Scomposizione degli effetti nell'analisi di mediazione: un esempio numerico

Effects decomposition in mediation analysis: a numerical example

Daniela Zugna,[1] Lorenzo Richiardi[1,2]

1 Dipartimento di scienze mediche, Università di Torino, Torino
2 Azienda ospedaliera universitaria Città della salute e della scienza di Torino e Centro di riferimento per l'epidemiologia e la prevenzione oncologica in Piemonte, Torino

Corrispondenza: Daniela Zugna; daniela.zugna@unito.it

RIASSUNTO

L'obiettivo principale di un'analisi di mediazione è la scomposizione dell'effetto totale dell'esposizione sull'esito nell'effetto diretto (non mediato) e indiretto (mediato da un mediatore). Quando si è interessati a capire se l'effetto dell'esposizione sull'esito vari in diversi sottogruppi della popolazione in studio o sotto diversi scenari, è necessario integrare l'analisi di mediazione con un'analisi d'interazione. Da qui, la necessità di scomporre l'effetto totale non solo in due componenti, diretta e indiretta, ma in altre due componenti legate all'interazione. L'interazione tra l'esposizione e il mediatore nel loro effetto sull'esito potrebbe, infatti, agire tramite l'effetto dell'esposizione sul mediatore o tramite il solo mediatore, qualora esso non sia totalmente spiegato dall'esposizione. Nell'articolo descriviamo diverse modalità, proposte in letteratura, di decomposizione dell'effetto totale attraverso un esempio ipotetico sull'effetto dell'età alla diagnosi del tumore sulla sopravvivenza, mediato o non mediato dall'approccio terapeutico, e un esempio numerico.

Parole chiave: mediazione, interazione, controfattuale

ABSTRACT

Mediation analysis aims to decompose the total effect of the exposure on the outcome into a direct effect (unmediated) and an indirect effect (mediated by a mediator). When the interest also lies on understanding whether the exposure effect differs in different sub-groups of study population or under different scenarios, the mediation analysis needs to be integrated with interaction analysis. In this setting it is necessary to decompose the total effect not only into two components, the direct and indirect effects, but other two components linked to interaction. The interaction between the exposure and the mediator in their effect on the outcome could indeed act through the effect of the exposure on the mediator or through the mediator when the mediator is not totally explained by the exposure. We describe options for decomposition, proposed in literature, of the total effect and we illustrate them through a hypothetical example of the effect of age at diagnosis of cancer on survival, mediated and unmediated by the therapeutical approach, and a numerical example.

Keywords: mediation, interaction, controfactual

INTRODUZIONE

Negli ultimi anni è stata data particolare enfasi alla formalizzazione degli approcci all'analisi di mediazione in epidemiologia.[1-6] In particolare, è stata estesa la definizione degli effetti diretti (non mediati da un mediatore specifico) e indiretti (mediati da un mediatore specifico) per permettere la presenza di non linearità e di interazione tra l'esposizione e il mediatore. Inoltre, si è sottolineata l'importanza dell'assunzione d'ipotesi, tra le quali l'assenza di confondimento tra l'esposizione e il mediatore e tra il mediatore e l'esito, al fine di permettere un'interpretazione causale degli effetti stimati.[7-8]
Questo articolo ha scopo esemplificativo ed è composto da due parti. Nella prima si descrive la scomposizione degli effetti totali in assenza/presenza di interazione[9-10] accompagnando le definizioni formali, basate sulla notazione controfattuale,[11-12] con un esempio ipotetico di uno studio sull'effetto dell'età alla diagnosi di un tumore sulla sopravvivenza, mediato o non mediato dall'approccio terapeutico. Nella seconda parte, un esempio proposto in un precedente lavoro del 2013[13] viene esteso al fine di mostrare numericamente come quantificare l'interazione tra esposizione e mediatore.

