



L'appropriatezza del percorso di follow-up in pazienti affetti da polipo adenomatoso con displasia: un'analisi di sopravvivenza

Michele Gobbato, Laura Rizzi, Elena Clagnan,
Francesca Valent

Novembre 2015

7 / 2015

L'appropriatezza del percorso di follow-up in pazienti affetti da polipo adenomatoso con displasia: un'analisi di sopravvivenza

Michele Gobatto¹, Laura Rizzi¹, Elena Clagnan², Francesca Valent²

1 - Dipartimento di Scienze Economiche e Statistiche - Università degli Studi di Udine

2 - Direzione centrale salute, integrazione socio-sanitaria e politiche sociali, Regione FVG

E-mail: laura.rizzi@uniud.it

Abstract

Questo studio vede l'applicazione di strumenti statistici per l'analisi della sopravvivenza finalizzati alla valutazione dei fattori che influenzano il rischio di colonscopia inappropriata durante il percorso di follow-up conseguente alla diagnosi incidente di polipo adenomatoso con displasia. L'analisi è stata effettuata, in termini retrospettivi, sulla coorte di tutti i casi incidenti di polipo registrati nel biennio 2007-2008 nella popolazione residente in Regione Friuli Venezia Giulia. La coorte utilizzata nello studio, dopo l'esclusione dei deceduti e di coloro con diagnosi di cancro colonrettale durante periodo 2007-2011, comprende 7259 pazienti, osservati nei 5 anni di follow-up. Sui dati della coorte si sono derivate le curve di Kaplan-Meier e si è stimato il modello Cox Proportional-Hazard. I risultati dell'analisi evidenziano che i fattori predittivi di un rischio significativamente maggiore di essere sottoposti a colonscopia inappropriata risultano essere: il genere maschile, la fascia d'età 55-74 anni e la residenza nelle Aziende Socio Sanitarie indicate con A e B.

Keywords: modelli di sopravvivenza, studio retrospettivo, probabilità di colonscopia inappropriata, diagnosi di polipo adenomatoso con displasia.

1 Introduzione

L'utilizzo degli strumenti classici per l'analisi della sopravvivenza, quali le curve di Kaplan-Meier (Kaplan and Meier, 1958) e il modello di Cox (Cox, 1972) é ampiamente documentato nella disciplina medica (Armitage *et al.*, 2002).

Esistono però altri ambiti ed altri obiettivi di ricerca che si possono analizzare ricorrendo a queste procedure statistiche e che escono dal contesto classico in cui la sopravvivenza rappresenta il periodo di tempo trascorso in vita condizionatamente alla presenza di una patologia o al verificarsi di un evento. In questa direzione vi sono gli studi relativi all'occupazione e disoccupazione in Lancaster and Nickell (1980) e Royston (1983), all'inflazione, alla durata di un prodotto, alla fedeltà di un cliente (si veda Danacica and AG (2010); Kiefer (1988)) e, in generale, a tutti quegli studi che implicano la valutazione della durata di una condizione e che sono caratterizzati dalla possibilità che la singola unità statistica esca dall'osservazione, risultando quindi censored. Sono diversi, infatti, i settori in cui si trovano applicazioni alternative dei modelli di sopravvivenza: economico e finanziario (Gepp and Kumar, 2008; Schumway, 2001), assicurativo (Andersen and Fledelisu, 1998; Czado and Rudolph, 2002) o delle telecomunicazioni (Lu, 2002).

Nel contesto sanitario, ad esempio, l'applicazione di tecniche di analisi della sopravvivenza nella misurazione dei costi di trattamento costituisce un ottimo strumento per il governo sanitario (Dudley *et al.*, 1993; Hallstrom and Sullivan, 1998), purché non ci sia una violazione dell'assunzione di censoring indipendente o non informativo. Ciò implica che, per ottenere stime non distorte nell'analisi, è necessario che i casi censored non rappresentino particolari sotto gruppi caratterizzati da un diverso rischio di censoring, si veda a riguardo Etzioni *et al.* (1999). Se tale assunzione è rispettata l'utilizzo di questi strumenti può considerarsi appropriato in relazione agli obiettivi dell'analisi. I vantaggi che queste tecniche permettono sono notevoli sia dal punto di vista descrittivo del campione su cui si effettuano le stime, che predittivo sulla popolazione di riferimento. Infatti, si possono effettuare confronti molto accurati e utilizzare rappresentazioni grafiche estremamente esplicative e di facile interpretazione. Inoltre il trattamento della variabile temporale

e la possibilità di effettuare delle stime che considerano i casi persi dall'osservazione permette la valutazione delle performance attraverso il monitoraggio di pazienti e, più in generale, l'analisi in contesti in cui l'osservazione di un fenomeno in un arco temporale prestabilito rappresenta il criterio alla base della valutazione.

