

Andamento temporale di incidenza, mortalità e sopravvivenza: esempi di analisi

Enzo Coviello
Registro Tumori Puglia
enzo.coviello@alice.it



Argomenti

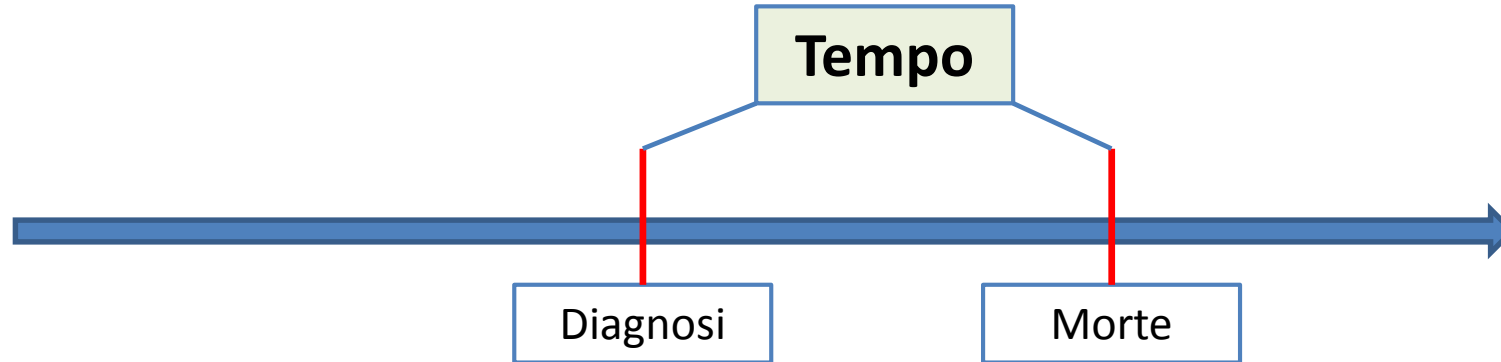
- Motivazione allo studio dei trend di 3 indicatori per valutare i progressi nella nostra capacità di contrasto al cancro
- Analizzare i trend di incidenza e mortalità
- Cenni all'analisi del trend della sopravvivenza



**Esempio di
analisi con
Stata**

Motivazione allo studio dei trend di 3 indicatori

Il ruolo del Tempo

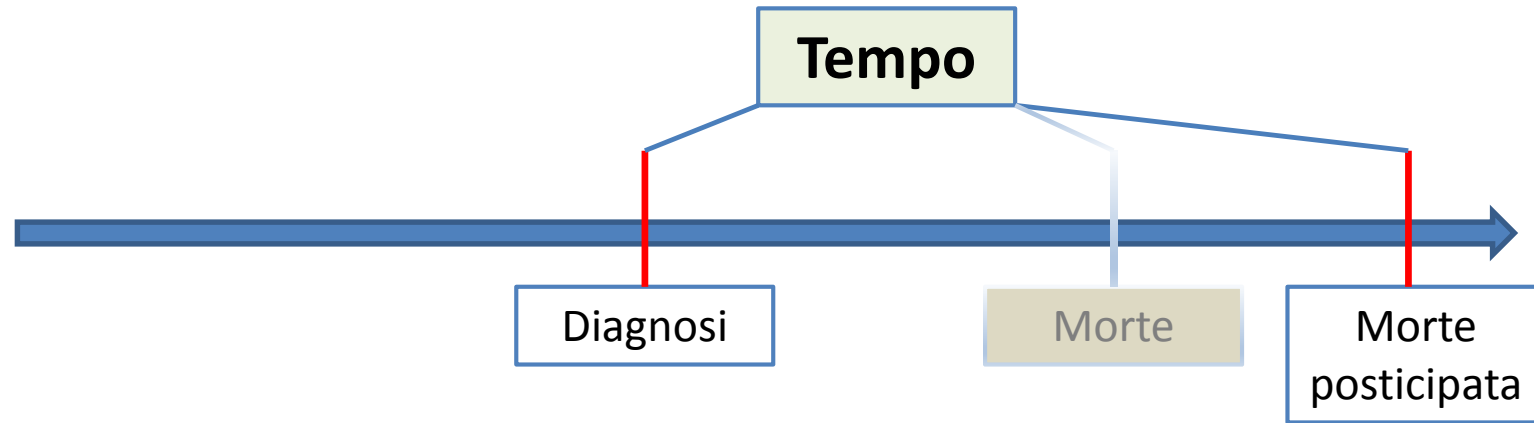


Per valutare i progressi nella lotta contro il cancro siamo interessati al tempo dalla diagnosi fino al momento del decesso.

L'analisi della sopravvivenza mette a disposizione una serie di metodi per studiare il tempo fino al verificarsi di un evento di interesse.

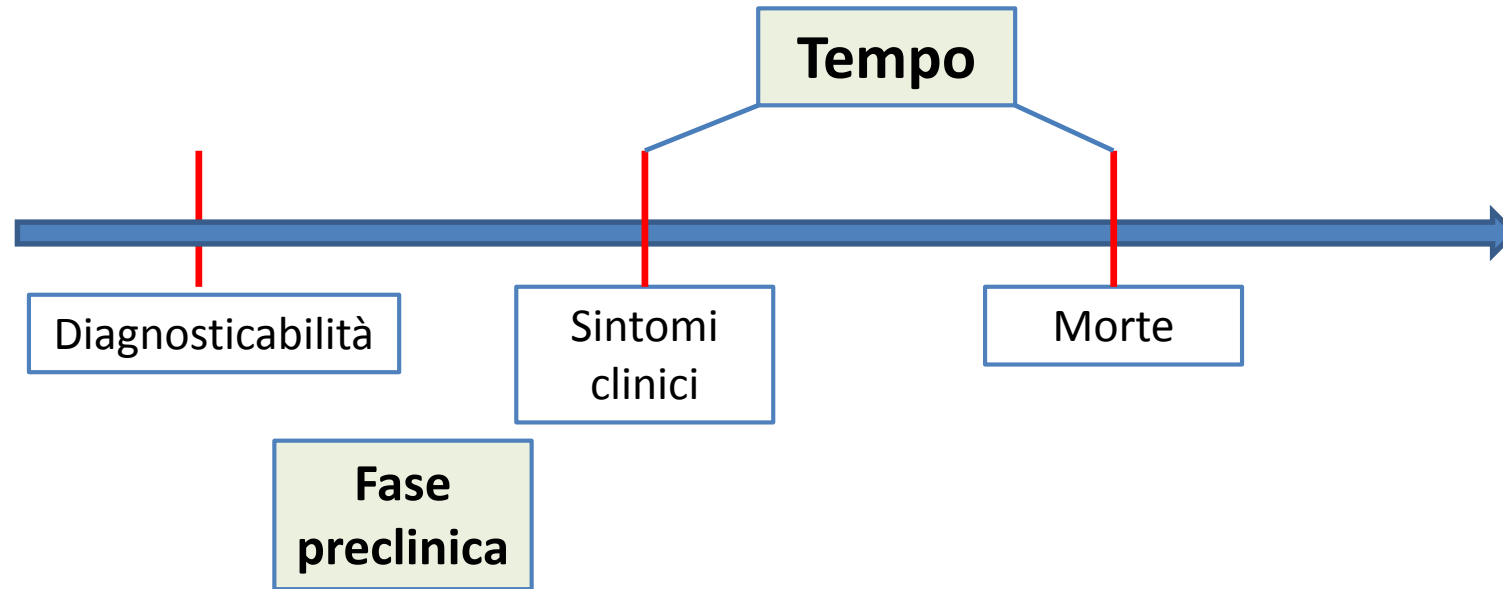
Tipici risultati di questa analisi sono **la stima della proporzione di persone vive decorso un certo tempo (probabilità di sopravvivenza)** e **la stima del tasso con cui i soggetti in osservazione muoiono (tasso di mortalità)**.

Indicatore di efficacia nel contrasto alla malattia



Studiamo la sopravvivenza per misurare la nostra capacità di allungare il tempo tra la diagnosi di una malattia (il tumore) e la morte (o un altro evento di interesse come recidive, metastasi, ecc.).

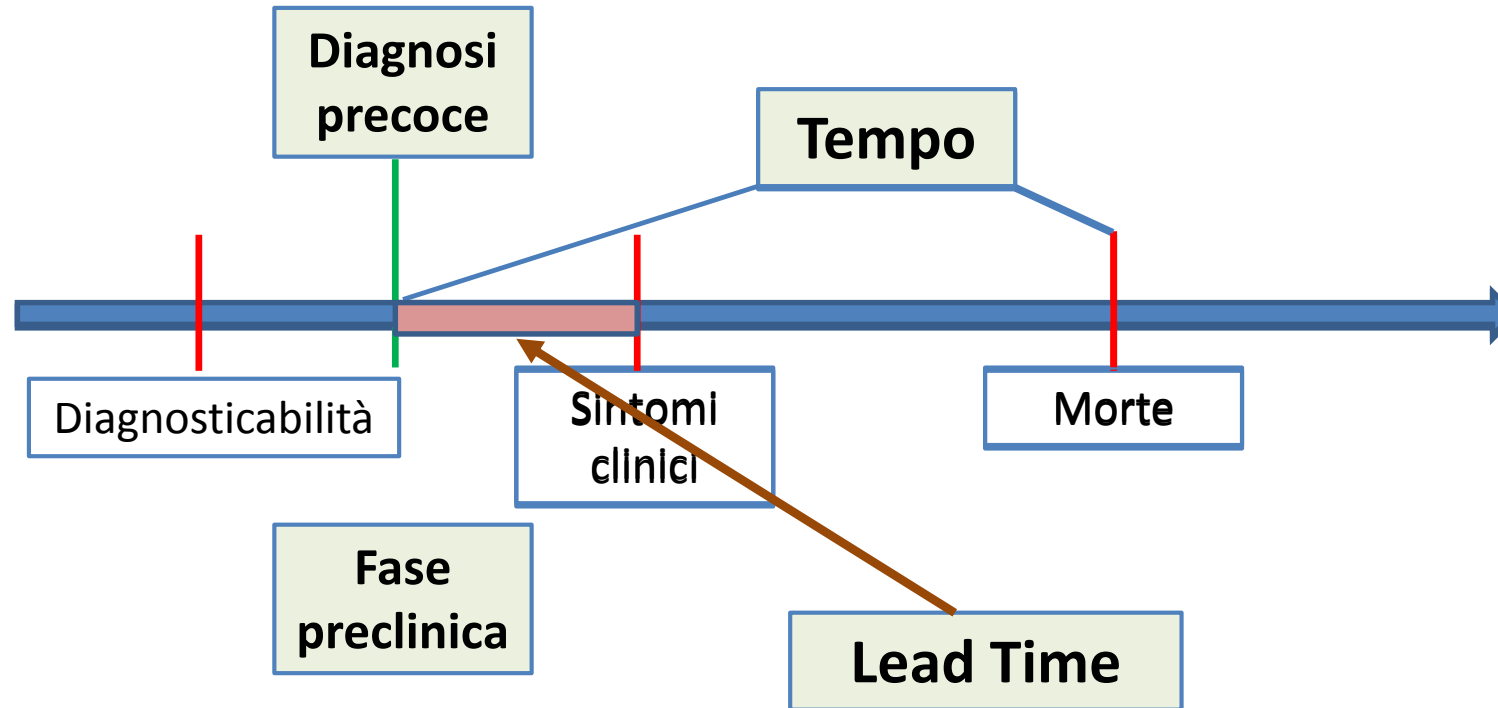
Momento della diagnosi



Il momento della diagnosi non è esattamente determinabile.

Può corrispondere all'insorgenza dei sintomi, ma può anche essere anticipato al momento dell'esecuzione di un'indagine in grado di rilevare la malattia.

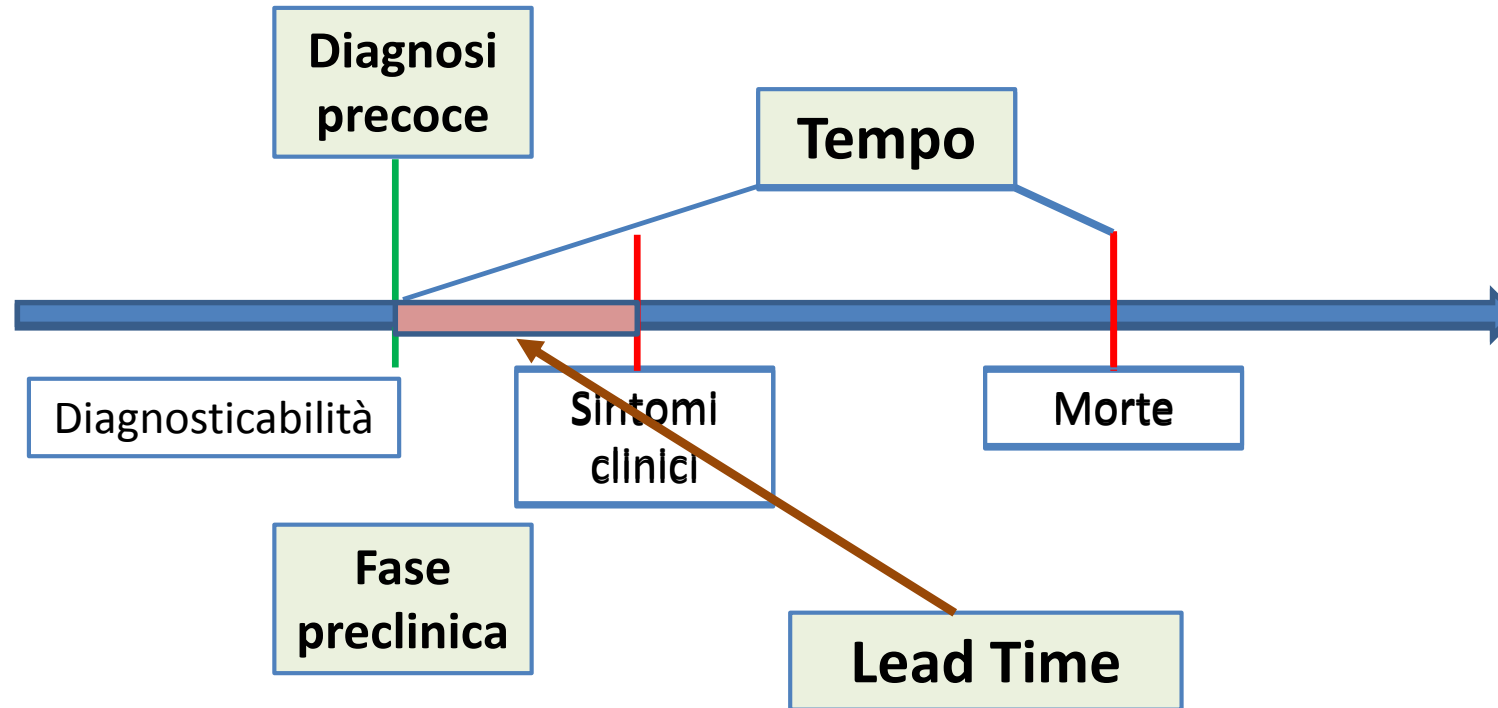
Lead time



Lo screening porta a un'anticipazione del momento della diagnosi del tumore rispetto al tempo in cui la malattia sarebbe stata riconosciuta per l'insorgenza dei sintomi.

La differenza tra i due tempi è detta **Lead Time**

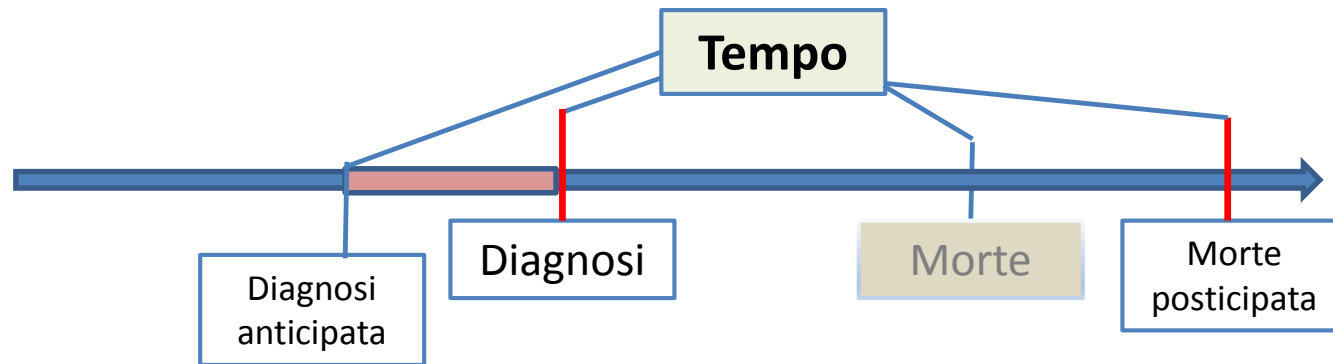
Non solo screening



L'anticipazione diagnostica in fase preclinica, anche al di fuori di uno screening, può verificarsi in diversi tumori come il tumore della prostata, il melanoma, il tumore della tiroide e in generale in quelle situazioni in cui aumentano le occasioni di contatto con il sistema sanitario.

Lead time bias o posticipazione della morte?

Lo studio della sopravvivenza porta spesso a interrogarsi sul fatto che le differenze osservate siano dovute a un reale miglioramento della prognosi (morte posticipata) e/o a un'anticipazione della diagnosi (lead time).



Una misura non ragionevole

Interpreting Cancer Survival Rates

The available data on survival are not a sensitive measure of progress in cancer control.