EFFETTI DIRETTI E INDIRETTI

Sia A l'esposizione, M il mediatore e Y l'esito d'interesse. Per semplicità, consideriamo le variabili A, M, Y binarie. Supponiamo per esempio che A rappresenti l'età alla diagnosi di tumore (anziano *vs.* giovane), M il trattamento (palliativo *vs.* curativo) e Y la mortalità a 5 anni dalla diagnosi (figura 1).
Ogni soggetto ha due esiti controfattuali: Y_1 sarebbe l'esito se il soggetto fosse stato esposto e Y_0 sarebbe l'esito se il soggetto non fosse stato esposto. Nell'esempio, Y_1 sarebbe lo stato in vita se il paziente fosse stato diagnosticato in età anziana e Y_0 lo stato in vita se il paziente fosse stato diagnosticato in età giovane. Gli esiti Y_0,Y_1 sono detti controfattuali, poiché solo uno dei due è realmente osservato per ogni paziente: Y_1 è osservato se il paziente è stato diagnosticato in età anziana, Y_0 è osservato se il paziente è stato diagnosticato in età giovane.
Analogamente M_1 e M_0 sono i controfattuali del mediatore nei due scenari di esposizione e di non esposizione del sogget-

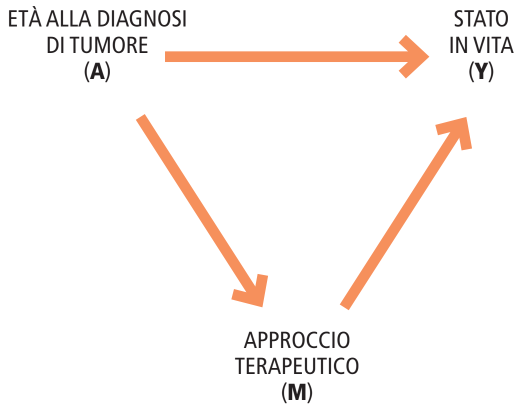

Figura 1. *Direct acyclic graph* (DAG) rappresentante il *pathway* di mediazione tra l'età alla diagnosi di tumore (A), l'approccio terapeutico (M) e la mortalità a 5 anni dalla diagnosi (Y).
Figure 1. Direct Acyclic Graph (DAG) representing the mediation pathway between the age at diagnosis of cancer (A), the therapeutical approach (M), and the mortality at 5 years after diagnosis (Y).

to ad A. Nell'esempio, M_1 sarebbe il trattamento assunto se il paziente fosse stato diagnosticato in età anziana e M_0 il trattamento assunto se il paziente fosse stato diagnosticato in età giovane.

Infine, Y_{am} è l'esito controfattuale per lo scenario A = a e M = m. Nell'esempio ci sono quattro controfattuali: Y_{11}, Y_{10}, Y_{01}, Y_{00}. Per esempio, Y_{11} sarebbe lo stato in vita se il paziente fosse stato diagnosticato in età anziana e fosse stato trattato con intento palliativo. Sebbene vi siano quattro esiti controfattuali in funzione di A e M, solo uno di loro è realmente osservato per ogni soggetto.

Utilizzando questa notazione controfattuale, di seguito definiamo l'effetto totale, l'effetto diretto controllato, l'effetto diretto e indiretto, puro e totale rispettivamente.

L'effetto totale causale (TE) di A su Y è definito come differenza tra i due esiti controfattuali, Y_{1M_1}-Y_{0M_0}. Tale effetto può essere stimato solo a livello di popolazione come differenza tra le medie dei due esiti controfattuali, $E(Y_{1M_1} - Y_{0M_0})$. Nell'esempio, l'effetto totale causale è l'effetto dell'età alla diagnosi sulla mortalità.

L'effetto diretto controllato (CDE) di A su Y per un valore di M = m è definito, a livello di popolazione, come $E(Y_{1m} - Y_{0m})$ e rappresenta l'effetto di A su Y non mediato da M, essendo quest'ultimo fissato per ciascun soggetto a un dato livello m. Nell'esempio, l'effetto diretto controllato $E(Y_{10} - Y_{00})$ è l'effetto dell'età alla diagnosi sulla mortalità se la popolazione fosse stata trattata con intento curativo, mentre l'effetto diretto controllato $E(Y_{11} - Y_{01})$ è l'effetto dell'età alla diagnosi sulla mortalità se la popolazione fosse stata trattata con intento palliativo.

Sotto determinate condizioni, il CDE può essere stimato come segue:

$$CDE(m) = E(Y_{1m} - Y_{0m}) = P(Y = 1|A = 1, M = m) - P(Y = 1|A = 0, M = m) \quad (1)$$

Il CDE può variare al variare di m, cioè, per esempio, l'effetto dell'età alla diagnosi sulla mortalità potrebbe essere diverso sotto lo scenario in cui tutta la popolazione fosse trattata con intento palliativo e quello in cui tutta la popolazione fosse trattata con intento curativo. Questo accade in presenza d'interazione tra A e M nel loro effetto su Y.