Bailey *et al.* (2008), inoltre, ha utilizzato lo stimatore di Kaplan-Meier (KM) e il modello di Cox per effettuare previsioni sui tassi di controllo dell'ipertensione stratificando per età e genere partendo da database amministrativi e osservando due popolazioni di pazienti da settembre 2001 a luglio 2003. Tassi di controllo significativamente diversi a seconda del genere dei pazienti e delle classi di età considerate sono stati rilevati nello studio.

In ambito ortopedico gli stessi strumenti sono stati applicati da Carr *et al.* (1993) per valutare la probabilità di sopravvivenza di un impianto di protesi per un determinato periodo di tempo. In questi studi, quindi l'evento considerato non è quindi il decesso della persona, ma la sostituzione di una protesi. Il fenomeno che approfondiamo in questo studio è relativo al monitoraggio, mediante esame di colonscopia, in pazienti affetti da polipo adenomatoso con displasia. Le linee guida internazionali prevedono un monitoraggio stratificato in base al rischio: si prevede una colonscopia ogni 3 anni in caso di adenoma avanzato, ogni 5-10 anni in caso di adenoma non avanzato (Liebermann *et al.*, 2012; Winawer *et al.*, 2006). Rimane comunque aperta la discussione sulla possibilità di estendere il follow-up oltre i 5 anni in Chung *et al.* (2011) e Miller *et al.* (2010)). Tuttavia la letteratura mette in luce come nella pratica comune le linee guida non sempre vengono rispettate, sia a livello internazionale (Schoen *et al.*, 2010) che nel territorio italiano (Radaelli *et al.*, 2012). Un esubero di coloscopie effettuate, rispetto alle linee guida, emerge anche da studi basati sulla compilazione di questionari da parte dei gastroenterologi (Mysliwiec *et al.*, 2004) in cui si evidenzia anche una parziale carenza nella conoscenza delle linee guida stesse (Saini *et al.*, 2009). L'obiettivo di questo studio è l'analisi, attraverso le tecniche di sopravvivenza, della presenza di coloscopie inappropriate in una coorte di pazienti affetti da polipo adenomatoso con displasia, seguiti nei 5 anni di follow-up dalla data incidente di diagnosi. Seguendo il nostro approccio si possono ottenere stime che tengono in considerazione il fattore temporale, de-

scrivendo come si evolve nel tempo la probabilità che i pazienti della coorte siano sottoposti a colonscopia non appropriata, controllando per altri fattori individuali e contestuali, e assumendo comunque non distorsione delle stime causata da casi censored non casuali.

2 Materiali e metodi

2.1 I dati utilizzati

In questa ricerca si è effettuato uno studio di coorte retrospettivo basato su database amministrativi. L'anonimato dei pazienti è stato garantito attraverso un codice identificativo univoco anonimo alfanumerico che cambia periodicamente, non permettendo in questo modo il riconoscimento della persona, e che permette comunque il linkage tra le diverse tabelle del Sistema Informativo Socio Sanitario Regionale del Friuli Venezia Giulia (Sissr).

La coorte analizzata è costituita da soggetti residenti in Regione Friuli Venezia Giulia (FVG) a cui, negli anni 2007 e 2008, è stata riscontrata una diagnosi incidente di polipo adenomatoso con displasia. I pazienti sono stati selezionati dai registri di anatomia patologica (ApSys) considerando i codici SNOMED tra 59000 e 59992 per l'asse topografica (T) e tra 72000 e 99701, con l'ultimo carattere uguale a 0 o 1, per l'asse morfologico (M). Sono stati poi esclusi dalla coorte, attraverso l'utilizzo di altre fonti dati, i casi di cancro colonrettale relativi al periodo 2007 al 2011 e successivi alla data di individuazione del polipo. In particolare sono stati esclusi in una prima fase i casi presenti nel registro tumori con codice ICD9-CM uguale a 153* o 154*. Per gli anni non coperti dal registro tumori, considerando quale fonte la Scheda di Dimissione Ospedaliera (SDO), sono stati estratti i pazienti ricoverati con codici di diagnosi, primaria o secondaria, ICD9-CM analoghi a quelli utilizzati in precedenza. L'ultimo passaggio di esclusione ha richiesto nuovamente l'utilizzo dei database di anatomia patologica e ha comportato l'eliminazione dei casi con diagnosi di cancro al colon. Infine, per garantire i 5 anni di osservazione successiva alla data incidente sono stati esclusi i pazienti deceduti nel periodo. Le

