Weighting approach per mediatori multipli nell'analisi di sopravvivenza

Francesca Fasanelli, Maria Teresa Giraudo, Fulvio Ricceri, Alex Francia e Daniela Zugna

19 Ottobre 2016

EPIDEMIOLOGIA PIEMONTE
DIPARTIMENTO DI MATEMATICA GIUSEPPE PEANO
UNIVERSITÀ DI TORINO
CPO
Centro di Riferimento per l'Epidemiologia e la Prevenzione Oncologica in Piemonte

Introduzione

- Il principale scopo dell'analisi di mediazione è studiare gli effetti diretti (non mediati dai mediatori) ed indiretti (mediati) di una esposizione su un outcome.
- Ad oggi la letteratura sull'analisi di mediazione nell'ambito di outcome di sopravvivenza, si è concentrata soprattutto su scenari con singolo mediatore.
- Gli approcci per mediatori multipli nell'analisi di sopravvivenza sviluppati in letteratura sono i seguenti:
    1. Tchetgen Tchetgen *Stat Med* 2013
    2. Lange T et al. *Am J Epidemiol* 2013
    3. Huang YT et al. *Biometrics* 2016

Introduzione

Obiettivi

1. Estendere matematicamente ad outcome di sopravvivenza il metodo proposto per mediatori multipli da Vanderweele et al.* nel 2013
2. Esplicitare i passi fondamentali dell'algoritmo di stima
3. Mostrare l'utilità del metodo utilizzando dati provenienti da uno studio sulla mortalità per tumore della prostata** (effetti di una variante del gene *DNMT3b* sulla mortalità attraverso la metilazione di più geni ed il punteggio Gleason)

*VanderWeele TJ and Vansteelandt S. **Mediation analysis with multiple mediators**. *Epidemiologic Methods* 2013; 2: 95-115.

Richiardi L, Fiano V, Vizzini L, et al. **Promoter methylation in APC, RUNX3 and GSTP1 and mortality in prostate cancer patients. *J Clin Oncol* 2009; 27: 3161-3168.

Effetti casuali nell'analisi di sopravvivenza

L'effetto causale totale

Definizione

L'effetto totale di una esposizione binaria A su un outcome tempo all'evento T è:

$$TCE = \frac{\lambda_{T^1}(t)}{\lambda_{T^0}(t)}. \tag{1}$$

Effetti casuali nell'analisi di sopravvivenza

L'effetto puro diretto

Definizione

L'effetto puro diretto di una esposizione binaria A su un outcome tempo all'evento T è:

$$PDE = \frac{\lambda_{T^{1,\mathbf{M}^0}}(t)}{\lambda_{T^{0,\mathbf{M}^0}}(t)}. \tag{2}$$

Effetti casuali nell'analisi di sopravvivenza

L'effetto naturale indiretto

Definizione

L'effetto naturale indiretto di una esposizione binaria A su un outcome tempo all'evento T è:

$$NIE = \frac{\lambda_{T^{1,\mathbf{M}^1}}(t)}{\lambda_{T^{1,\mathbf{M}^0}}(t)}. \tag{3}$$

Effetti casuali nell'analisi di sopravvivenza

Si può mostrare che l'effetto causale totale è il prodotto dell'effetto naturale indiretto e dell'effetto puro diretto:

Proposizione

Sotto la **composition assumption***:*

$$T^a = T^{a,\mathbf{M}^a} \tag{4}$$

cioè il valore di T che si avrebbe se $A = a$ è uguale al valore di T che si avrebbe se $A = a$ e $\mathbf{M}$ assumesse il valore che avvrebbe se $A = a$,

$$TCE = NIE \cdot PDE. \tag{5}$$

Estensione matematica del metodo di Vanderweele

Teorema

Sotto le seguenti assunzioni:

1. $T^{a\mathbf{m}} \coprod A \,|\, \mathbf{C}$ *assenza di confondenti della relazione esposizione-outcome*
2. $T^{a\mathbf{m}} \coprod \mathbf{M} \,|\, (A, \mathbf{C})$ *assenza di confondenti delle relazioni mediatori-outcome*
3. $\mathbf{M}^{a} \coprod A \,|\, \mathbf{C}$ *assenza di confondenti delle relazioni esposizione-mediatori*
4. $T^{a\mathbf{m}} \coprod \mathbf{M}^{a^*} \,|\, \mathbf{C}$ *nessun effetto dell'esposizione che confonde le relazioni mediatori-outcome*

$$\lambda_{T^{a,\mathbf{M}^{a^*}}}(t) = \frac{\mathbb{E}_{[C,\mathbf{M}]^{a^*}}\left[\frac{P(A=a^*)}{P(A=a^*|C)}\, f_T(t \mid A = a, \mathbf{M}, C)\right]}{\mathbb{E}_{[C,\mathbf{M}]^{a^*}}\left[\frac{P(A=a^*)}{P(A=a^*|C)}\, S_T(t \mid A = a, \mathbf{M}, C)\right]} \tag{6}$$

dove $\mathbb{E}_{[C,\mathbf{M}]^{a^*}}$ *indica l'attesa rispetto alla densità congiunta di* $\mathbf{M}$ *e* C *condizionata ad* $A = a^*$.