James E. Enstrom and Donald F. Austin

Science 04 Mar 1977:
Vol. 195, Issue 4281, pp. 847-851
DOI: [10.1126/science.841309](https://doi.org/10.1126/science.841309)

Dobbiamo riconsiderare la mortalità?

BMJ 2010;341:c4112

UK cancer survival statistics

Are misleading and make survival look worse than it is

RESEARCH, p 335

Valerie Beral professor of epidemiology, Cancer Epidemiology Unit, University of Oxford, Oxford OX 37LF

va.beral@ou.ox.ac.uk

Richard Peto professor of medical statistics and epidemiology, Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), University of Oxford, Oxford OX 37LF

Competing interests: Both authors have completed the Unified Competing Interest form at www.icmje.org/ucoi_disclosure.pdf (available on request from either

In the linked article, Auzier and colleagues report that (population based) breast cancer mortality rates have fallen over the past two decades in many European countries, with a greater decline in the United Kingdom than in any other large country.¹ That the UK is leading Europe in the speed with which national breast cancer mortality rates are falling is in stark contrast to, and at first sight difficult to reconcile with, claims that survival after breast cancer onset is worse in the UK than elsewhere in western Europe.²

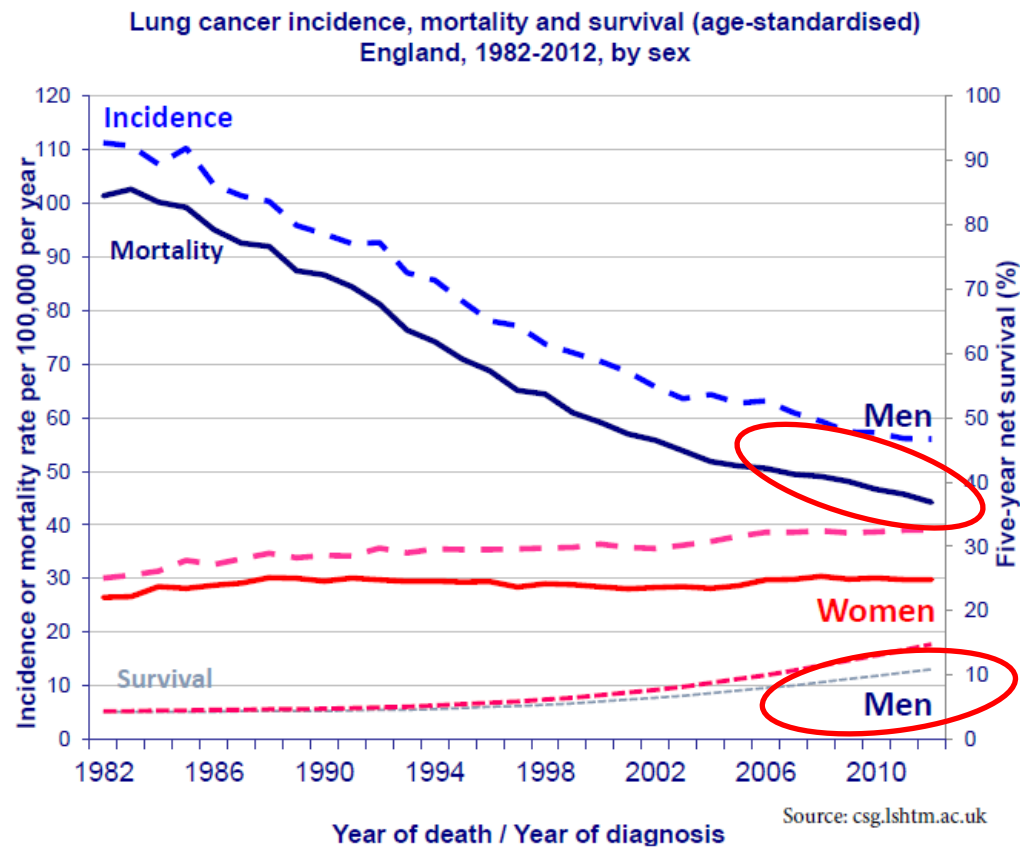
The unpromising UK cancer survival estimates are, however, misleading. In contrast, population based mortality trends are reasonably reliable (at least in middle age, for example, people aged 35-69 years) because a death certificate is legally required before someone can be buried

and vital calculations based on registry data make UK cancer survival rates seem significantly worse than they really are.

Information in cancer registries on deaths from cancer is virtually complete because every death certificate that mentions cancer is automatically sent to one of the regional registries that, between them, cover the UK. That cancer is then registered, and further information is sought (not always successfully) from medical records. Death certificates have for decades played an important role in the way UK registries identify people with cancer. Without this source of information, many such cancers could have been missed; even with it, many people who die of cancer may have no record other than the death certificate ever traced by the registry ("death certificate only" cases) or may have had only the last phase

... è preferibile comparare i trend dei tassi di mortalità nella popolazione di mezza età conosciuti in modo affidabile

Limite della mortalità come indicatore di progresso



La mortalità rapporta il numero dei casi a tutta la popolazione.

Di conseguenza si riduce se il numero di nuovi casi (incidenza) si riduce senza un rilevante miglioramento della sopravvivenza dei malati.

Incidenza Mortalità e Sopravvivenza

When Do Changes in Cancer Survival Mean Progress? The Insight From Population Incidence and Mortality

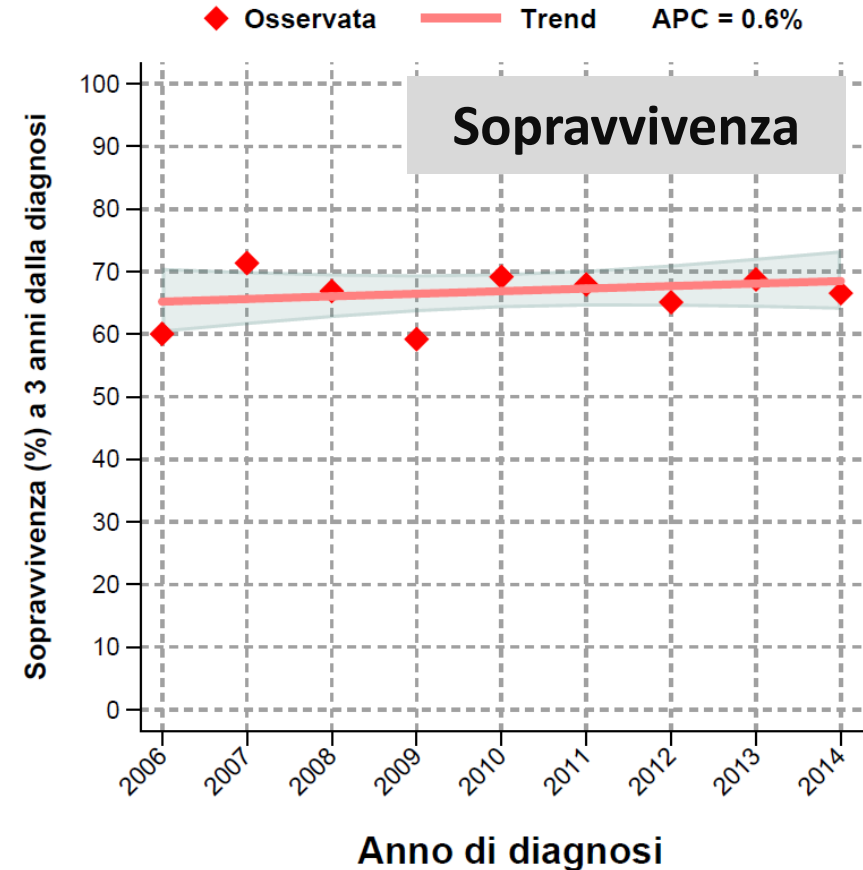
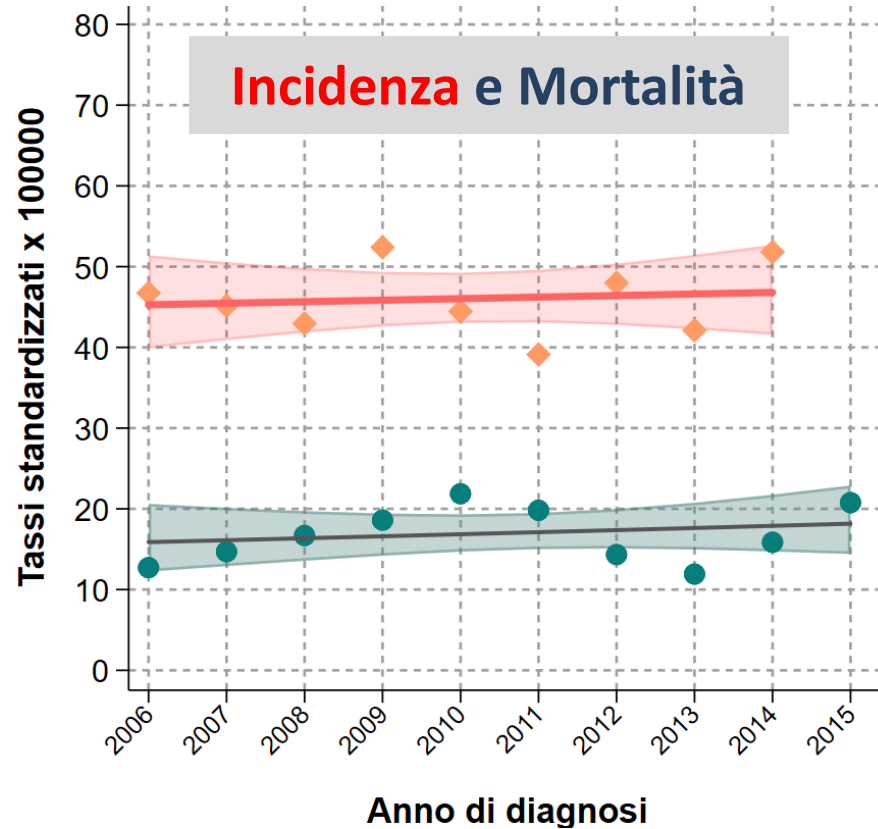
Hyunsoon Cho, Angela B. Mariotto, Lisa M. Schwartz, Jun Luo, Steven Woloshin

Conclusions

Changes in survival must be interpreted in the context of incidence and mortality. Increased survival only represents progress when accompanied by a reduction in incidence, mortality, or ideally both.

J Natl Cancer Inst Monogr 2014;49:187-197

Progressi nella lotta contro il cancro

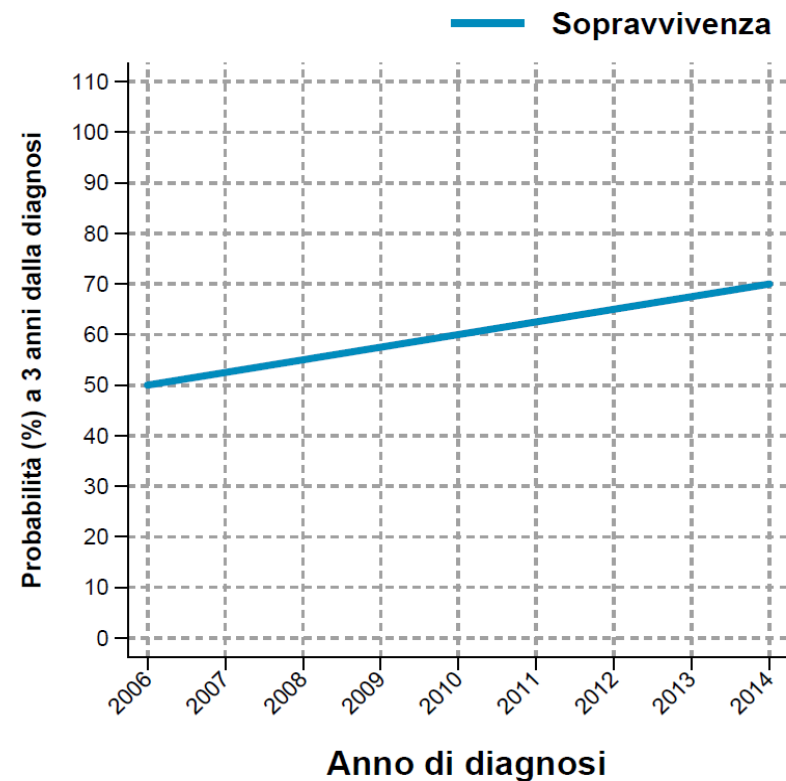
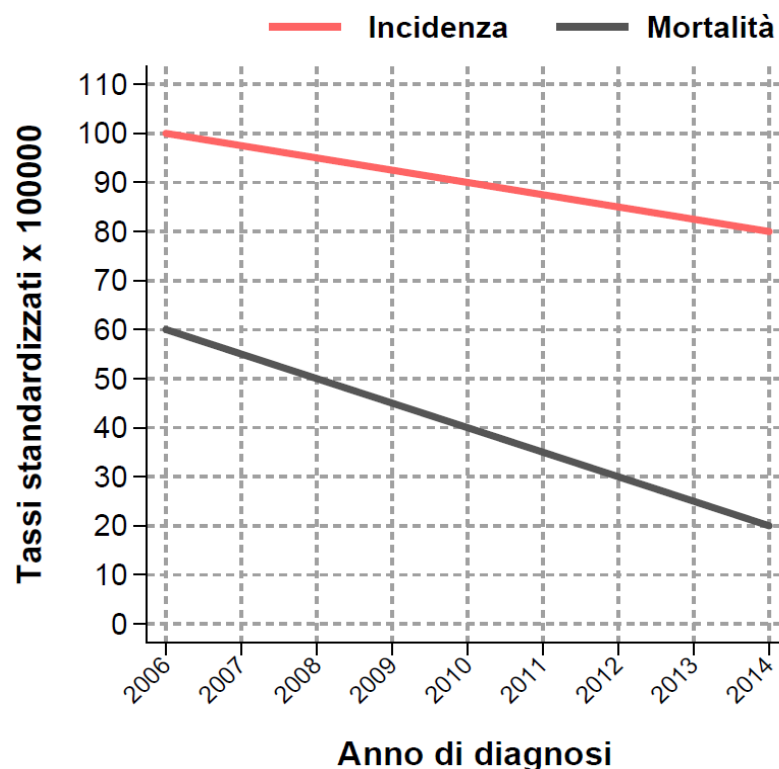


Usando insieme gli andamenti temporali di Incidenza, Mortalità e Sopravvivenza è possibile **Valutare i progressi prodotti da prevenzione primaria, screening, diagnosi precoce e terapie nella lotta contro il cancro**

Schema semplificato per leggere i due grafici

Progresso Ottimale

Linfoma di Hodgkin

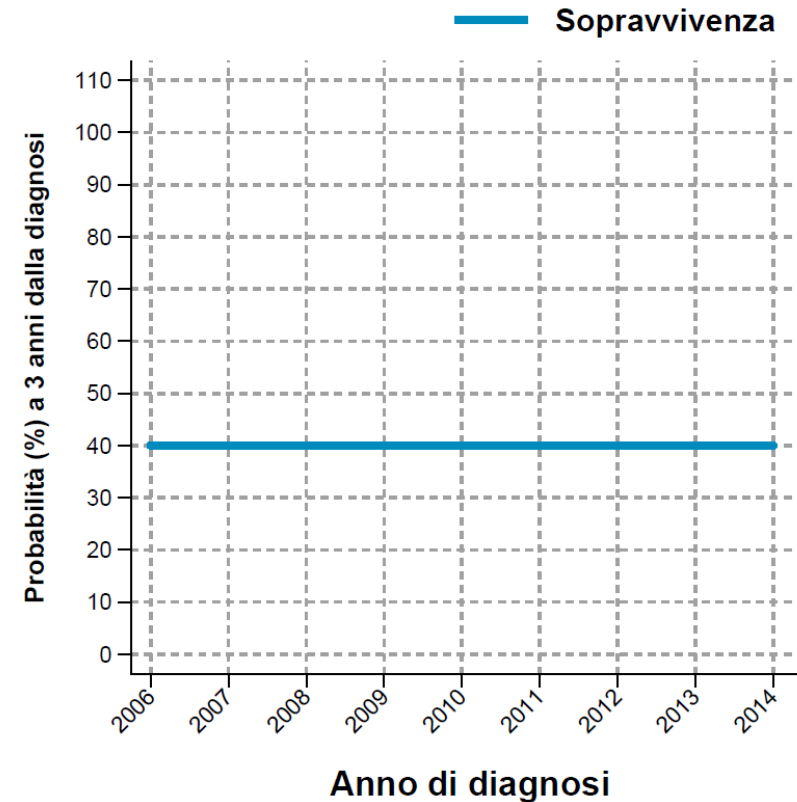
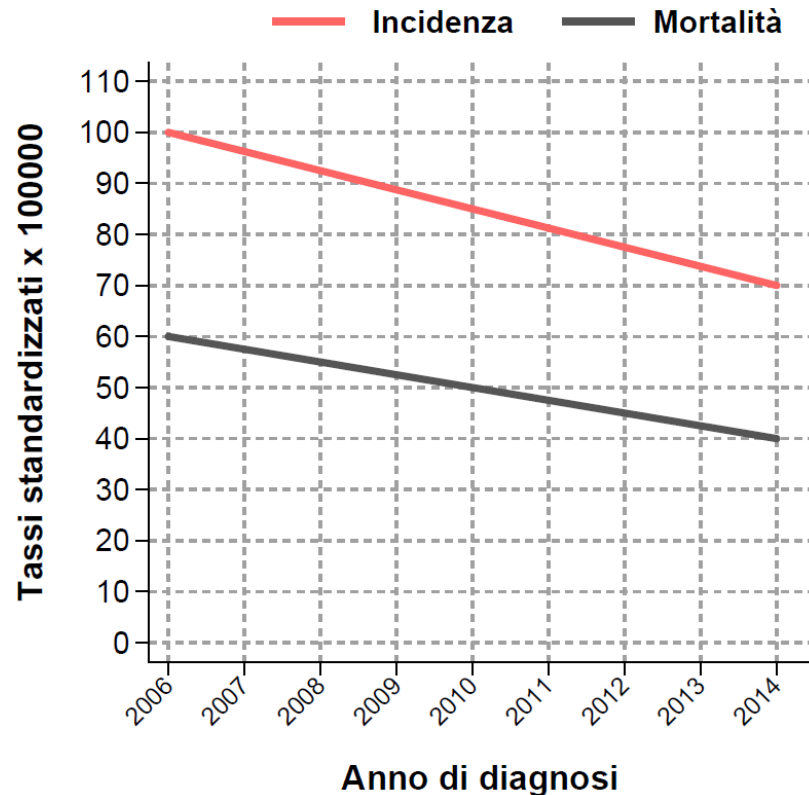