variabili anagrafiche dei pazienti considerate nello studio sono: l'età, il genere e l'Azienda per i Servizi Sanitari (ASS) di residenza al momento dell'individuazione del polipo.

I pazienti con diagnosi di polipo adenomatoso con displasia, individuati nei due anni di riferimento, erano inizialmente 9515, ma 652 sono stati esclusi dallo studio per decesso nel periodo di follow-up mentre 680 perchè al momento dell'evento incidente non risultavano residenti in regione FVG. Successivamente si sono esclusi 924 casi per i quali è stato registrato un cancro colonrettale nei 5 anni successivi alla data incidente. La coorte finale analizzata è stata incrociata alle basi dati delle prestazioni ambulatoriali individuando, attraverso i codici del nomenclatore regionale, le colonscopie effettuate nei 5 anni di follow-up.

2.2 Gli strumenti statistici

Le curve di Kaplan-Meier sono state utilizzate per rappresentare l'evoluzione nel tempo della probabilità di essere sottoposti ad una colonscopia inappropriata, in base alle variabili individuali dei soggetti. E' stato poi stato stimato il modello Cox Proportional-Hazard, e i relativi Hazard Ratio (HR), per valutare il rischio di esami endoscopici all'interno del periodo di follow-up.

Il modello di Cox utilizzato nello studio è rappresentato dalla seguente formula:

$$h(t, X) = h_0(t) \exp \sum_{i=0}^p \beta_i x_i$$

dove $i = 1, 2, 3$, la funzione h_0 è chiamata baseline hazard ed è tempo-dipendente. Inoltre, la matrice X è stata costruita considerando le seguenti variabili esplicative:

- x_1 variabile categorica che indica la classe di età del paziente.
- x_2 variabile indicatrice per il genere.
- x_3 variabile categorica che indica l'azienda socio sanitaria di residenza.

L'HR viene definito come il rapporto degli hazard stimati dal modello per le

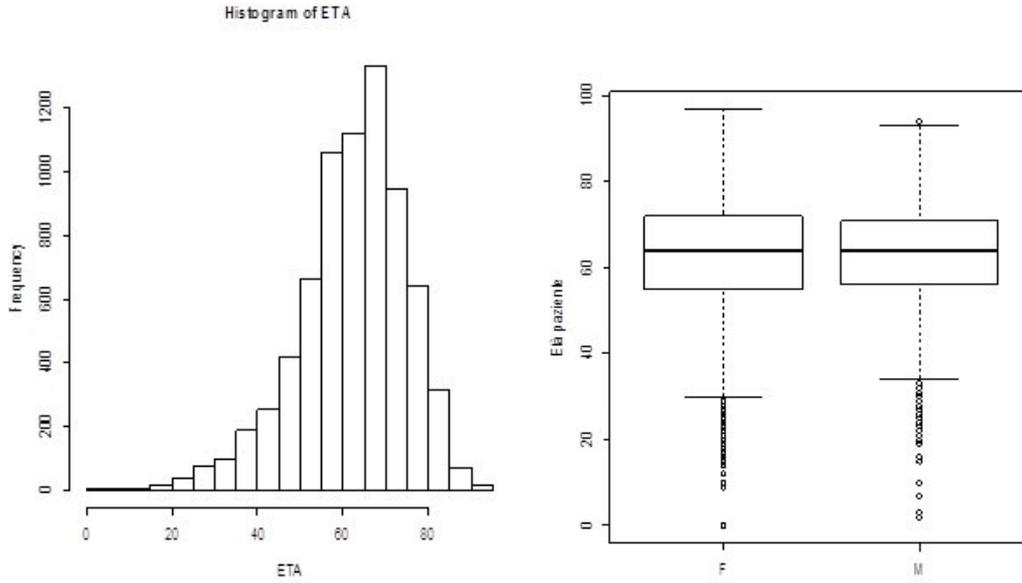


Figure 1: Distribuzione della coorte in base all'età e boxplot dell'età per genere del paziente.