L'effetto puro diretto (PDE) di A su Y è definito, a livello di popolazione, come $E(Y_{1M_0} - Y_{0M_0})$ e rappresenta l'effetto di A su Y non mediato da M essendo quest'ultimo fissato per ciascun soggetto al livello che M assumerebbe naturalmente in assenza dell'esposizione. Nell'esempio, il PDE è l'effetto dell'età alla diagnosi sulla mortalità se l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se il tumore fosse stato diagnosticato in età giovane. Sotto determinate condizioni, il PDE può essere stimato come media pesata dei CDE con pesi uguali alla probabilità di M = m nei non esposti:

$$\begin{aligned} PDE &= E(Y_{1M_0} - Y_{0M_0}) = \textstyle\sum_m E[Y_{1m} - Y_{0m}]P(M = m|\, A = 0) = \\ &[P(Y = 1|\, A = 1, M = 1) - P(Y = 1|\, A = 0, M = 1)]\, P(M = 1|\, A = 0) + \\ &\quad [P(Y = 1|\, A = 1, M = 0) - P(Y = 1\, |A = 0, M = 0)]\, P(M = 0|\, A = 0) \end{aligned} \quad (2)$$

L'effetto totale diretto (TDE) di A su Y differisce dal PDE, in quanto M è fissato per ciascun soggetto al livello che M assumerebbe naturalmente in presenza dell'esposizione. Nell'esempio, il TDE è l'effetto dell'età alla diagnosi sulla mortalità se l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se fosse stato diagnosticato in età anziana. Sotto determinate condizioni, il TDE può essere stimato come media pesata dei CDE con pesi uguali alla probabilità di M = m negli esposti:

Epidemiol Prev 2018; 42 (2):127-133. doi: 10.19191/EP18.2.P127.041

$$
\begin{aligned}
&TDE = E(Y_{1M_1} - Y_{0M_1}) = \\
&\Sigma_m E[Y_{1m} - Y_{0m}]P(M = m| A = 1) = \\
&\quad [P(Y = 1| A = 1, M = 1) - P(Y = 1| A = 0, M = 1)]\ P(M = 1| A = 1) + \\
&\quad [P(Y = 1| A = 1, M = 0) - P(Y = 1 |A = 0, M = 0)]\ P(M = 0| A = 1)
\end{aligned} \tag{3}
$$

Il PDE e il TDE non sono nulli se A ha un effetto su Y non mediato da M, $P(Y = 1|A = 1, M = m) \neq P(Y = 1| A = 0, M = m)$.

L'effetto puro indiretto (PIE) di A su Y è definito, a livello di popolazione, come $E(Y_{0M_1}\text{-}Y_{0M_0})$ e rappresenta l'effetto di A su Y mediato da M, essendo A fissata come assente e confrontando gli scenari in cui M assumerebbe il valore che avrebbe assunto naturalmente in presenza e in assenza di A rispettivamente. Nell'esempio, il PIE è l'effetto dell'età alla diagnosi sulla mortalità confrontando la mortalità se i pazienti fossero stati diagnosticati a un'età giovane e l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età anziana e in età giovane, rispettivamente. Sotto determinate condizioni, il PIE può essere stimato come segue:

$$
\begin{aligned}
&PIE = E(Y_{0M_1} - Y_{0M_0}) = \\
&\Sigma_m E\ [Y_{0m}][P(M = m \mid A = 1) - P(M = m \mid A = 0)] = \\
&\quad P(Y = 1 \mid A = 0, M = 1)[P(M = 1| A = 1) - P(M = 1| A = 0)] + \\
&\quad P(Y = 1 \mid A = 0, M = 0)[P(M = 0| A = 1) - P(M = 0| A = 0)] = \\
&\quad [P(Y = 1| A = 0, M = 1) - P(Y = 1 \mid A = 0, M = 0)]\ [P(M = 1| A = 1) - P(M = 1| A = 0)]
\end{aligned} \tag{4}
$$