I nuovi casi diminuiscono per la minore esposizione a fattori di rischio. La sopravvivenza migliora perché curiamo meglio il cancro. La mortalità è in netta diminuzione

Schema semplificato per leggere i due grafici

Polmone Maschi

Progresso per minor carico di malattia

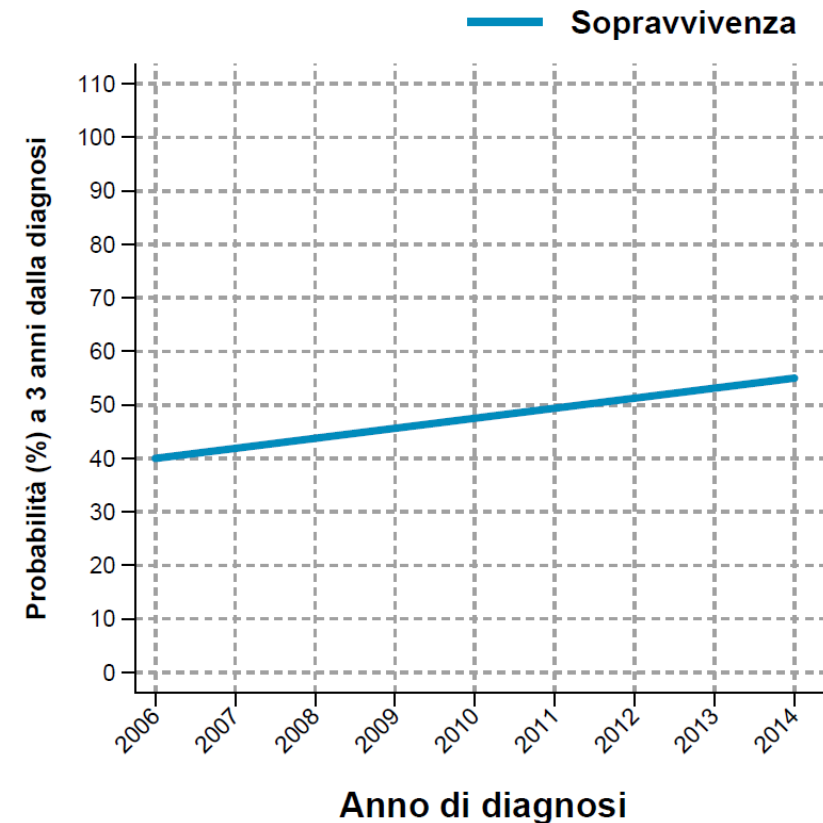
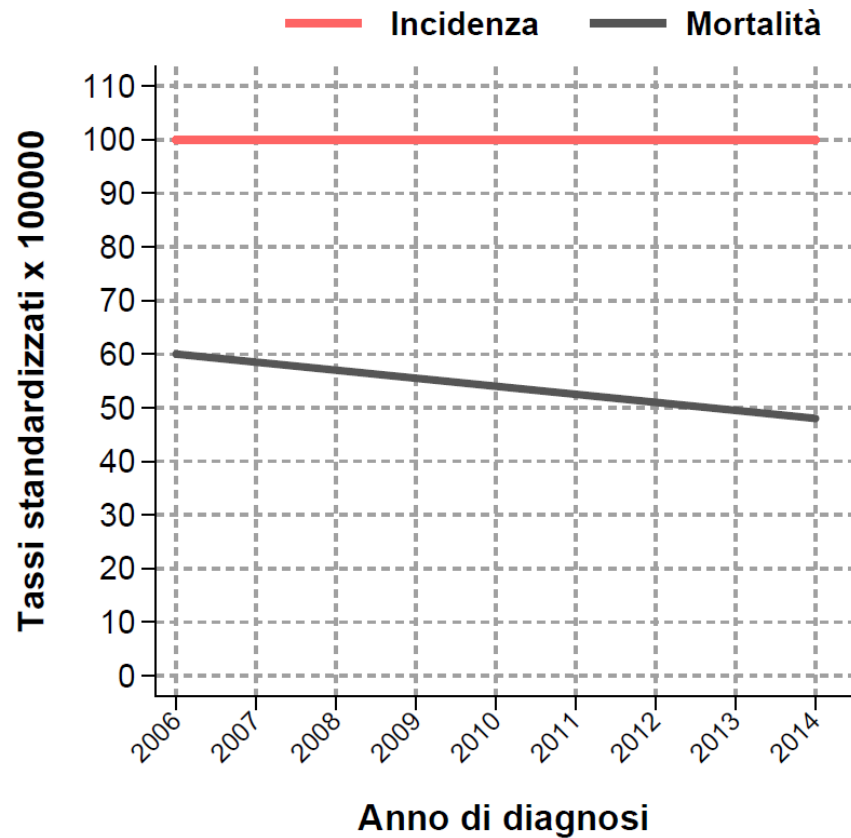


I nuovi casi diminuiscono per la minore esposizione a fattori di rischio. La sopravvivenza non migliora. La mortalità si riduce per il minor numero di persone colpite dal cancro

Schema semplificato per leggere i due grafici

Testa e collo

Progresso

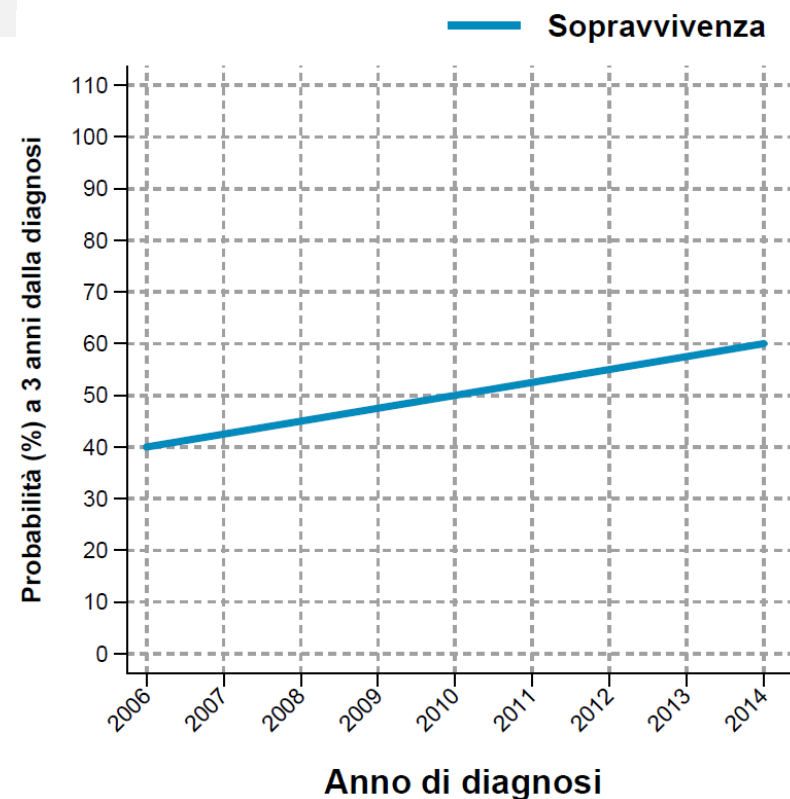
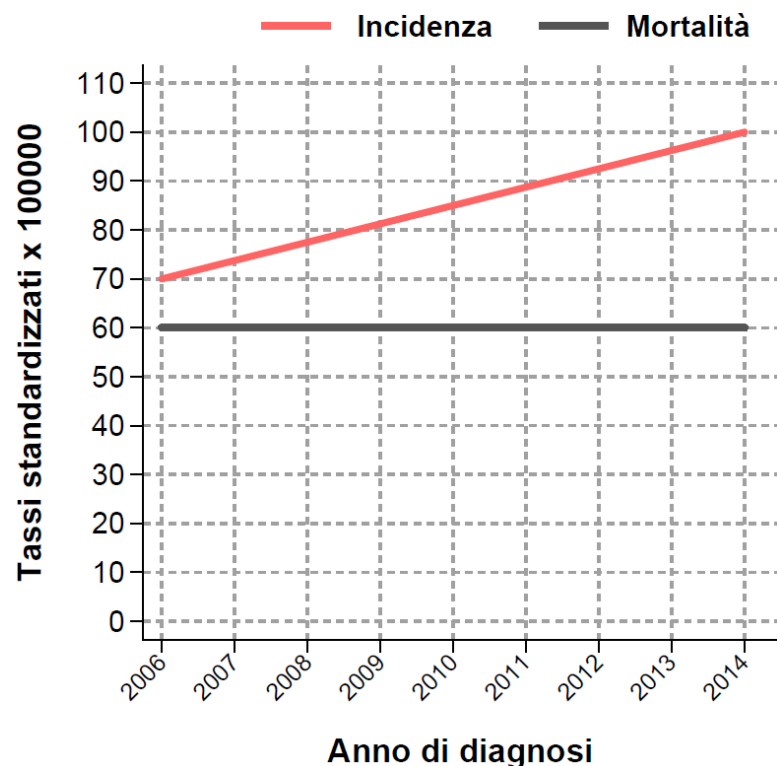


I nuovi casi non subiscono variazioni. La sopravvivenza migliora perché curiamo meglio il cancro. La mortalità è in diminuzione

Schema semplificato per leggere i due grafici



Nessun Progresso (falso progresso) oppure
Progresso ma con maggior carico di malattia

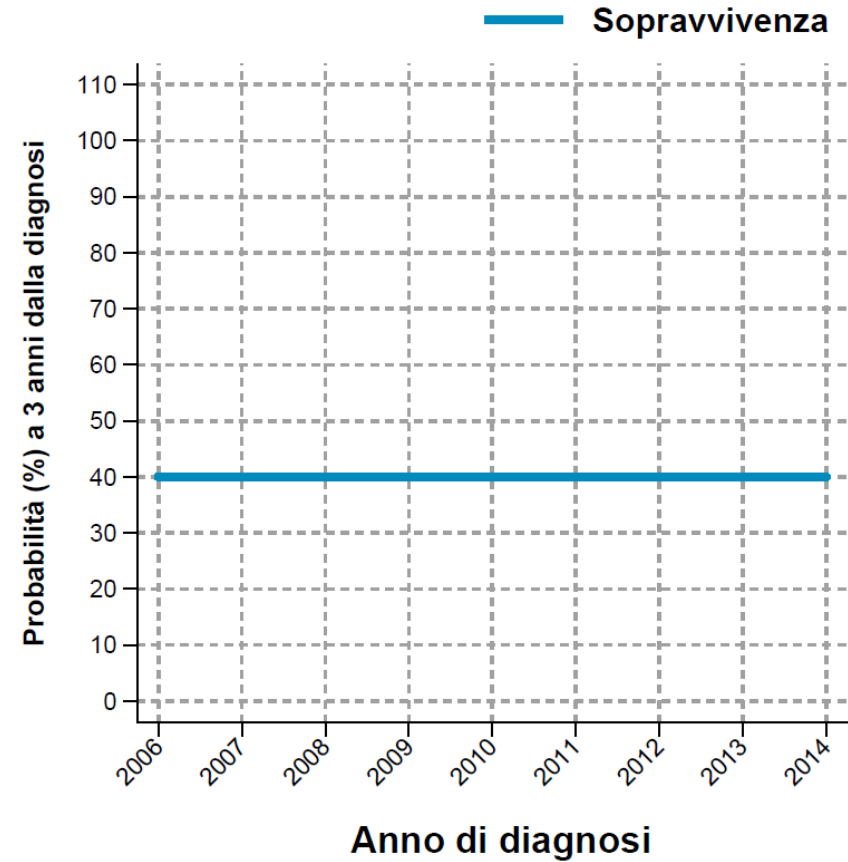
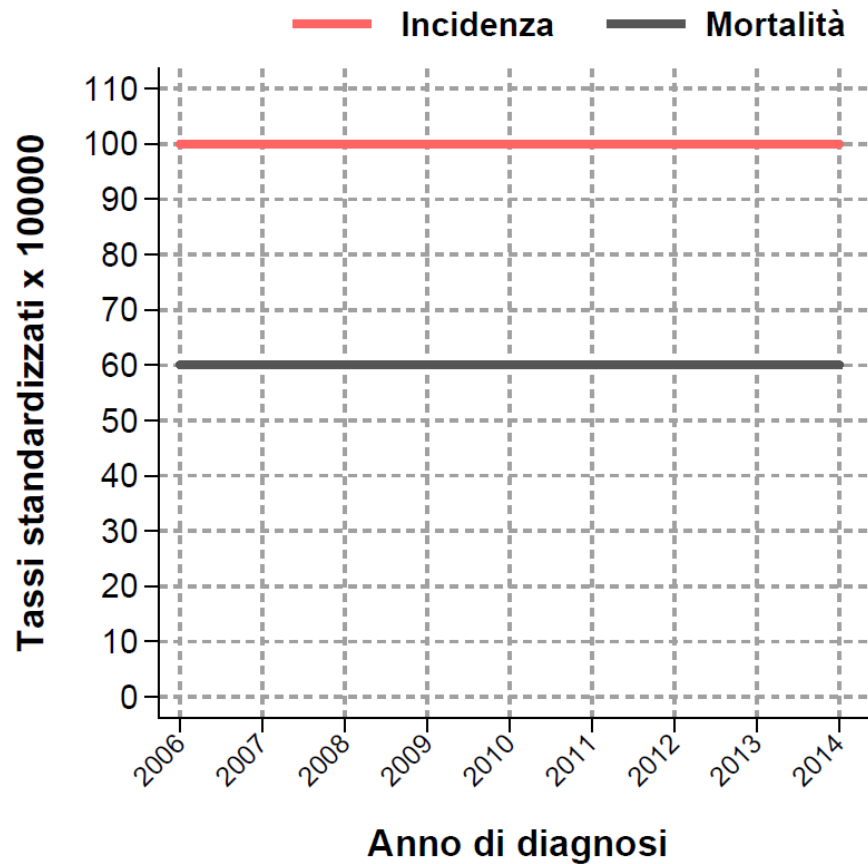


Il numero di nuovi casi aumenta. La sopravvivenza migliora. La mortalità non si modifica

