Survivalanalyse bei mehreren Todesursachen

20. Juni 2017

Amelie Forkel

Gliederung

Einleitung

- Umfangreiche elektronische medizinische Datenquellen
- Weiterentwicklung wissenschaftlicher Fragestellungen
  - Tod aufgrund <u>einer</u> bestimmten Ursache
  - Zusammenhang mit möglichen anderen Ursachen
- Bisher nur Ansätze mit nur einer Todesursache
- Langjährigen Bemühungen um internationale Standards für „cause of death“ Daten

Einleitung

Sterbeurkunde

Cause of death		Approximate interval between onset and death
I Disease or condition directly leading to death *)	a) Lungenembolie due to (or as a consequence of)	Minuten
Antecedent causes Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last	b) Beinvenenthrombose due to (or as a consequence of)	Wochen
	c) Unbeweglichkeit due to (or as a consequence of)	Wochen
	d) Oberschenkelbruch	Wochen
II Other significant conditions contributing to the death, but not related to the disease or conditions causing it	Osteoporose	Jahre

**This does not mean the mode of dying, e.g. heart failure, respiratory failure. It means the disease, injury, or complication that caused death.*

Gliederung

Vorüberlegungen

Ziele und Ansprüche an die Methoden

Ziel 1:

Quantifizieren der Mortalität, die jeder Krankheit zuzuordnen ist

Ziel 2:

Quantifizieren der Wirkung bestimmter Faktoren auf die von jeder Krankheit ausgehenden Mortalitätskraft (egal, ob andere separate Krankheiten zeitgleich vorhanden sind, oder nicht)

Vorüberlegungen

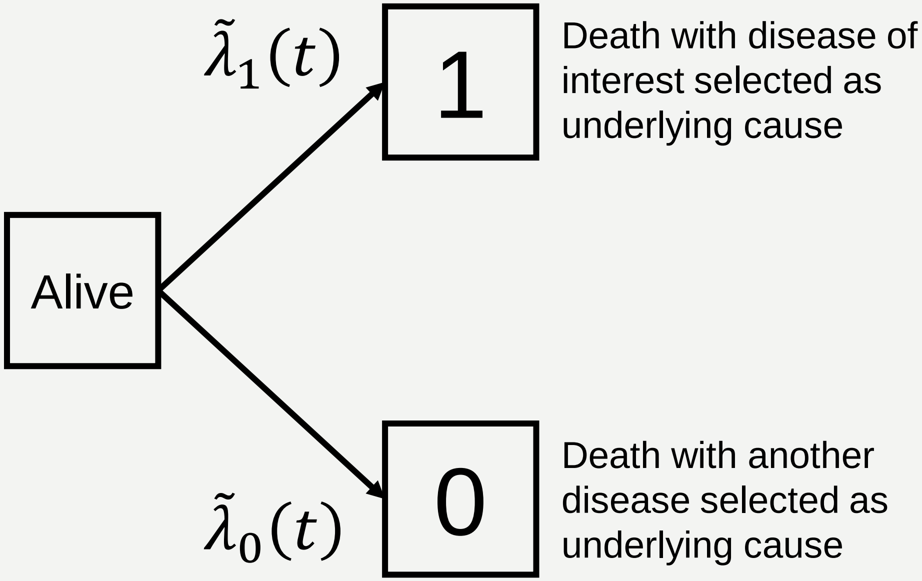
Ziele und Ansprüche an die Methoden

Die Methode sollte...

1. ... die multiplen Krankheiten, die zum Tod beitragen, so einbeziehen, wie sie auf der Sterbeurkunde stehen.
2. ... jeden Tod nur einmal zählen, ungeachtet dessen, wie viele Krankheiten auf der Sterbeurkunde vermerkt sind.
3. ... die relative Bedeutung jeder Krankheit für das Eintreten des Todes widerspiegeln.

Single-cause

Jeder Tod wird komplett einer Krankheit zugeschrieben (underlying cause)



$\tilde{\lambda}_k$:

- Zustandsspezifischer hazard zur Zeit $t > 0$ für einen Zustand $k \epsilon \{0,1\}$
- Unmittelbare Hazardrate für den Übergang in einen der Zustände

Vorüberlegungen
Ansatz II

Any-mention

- Jede Nennung der interessierenden Krankheit auf der Sterbeurkund wird als Ereignis betrachtet
- Statistische Einheiten sind hier Einträge in der Sterbeurkunde, anstatt Todesfälle