L'effetto totale indiretto (TIE) di A su Y differisce dal PIE, in quanto A è fissata come presente. Nell'esempio, il TIE è l'effetto dell'età alla diagnosi sulla mortalità confrontando la mortalità se i pazienti fossero stati diagnosticati a un'età anziana e l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età anziana e in età giovane rispettivamente. Sotto determinate condizioni, il TIE può essere stimato come segue:

$$
\begin{aligned}
&TIE = E(Y_{1M_1}\text{-}Y_{1M_0}) = \\
&\Sigma_m E[Y_{1m}][P(M = m \mid A = 1) - P(M = m \mid A = 0)] = \\
&\quad P(Y = 1 \mid A = 1, M = 1)[P(M = 1| A = 1) - P(M = 1| A = 0)] + \\
&\quad P(Y = 1 \mid A = 1, M = 0)[P(M = 0| A = 1) - P(M = 0| A = 0)] = \\
&\quad [P(Y = 1| A = 1, M = 1) - P(Y = 1 \mid A = 1, M = 0)]\ [P(M = 1| A = 1) - P(M = 1| A = 0)]
\end{aligned} \tag{5}
$$

Il PIE e il TIE non sono nulli se A ha un effetto su M ($P(M = 1| A = 1) \neq P(M = 1| A = 0)$) in grado di provocare, a sua volta, un effetto su Y ($P(Y = 1| A = a, M = 1) \neq P(Y = 1| A = a, M = 0)$).

SCOMPOSIZIONE DEGLI EFFETTI

Per semplicità, indichiamo con $p_{am} = P(Y = 1| A = a, M = m)$ la probabilità dell'esito condizionatamente all'esposizione e al mediatore. L'effetto totale di A su Y può essere scomposto nella somma dell'effetto puro diretto e dell'effetto totale indiretto:[7]

$$
TE = PDE + TIE = (p_{11} - p_{01})\ P(M = 1| A = 0) + (p_{10} - p_{00})P(M = 0| A = 0) + (p_{11} - p_{10})\ [P(M = 1| A = 1) - P(M = 1| A = 0)] \tag{6}
$$

o nella somma dell'effetto totale diretto e dell'effetto puro indiretto:

$$
TE = TDE + PIE = (p_{11} - p_{01})\ P(M = 1| A = 1) + (p_{10} - p_{00})\ P(M = 0| A = 1) + (p_{01} - p_{00})\ [P(M = 1| A = 1) - P(M = 1| A = 0)] \tag{7}
$$

In assenza d'interazione tra A e M, TDE e PDE sono uguali e, analogamente, TIE è uguale a PIE. In presenza d'interazione tra A e M, la differenza tra gli effetti totali e quelli puri (diretti e indiretti) corrisponde all'interazione tra A e M. L'interazione mediata è, quindi, definita come differenza tra l'effetto totale e quello puro, indiretto o diretto:

$$
INT_{med} = (p_{11} - p_{10} - p_{01} + p_{00})\ [P(M = 1| A = 1) - P(M = 1| A = 0)] \tag{8}
$$

L'espressione $(p_{11} - p_{10} - p_{01} + p_{00})$ corrisponde alla definizione classica dell'interazione su scala additiva ed è la differenza

Epidemiol Prev 2018; 42 (2):127-133. doi: 10.19191/EP18.2.P127.041

tra l'effetto congiunto di A e M, $(p_{11} - p_{00})$, e la somma degli effetti dovuti alla sola A, $(p_{10} - p_{00})$ e al solo M, $(p_{01} - p_{00})$.[14] Nell'esempio, $(p_{11} - p_{00})$ è l'effetto sulla mortalità dell'essere diagnosticati in età anziana e trattato con intento palliativo rispetto all'essere diagnosticati in età giovane e trattato con intento curativo, $(p_{10} - p_{00})$ è l'effetto sulla mortalità dell'essere diagnosticati in età anziana rispetto all'essere diagnosticati in età giovane quando si è trattati con intento curativo e $(p_{01} - p_{00})$ è l'effetto sulla mortalità dell'essere trattato con intento palliativo rispetto a curativo quando si è diagnosticati in età giovane. C'è interazione mediata se:

- l'età alla diagnosi influisce sulla tipologia di trattamento;
- l'effetto congiunto dell'età e del trattamento è diverso dalla somma dei loro effetti marginali.