Quantità di interesse (esposizione binaria)

$$PDE = \frac{\lambda_{T^{1,\mathbf{M}^0}}(t)}{\lambda_{T^{0,\mathbf{M}^0}}(t)}$$

$$NIE = \frac{\lambda_{T^{1,\mathbf{M}^1}}(t)}{\lambda_{T^{1,\mathbf{M}^0}}(t)}$$

$$TOT = PDE \cdot NIE$$

1. $\lambda_{T^{1,\mathbf{M}^0}}(t)$
2. $\lambda_{T^{0,\mathbf{M}^0}}(t)$
3. $\lambda_{T^{1,\mathbf{M}^1}}(t)$

Algoritmo di stima

STEP 1

1. Stima di $\lambda_{T^{1,\mathbf{M}^0}}(t)$:

$$\lambda_{T^{1,\mathbf{M}^0}}(t) = \frac{\mathbb{E}_{[C,\mathbf{M}]^0}\left[\frac{P(A=0)}{P(A=0|C)}\, f_T(t \mid A = 1, \mathbf{M}, C)\right]}{\mathbb{E}_{[C,\mathbf{M}]^0}\left[\frac{P(A=0)}{P(A=0|C)}\, S_T(t \mid A = 1, \mathbf{M}, C)\right]}$$

- per ogni soggetto con $A = 0$ la funzione hazard viene modellata per ottenere una stima predetta della densità e della sopravvivenza forzando $A = 1$ piuttosto che $A = 0$, ma utilizzando i valori veri dei mediatori e delle covariate
- due medie pesate dei valori predetti sono calcolate separatamente per densità e sopravvivenza per i soggetti con $A = 0$ (a ogni soggetto i viene assegnato il peso $\frac{P(A=0)}{P(A=0|\mathbf{C}_i)}$ dove $\mathbf{C}_i$ denota i valori veri delle covariate per il soggetto i)
- si calcola infine il rapporto delle due medie pesate sopra descritte

Algoritmo di stima
STEP 2

2. Stima di $\lambda_{T^{0,\mathbf{M}^0}}(t)$:

$$\lambda_{T^{0,\mathbf{M}^0}}(t) = \frac{\mathbb{E}_{[C,\mathbf{M}]^0}\left[\frac{P(A=0)}{P(A=0|C)} f_T(t \mid A=0, \mathbf{M}, C)\right]}{\mathbb{E}_{[C,\mathbf{M}]^0}\left[\frac{P(A=0)}{P(A=0|C)} S_T(t \mid A=0, \mathbf{M}, C)\right]}$$

- per ogni soggetto con $A = 0$ la funzione hazard viene modellata per ottenere una stima predetta della densità e della sopravvivenza utilizzando i valori veri dell'esposizione, dei mediatori e delle covariate
- due medie pesate dei valori predetti sono calcolate separatamente per densità e sopravvivenza per i soggetti con $A = 0$ (a ogni soggetto i viene assegnato il peso $\frac{P(A=0)}{P(A=0|\mathbf{C}_i)}$ dove $\mathbf{C}_i$ denota i valori veri delle covariate per il soggetto i)
- si calcola infine il rapporto delle due medie pesate sopra descritte

Algoritmo di stima

STEP 3

3. Stima di $\lambda_{T^{1,\mathbf{M}^1}}(t)$:

$$\lambda_{T^{1,\mathbf{M}^1}}(t) = \frac{\mathbb{E}_{[C,\mathbf{M}]^1}\left[\frac{P(A=1)}{P(A=1|C)}\, f_T(t \mid A=1, \mathbf{M}, C)\right]}{\mathbb{E}_{[C,\mathbf{M}]^1}\left[\frac{P(A=1)}{P(A=1|C)}\, S_T(t \mid A=1, \mathbf{M}, C)\right]}$$

- per ogni soggetto con $A = 1$ la funzione hazard viene modellata per ottenere una stima predetta della densità e della sopravvivenza utilizzando i valori veri dell'esposizione, dei mediatori e delle covariate
- due medie pesate dei valori predetti sono calcolate separatamente per densità e sopravvivenza per i soggetti con $A = 1$ (a ogni soggetto i viene assegnato il peso $\frac{P(A=1)}{P(A=1|\mathbf{C}_i)}$ dove $\mathbf{C}_i$ denota i valori veri delle covariate per il soggetto i)
- si calcola infine il rapporto delle due medie pesate sopra descritte