Schema semplificato per leggere i due grafici

Colon Retto nel Sud

Nessun Progresso



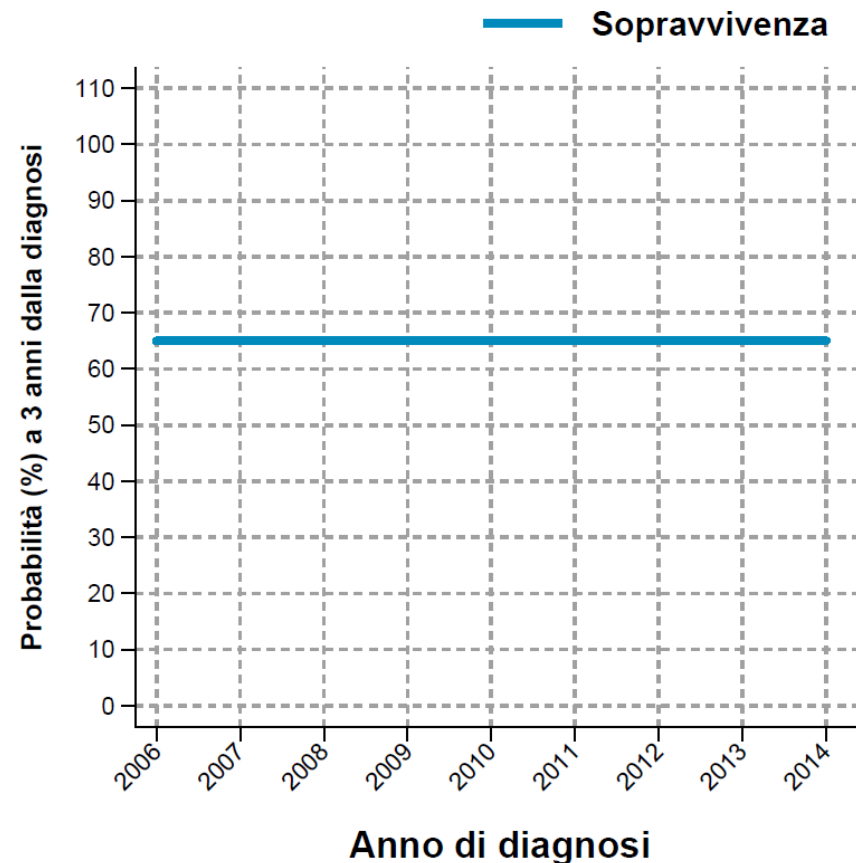
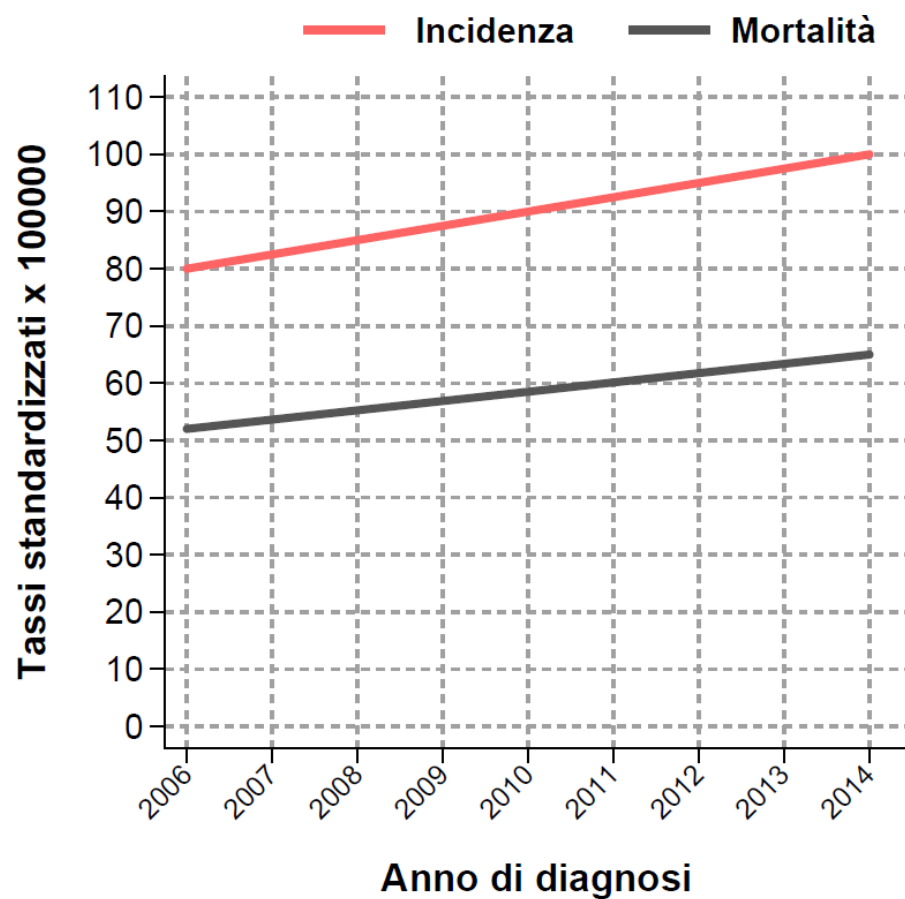
Il numero di nuovi casi, la sopravvivenza e la mortalità non subiscono variazioni

Griglia di interpretazione

Incidenza	Mortalità	Sopravvivenza	Interpretazione	Esempi
In diminuzione	In diminuzione	In aumento	Progresso ottimale	Linfoma di Hodgkin
In diminuzione	In diminuzione	Stabile	Progresso dovuto al minor carico di malattia	Cancro del polmone negli uomini
Stabile o in diminuzione	In diminuzione	In aumento	Progresso	Cancro della Mammella
In aumento	Stabile	In aumento	Progresso, ma con maggior carico di malattia oppure Nessun progresso	Cancro della tiroide
In aumento	In aumento	In leggero aumento o stabile	Nessun Progresso	Cancro del polmone nelle donne

Cattivo andamento

Regresso



I nuovi casi aumentano per la maggiore esposizione a fattori di rischio. La sopravvivenza non si modifica. La mortalità è in aumento.

Caveat

- Rispetto agli scenari-tipo mostrati in precedenza, la realtà presenta un'ampia serie di combinazioni diverse di andamenti di incidenza, mortalità e sopravvivenza.
- L'incidenza precede di un certo numero di anni la mortalità collegata al tumore. Quindi, variazioni dell'andamento dell'incidenza si riflettono in variazioni dell'andamento della mortalità con un ritardo diverso a seconda del tipo di tumore.
- Per alcuni tumori l'esame congiunto di incidenza, mortalità e sopravvivenza non offre un quadro univocamente interpretabile della nostra capacità di contrasto al cancro ma invita ad approfondimenti e apre a ipotesi di studio.

Analizzare i trend di incidenza e mortalità

I dati

File LNH.dta

Il primo obiettivo è quello di calcolare i tassi standardizzati diretti per anno di incidenza o per anno di decesso.

I dati di partenza devono contenere il numero di casi-decessi e la popolazione per ciascun anno, classe di età e opzionalmente per sesso e altre variabili.

E' importante che siano rappresentate anche le classi di età dove non sono stati osservati casi-decessi.

anno	sesso	agecla	eventi	pop
2006	M	0	0	24712
2007	M	0	0	24598
2008	M	0	0	24690
2009	M	0	0	24677
2010	M	0	0	24362
2011	M	0	0	23929
2012	M	0	0	23530
2013	M	0	1	23083
2014	M	0	0	22535
2015	M	0	1	21705
2006	M	1	1	26513
2007	M	1	1	26296
2008	M	1	0	26111
2009	M	1	1	25926
2010	M	1	1	25734

Stima dei tassi standardizzati diretti (TSD)

Per la stima è utile un comando user-written scaricabile dalla rete:

```
ssc install distrate
```

```
distrate eventi pop using popstEU2013.dta, standstrata(agecla) by(anno) mult(100000) ///
```

```
saving(LNH_standrate)
```

anno	eventi	N	crude	rateadj	lb_gam	ub_gam	se_gam	srr	lb_srr	ub_srr
2006	55	969416	5.6735189	6.6061022	4.9540923	8.6175425	.91124163	1	.	.
2007	48	971369	4.9414795	5.3527838	3.9419064	7.1026447	.7836662	.81027868	.5374968	1.2201534
2008	44	973598	4.5193191	4.9734972	3.6001308	6.691359	.76628627	.7528641	.49280857	1.145059
2009	52	974947	5.3336233	5.8786519	4.3676022	7.7346832	.83632305	.88988207	.59428482	1.3315218
2010	62	975936	6.3528756	6.7974308	5.1942358	8.7343405	.88029196	1.0289624	.7016264	1.5145807
2011	68	976143	6.9661925	7.2705508	5.6372412	9.2281378	.8932626	1.100581	.75831241	1.6068493
2012	55	975618	5.6374524	5.917142	4.4517125	7.7077934	.80844927	.89570852	.60382834	1.3310804
2013	52	979646	5.3080398	5.4416897	4.0619182	7.1390729	.76322253	.82373683	.55205508	1.230433
2014	61	983392	6.2030198	6.178063	4.7201274	7.9460431	.80107445	.93520548	.6376154	1.3778448
2015	64	981161	6.5228846	6.3743338	4.9007318	8.1550214	.80829697	.96491601	.66085203	1.4165528

Comando official per la stima dei TSD

Il comando official -dstdize- per la stima dei tassi standardizzati diretti non è consigliabile:

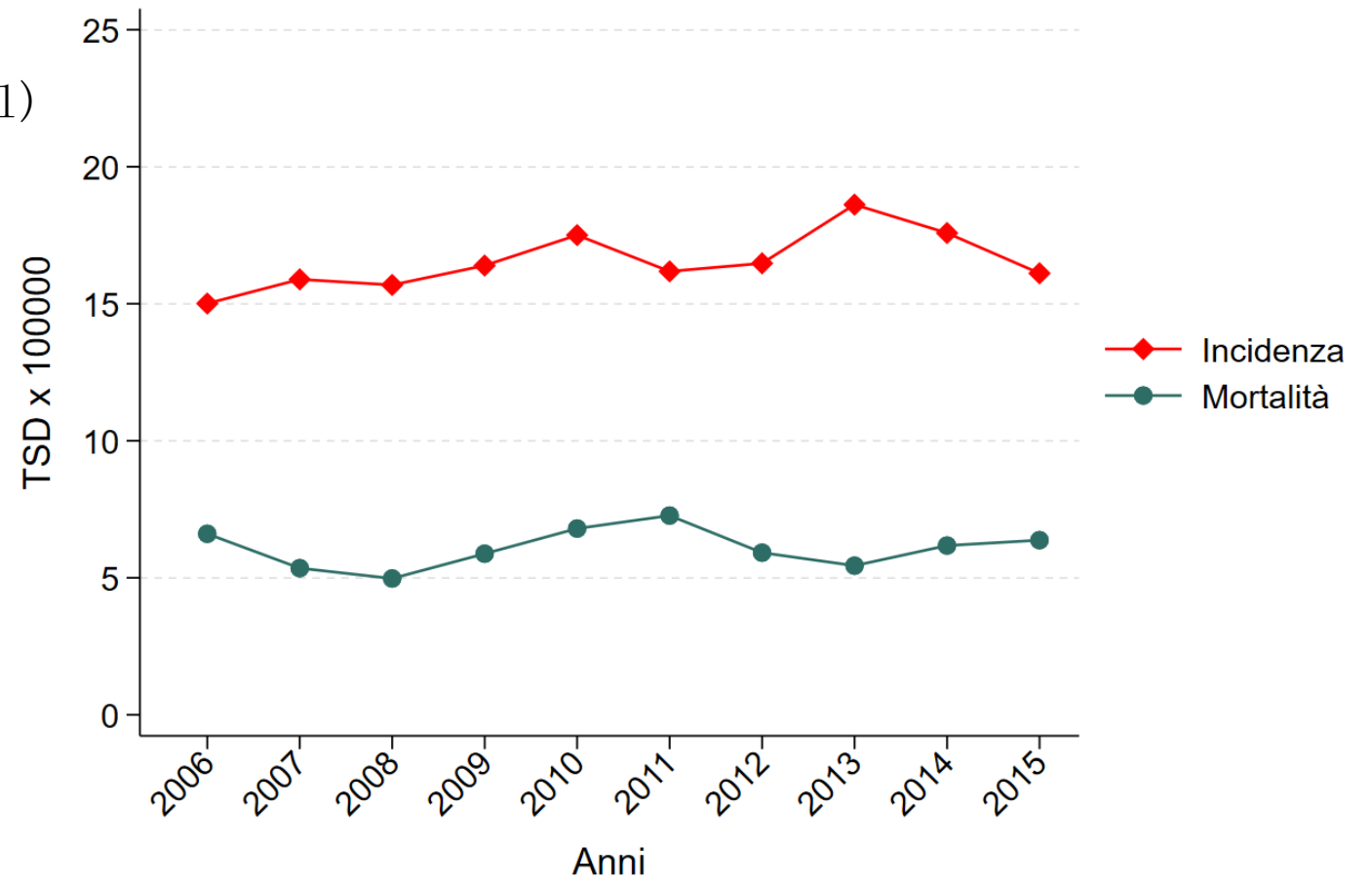
- 1- Non consente di salvare le stime in un file da utilizzare per i grafici
- 2- Non stima l'errore standard dei TSD necessario per la stima del trend
- 3- Stima intervalli di confidenza «liberal» in caso di tumori rari o di popolazioni poco numerose

```
dstdize eventi pop cleta, by(anno) using(C:\AIE2019\dati\popstEU2013)
```

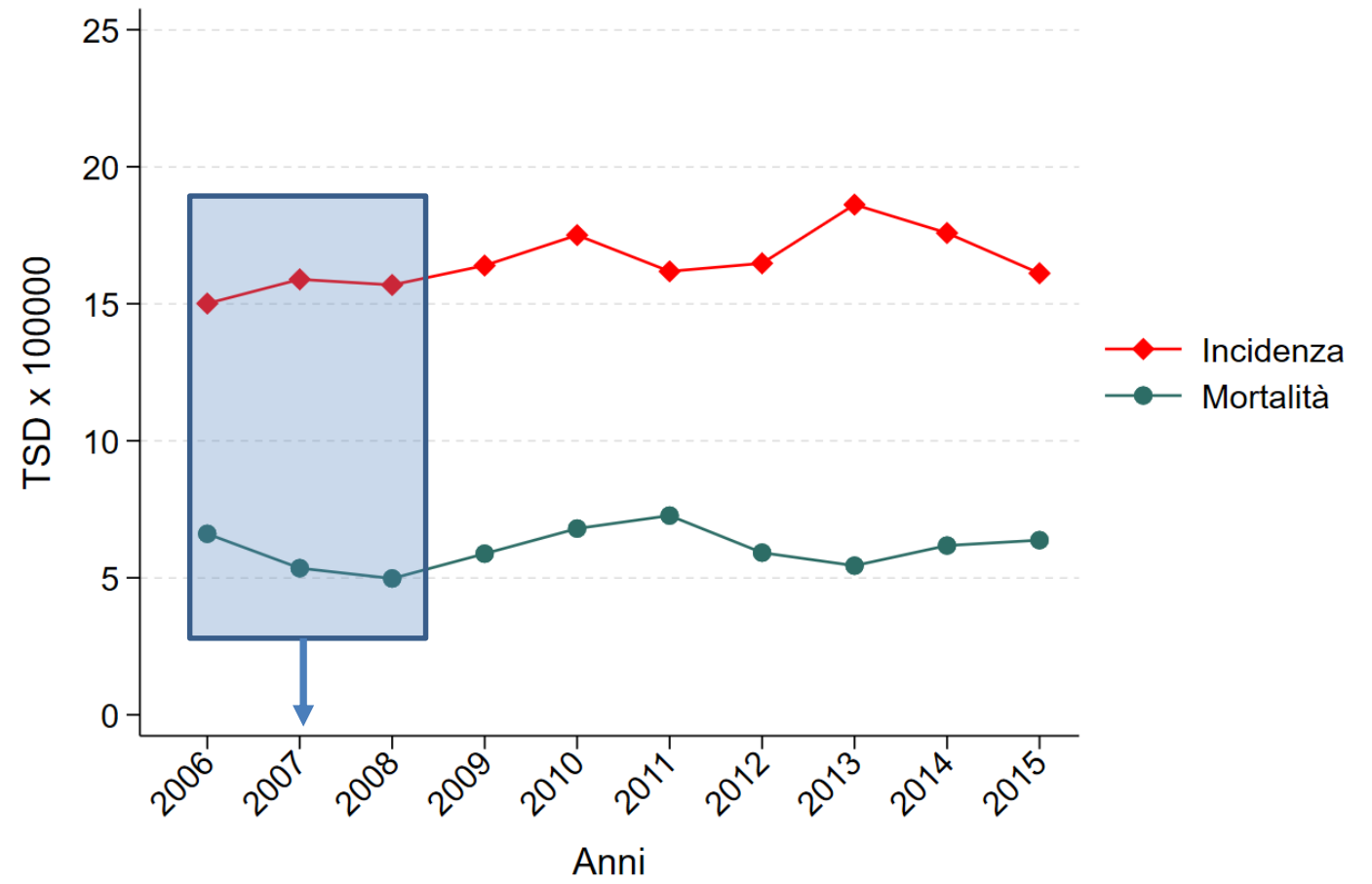
distrata output							distdize output			
anno	eventi	N	crude	rateadj	lb_gam	ub_gam				
2006	55	969416	.00005674	.00006606	.00004954	.00008618	0.000057	0.000066	[0.000048, 0.000084]
2007	48	971369	.00004941	.00005353	.00003942	.00007103	0.000049	0.000054	[0.000038, 0.000069]
2008	44	973598	.00004519	.00004973	.000036	.00006691	0.000045	0.000050	[0.000035, 0.000065]
2009	52	974947	.00005334	.00005879	.00004368	.00007735	0.000053	0.000059	[0.000043, 0.000075]
2010	62	975936	.00006353	.00006797	.00005194	.00008734	0.000064	0.000068	[0.000051, 0.000085]