diverse modalità di ogni variabile categorica e la modalità di riferimento.

$$\begin{aligned}
 HR &= \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \frac{\hat{h}_0 \exp \sum_{i=0}^p \hat{\beta}_i x_i^*}{\hat{h}_0 \exp \sum_{i=0}^p \hat{\beta}_i x_i} \\
 &= \exp \sum_{i=0}^p \hat{\beta}_i (x_i^* - x_i)
 \end{aligned}$$

3 I risultati

La coorte finale comprende 7259 pazienti con età media pari a 62 anni (mediana=64) e un numero di donne pari a 2939, rappresentanti il 40% della coorte. Il test t di Student non ha evidenziato una differenza statisticamente significativa in termini di età tra i maschi e le femmine (p-value 0.21), riportati in 1. Le classi di età sono state create sulla base dei quartili della distribuzione e le frequenze ottenute sono rappresentate nella figura 2.

Le frequenze assolute nelle diverse classi d'età sono le seguenti: 1778 per la classe 0-54, 2219 per la classe 55-64, 2320 nei casi con età compresa tra 65 e 74 anni e 1073 con 75 anni o più.

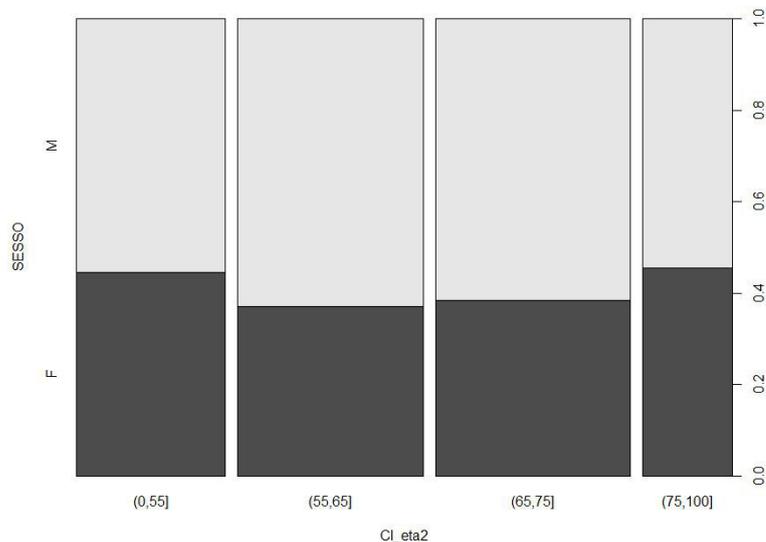


Figure 2: Frequenze relative per classe d'età e genere.

Dei 7259 pazienti che compongono la coorte 1789 hanno eseguito una colonoscopia nei 5 anni di osservazione (24%). L'analisi attraverso gli strumenti di sopravvivenza permette in primo luogo di valutare la probabilità che una persona affetta da polipo adenomatoso con displasia superi i 5 anni di follow-up senza incorrere in esami coloscopici inappropriati.

Come si osserva in figura 3 la probabilità di superare il follow-up senza incorrere in una colonoscopia inappropriata risulta circa pari all'80%. Il fenomeno risulta eterogeneo a tra le Aziende Sanitarie territoriali (Ass) di residenza del paziente. Infatti, dalle curve riportate in figura 3, è possibile identificare 2 gruppi con diverso rischio determinati dalla variabile territoriale. Un gruppo a basso rischio di essere sottoposti ad esame inappropriato (i soggetti residenti nelle Ass indicate con F, C, D) e un gruppo a alto rischio (i soggetti residenti nelle Ass indicate con A, B). La curva relativa all'Ass E si posiziona in posizione intermedia rispetto alle curve due gruppi, per capire a quale gruppo appartiene è necessario valutare i risultati del modello di Cox. L'azienda A, inoltre, risulta caratterizzata da una probabilità più alta che i pazienti in essa residenti vengano sottoposti a un esame coloscopico en-

