

# 直接効果・間接効果の推定および未測定の変動の交絡に対する感度解析

田栗 正隆<sup>†</sup>

(受付 2013 年 9 月 12 日；改訂 2014 年 1 月 2 日；採択 1 月 6 日)

## 要 旨

疫学研究の1つの目的は、興味のある曝露の疾病発生に対する因果効果を推定することである。曝露と疾病の間に因果関係が示唆された場合、どのようなメカニズムで効果があるのかについての知見を得ることに興味を持たれる場合がある。この問題に対しての1つのアプローチは、曝露の疾病に対する影響(総合効果)を、中間変数を介しない直接効果と、中間変数を介した間接効果に分解することである。本論文では、潜在結果変数モデルに基づく直接効果・間接効果の定義と識別のための仮定、識別式についてまとめる。また、識別のための、未測定の変動がないという仮定が崩れた場合の感度解析方法を紹介する。紹介した方法の適用事例として、全米保健医療統計センターが公表している米国の出生証明書および乳児死亡に関するデータ解析結果を報告する。

キーワード：因果推論、感度解析、効果の分解、直接効果・間接効果、未測定の変動の交絡。

## 1. はじめに

疾病発生を前向きに追跡する縦断的な疫学研究の1つの目的は、興味のある曝露の疾病発生に対する因果効果を推定することである。例えば、喫煙者が禁煙をすることによって、心筋梗塞の発症リスクをどの程度減らすことができるであろうか。また、曝露と疾病の間に因果関係が示唆された場合、どのようなメカニズムで効果があるのかについての知見を得ることに興味を持たれる場合がある。先ほどの例では、喫煙の心筋梗塞に対する影響のうち、どの程度の割合を血清 LDL コレステロール値を低下させる効果で説明できるかといった仮説である。後者の問題については、曝露(喫煙)の疾病(心筋梗塞)に対する影響を、中間変数(LDL コレステロール値)を介した間接効果(indirect effect)と、それを介しない直接効果(direct effect)に分解することが、統計解析の目的となる。

直接効果・間接効果の推定問題は、伝統的には曝露変数と結果変数の関係を中間変数を調整した上で検討すること、および、中間変数と結果変数の関係を検討することによって行われてきた(Baron and Kenny, 1986; MacKinnon, 2008)。Baron and Kenny (1986)の方法では、 $A$  を曝露変数、 $M$  を中間変数、 $Y$  を結果変数としたとき、以下のような2つの線形回帰モデルを考える。

$$(1.1) \quad E[Y|A = a, M = m] = \beta_0 + \beta_1 a + \beta_2 m,$$

$$(1.2) \quad E[M|A = a] = \gamma_0 + \gamma_1 a,$$

<sup>†</sup>横浜市立大学 学術院医学群臨床統計学・疫学：〒236-0004 神奈川県横浜市金沢区福浦 3-9

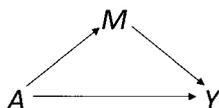


図 1. Baron and Kenny (1986) により提案されたモデル.

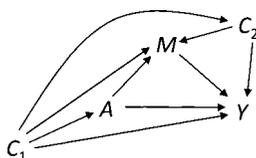


図 2. 識別のための仮定を満たす因果グラフ.