Grafico scatter dei TSD

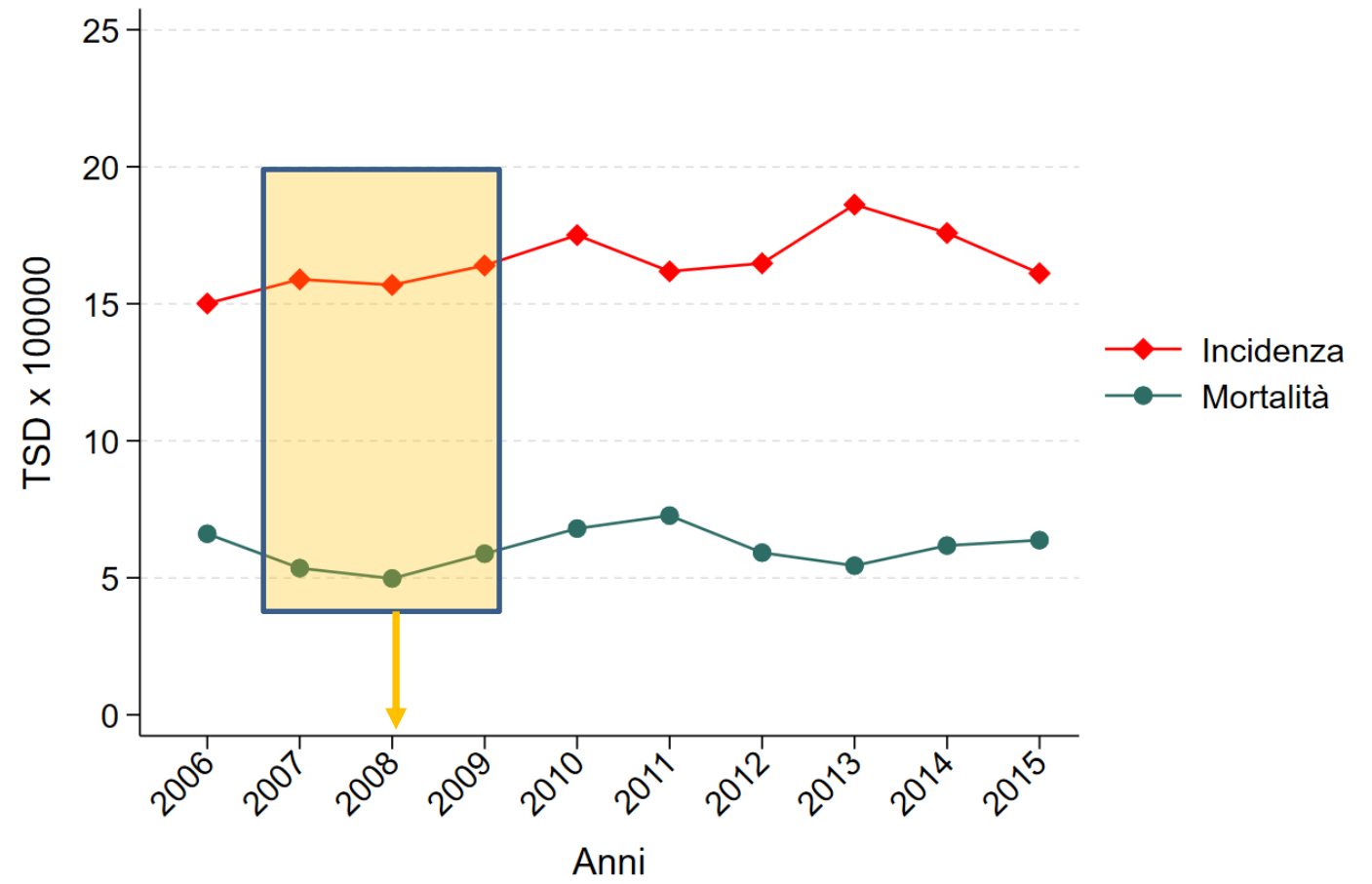
```
use LNH_standrate, clear
twoway (scatter rateadj anno ..., c(1)
..
```



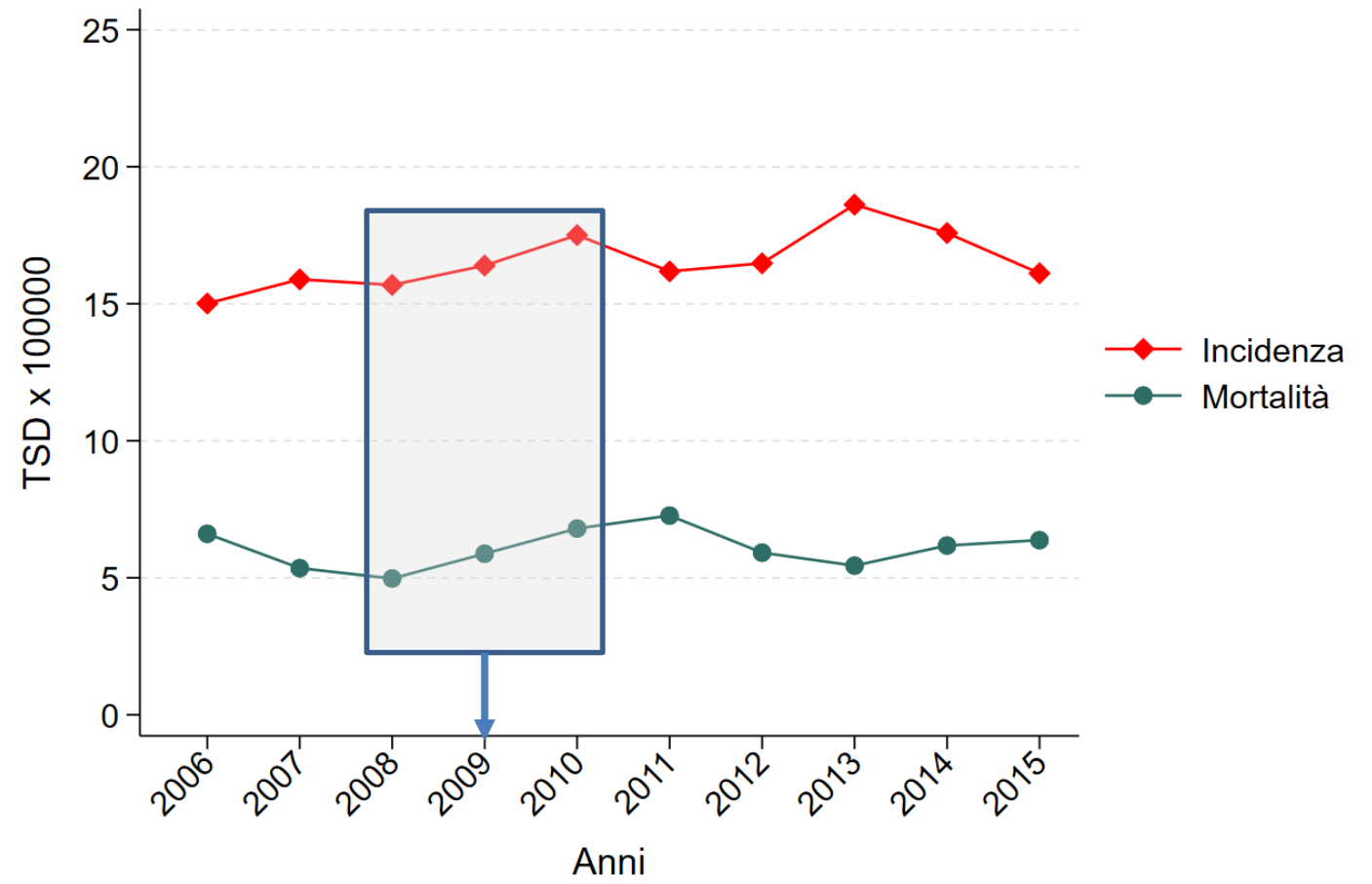
Smoothing TSD: medie mobili – finestra 3 anni



Smoothing TSD: medie mobili – finestra 3 anni

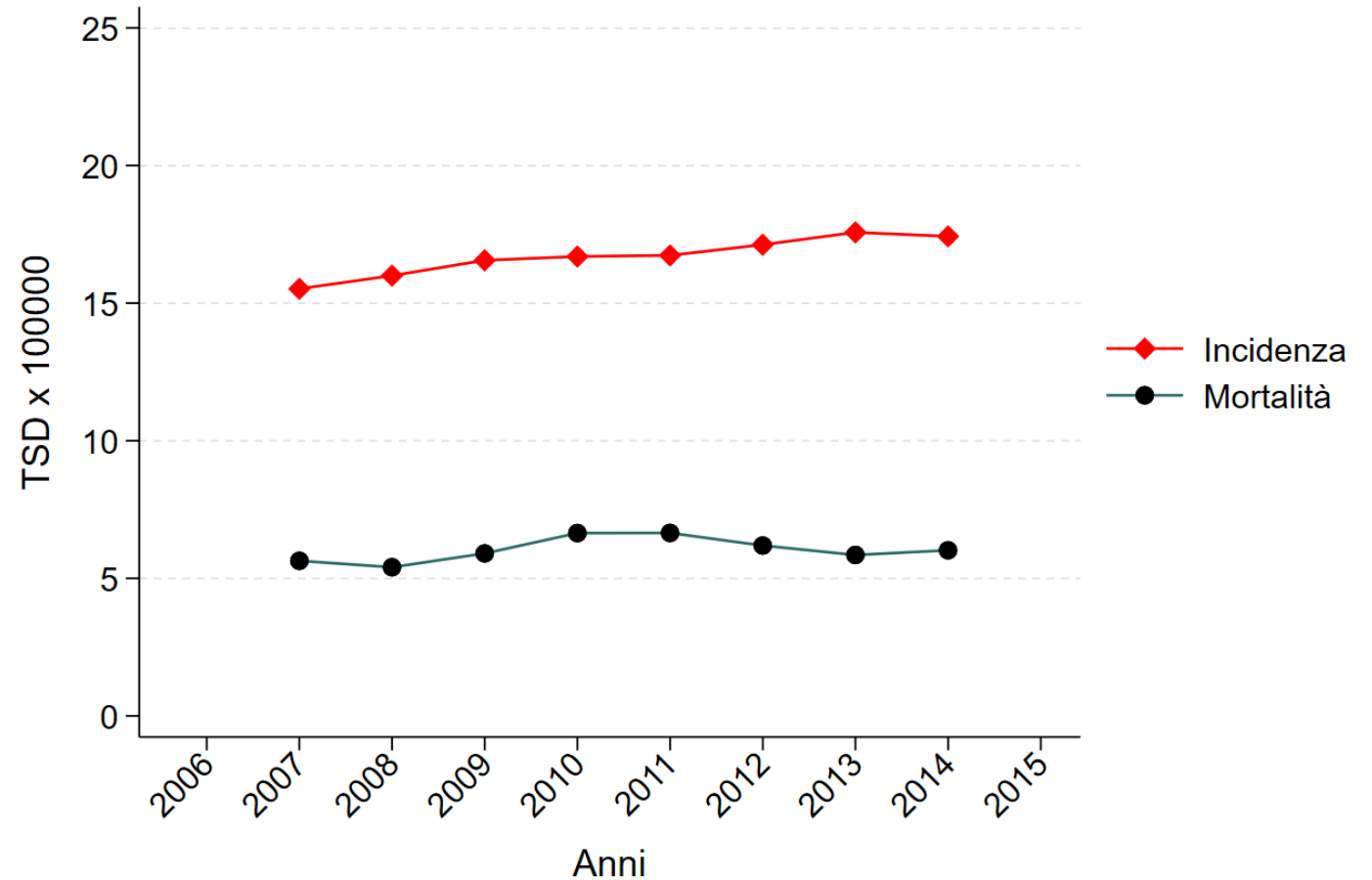


Smoothing TSD: medie mobili – finestra 3 anni



Smoothing TSD: medie mobili – finestra 3 anni

```
use LNH, clear
tempname memhold
tempfile app
postfile `memhold' anno tsd tsd_lci tsd_uci /
        using mediemobili,replace
forval i=2006/2013 {
    preserve
    distrate eventi pop using popEU2013.dta ///
        if anno>=`i' & anno<`i'+3, ///
        standstrata(cleta) saving(`app',replace)
    use `app',clear
    local y = `i'+1
    local tsd = rateadj[1]
    local tsd_lb = lb_gam[1]
    local tsd_ub = ub_gam[1]
    post `memhold' (`y') (`tsd') (`tsd_lb') `tsd'
    restore
}
postclose `memhold'
use mediemobili,clear
```



Trend lineare TSD

Il trend lineare del tasso standardizzato diretto è stimato con una regressione lineare in cui la variabile dipendente è il logaritmo del tasso standardizzato pesata per l'inverso della varianza

```
use LNH_standrate, clear
generate log_rate = log(rateadj)
generate wei = rateadj^2/se_gam^2
regress log_rate anno [aw=wei]
```

log_rate	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
anno	.0127807	.0060731	2.10	0.068	-.0012239 .0267854
_cons	-34.40186	12.21152	-2.82	0.023	-62.56168 -6.242042

La variazione percentuale annuale media del tasso di incidenza standardizzato degli LNH per tutti gli anni dal 2006 al 2015:

$$\text{APC Incidenza} = (\exp(.0127807)-1)*100 = 1.29\% \quad p = .0685$$

Trend lineare TSD

Il trend lineare del tasso standardizzato diretto è stimato con una regressione lineare in cui la variabile dipendente è il logaritmo del tasso standardizzato pesata per l'inverso della varianza

```
use LNH_standrate, clear
g log_rate = log(rateadj)
g wei = rateadj^2/se_gam^2
regress log_rate anno [aw=wei]
```

MORTALITA'

log_rate	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
anno	.0060601	.0133867	0.45	0.663	-.0248097 .03693
_cons	-21.88656	26.917	-0.81	0.440	-83.95727 40.18415

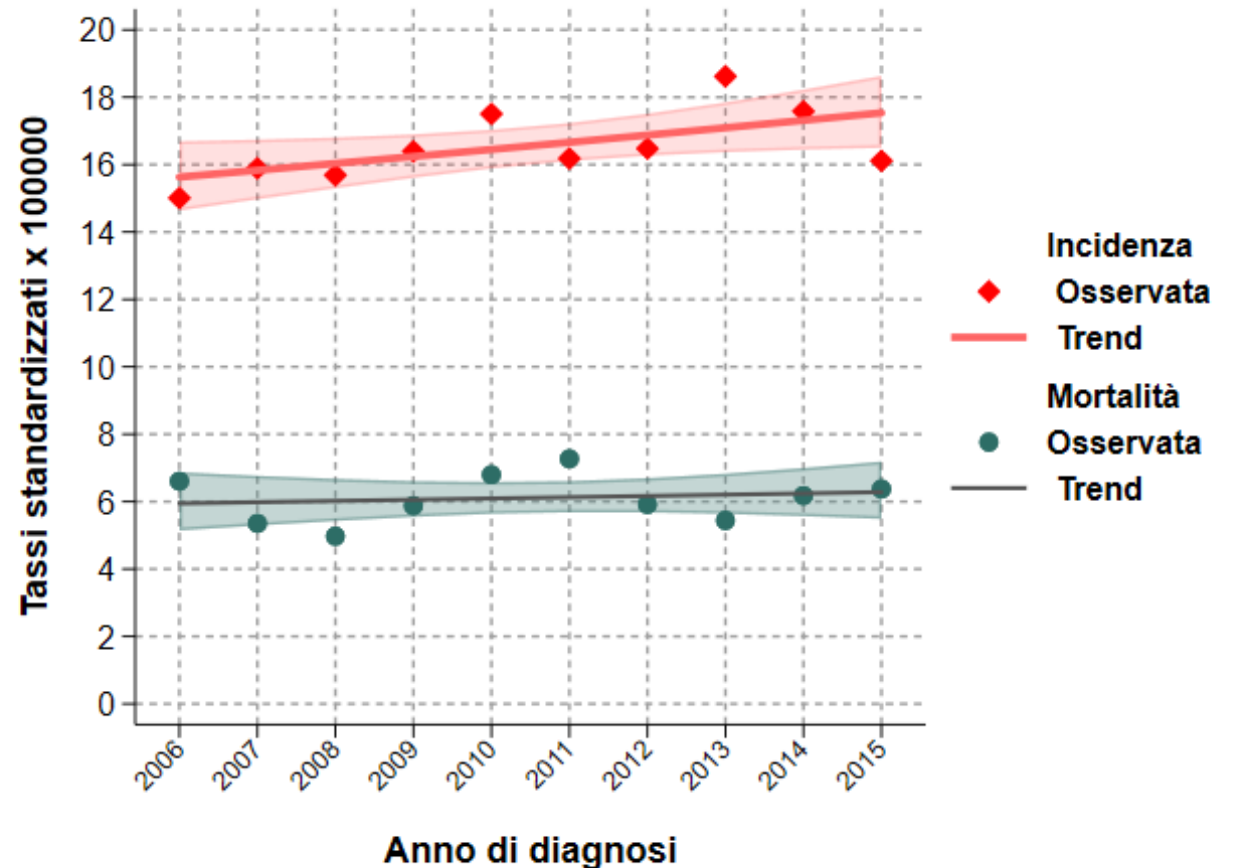
La variazione percentuale annuale media del tasso di incidenza standardizzato degli LNH per tutti gli anni dal 2006 al 2015:

$$\text{APC Mortalità} = (\exp(.0060601)-1)*100 = 0.61 (-2.39 - 3.7) \quad p = 0.6628$$

Trend lineare TSD

Il trend lineare del tasso standardizzato diretto è stimato con una regressione lineare in cui la variabile dipendente è il logaritmo del tasso standardizzato pesata per l'inverso della varianza

```
use LNH_standrate, clear
g log_rate = log(rateadj)
g wei = rateadj^2/se_gam^2
regress log_rate anno [aw=wei]
predict tassipred
predict seinc, stdp
g lo_tx = exp(tassipred - 1.96*seinc) *100000
g hi_tx = exp(tassipred + 1.96*seinc) *100000
replace tassipred =exp(tassipred) * 100000
```



Jointpoint regression

La regressione jointpoint è un metodo molto diffuso per l'analisi dei trend temporali dei tassi standardizzati diretti.

Il metodo adatta una regressione lineare, pesata in base all'inverso della varianza, in cui la variabile dipendente è il logaritmo del tasso standardizzato diretto.