  → Erschwert die Interpretation der Schätzer

Vorüberlegungen

Ansatz III

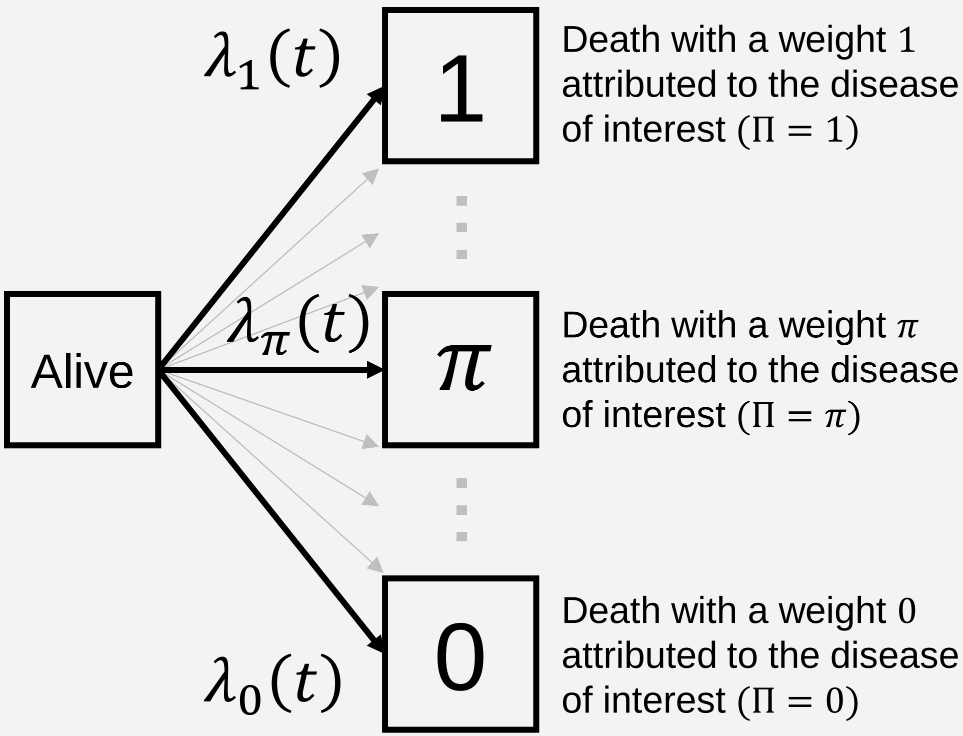
Multiple-cause

- Jeder Krankheit wird ein positives Gewicht zugewiesen, so dass die Summe der Gewichte eins ist
- Alle genannten Krankheiten werden in die Analyse aufgenommen
- Idealerweise spiegelt das zugewiesene Gewicht jeder Krankheit ihren Beitrag zum Tod wider

Vorüberlegungen

Ansätze III

Multiple-cause



$\lambda_\pi(t)$: Unmittelbare Hazardrate

Vorüberlegungen

Strategien der Gewichtszuweisung

Idee eines Satzes von Gewichten, die die Anforderungen erfüllen und alle Informationen aus der Sterbeurkunde nutzen

ABER:

- trotzdem nur Annäherung an die tatsächlichen kausalen Zusammenhänge, die zum Tod führen
- subjektiv

→ internationaler Konsens

Vorüberlegungen

Strategien der Gewichtszuweisung

Regel 1: $\Pi = 1$

Wenn KH = underlying cause & keine weiteren KHen in Part II genannt sind

Regel 2: $\Pi = \omega$

Wenn KH = underlying cause & und weitere KHen in Part II genannt sind

Regel 3: $\Pi = (1 - \omega)/(k_2 + 1)$

Wenn KH ≠ underlying cause, aber zusammen mit $k_2 \geq 2$ anderen KHen in Part II genannt ist

Regel 4: $\Pi = 0$

Wenn KH entweder gar nicht, oder in Part I, aber nicht als underlying cause genannt ist

Vorüberlegungen

Lösungsansätze

<u>Ziel 1:</u>

Totale Mortalität der interessierenden KH kann mit Hilfe der Bruchteile der Tode, die dieser KH zugeschrieben sind, berechnet werden:

$$\frac{\sum_{\pi>0} \pi N_\pi}{Total\ person - time - at - risk}$$

$Total\ person - time - at - risk$:
gesamte Zeit, die alle Individuen der Risikogruppe angehören

N_π:
Anzahl der Tode für die $\Pi = \pi$ ist

Vorüberlegungen

Lösungsansätze

<u>Ziel 2:</u>

- Pure hazards: $\lambda_1(t),\ \lambda_0(t)$

- Mögliche Annahme für π, $(0 \leq \pi \leq 1)$:

$$\lambda_\pi(t|X,Z) = \pi\lambda_1(t|X) + (1-\pi)\lambda_0(t|Z) \quad \text{(A0)}$$