Algoritmo di stima

STEP 4 E STEP 5

4. L'effetto puro diretto, naturale indiretto e totale possono essere calcolati a partire dai rapporti di medie pesate descritti in precedenza nel modo seguente:

$$PDE = \frac{\lambda_{T_{1M_0}}(t)}{\lambda_{T_{0M_0}}(t)},$$

$$NIE = \frac{\lambda_{T_{1M_1}}(t)}{\lambda_{T_{1M_0}}(t)},$$

$$TCE = NIE \cdot PDE.$$

5. Gli intervalli di confidenza sono calcolati utilizzando la tecnica del bootstrap.

Applicazione su dati reali

Introduzione

Cancer Causes Control (2012) 23:1549–1555
DOI 10.1007/s10552-012-0032-9

ORIGINAL PAPER

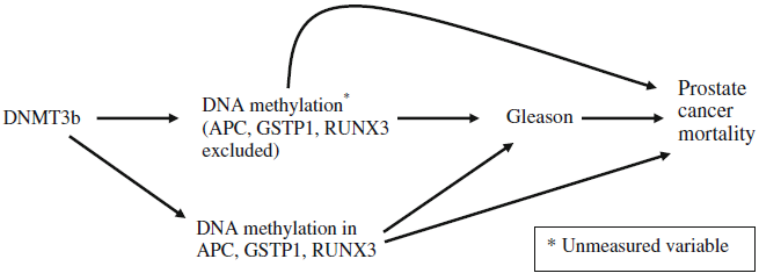
DNA methyltransferase 3b (DNMT3b), tumor tissue DNA methylation, Gleason score, and prostate cancer mortality: investigating causal relationships

Anna Gillio-Tos · Valentina Fiano · Daniela Zugna · Loredana Vizzini · Neil Pearce · Luisa Delsedime · Franco Merletti · Lorenzo Richiardi

Studio della relazione tra il genotipo del gene *DNMT3b*, la metilazione dei geni *GSTP1*, *APC* e *RUNX3*, l'aggressività tumorale (misurata dal punteggio Gleason) e la mortalità per tumore della prostata.

Applicazione su dati reali

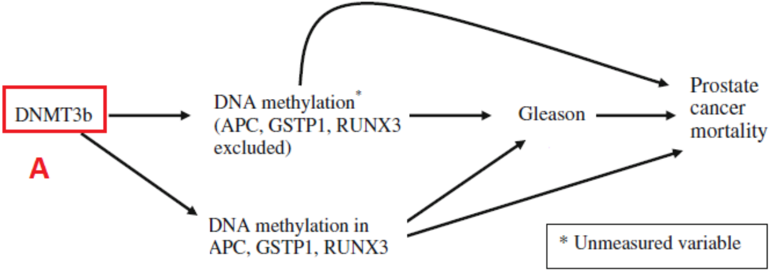
Introduzione



L'attività del gene *DNMT3b* ha effetti sulla mortalità direttamente e indirettamente attraverso la metilazione del tessuto tumorale e il punteggio Gleason.

Applicazione su dati reali

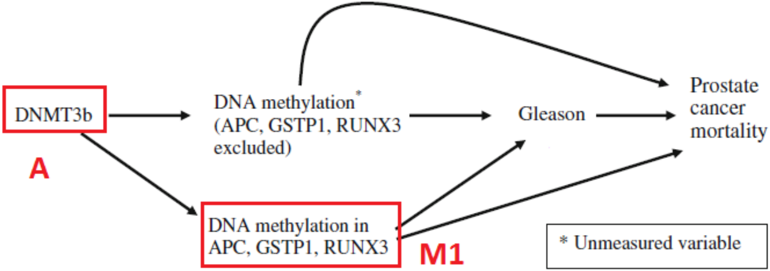
Introduzione



L'attività del gene *DNMT3b* ha effetti sulla mortalità direttamente e indirettamente attraverso la metilazione del tessuto tumorale e il punteggio Gleason.

Applicazione su dati reali

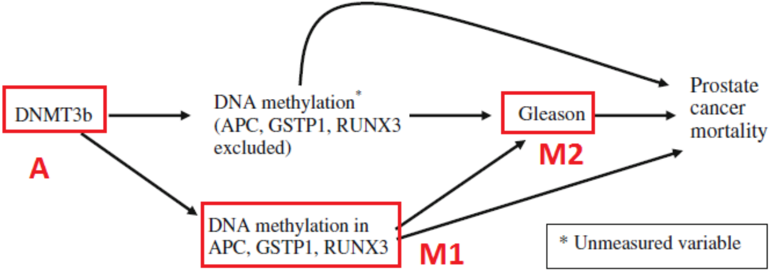
Introduzione



L'attività del gene *DNMT3b* ha effetti sulla mortalità direttamente e indirettamente attraverso la metilazione del tessuto tumorale e il punteggio Gleason.