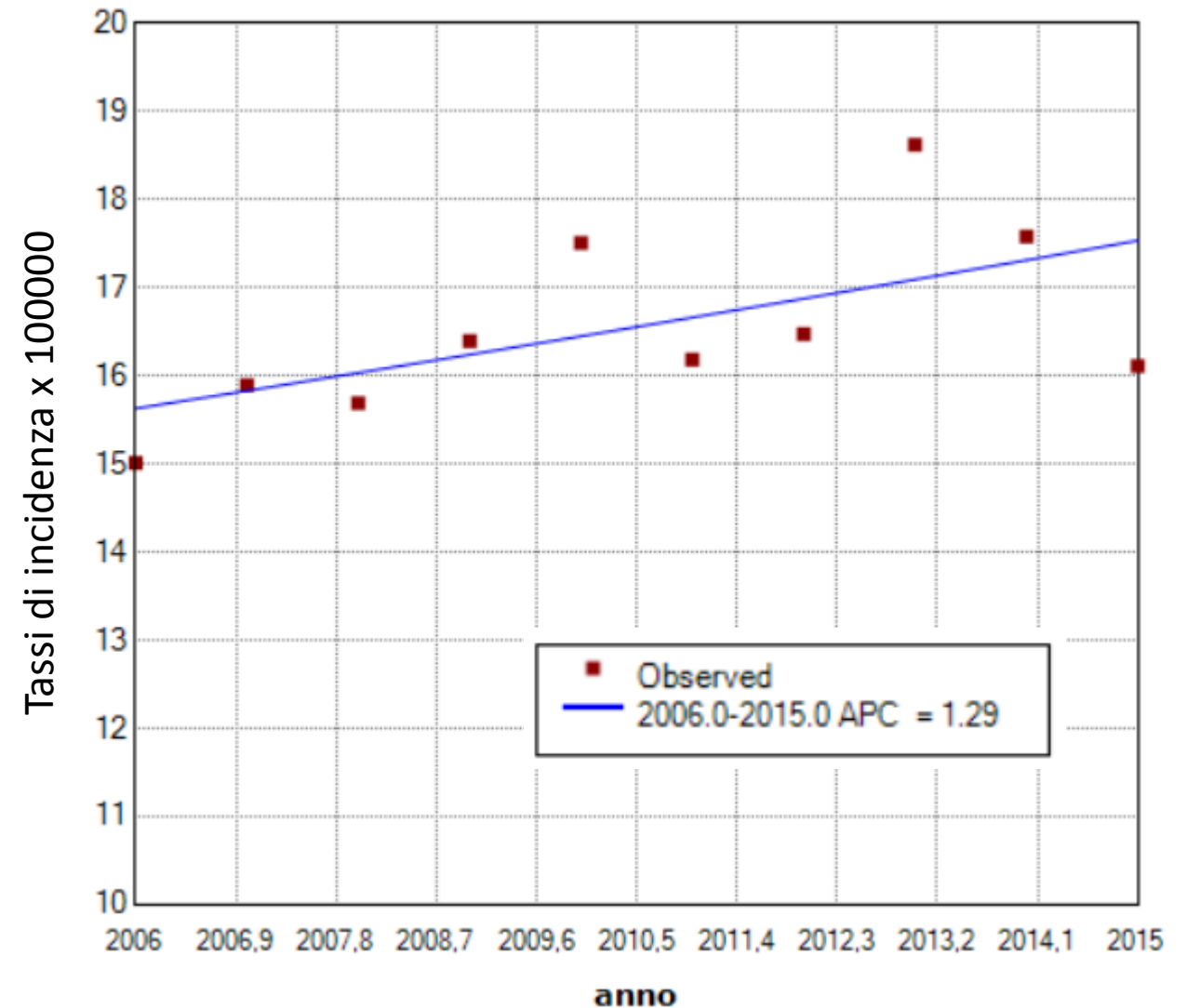
Con un metodo basato sulle permutazioni il **modello adatta l'andamento temporale (trend)** dei tassi e identifica l'eventuale presenza di «**punti temporali**» (**jointpoint**) in cui si è verificata una variazione del trend. Tra un anno di inizio e di fine selezionato dal modello, si stima il **cambiamento percentuale annuo (APC)** costante del tasso in un anno rispetto a quello dell'anno precedente e i suoi limiti di confidenza (95%).

Con dati relativi a 10 anni, il numero massimo di jointpoint proposto dal modello è 1 e la posizione viene ricercata con il metodo della griglia potendosi stabilire altri tempi intermedi tra i valori osservati dei tassi annuali*. (*Kim HJ et al. 2000*)

Output di Jointpoint

Nel caso dei LNH la regressione jointpoint seleziona un modello senza jointpoint, una regressione lineare pesata con risultato del tutto identici a quelli stimati con i comandi di Stata.

Con Stata è stato possibile mostrare in un unico grafico il trend dell'incidenza e della mortalità e aggiungere gli intervalli di confidenza

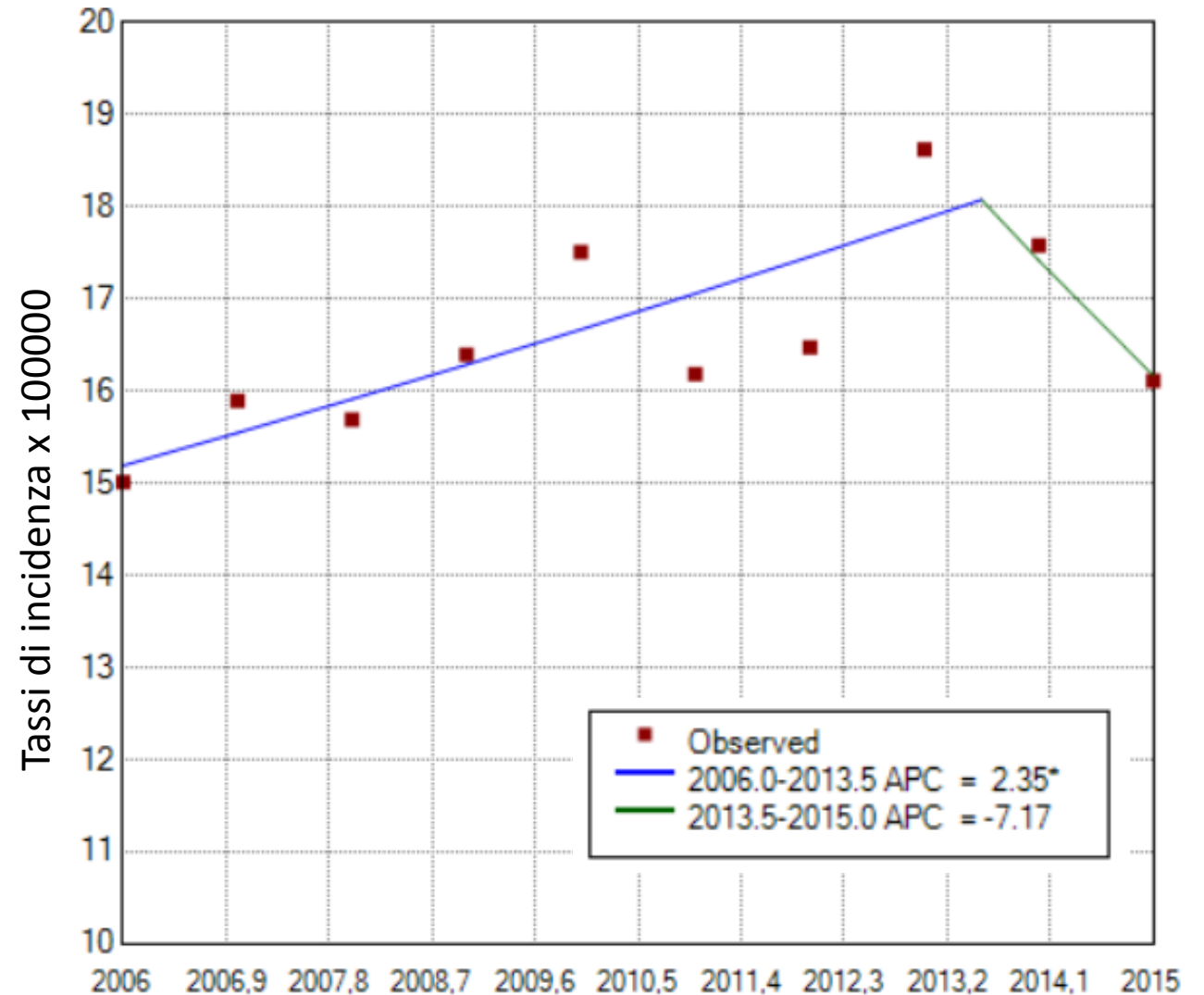


Trend con 1 jointpoint

Ipotizziamo che invece fosse stato selezionato un modello con 1 jointpoint.

In questo caso dal 2006 fino alla metà 2013 i tassi aumenterebbero del 2.35% all'anno (APC), ma nell'ultimo anno e mezzo ci sarebbe una brusca flessione del 7.17% all'anno.

Si ricorda che questo andamento poco plausibile non è quello selezionato dal metodo, ma è stato utilizzato come esempio per mostrare un andamento del tasso quando il metodo seleziona la presenza di 1 jointpoint.



Trend flessibili

La jointpoint regression individua gli anni in cui si verificano brusche variazioni dell'incidenza o della mortalità. Nell'esempio precedente l'andamento dei tassi di incidenza è stimato in aumento del 2.3% all'anno fino alla metà del 2013 e successivamente in diminuzione del 7.2%.

Nella realtà è invece plausibile che le variazioni di incidenza o di mortalità non si verifichino bruscamente ma si realizzino nell'arco di uno o più anni.

Un modo più vantaggioso e aderente alla realtà per descrivere l'andamento temporale di un tasso è quindi quello di ricorrere ad altri modelli in cui l'effetto dell'anno di calendario è adattato in modo flessibile e descritto da una curva piuttosto che da semirette unite in un angolo.

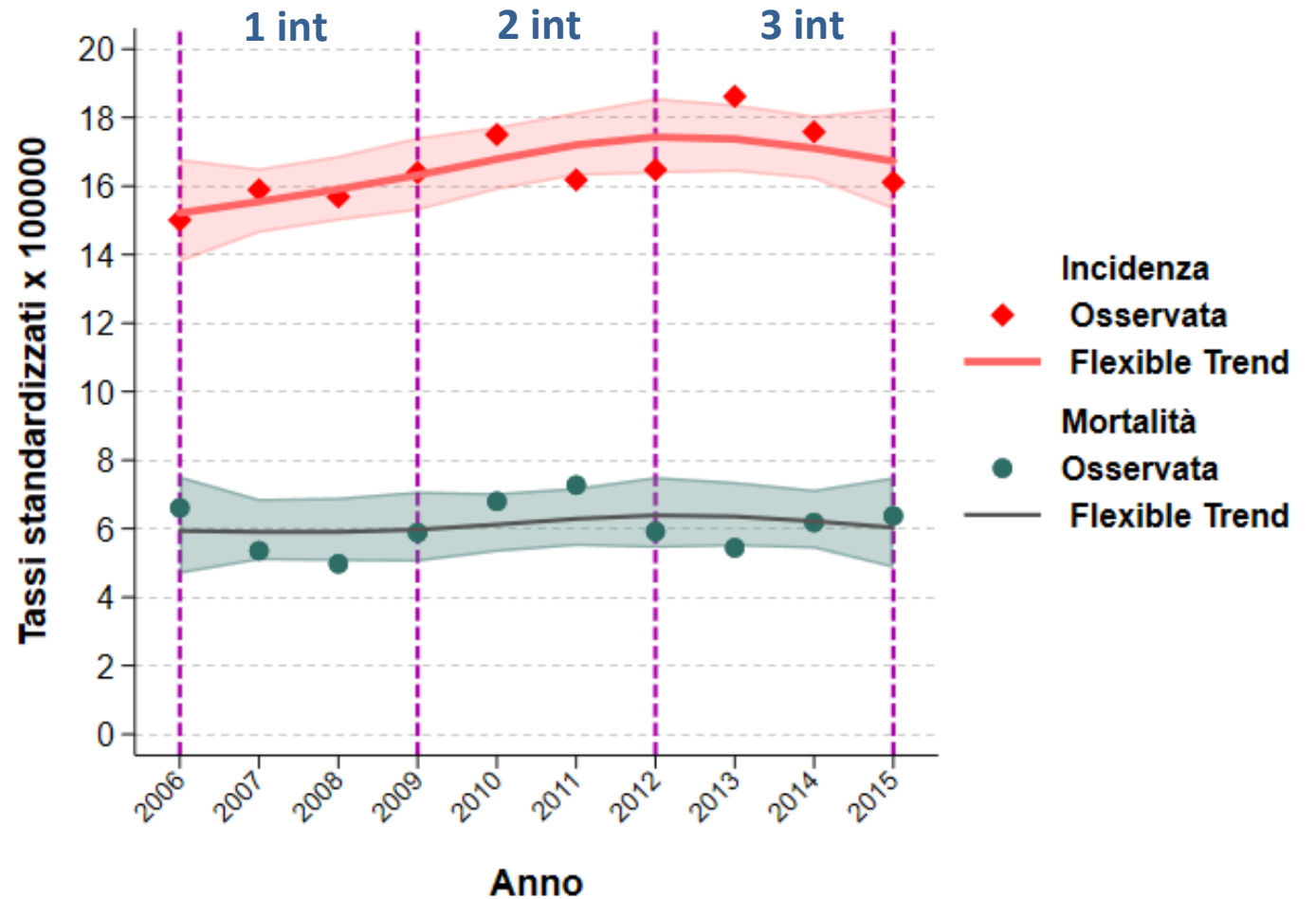
Un modo semplice per adattare l'effetto dell'anno di calendario in modo flessibile è quello di ricorrere a una «spline» regression.

Spline in azione

In una spline regression si divide il periodo indagato dal 2006 al 2015 in un certo numero di intervalli, in questo caso 3 intervalli.

Il secondo passaggio è che si adatta una regressione «polinomiale» per ciascun intervallo. In questo caso una per il primo intervallo, una per il secondo e una per il terzo.

Il terzo step è che i risultati delle tre regressioni sono collegati tra loro in modo che il trend non subisca interruzioni in corrispondenza dei confini di ciascun intervallo.



`-rcsgen-` [-mkspline- , -splinegen-]

Per mettere in opera questo meccanismo è necessario generare delle variabili appropriate attraverso l'uso di un apposito comando:

```
rcsgen anno, df(3) gen(year_rcs) orthog
```

Con questo comando Stata ha generato 3 variabili, `-year_rcs1-` `-year_rcs2-` e `-year_rcs3-`, che possiamo inserire nella regressione al posto della variabile `-anno-` per adattare l'effetto dell'anno di calendario in modo flessibile:

```
regress log_rate year_rcs1 year_rcs2 year_rcs3 [aw=wei]
```

Per la stima del trend lineare avevamo adattato

```
regress log_rate anno [aw=wei]
```


Predizioni flessibili

```
regress log_rate year_rcs1 year_rcs2 year_rcs3 [aw=wei]
```

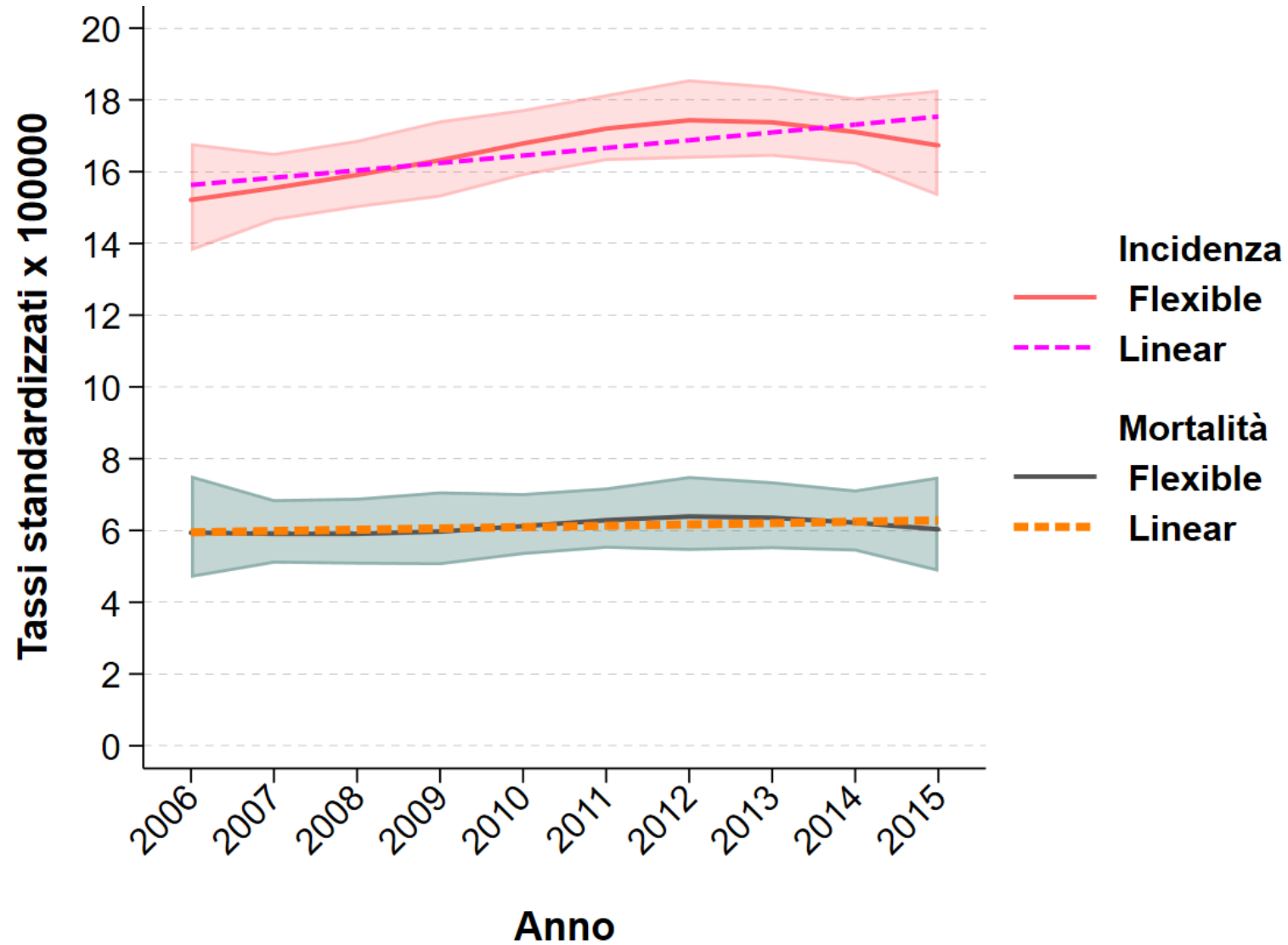
Dopo la aver adattato la spline regression, il meccanismo per la generazione dei tassi predetti dal modello è identica a quella usata dopo la regressione del trend lineare:

```
predict tassipred
```

E per gli intervalli di confidenza

```
predict seinc, stdp  
gen lo_tassipred = exp(tassipred - 1.96*seinc) * 100000  
gen hi_tassipred = exp(tassipred + 1.96*seinc) * 100000  
replace tassipred =exp(tassipred) * 100000
```

Trend flessibile e lineare



Test per l'ipotesi di linearità del trend dei tassi

Dopo aver adattato un spline regression è possibile stimare la significatività della linearità del trend. Infatti, la variabile `year_rcs1` corrisponde all'anno di incidenza/mortalità. Potremmo adattare il trend lineare includendo solo questa variabile nella regressione.