È possibile, quindi, scomporre ulteriormente l'effetto totale di A su Y in tre componenti:[9] l'effetto puro diretto, l'effetto puro indiretto e l'interazione mediata:

$$TE = PDE + PIE + INT_{med} = (p_{11} - p_{01})\, P(M = 1|\, A = 0) + (p_{10} - p_{00}) P(M = 0|\, A = 0) + \\ (p_{01} - p_{00})\, [P(M = 1|\, A = 1) - P(M = 1|\, A = 0)] + (p_{11} - p_{10} - p_{01} + p_{00})\, [P(M = 1|\, A = 1) - P(M = 1|\, A = 0)] \quad (9)$$

In assenza d'interazione tra A e M, il PDE è uguale al CDE per qualsiasi valore di M. In presenza d'interazione, il PDE è composto da due componenti, il CDE in assenza di M (controllato per m = 0) e l'interazione di riferimento così definita:

$$INT_{ref} = (p_{11} - p_{10} - p_{01} + p_{00})\, P(M = 1|\, A = 0) \quad (10)$$

Questa interazione su scala additiva $(p_{11} - p_{10} - p_{01} + p_{00})$ opera solo se M è presente anche in assenza di A $(P(M = 1| A = 0) \neq 0)$. Nell'esempio, c'è interazione di riferimento se:

- la prevalenza del trattamento palliativo nei diagnosticati in età giovane è diversa da zero;
- l'effetto congiunto dell'età e del trattamento è diverso dalla somma dei loro effetti marginali.

Mentre l'interazione mediata è conseguenza dell'effetto di A su M, l'interazione di riferimento opera nella misura in cui M non è spiegato da A (nell'esempio, si tratta di pazienti che avrebbero comunque ricevuto il trattamento palliativo indipendentemente dall'età). È possibile, quindi, scomporre ulteriormente l'effetto totale di A su Y in quattro componenti: l'effetto diretto controllato in assenza del mediatore, l'interazione di riferimento, l'effetto puro indiretto e l'interazione mediata:[10]

$$TE = CDE(m = 0) + INT_{rif} + PIE + INT_{med} = \\ (p_{10} - p_{00}) + (p_{11} - p_{10} - p_{01} + p_{00})\, P(M = 1|\, A = 0) + \\ (p_{01} - p_{00})\, [P(M = 1|\, A = 1) - P(M = 1|\, A = 0)] + (p_{11} - p_{10} - p_{01} + p_{00})\, [P(M = 1|\, A = 1) - P(M = 1|\, A = 0)] \quad (11)$$

Le quattro modalità di scomposizione del TE, con riferimento all'esempio, sono riportate in tabella 1. Una volta scomposto l'effetto totale, è possibile calcolare:

- la proporzione totale mediata, che è data dal rapporto tra la somma del PIE e l'INT_{med} al numeratore e il TE al denominatore;
- la proporzione totale attribuibile all'interazione, data dal rapporto tra la somma dell'INT_{rif} e l'INT_{med} al numeratore e il TE al denominatore.

ESEMPIO NUMERICO

Consideriamo l'esempio ipotetico di Richiardi et al.[13] che riporta il caso di un'esposizione dicotomica A che interagisce con un mediatore dicotomico M nel causare l'esito dicotomico Y (tabella 2).
Su scala additiva, l'esposizione aumenta il rischio di Y di 0,048:

