli assuntori a quella dei non assuntori e viceversa; esempi di questo genere sono approfonditi nel capitolo 16 sulla farmacoepidemiologia.

In questi due esempi, gli eventi che si verificano vengono assegnati al gruppo degli esposti o a quello dei non esposti, a seconda che essi insorgano nel corso del tempo-persona di esposizione o del tempo-persona di non esposizione.

SVANTAGGI DELL'UTILIZZO DEGLI ANNI-PERSONA

Il principale svantaggio dell'utilizzo degli anni-persona deriva dall'assunzione che il rischio aumenti proporzionalmente all'aumentare dell'intervallo di tempo di osservazione.

Sebbene quella illustrata nella tabella 10.3 rappresenti la maniera corretta per ottenere il denominatore nel calcolo di un tasso di incidenza, essa presenta tuttavia un inconveniente: un soggetto seguito per 10 mesi dà lo stesso contributo al denominatore di 10 soggetti seguiti per un mese, cioè il rischio viene considerato uguale in ogni unità di tempo-persona. Ciò può essere non vero nel caso in

cui si studi un'esposizione che causi un rischio cumulativo (un soggetto che fuma da 20 anni ha un rischio maggiore di tumore ai polmoni rispetto a 20 persone che fumano da un anno), oppure quando solo un determinato periodo all'interno del follow-up risulti a rischio (in caso di infarto il periodo con un maggiore rischio di morte è quello immediatamente successivo all'evento, mentre nelle settimane o nei mesi successivi all'infarto un soggetto ha un rischio decrescente di morte).

Nel caso in cui non si sia sicuri della fondatezza dell'assunzione sull'aumento del rischio proporzionale all'aumento del tempo di osservazione (ad esempio, nel caso di un rischio cumulativo) è preferibile analizzare i dati utilizzando l'analisi della sopravvivenza piuttosto che gli anni-persona.

Bibliografia

1. Reed LJ, Merrell M. A short method for constructing an abridge life table. 1939. *Am J Epidemiol* 1995; 141 (11): 993-1022.
2. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Am J Epidemiol* 1958; 53: 457-81.
3. Marubini E, Valsecchi MG. Analysing survival data from clinical trial and observational studies. Chichester: John Wiley & Sons, 1998.
4. Kahn HA, Sempos CT. Statistical methods in epidemiology. New York: Oxford University Press, 1989.