Saggiare l'ipotesi della linearità del trend equivale quindi a saggiare l'ipotesi che i coefficienti delle altre due variabili generate da `-rcsgen-`, **`year_rcs2`** e **`year_rcs3`**, siano uguali a 0.

INCIDENZA

```
. testparm year_rcs2 year_rcs3
( 1)  year_rcs2 = 0
( 2)  year_rcs3 = 0
      F( 2,      6) =    1.08
      Prob > F =    0.3968
```

MORTALITA'

```
. testparm year_rcs2 year_rcs3
( 1)  year_rcs2 = 0
( 2)  year_rcs3 = 0
      F( 2,      6) =    0.14
      Prob > F =    0.8757
```

Per l'incidenza e la mortalità l'ipotesi nulla che `year_rcs2` e `year_rcs3` siano diverse da 0 non è rigettata.

Possiamo quindi concludere che il trend del tasso standardizzato di incidenza e di mortalità può essere adattato solo con un effetto lineare

Cenni all'analisi dei trend di sopravvivenza

Sopravvivenza relativa

La sopravvivenza relativa è ottenuta dal rapporto tra la sopravvivenza osservata per tutte le cause dei casi di tumore e la sopravvivenza di un gruppo comparabile della popolazione generale che si immagina privo della malattia in studio (sopravvivenza attesa):

$$\text{Rel Surv} = \frac{S_O}{S_E}$$

La sopravvivenza relativa ha il vantaggio di non aver bisogno dell'informazione sulla causa di morte e per questo è in genere preferita negli studi di sopravvivenza in cui bisogna stimare questo indicatore su casistiche molto numerose come quelle dei registri tumori (studi di popolazione).

Attualmente l'unico metodo utilizzato per calcolare la sopravvivenza relativa è il metodo di Ederer II che però è affetto da una leggera distorsione. Pohar Perme ha proposto un metodo diverso che produce stime di sopravvivenza non distorte.

File necessari

Per stimare la sopravvivenza relativa abbiamo bisogno di **due file**.

Il primo è quello dei pazienti in studio che ci permette di calcolare la sopravvivenza osservata.

In questo file devono essere contenute:

- Data di diagnosi
- Data di fine osservazione
- Stato in vita
- Sesso
- Età alla diagnosi
- Opzionalmente altre variabili come la provincia di residenza

Il secondo file è necessario per calcolare la sopravvivenza attesa di una popolazione di riferimento simile a quella da cui derivano i casi di tumore ma non affetta dalla malattia.

Tavole di sopravvivenza della popolazione generale

Il file necessario per stimare la sopravvivenza attesa deve essere così strutturato:

```
+-----+
| sesso  _year  _age  prob  |
|-----|
|   1    2006   0    .96429 |
|   1    2006   1    .99639 |
|   1    2006   2    .99783 |
|   1    2006   3    .99842 |
|   1    2006   4    .99882 |
| .....|
|   2    2015  91    .97893 |
|   2    2015  92    .96253 |
|   2    2015  93    .94905 |
|   2    2015  94    .93320 |
|   2    2015  95    .92131 |
+-----+
```

Un file con queste caratteristiche può essere scaricato dal sito Demo Istat <http://demo.istat.it/tvm2016/index.php?lingua=ita>

Per semplificare l'uso del comando per la stima della sopravvivenza relativa è opportuno che le variabili in questo file siano designate nel modo indicato:

- **_year** per l'anno di calendario,
- **_age** per l'età,
- **prob** per la probabilità di sopravvivenza.

Stima della sopravvivenza relativa con Stata : -strs-

```
strs using lifetabPuglia, breaks(0(.25)3) mergeby(_year sex _age)
```

- using nomefile specifica il file, in questo caso **lifetabPuglia**, da utilizzare per il calcolo della sopravvivenza attesa.
- `breaks(numlist)` . Per stimare la sopravvivenza il tempo di osservazione è diviso in intervalli. Specificando `breaks(0(0.25)3)` si è indicato di dividere il tempo di osservazione fino a 3 anni dalla diagnosi in intervalli trimestrali (un quarto di anno).
- `mergeby(provincia _year sex _age)` specifica le variabili per cui è stratificata la probabilità di sopravvivenza nel file -lifetabPuglia-.

Tabella prodotta da -strs-

start	end	cp	cr_e2	lo_cr_e2	hi_cr_e2	cns_pp	lo_cnspp	hi_cnspp
0	.25	0.8932	0.8983	0.8766	0.9166	0.8979	0.8760	0.9160
.25	.5	0.8333	0.8425	0.8168	0.8652	0.8415	0.8156	0.8641
.5	.75	0.7894	0.8022	0.7742	0.8274	0.8009	0.7727	0.8261
.75	1	0.7519	0.7680	0.7384	0.7950	0.7659	0.7360	0.7930
1	1.25	0.7262	0.7453	0.7147	0.7734	0.7433	0.7124	0.7714
1.25	1.5	0.7090	0.7311	0.6999	0.7600	0.7284	0.6967	0.7574
1.5	1.75	0.6875	0.7123	0.6804	0.7420	0.7098	0.6774	0.7396
1.75	2	0.6735	0.7012	0.6689	0.7315	0.6990	0.6661	0.7294
2	2.25	0.6506	0.6807	0.6477	0.7117	0.6783	0.6447	0.7095
2.25	2.5	0.6360	0.6687	0.6353	0.7003	0.6653	0.6311	0.6971
2.5	2.75	0.6268	0.6622	0.6284	0.6942	0.6589	0.6243	0.6912
2.75	3	0.6149	0.6528	0.6186	0.6852	0.6486	0.6133	0.6815

Utility di -strs-

```
strs using lifetabPuglia [iw=standwei], mergeby(_age sesso _year) br(0(.25)3) ///
```

```
pohar by(yydx) standstrata(agecla) savstand(SurvAnno, replace)
```

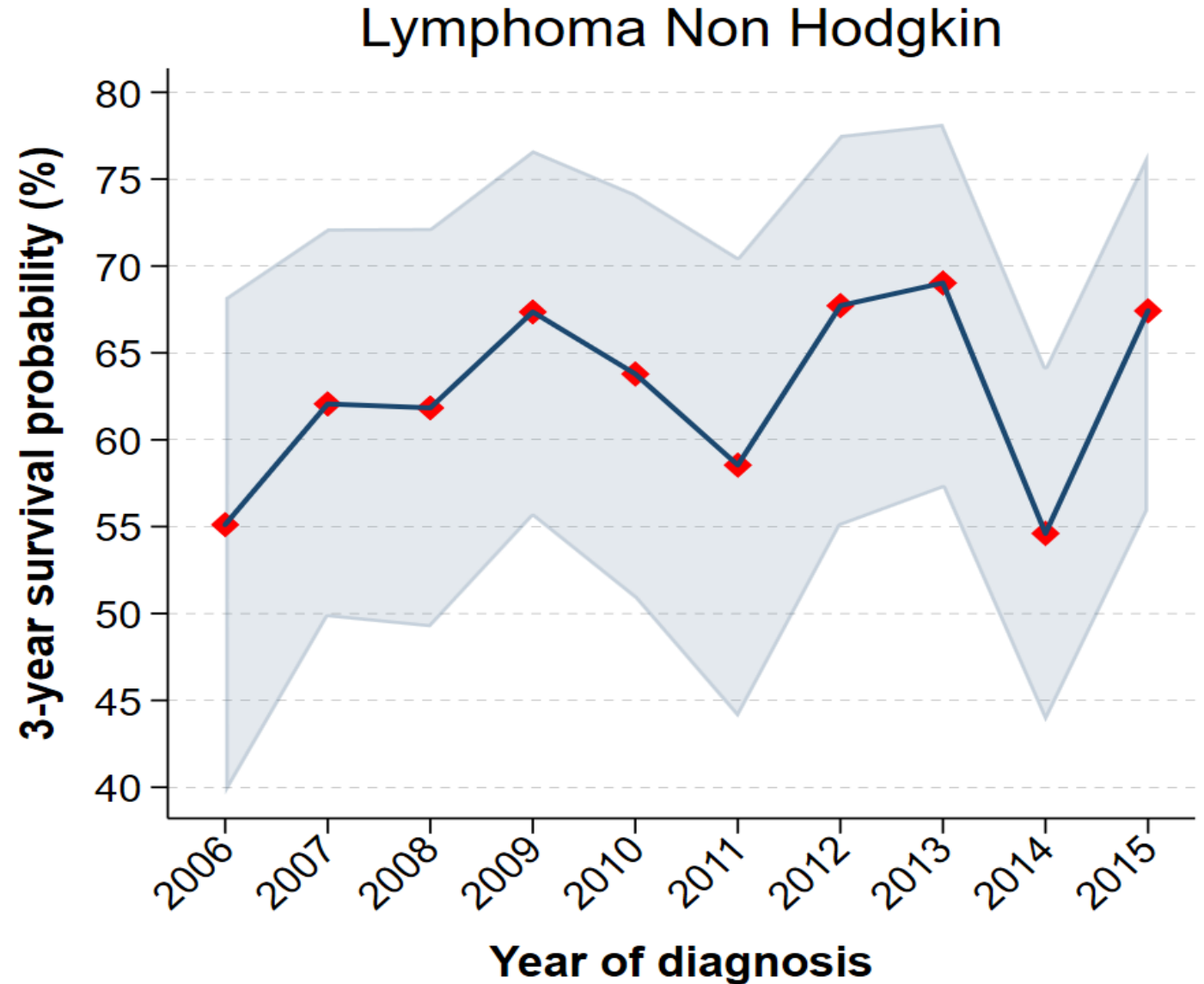
Alcune opzioni del comando consentono di stimare rapidamente il trend della sopravvivenza a un certo tempo dalla diagnosi.

In nero grassetto sono indicate le istruzioni per stimare la sopravvivenza standardizzata per età.

In rosso sono indicate le istruzioni per stimare la sopravvivenza per ciascun anno di incidenza dal 2006 al 2015 **[by(yydx)]** e salvare le stime standardizzate in un file **[savstand(SurvAnno, replace)]**

Grafico della sopravvivenza per anno di incidenza

```
use SurvAnno, clear
keep if end==3 // Sopravvivenza a 3 anni
twoway (rarea lo_cns_pp hi_cns_pp yydx) ///
      (line cns_pp yydx, lc(navy) lp(1)) ...
```



Media mobile della sopravvivenza a 3 anni dalla diagnosi

* Inizializza il file MediaMobile con primo triennio incidenza 2006-2008

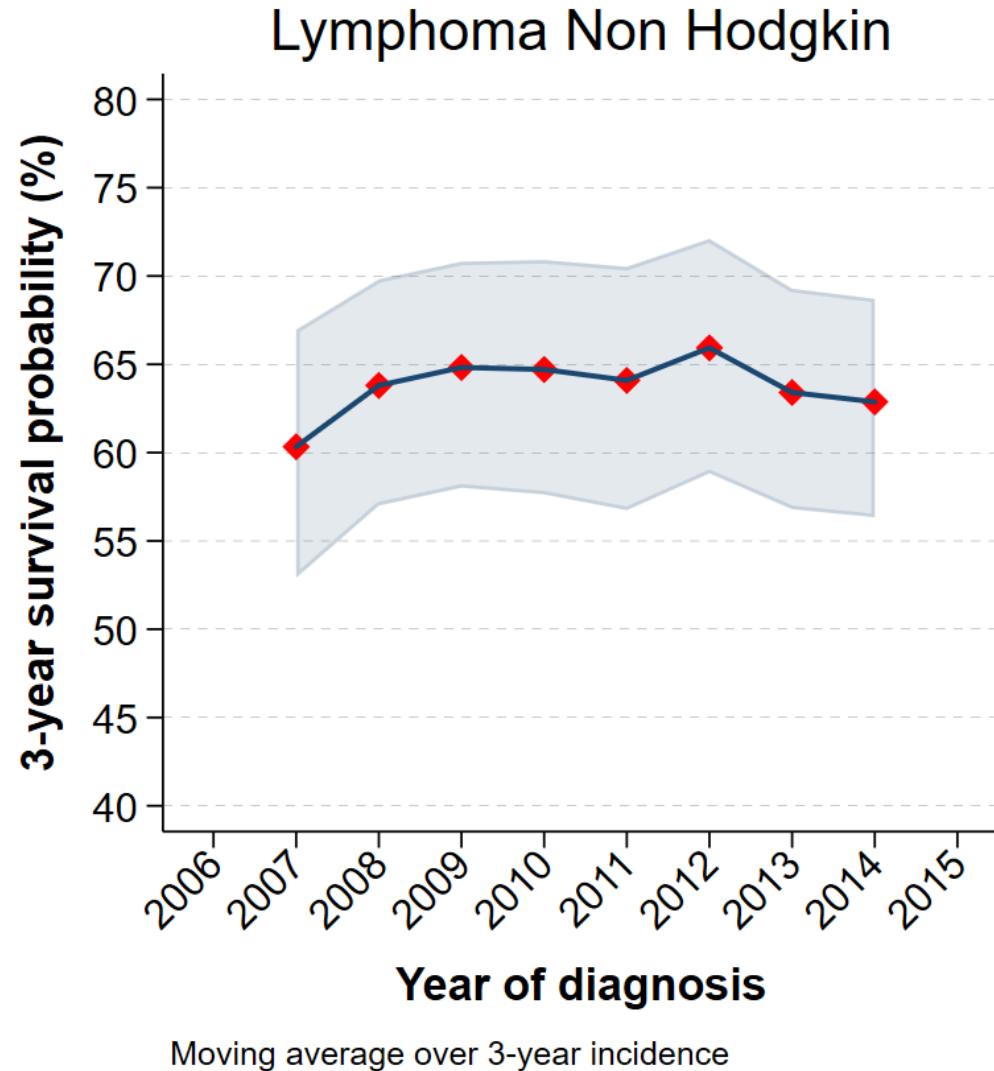
```
strs using lifetabPuglia if yydx<2009, mergeby(idrt _age sesso _year) br(0(.25)3) pohar savgr(Temp,replace)
use Temp, clear
g anno_est = 2007
keep if end==3
save MediaMobile, replace
```

* Stima sopravvivenza per i trienni successivi 2007-2009, 2008-2010,

```
forval i=2007/2012 {
    use LNH, clear
    keep if yydx>=`i' & yydx<`i'+3
    qui strs using lifetabPuglia, mergeby(idrt _age sesso _year) br(0(.25)3) ///
        pohar savgr(Temp,replace)
    use Temp, clear
    g anno_est = `i' + 1
    keep if end==3 // Sopravvivenza a 3 anni
    append using MediaMobile
    save MediaMobile, replace
}
```

Media mobile della sopravvivenza a 3 anni dalla diagnosi

```
use MediaMobile, clear
twoway (rarea lo_cns_pp hi_cns_pp anno_est) ///
      (line cns_pp anno_est, lc(navy) lp(1)) ...
```



Trend lineare della sopravvivenza a 3 anni dalla diagnosi

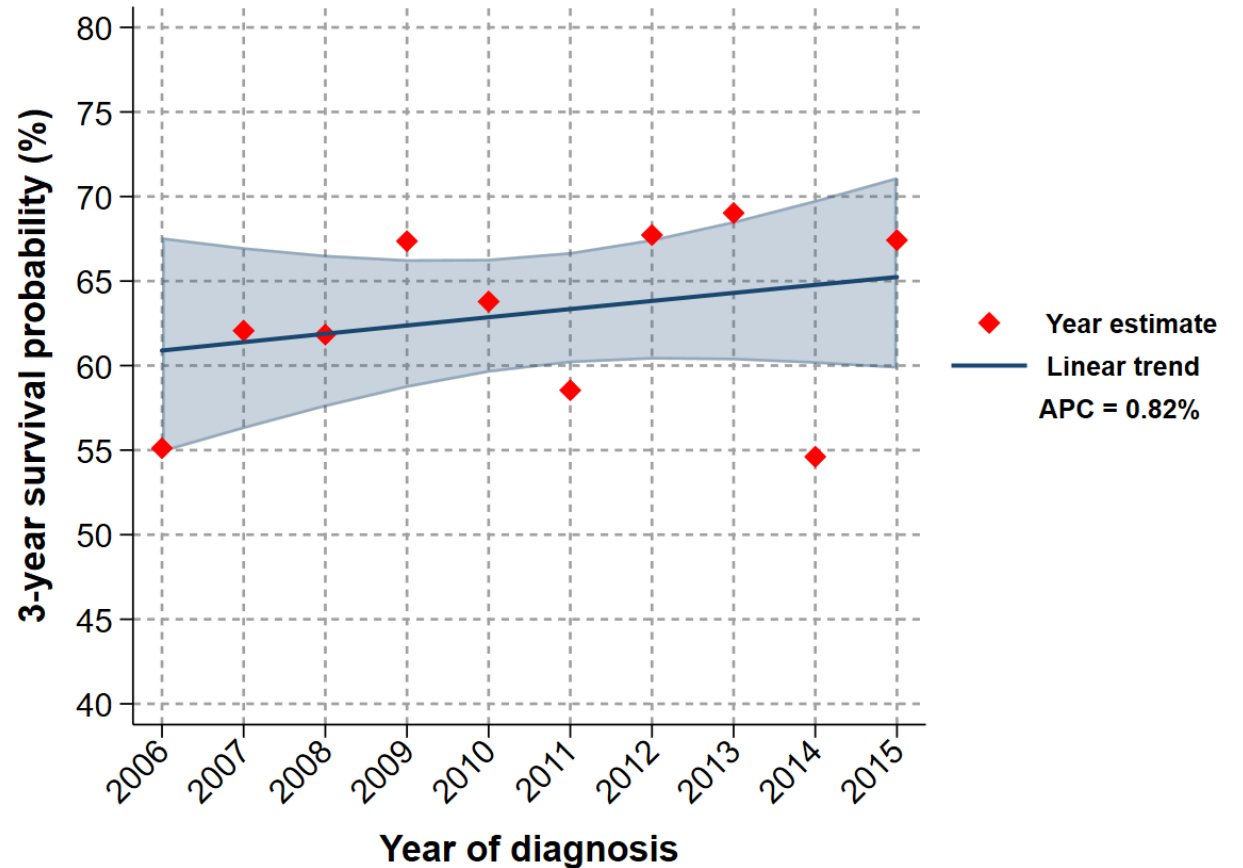
La stima del trend lineare e del trend flessibile della sopravvivenza relativa richiede l'uso di modelli di sopravvivenza parametrici flessibili.