Variabile	Coef	HR	HR - IC (95%)	p-value
<i>Classe d'età</i>				
0-54		1.00		
55-64	0.35	1.42	1.25 – 1.62	1.33×10^{-7}
65-74	0.31	1.36	1.20 – 1.55	3.04×10^{-6}
Over 75	-0.05	0.95	0.80 – 1.12	0.544
<i>Genere</i>				
Femmine		1.00		
Male	0.10	1.11	1.00 – 1.22	0.039
<i>Azienda di residenza</i>				
A		1.00		
B	-0.09	0.92	0.75 – 1.11	0.3702
C	-0.82	0.44	0.36 – 0.54	9.66×10^{-15}
D	-0.80	0.45	0.38 – 0.52	$< 2 \times 10^{-16}$
E	-0.35	0.71	0.59 – 0.84	0.0001
F	-1.08	0.34	0.29 – 0.40	$< 2 \times 10^{-16}$

Table 1: Risultati della stima del modello di Cox

tro un anno dall'individuazione del polipo. Questo risultato è visibile dalla discesa rapida della funzione di sopravvivenza nel primo periodo d'osservazione, la curva relativa presenta infatti un andamento che decresce molto velocemente. Anche per quanto riguarda l'età si individuano due gruppi di rischio. Le classi d'età sotto i 55 e sopra i 74 anni presentano una probabilità significativamente più bassa di incorrere nell'esame inappropriato, mentre i pazienti con età compresa tra 55 e 74 anni rappresentano invece il gruppo ad alto rischio. L'analisi con il modello di Cox 1 quantifica le differenze descritte in precedenza. Non si riscontrano differenze statisticamente significative negli Hazard Ratio tra le classi d'età 0-54 e 75+, così come non si riscontrano per le aziende A e B. Questo conferma la presenza dei gruppi a diverso rischio già identificati dalle analisi delle curve di Kaplan-Meier. L'azienda E, che non era facilmente valutabile dall'analisi grafica, si posiziona nel gruppo a basso rischio (HR=0.71 rispetto all'Ass A).

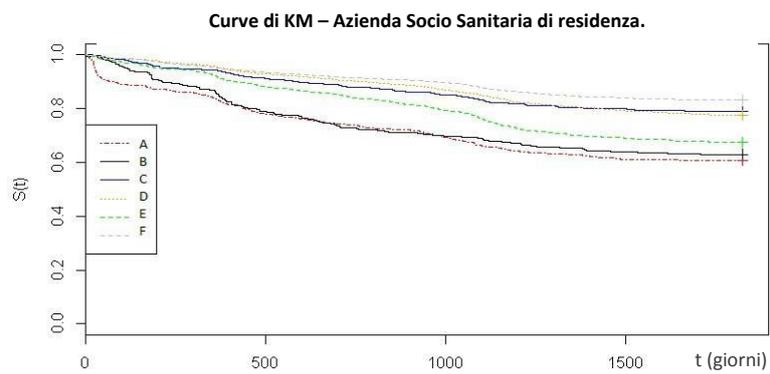
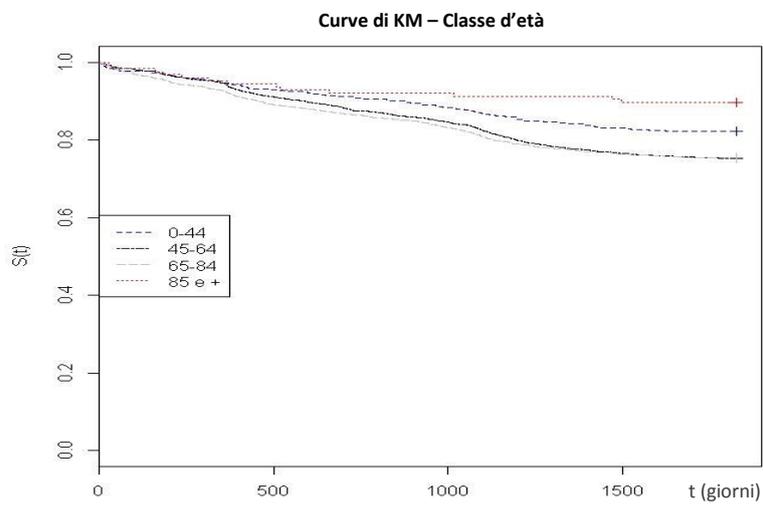
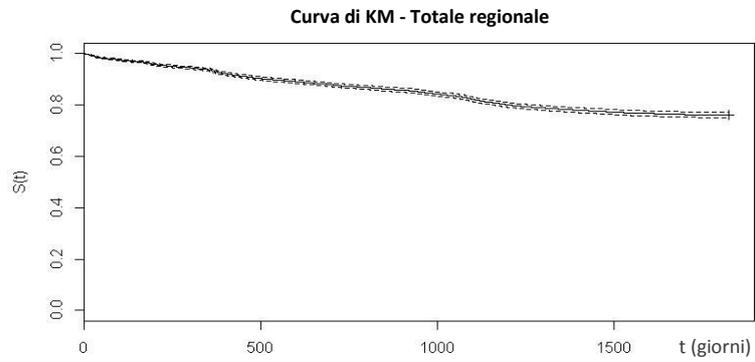