- Fitten von Regressionsmodellen unter der Annahme (A0) sichert, dass alle „Tode“ Informationen zur Schätzung des Effektparameters ρ beisteuern

Gliederung

Cox-Regression
Daten

Für jedes Individuum:

1. Minimum aus Zeit-bis-zum-Tod und rechtszensierter Zeit
   $\widetilde{T}_i = \min\{T_i, C_i\}$
2. Zensurindikator $U_i = 1\ (C_i \leq T_i)$
   $U_i = 0\ (C_i > T_i) \rightarrow \pi_i$
3. $\mathrm{X_i}, \mathrm{Z_i}$

Notwendigkeit einige Individuen in den reinen Zuständen (1 oder 0) zu untersuchen

Cox-Regression
Modell

- iid Daten
- Unabhängige Rechtszensierung
- Beide pure hazards folgen dem Cox-Modell & sind verwandt

$$\lambda_1(t|X) = \lambda_{10}(t)\exp(\rho'X) \qquad \text{(A1)}$$

$$\lambda_0(t|X) = \lambda_{00}(t)\exp(\phi'Z) \qquad \text{(A2)}$$

$$\lambda_{00}(t|X) = \lambda_{10}(t)\exp(-\xi(t)) \qquad \text{(A3)}$$

$$\overset{(A0)}{\Longrightarrow} \lambda_\pi(t|X,Z) = \lambda_{10}(t)\pi[exp\ (\rho'X) + (1-\pi)\exp(-\xi(t) + \phi'Z]$$

- Zielparameter ρ: Effekt von X auf den pure hazard
- $\xi(t)$: log-ratio der pure baseline hazards

Gliederung

Theorie

Zustandsspezifische Häufigkeitsfunktion

Multiple-cause Modell:

- $F_\pi(t)$ für $t > 0$ und $0 \leq \pi \leq 1$:

  Wahrscheinlichkeit für den Übergang in den Zustand π nach Zeit t

- Unter Annahme (A0) bestimmt durch: $\lambda_1(t|X)$ und $\lambda_0(t|X)$

- $$\begin{aligned} F_\pi(t|X,Z) &= \int_0^t S(u|X,Z)\,\{\pi\lambda_1(u|X) + (1-\pi)\lambda_0(u|Z)\}du \\ &= \pi F_1(t|X,Z) + (1-\pi)F_0(t|X,Z) \end{aligned}$$

  Mit $S(t) = P(T \geq t)$ Survivalfunktion und F_0 und F_1 reine kumulierte Häufigkeitsfunktionen

Theorie
Schätzgleichungen

- Parameterschätzung für das Cox-Modell für pure hazards, basierend auf den Annahmen (A0)-(A3)

- Annahme: $\xi(t) = \xi$ konstant

- ξ bekannt: Schätzung von ρ und ϕ basierend auf partieller Likelihood

$$L(\rho, \phi, \xi) = \prod_{\pi \epsilon \{\pi^{(1)} \ldots \pi^{(J)}\}} \prod_{i \epsilon D_\pi} \frac{\pi \exp(\rho' X_i) + (1-\pi)\exp(-\xi + \phi' Z_i)}{\sum_{j \epsilon R(\widetilde{T_i})} \pi \exp\left(\rho' X_j\right) + (1-\pi)\exp(-\xi + \phi' Z_j)}$$

Wobei $\pi^1 := 1, \pi^2, \ldots, \pi^J \coloneqq 0$ die eindeutigen Werte von π sind, die in der Studie beobachtet wurden

Theorie

Schätzgleichungen

- ξ unbekannt: Schätzung von ξ mit Likelihood geschätzt, die sich jedem Ereignistyp anpasst

$$L^*(\rho, \phi, \xi) = \prod_{\pi \epsilon \{\pi^{(1)} \ldots \pi^{(J)}\}} \prod_{i \epsilon D_\pi} \frac{\pi \exp(\rho' X_i) + (1-\pi)\exp(-\xi + \phi' Z_i)}{\sum_{j \epsilon R(\widetilde{T_i})} \tilde{\pi} \exp\left(\rho' X_j\right) + (J - \tilde{\pi})\exp(-\xi + \phi' Z_j)}$$

Wobei $\tilde{\pi} := \sum_{j=1}^{J} \pi^{(j)}$ ist und der Nenner aus (A0)- (A3) und der absoluten Hazardrate

$$\lambda(u|X_i, Z_i) = \tilde{\pi}\lambda_1(u|X_i) + (J - \tilde{\pi})\lambda_0(u|Z_i) \text{ folgt}$$