この時、 $\beta_1$  は曝露  $A$  の結果変数  $Y$  に対する直接効果(図 1 における  $A \rightarrow Y$  の経路を通じた効果)、 $\beta_2\gamma_1$  は曝露  $A$  の結果変数  $Y$  に対する間接効果(図 1 における  $A \rightarrow M \rightarrow Y$  の経路を通じた効果)と解釈される (Baron and Kenny, 1986). しかしながら、この方法には少なくとも以下に挙げる 2 つの欠点が存在する. 第 1 に、モデル (1.1) において  $A$  と  $M$  の間に交互作用が存在する場合、あるいはモデル (1.1) やモデル (1.2) のような線形モデルではなくロジスティック回帰モデルなどの非線形モデルが当てはまる場合、直接効果および間接効果の定義をどう拡張するべきかという疑問が残る点である. これは本質的には、ある特定の回帰モデル (1.1) および (1.2) のパラメータに基づいて、直接効果や間接効果を定義したことによる欠点である. 第 2 に、交絡の問題が挙げられる. 図 2 の様に、中間変数  $M$  と結果変数  $Y$  の両者に影響を与える様な交絡因子  $C_2$  が存在するとき、 $C_2$  を解析時に調整しない場合は、(例え曝露  $A$  がランダム割り付けされていたとしても)  $\beta_1$  はもはや直接効果と解釈できず、推定結果はバイアスを伴ったものとなる. この理由は、モデル (1.1) において中間変数  $M$  を条件付けることによって、 $A \rightarrow M \leftarrow C_2 \rightarrow Y$  の経路において合流点層別バイアス (collider stratification bias) と呼ばれるバイアスが生じてしまうためである (Cole and Hernán, 2002; Cole et al., 2010). この時、 $A$  と  $Y$  の間に因果的関連が存在しなかったとしても、 $A$  と  $Y$  の間には非因果的な関連が生じてしまう.

古典的なアプローチの問題に対して、Robins and Greenland (1992) と Pearl (2001) は潜在結果変数モデル (Rubin, 1974) に基づく直接効果・間接効果の定義およびその推定方法を提案した. 潜在結果変数モデルに基づくアプローチでは、データのみからは識別可能でない量に基づいて直接効果や間接効果が定義されるため、識別条件とその成立に注意が必要になるが、(1.1)、(1.2) のような線形モデルが成立しない状況においても、直接効果・間接効果への分解が可能である. また、識別条件のうちの 1 つには中間変数-結果変数間の交絡に関する条件も含まれるが、どのような状況下で推定結果を因果的に解釈できるかを明らかにしている点は、このアプローチの 1 つの利点である.

本論文では、潜在結果変数モデルに基づいて、直接効果および間接効果の検討に用いることのできる統計手法のレビューを行う. まず 2 節でモチベーションとなるいくつかの事例を紹介する. 続いて 3 節で観察データに関する記法と潜在結果変数モデルに基づく直接効果・間接効果の定義を述べる. 4 節では直接効果・間接効果を推定するための十分条件となる仮定と識別について述べる. 5 節では 4 節で述べる仮定のうちのいくつかを満たされない場合の感度解析方法を紹介する. 6 節では議論した方法の適用事例を紹介し、最後に 7 節でまとめを行う.

## 2. 事例

### 2.1 周産期疫学研究

全米保健医療統計センター(National Center for Health Statistics; NCHS)は、同センターのホームページにおいて、全米の出生証明書のデータと乳児死亡(生後1年未満の死亡)のデータをリンク可能な状態で公表している。先行研究の結果から、妊婦の適切な出生前ケア( $A = 1$ : 適切,  $A = 0$ : 不適切; 検診の密度等で定義)によって乳児死亡( $Y = 1$ : 死亡,  $Y = 0$ : 生存)がある程度予防できることが示唆されているが(Partridge et al., 2012), その予防効果のうちのどの程度が正期産( $M = 1$ : 正期産,  $M = 0$ : 早産)の増加(早産の減少)を介する間接効果で説明できるであろうか。これらの変数間の因果的関連を検討するための付加的なデータとして、出生証明書において母親の年齢, 母親の人種, 母親の教育歴, 母親の婚姻状態などが測定されている。