Questi modelli hanno un'implementazione molto ampia in Stata.

Il comando fondamentale è `-stpm2-`

L'introduzione all'uso di questi modelli non può essere effettuata nell'arco di questo incontro.

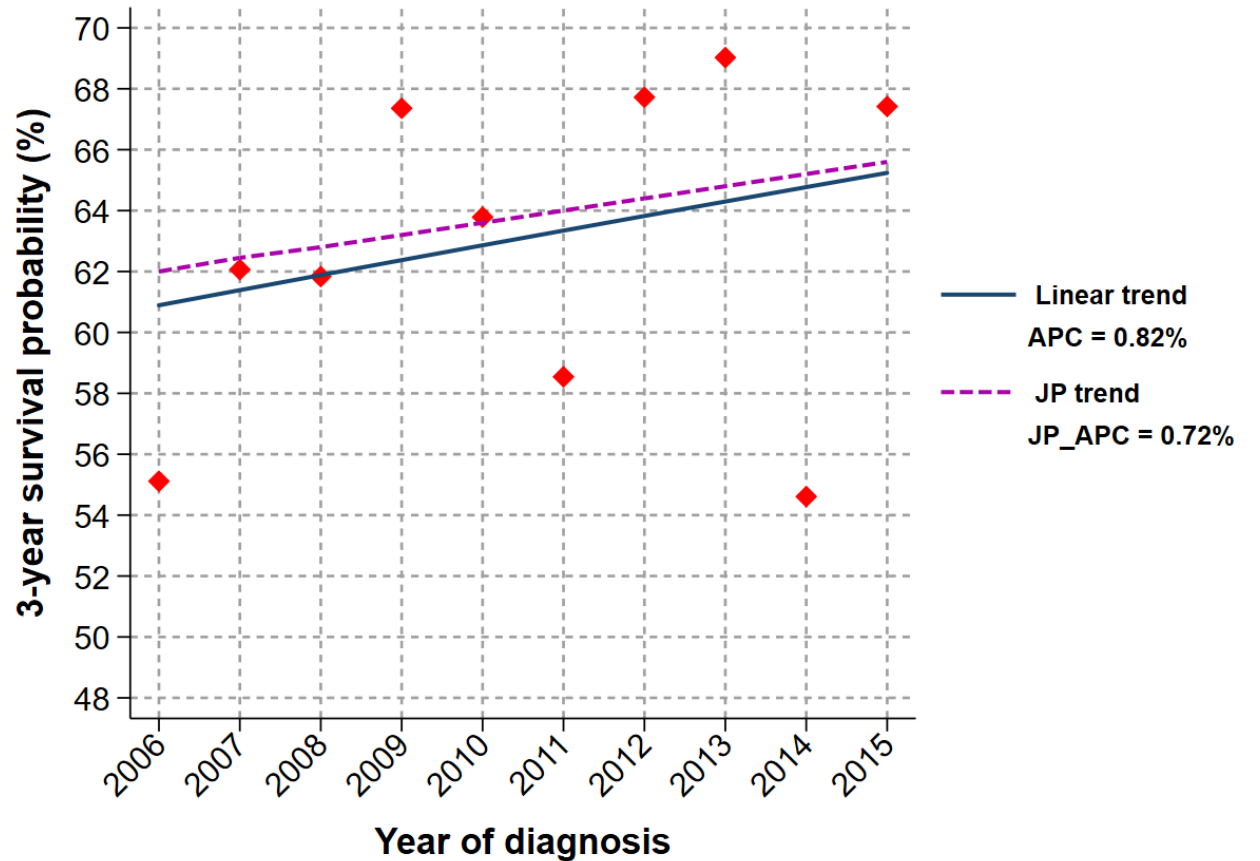
Si mostra solo il grafico del trend lineare della sopravvivenza relativa a 3 anni ottenuto applicando un modello di questo tipo.



Modello JP per il trend della sopravvivenza

Utilizzando il file con le proporzioni dei casi sopravvissuti diagnosticati in un particolare anno di incidenza (file SurvAnno) si può utilizzare anche una jointpoint regression per stimare il trend delle probabilità di sopravvivenza a un certo tempo dalla diagnosi.

Questo approccio, seppur utilizzato, non è appropriato perché la stima del trend della sopravvivenza richiede l'uso di dati individuali.

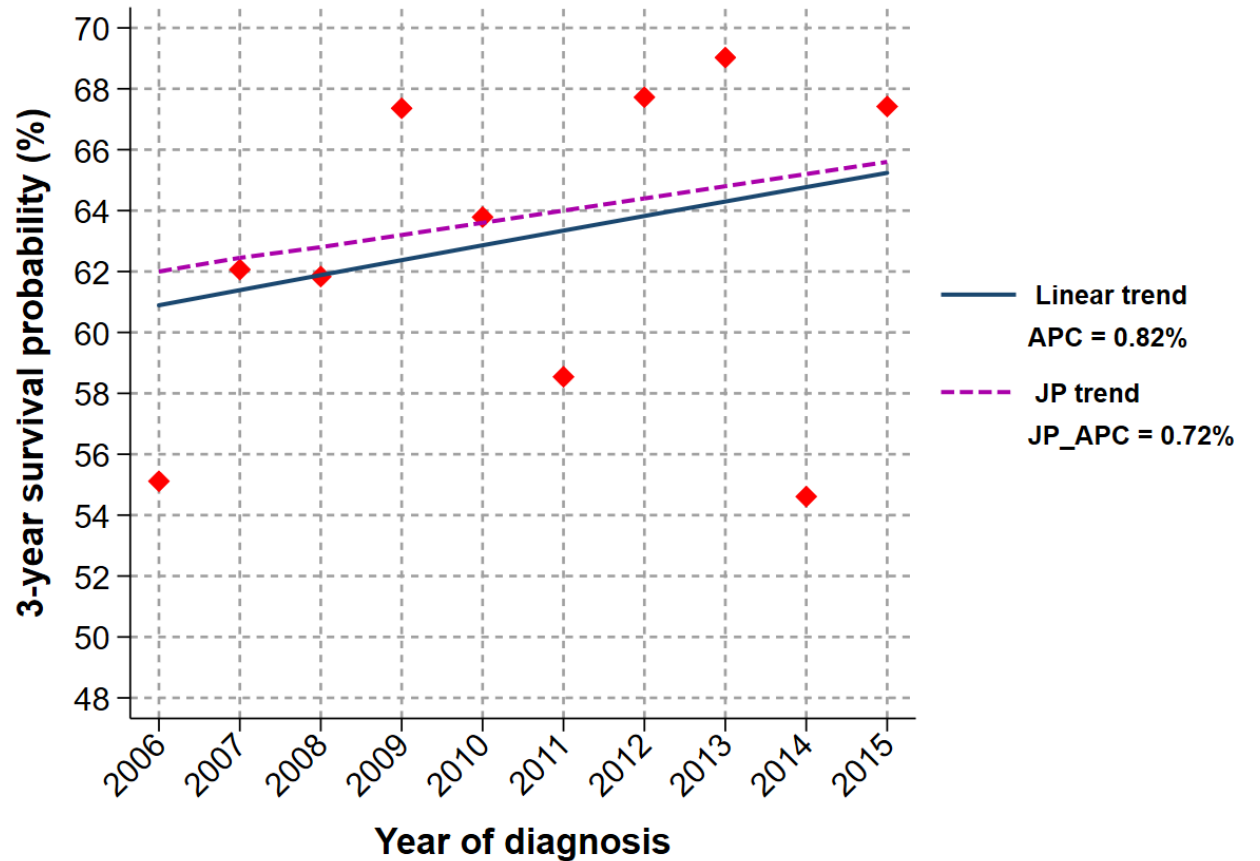


Modello JP per il trend della sopravvivenza

Il grafico a destra confronta i trend ottenuti con un modello parametrico flessibile che usa dati individuali e una jointpoint regression che utilizza le proporzioni di sopravvissuti a 3 anni dalla diagnosi (file SurvAnno).

Le rette dei trend lineari adattate dai due modelli sono diverse e quella corretta è la linea blu.

Il National Cancer Institute ha sviluppato **JPSurv**, uno specifico software che adatta una jointpoint survival regression con dati individuali, diversa dalla jointpoint regression adattata da Jointpoint con le proporzioni dei casi sopravvissuti.



Progresso nella lotta ai Linfomi Non Hodgkin. Anni 2006 - 2015

In questo esempio è stato stimato un incremento dell'incidenza di circa l'1.3% all'anno, un dato che indica un maggior carico di patologia. La mortalità resta stabile e coerentemente la sopravvivenza a 3 anni dalla diagnosi mostra un aumento di poco più dello 0.8% all'anno.

Una situazione simile è osservata per il cancro della tiroide e, in alcuni casi, per il cancro della prostata. L'interpretazione più accreditata è che questo dipenda dalla sovradiagnosi cioè dalla diagnosi di tumori ricercati con accentuata attenzione diagnostica, ma che non provocheranno mai disturbi clinici o la morte [*Nessun Progresso*].

Nel caso dei linfomi Non Hodgkin questa interpretazione non è ipotizzabile. Resta quindi da considerare l'ipotesi che in questa popolazione possa esservi un aumentato carico di patologia a cui però si associa un miglioramento delle capacità terapeutiche [*Progresso ma con maggior carico di malattia*].

Un quadro controverso che stimola a ricerche e approfondimenti

Conclusioni

La disponibilità di lunghe serie temporali di dati mette i registri nelle condizioni di analizzare le variazioni dei tre principali indicatori dell'epidemiologia dei tumori.

Questa analisi consente di interpretare se i cambiamenti verificatisi in periodo di tempo abbiano significato un reale (misurato a livello di popolazione) progresso nella nostra capacità di contrasto al cancro.

Importanti progressi si possono raggiungere con interventi di prevenzione che determinano consistenti riduzioni di incidenza e mortalità dei tumori, ma solo l'incremento della sopravvivenza indica se è migliorata la nostra capacità di cura del cancro.

Tuttavia, non sempre i miglioramenti della sopravvivenza sono interpretabili come progresso nella lotta contro il cancro né sono sempre sufficienti a determinare una riduzione della mortalità causata da un particolare tumore.

In ogni caso, l'analisi congiunta di incidenza, mortalità e sopravvivenza mette a disposizione un quadro più informativo e interpretabile di quello desumibile dall'analisi dei singoli indicatori e consente di definire meglio i motivi di attenzione e gli obiettivi su cui orientare le nostre future ricerche.



In conclusion, survival is an important cancer prognosis measure for clinicians and their patients.

To be correctly used as a cancer policy measure and to correctly understand temporal changes in survival, changes in incidence and mortality need to be considered simultaneously.

Improved survival does represent progress when accompanied by a reduction in mortality, incidence, or ideally both.

Hyunsoon Cho, Angela Mariotto, and coll.

Grazie

Rapporto di tassi rispetto a un anno di riferimento

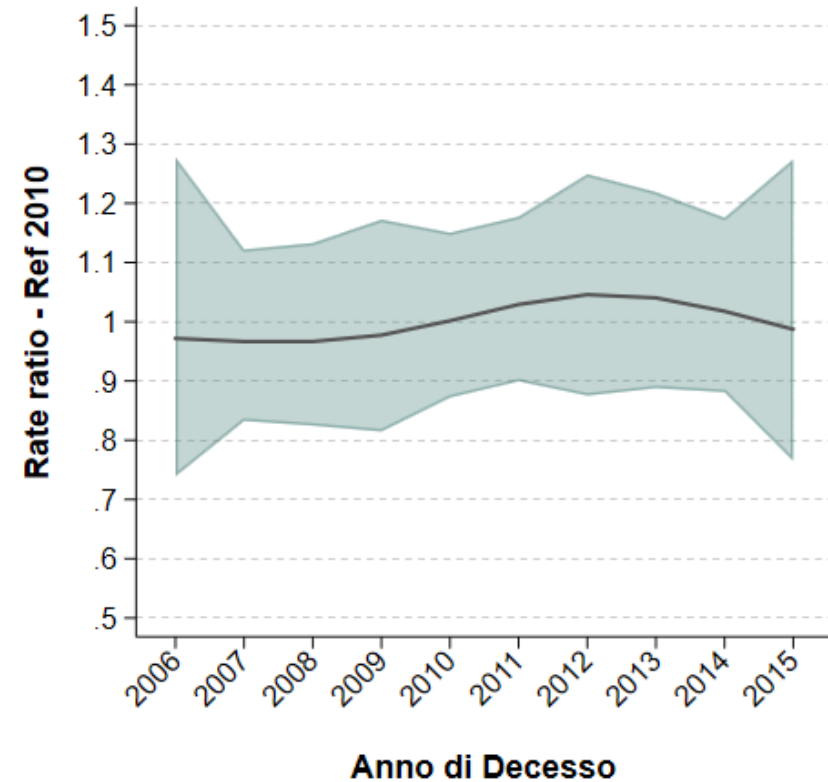
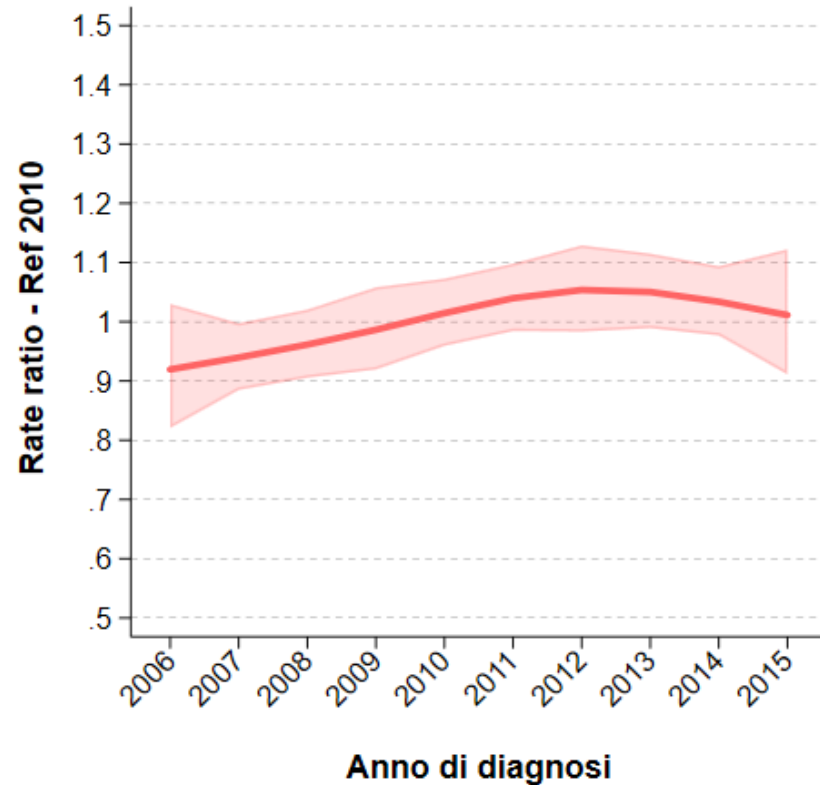
In presenza di spline non è possibile stimare una variazione costante del tasso nel corso del periodo indagato. Possiamo però stimare il rapporto tra il tasso di un anno e quello di un anno di riferimento

```
partpred RR_I, for(year_rcs*) ci(RR_Ilci RR_Iuci) eform ref(anno 2010)
```

Con questa istruzione Stata calcola il Rate Ratio che confronta il tasso adattato dal modello rispetto al tasso del 2010

Con l'opzione `-ci()` sono stati stimati anche gli intervalli di confidenza

Rapporto di tassi rispetto a un anno di riferimento



Questo grafico e il precedente mostrano un andamento in crescita dell'incidenza fino al 2012-2013 seguita da una leggera fino al 2015, anno in cui i tassi tornano agli stessi livelli del 2010.

La mortalità mostra lo stesso andamento ma con variazioni più limitate.