- Schätzungen für $\Theta = (\xi, \rho, \phi)$ durch Iterieren zwischen Maximierungen von L, wenn ξ bekannt, und L^*, wenn ρ und ϕ bekannt

Gliederung

Beispielstudie

Ziel: Messen relativer Bildungsungleichheiten in der krankheitsbedingten Mortalität

- Kohorte von n= 148.384 Männern, $\geq$ 30 Jahre (01.01.2000)
- Kovariable:Sozioökonomischer Rang:

  $0 =$ „am meisten gebildet“ bis $1 =$ „am wenigsten gebildet“
- Cox-Modelle separat für jede Krankheitsgruppe
- Interessierender Parameter: relativer Index der Ungleichheit
- Zeitskala: Alter

Beispielstudie

Disease Group[a]	Any Mention	Single Cause $\omega = 1$	$\omega = 0.75$	Multiple Cause $\omega = 0.5$
Neoplasms	1.8 (1.6, 1.9)	1.8 (1.6, 2.0)	1.7 (1.6, 1.9)	1.7 (1.6, 1.9)
Cardiovascular	2.0 (1.9, 2.2)	2.1 (1.9, 2.3)	2.0 (1.8, 2.3)	2.0 (1.8, 2.3)
Digestive	2.7 (2.2, 3.2)	3.3 (2.5, 4.4)	4.7 (3.3, 6.7)	4.9 (3.4, 7.0)
Nervous/sense	1.3 (1.1, 1.6)	1.0 (0.7, 1.3)	1.0 (0.7, 1.4)	1.0 (0.7, 1.4)
Musculoskeletal	2.1 (1.3, 3.6)	2.0 (0.8, 4.9)	2.0 (0.6, 6.2)	2.1 (0.7, 6.5)
Respiratory	2.4 (2.1, 2.7)	3.1 (2.4, 4.0)	3.5 (2.5, 4.8)	3.6 (2.6, 5.0)
Endocrine/nutritional	2.2 (1.8, 2.6)	2.4 (1.7, 3.4)	2.3 (1.5, 3.5)	2.2 (1.5, 3.4)
Mental	3.5 (2.9, 4.2)	2.9 (2.0, 4.3)	3.5 (2.2, 5.4)	3.9 (2.5, 6.0)
Infectious	2.0 (1.6, 2.5)	1.9 (1.2, 3.0)	2.6 (1.4, 5.0)	2.7 (1.4, 5.0)
Genitourinary	1.9 (1.5, 2.4)	2.5 (1.5, 4.2)	2.8 (1.4, 5.7)	2.8 (1.4, 5.7)
Skin	2.8 (1.7, 4.7)	5.8 (1.2, 27.4)	5.8 (0.7, 44.9)	3.5 (0.5, 23.9)
Blood	2.1 (1.4, 3.1)	3.8 (1.2, 12.1)	8.9 (1.6, 48.5)	8.9 (1.6, 48.3)
Other	2.0 (1.9, 2.2)	1.9 (1.6, 2.2)	1.9 (1.6, 2.2)	1.9 (1.6, 2.2)

[a]Disease groups sorted as in Table 1.

Gliederung

Diskussion

- Satz von Gewichten nur Annäherung an tatsächliche kausale Zusammenhänge
- Problem der Gewichtszuweisung entspricht dem klassischen Problem der Epidemiologie:
  Belastungen kausale Verantwortungen zuweisen, wenn mehrere Belastungen interagieren
- Internationaler Konsens über Wahl der Gewichte
- Gewichte schätzen (Problem: Datenquelle)
- Multiple-cause Ansatz löst nicht das „competing risks“-Problem

  ABER: Identifizierung von Wechselbeziehungen zwischen Krankheiten

Fazit

- Ansätze ermöglichen das Nutzen der Daten zum besseren Verstehen von Ursachen und krankheitsbedingter Sterblichkeit
- Trotz Mankos bleiben Datenquellen zur Sterblichkeit eine Referenz für die Überwachung der öffentlichen Gesundheit
- Weitere Forschung bzgl. des statistischen Modellierens von mehreren Todesursachen berechtigt

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Vielen Dank für die Aufmerksamkeit.

Quellen

Haneuse, S., „Multiple Causes of Death, The Importance of Substantive Knowledge in the Big Data Era“.In Epidemiology, Volume 28 (2017), S. 28-29.

Moreno-Betancur, M., Sadaoui, H., Piffaretti, C., Rey, G., „Survival Analysis with Multiple Causes of Death, Extending the Competing Risks Model“.In Epidemiology, Volume 28 (2017), S. 12-19.