Applicazione su dati reali

Introduzione



L'attività del gene *DNMT3b* ha effetti sulla mortalità direttamente e indirettamente attraverso la metilazione del tessuto tumorale e il punteggio Gleason.

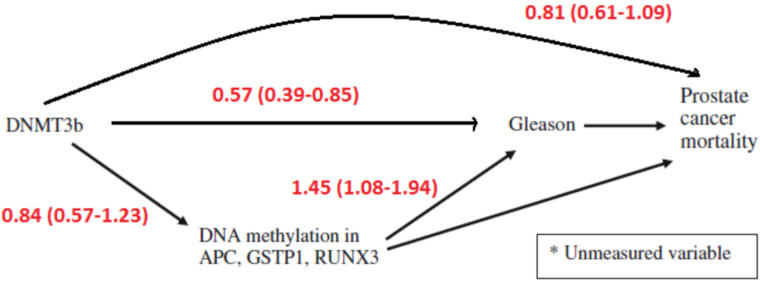
Applicazione su dati reali

Introduzione

- esposizione A: la variante del gene *DNMT3b* (portatori di almeno una T)
- primo mediatore M_1: metilazione del DNA (0-1, 2, 3 geni metilati)
- secondo mediatore M_2: punteggio Gleason (≥ 8 o viceversa)
- outcome T: tempo alla morte per tumore alla prostata
- <u>confondenti</u>: età alla diagnosi, periodo di diagnosi, fonte utilizzata per la tipizzazione del tessuto tumorale
- assunzione di assenza di altri confondenti non misurati

Applicazione su dati reali

Analisi preliminari



Applicazione su dati reali

Analisi preliminari

Approccio standard

- modello senza mediatori:
  HR: 0.81, 95% CI: 0.61 – 1.09
- modello con metilazione del DNA (primo mediatore):
  HR: 0.80, 95% CI: 0.59 – 1.09
- modello con entrambi i mediatori:
  HR: 0.90, 95% CI: 0.66 – 1.23

Ma quando l'outcome non è raro, l'approccio standard non può essere utilizzato per quantificare l'effetto mediato dai mediatori*!

*Tein J-Y and MacKinnon DP. **Estimating Mediated Effects with Survival Data**. In: Yanai H, Rikkyo AO, Shigemasu K, Kano Y and Meulman JJ (eds) *New Developments on Psychometrics*. Tokyo: Springer Japan, 2003, pp.405-412.

Applicazione su dati reali

Analisi di mediazione

	DNMT3b rs406193	
	CC	CT+TT
PDE	1	0.81 (0.55 - 1.07)
NIE	1	1.00 (0.77 - 1.27)
(attraverso M_1)		
TCE	1	0.83 (0.50 - 1.16)
PDE	1	0.90 (0.58 - 1.23)
NIE	1	0.79 (0.54 - 1.04)
(attraverso M_1 e M_2)		
TCE	1	0.71 (0.31 - 1.11)

Tabella : Risultati dell'analisi di mediazione.

L'analisi suggerisce che l'effetto totale della variante sulla mortalità per cancro alla prostata è attribuibile maggiormente all'effetto indiretto. Sebbene l'effetto indiretto non possa essere ripartito tra i mediatori, il punteggio Gleason sembra essere il mediatore più rilevante della relazione studiata.

Conclusioni

I principali vantaggi del metodo sono i seguenti:

- vasta applicabilità a contesti in cui l'outcome non è raro
- i mediatori considerati possono dipendere l'uno dall'altro
- non è richiesta alcuna modellizzazione dei mediatori (così si riduce il problema della mispecificazione del modello)
- il metodo può essere facilmente implementato in standard software
- il metodo consente di introdurre interazioni esposizione-mediatori e mediatore-mediatore

Conclusioni

I principali limiti del metodo sono i seguenti:

- la performance della stima dipende dalla validità delle ipotesi alla base dell'analisi di mediazione
- incapacità di caratterizzare gli effetti path-specifici
- procedura basata sul calcolo di pesi che possono essere instabili
- l'esposizione può essere solo binaria o categorica

Il metodo proposto è un utile strumento nell'ambito dell'analisi di mediazione in presenza di molteplici mediatori ed outcome di sopravvivenza

Grazie a...

Carlotta Sacerdote
Anna Gillio-Tos
Valentina Fiano
Loredana Vizzini
Neil Pearce
Luisa Delsedime
Franco Merletti
Lorenzo Richiardi

Grazie a voi per l'attenzione!

40 anni

EPIDEMIOLOGIA
& PRECARIATO

Nessun conflitto di interessi!