Figure 3: Curve di Kaplan-Meier sulla probabilità di colonscopia inappropriata.

4 Conclusioni

L'analisi di sopravvivenza si dimostra efficace per la valutazione in ambito regionale del fenomeno dell'inappropriatezza delle coloscopie durante il follow up della coorte dei casi incidenti di polipo adenomatoso con displasia. Si conferma quindi l'utilità nell'utilizzo degli strumenti di sopravvivenza in ambiti di ricerca come quello da noi considerato, aspetto evidenziato inizialmente considerando gli studi effettuati in contesti e discipline eterogenei.

L'analisi ha evidenziato, innanzitutto, che la quota di pazienti affetti dalla patologia considerata, identificata negli anni 2007-2008, e sottoposti a una coloscopia inappropriata durante i 5 anni di follow-up è pari al 24%. Le caratteristiche individuali dei pazienti sono fattori predittivi del rischio di esame inappropriato, infatti i maschi presentano una probabilità di essere sottoposti a esame inappropriato pari al 10%, superiore rispetto a quella delle donne. L'Hazard Ratio nei pazienti con età compresa tra 55 e 64 anni è risultato maggiore rispetto a quello dei pazienti più giovani (meno di 54) o più anziani (oltre i 75), con una differenza statisticamente significativa. Le due classi d'età 55-64 e 65-74 formano quindi un gruppo a maggior rischio di inappropriatezza, con una probabilità media di esami inadeguati di circa il 40% superiore ai pazienti delle altre due classi.

Dall'analisi dei risultati si riscontra un'effetto della componente geografica. Si individua infatti un'eterogeneità territoriale: le persone residenti nelle Aziende Sanitarie indicate con A e B presentano una probabilità significativamente superiore rispetto a quelle residenti nelle altre aziende.

I risultati del modello di Cox 1 ci dicono infatti che, nei cinque anni di osservazione, la probabilità di incorrere in una coloscopia precedente al periodo di follow-up nelle altre aziende è notevolmente inferiore, in un caso anche del 66% (azienda F). Le curve di Kaplan-Meier permettono di rappresentare la storia del fenomeno nei 5 anni di osservazione rappresentando l'uscita dei soggetti nei diversi istanti temporali (giorni), mettendo in evidenza le criticità. Questo permette di constatare come le aziende A e B arrivino a fine periodo di osservazione con un numero inferiore di pazienti che conclude regolarmente il periodo di follow up. Inoltre si riscontra una discesa molto veloce della curva relativa ai residenti nell'azienda A nei primi

mesi successivi all'evento. L'utilizzo di tale strumento grafico si può rivelare molto utile per gli organi decisionali, sia a livello regionale che di direzione aziendale. Identificare un'eterogeneità territoriale, seppur per azienda di residenza, consente infatti interventi sull'azienda territoriale per uniformare la realtà regionale verso gli standard riscontrati dalla best practice in coerenza con le linee guida internazionali (Azienda F). Lo studio necessita ulteriori approfondimenti sull'eterogeneità territoriale, concentrando l'attenzione sulle Aziende con un rischio di inappropriata più elevato e considerando anche l'attività dei medici prescrittori degli esami durante il follow up.