Moreno-Betancur, M., Sadaoui, H., Piffaretti, C., Rey, G., eAppendix for: „Survival Analysis with Multiple Causes of Death, Extending the Competing Risks Model“.

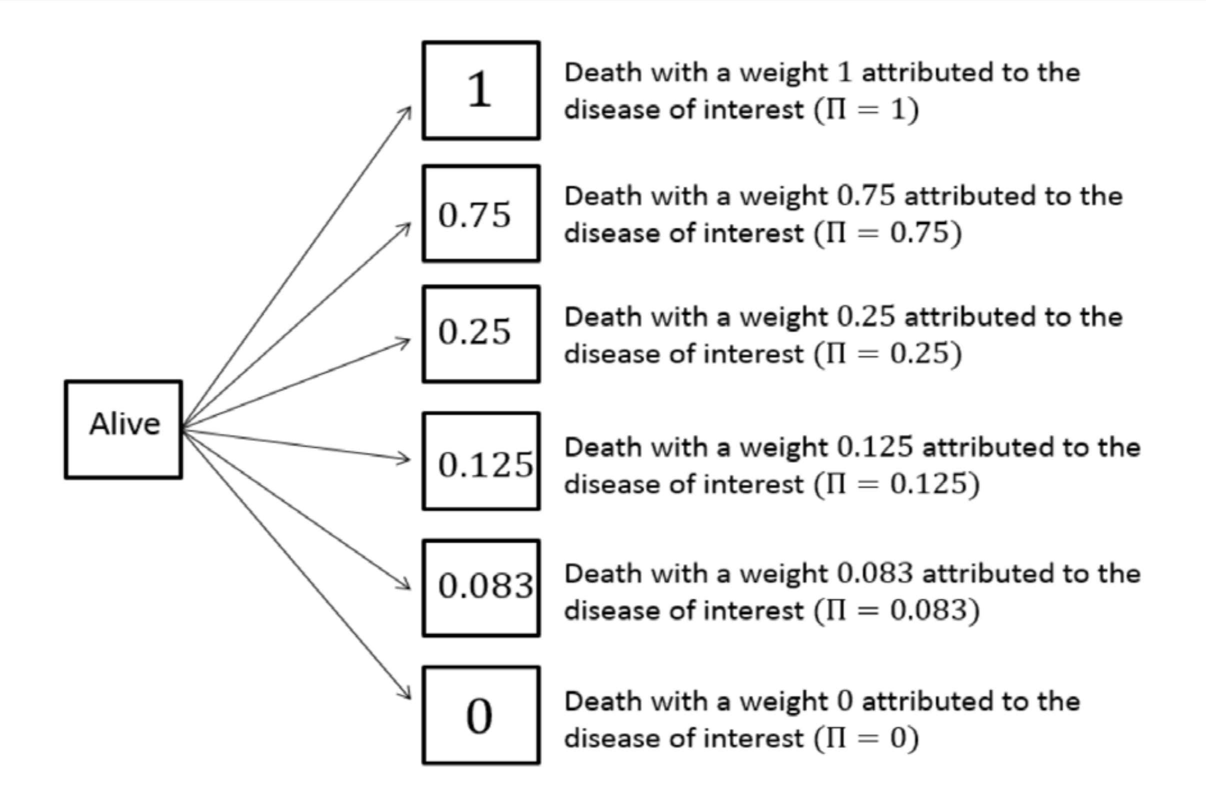
Mackenbach, J.P., Kunst, A.E., Lautenbach, H., Bijlsma, F. and Oei, Y.B., „Competing Causes of Death: An Analysis using Multiple-Cause-of-Death Data from the Netherlands” in: American Journal of Epidemiology, Vol. 141, No.5 (1995), S. 466-475

ICD-10 Second Edition Volume 2 - World Health Organization [18.06.2017]

http://han.vermont.gov/hc/death_certificate/mistakes.htm [18.06.2017]

Simulationsstudie

Multi-state Modell mit 6 Ereignistypen
Generierte Daten, n=4000



Simulationsstudie

Annahmen (A0)-(A3) und konstanter log ratio der pure baseline hazards $\xi(t) = -1$

- $\phi = 0$

Für jedes π:

$\lambda_\pi(t|X,Z) = 0{,}002t\pi[exp\ (\rho' X) + (1-\pi)\exp(-\xi(t) + \phi' Z]$

- X binär
- 30% rechtszensiert
- Vergleich mehrere Ansätze um den Zielparameter ρ zu schätzen

Simulationsstudie