$$TE = 350/6.000 - 110/10.500 = 0{,}048$$

Epidemiol Prev 2018; 42 (2):127-133. doi: 10.19191/EP18.2.P127.041

MODALITÀ DI DECOMPOSIZIONE DEL TE	DESCRIZIONE
TE = PDE + TIE	Effetto dell'età alla diagnosi sulla mortalità = Effetto dell'età alla diagnosi sulla mortalità se l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età giovane + Effetto dell'età alla diagnosi sulla mortalità confrontando la mortalità se i pazienti fossero stati diagnosticati a un'età anziana e l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età anziana o in età giovane, rispettivamente
TE = TDE + PIE	Effetto dell'età alla diagnosi sulla mortalità = Effetto dell'età alla diagnosi sulla mortalità se l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se fosse stato diagnosticato in età anziana + Effetto dell'età alla diagnosi sulla mortalità confrontando la mortalità se i pazienti fossero stati diagnosticati a un'età giovane e l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età anziana o in età giovane, rispettivamente
TE = PDE + PIE + INT_{med}	Effetto dell'età alla diagnosi sulla mortalità = Effetto dell'età alla diagnosi sulla mortalità se l'approccio al trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età giovane + Effetto dell'età alla diagnosi sulla mortalità confrontando la mortalità se i pazienti fossero stati diagnosticati a un'età giovane e l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età anziana e in età giovane, rispettivamente + Interazione su scala additiva tra l'età e l'approccio al trattamento in misura proporzionale all'effetto dell'età alla diagnosi sull'approccio al trattamento
TE = CDE(m=0) + INT_{rif} + PIE + INT_{med}	Effetto dell'età alla diagnosi sulla mortalità = Effetto dell'età alla diagnosi sulla mortalità se la popolazione fosse stata trattata con intento curativo + Interazione su scala additiva tra l'età e l'approccio al trattamento in misura proporzionale alla prevalenza del trattamento palliativo nei giovani + Effetto dell'età alla diagnosi sulla mortalità confrontando la mortalità se i pazienti fossero stati diagnosticati a un'età giovane e l'approccio di trattamento per ogni paziente fosse stato quello che il paziente avrebbe ricevuto se diagnosticato in età anziana o in età giovane rispettivamente + Interazione su scala additiva tra l'età e l'approccio di trattamento in misura proporzionale all'effetto dell'età alla diagnosi sull'approccio al trattamento

Tabella 1. Scomposizione dell'effetto totale (TE).
Table 1. Decomposition of the total effect (TE).

A	M	RISCHIO	Y=1	Y=0	TOTALE
0	0	1%	100	9.900	10.000
1	0	3%	150	4.850	5.000
0	1	2%	10	490	500
1	1	20%	200	800	1.000

A: esposizione / *exposure* **M:** mediatore / *mediator* **Y:** esito / *outcome*

Tabella 2. Esempio numerico.
Table 2. Numerical example.

Gli effetti diretti e indiretti si ottengono come segue (equazioni 2-5):

$$\text{CDE}(m = 0) = 0{,}030 - 0{,}010 = 0{,}020$$
$$\text{CDE}(m = 1) = 0{,}200 - 0{,}020 = 0{,}180$$
$$\text{PDE} = 0{,}180 * 500/10.500 + 0{,}020 * 10.000/10.500 = 0{,}028$$
$$\text{TDE} = 0{,}180 * 1.000/6.000 + 0{,}020 * 5.000/6.000 = 0{,}047$$
$$\text{PIE} = (0{,}020 - 0{,}010) * (1.000/6.000 - 500/10.500) = 0{,}001$$
$$\text{TIE} = (0{,}200 - 0{,}030) * (1.000/6.000 - 500/10.500) = 0{,}020$$

La diversità tra i CDE, tra PDE e TDE e tra PIE e TIE suggeriscono la presenza di interazione tra A e M nel loro effetto su Y. Per la (8), l'INT_{med} è uguale a:

$$\text{INT}_{med} = (0{,}200 - 0{,}020 - 0.030 + 0.010) * (1.000/6.000 - 500/10.500) = 0{,}019$$

Essa corrisponde anche alla differenza tra gli effetti totali e puri, diretti o indiretti:

$$\text{TDE} - \text{PDE} = 0{,}047 - 0{,}028 = 0{,}019$$
$$\text{TIE} - \text{PIE} = 0{,}020 - 0{,}001 = 0{,}019$$

Epidemiol Prev 2018; 42 (2):127-133. doi: 10.19191/EP18.2.P127.041

Per la (10), l'INT_{rif} è uguale a:

INTrif = (0,200 – 0,020 – 0,030 + 0,010) * (500/10.500) = 0,008

Il PDE è quindi dato dalla somma del CDE(m = 0) e dell'INT_{rif}:

PDE = 0,020 + 0,008 = 0,028

Il TE può quindi essere decomposto

- in due componenti (equazioni 6-7):
  TE = PDE + TIE = 0,028 + 0,020 = 0,048
  TE = TDE + PIE = 0,047 + 0,001 = 0,048

- in tre componenti (equazione 9):
  TE = PDE + PIE + INT_{med} = 0,028 + 0,001 + 0,019 = 0,048