References

- Andersen, C. and Fledelisu, P. (1998) The cox regression model for clais data in non-life insurance. **28**, 95–118.
- Armitage, P., Berry, G. and Mattheows, J. (2002) *Statistical Methods in Medicine Research*. Blackwell Science, Oxford.
- Bailey, K., Frossard, B. and Graves, J. (2008) Novel use of kaplan-meier methods to explain age and gender differences in hypertension control rates. *Hypertension*, **51**, 841–847.
- Carr, A., Morris, R., Murray, D. *et al.* (1993) Survival analysis in joint replacement surgery. *The Journal of Bone and Joint Surgery*, **75**, 178–182.
- Chung, S., Kim, Y., Yang, S. *et al.* (2011) Five-year risk for advanced colorectal neoplasia after initial colonoscopy according to the baseline risk stratification: a prospective studio in 2452 asymptomatic koreans. *Gut*, **60**, 1537–1543.
- Cox, D. (1972) Regression models and life tabls (with discussion). *Journal of the Royal Statistical Society. Series B.*, **34**, 187–220.
- Czado, C. and Rudoplh, F. (2002) Application of survival analysis methods to long term care insurance. *Insurance: Mathematics and Economics*, **31**, 395–413.

- Danacica, D. and AG, B. (2010) Using survival analysis in economics. *Analele Stiintifice ale Universitatii "Alexandru Ioan Cuza" din Iasi*, **57**, 439–450.
- Dudley, R., Harrell, F. J., Smith, L. *et al.* (1993) Comparison of analytic model for estimating the effect of clinical factors on the cost of coronary artery bypass graft surgery. *Journal of Clinical Epidemiology*, **46**, 261–271.
- Etzioni, R., Feuer, E., Sullivan, S. *et al.* (1999) On the use of survival analysis techniques to estimate medical care costs. *Journal of Health Economics*, **18**, 365–380.
- Gepp, A. and Kumar, K. (2008) The role of survival analysis in financial distress prediction. *International Research Journal of Finance and Economics*, **16**.
- Hallstrom, A. and Sullivan, S. (1998) On estimation costs for economic evaluation in failure time studies. *Med care*, **36**, 433–436.
- Kaplan, E. and Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.
- Kiefer, N. (1988) Economic duration data and hazard function. *Journal of Economic Literature*, **26**, 646–679.
- Lancaster, N. and Nickell, S. (1980) The analysis of re-employment probabilities for the unemployed. *Journal of the Royal Statistical Society. Series A.*, **143**, 141–165.
- Libermann, D., Rex, D., Winawer, S. *et al.* (2012) Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the us multi-society task force on colorectal cancer. *Gastroenterology*, **143**, 844–857.
- Lu, J. (2002) Predicting customer churn in telecommunications industry an application of survival analysis modeling using sas. *SAS User Group International (SUGI) Paper*, 114–127.
- Miller, H., Mukherjee, R., J, T. *et al.* (2010) Colonoscopy surveillance after polypectomy may be extended beyond five years. *J Clin Gastroenterol*, **44**, 162–166.

- Mysliwiec, P., Brown, M. and Klabunde, C. (2004) Are physicians doing too much colonoscopy? a national survey of colorectal surveillance after polypectomy. *Ann Intern Med*, **141**, 264–271.
- Radaelli, F., Paggi, S., Bortoli *et al.* (2012) Overutilization of post-polypectomy surveillance colonoscopy in clinical practice: a prospective, multicentre study. *Dig Liver Dis*, **44**, 748–753.
- Royston, G. (1983) Wider application of survival analysis: An evaluation of an unemployment benefit procedure. *Journal of the Royal Statistical Society. Series D (The Statistician)*, **32**, 301–306.
- Saini, S., Nayak, R., Kuhn, L. and Schoenfeld, P. (2009) Why don't gastroenterologists follow colon polyp surveillance guidelines? results of a national survey. *J Clin Gastroenterol*, **43**, 554–558.
- Schoen, R., Pinsky, P., Weissfed, J. *et al.* (2010) Utilization of surveillance colonoscopy in community practice. *Gastroenterology*, **138**, 73–81.
- Schumway, T. (2001) Forecasting bankruptcy: A simple hazard model. *The Journal of Business*, **74**, 101–124.
- Winawer, S., Zauber, A., Fletcher, R. *et al.* (2006) Guidelines for colonoscopy surveillance after polypectomy: a census update by the US multi-society task force on colorectal cancer and the American Cancer Society. *Gastroenterology*, **130**, 1872–1875.