- e in quattro componenti (eq 11):
  TE = CDE(m = 0) + INT_{ref} + INT_{med} + PIE = 0,020 + 0,008 + 0,019 + 0,001 = 0,048

Nell'esempio numerico, quindi, le due componenti con maggiore impatto sull'effetto totale sono l'effetto diretto controllato di A su Y in assenza di M e l'interazione mediata. Quest'ultima è maggiore dell'interazione di riferimento. Infatti, il termine d'interazione è il medesimo, ma l'effetto di A su M, espresso come differenza nella prevalenza di M in presenza e assenza di A (1.000/6.000 – 500/10.500 = 0,119), è maggiore della prevalenza di M in assenza di A (500/10.500 = 0,048). La proporzione totale mediata è pari al 42%:

(PIE + INT_{med})/TE = (0,001 + 0,019)/0,048 = 0,417

e la proporzione totale attribuibile all'interazione è pari al 56%:

(INT_{rif} + INT_{med})/TE = (0,008 + 0,019)/0,048 = 0,562

DISCUSSIONE

Questo articolo presenta un esercizio semplificato per la scomposizione dell'effetto totale di un'esposizione su un esito in molteplici componenti. L'articolo ha uno scopo didattico e illustrativo. Abbiamo presentato la scomposizione dell'effetto totale su scala additiva, ma sono state date corrispondenti definizioni su scala moltiplicativa. Sia l'esempio pratico sia l'esempio numerico non hanno considerato il potenziale errore di misura e la presenza di confondenti delle associazioni esposizione-esito, mediatore-esito ed esposizione-mediatore. In pratica, abbiamo assunto che non ci fosse confondimento, sebbene sia facile immaginare numerosi confondenti dell'effetto dell'approccio terapeutico sulla sopravvivenza per tumore. Le definizioni date nell'articolo possono essere estese per stimare gli effetti, diretti e indiretti, condizionatamente o marginalmente ai confondenti noti e misurati. Aspetti che esulano dallo scopo di questo articolo illustrativo, quali per esempio l'identificazione degli effetti e la modellistica, sono stati trattati nella letteratura specifica.[9,10,15] A nostra conoscenza, in SAS è stata implementata la scomposizione dell'effetto totale nelle quattro componenti tramite un approccio basato sui modelli di regressione sotto molteplici scenari, con l'esito e il mediatore continuo, binario e la loro combinazione.[10]

Conflitti di interesse dichiarati: nessuno.

Finanziamenti: il lavoro è stato parzialmente finanziato dal programma dell'Unione europea per la ricerca e l'innovazione "Horizon 2020" nell'ambito della convenzione di finanziamento n. 733206, progetto LIFE-CYCLE.

BIBLIOGRAFIA

1. Cole SR, Hernan MA. Fallibility in estimating direct effects. Int J Epidemiol 2002; 31(1):163-65.
2. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Statist 1978;6(1):34-58.
3. Petersen ML, Sinisi SE, Van der Laan MJ. Estimation of direct causal effects. Epidemiology 2006;17(3):276-84.
4. VanderWeele TJ. Simple relations between principal stratification and direct and indirect effects. Stat Prob Lett 2008;78:2957-62.
5. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550-60.
6. Daniel RM, De Stavola BL, Cousens SN. gformula: estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula. Stata Journal 2011;11(4):479-517.
7. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology 1992;3(2):143-55.
8. Pearl J. Direct and indirect effects. Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence. San Francisco, Morgan Kaufmann, 2001; pp. 411-20.
9. VanderWeele TJ. A three-way decomposition of a total effect into direct, indirect, and interactive effects. Epidemiology 2013;24(2):224-32.
10. VanderWeele TJ. A unification of mediation and interaction: a 4-way decomposition. Epidemiology 2014;25(5):749-61.
11. Splawa-Neyman J. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. Statistical Science 1990;5(4):465-72.
12. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol 1974;66(5):688-701.
13. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol 2013;42(5):1511-19.
14. Rothman KJ. Modern Epidemiology. 1st edition. Boston (MA), Little, Brown and Company, 1986.
15. VanderWeele, TJ. Explanation in causal Inference. Methods for mediation and Interaction. Oxford University Press 2015.

 Epidemiol Prev 2018; 42 (2):127-133. doi: 10.19191/EP18.2.P127.041