### 2.2 母乳育児に関する介入研究

Otsuka et al.(2014)は、妊婦に対して、母乳育児に対する自己効力感を高めることを目的としたワークブックを用いた介入を行うことで実際の母乳育児割合が改善するかどうかの比較介入研究を行った。その結果、赤ちゃんにやさしい病院(baby friendly hospitals)では、ワークブックによる介入( $A$ )は、調査票により測定された産後退院時の自己効力感( $M$ )を統計学的に有意に高めるとともに( $p = 0.037$ ), 産後4週間後の完全母乳割合( $Y$ )を有意に高める(介入に対する調整オッズ比: 2.32;  $p = 0.047$ )ことが示唆された。介入効果のうちの大部分は、当初の狙い通り自己効力感の増加を通じた間接効果で説明できるであろうか。それとも、例えば産後うつ等の発症割合の低下などによる、 $A \rightarrow M \rightarrow Y$ の経路を介しない介入の直接効果も存在するであろうか。なお、現実的には調査票により測定された自己効力感には測定誤差が伴うと考えられるが、ここでは例示のためにその影響はないものと考えた。

### 2.3 遺伝子および喫煙と肺癌に関する疫学研究

染色体 15q25.1 上の遺伝子変異は肺癌の発症と関連することが知られている。一方で、その遺伝子変異は喫煙習慣(ニコチン依存)とも関連することが知られている。VanderWeele et al.(2012)は、染色体 15q25.1 上の rs8034191 と rs1051730 の2つの一塩基多型それぞれ( $A$ )が肺癌発症( $Y$ )に与える影響のうち、どの程度が喫煙割合( $M$ )の増加を介した間接効果で説明されるかを検討している。その結果、間接効果/直接効果が  $[0,1]$  に収まるという仮定のもとで計算した総合効果に占める間接効果の割合は rs8034191 で 3.2%, rs1051730 で 2.3%と推定され、遺伝子変異と肺癌発症の関連は喫煙の増加を介しない直接効果で大部分が説明されることが示唆された。

## 3. 記法と定義

$Y$  を決められた観察期間内の疾病発生を表す 2 値の結果変数(1: 疾病あり, 0: 疾病なし),  $A$  を 2 値の曝露変数(1: 曝露あり, 0: 曝露なし),  $M$  を中間変数とする。  $C$  を曝露変数-結果変数, 曝露変数-中間変数, 中間変数-結果変数間の関連に関して測定された交絡因子とする。交絡因子  $C$  は、曝露変数の影響を受けないことが分かっているものとする(例えばランダム化比較試験のベースライン共変量である場合)。本論文では基本的に  $C$  はスカラーの離散型確率変数であると仮定するが、 $C$  が連続量やベクトルであってもほぼ同様の議論が成立する。中間変数は研究開始後のある時点において測定され、結果変数である疾病発生は中間変数の測定前には起こらないものとする。

直接効果・間接効果を定義するために、潜在結果変数モデルに基づく記法の導入を行う。  $Y(a)$

と  $M(a)$  を、対象者が  $A = a$  を受けた場合に観察されたであろう潜在的な結果変数および中間変数の値とする。同様に、 $Y(a, m)$  を  $A = a, M = m$  を受けた場合に観察されたであろう潜在的な結果変数の値とする。また、観察データと潜在データを関連付けるために、以下に述べる一貫性の仮定 (consistency assumption) と構成性の仮定 (composition assumption) を置く。一貫性の仮定は、 $A = a$  かつ  $M = m$  が観察されたサブグループでは、結果変数  $Y$  が潜在結果変数  $Y(a, m)$  と一致するという仮定である (VanderWeele, 2009a)。同様に、 $A = a$  が観察されたサブグループでは、中間変数  $M$  が潜在変数  $M(a)$  と一致することを仮定する。また、構成性の仮定は、 $Y(a) = Y(a, M(a))$  を意味する (Pearl, 2009)。これらの仮定に関する議論については VanderWeele and Vansteelandt (2009) が参考になる。また、Pearl の提案した構造的因果モデル (structural causal model) の枠組みでは、構成性の仮定は構造的因果モデルのデータ生成過程 (data generating process) から導かれる定理であり、一貫性の仮定は構成性から導かれる系である (Pearl, 2009)。Pearl の構造的因果モデルについては、日本語の解説として黒木・小林 (2012) で論じられている。

以上のもとで、曝露の総合効果 (total effect; TE) は、研究対象集団全体が曝露を受けた場合 ( $A = 1$ ) と受けなかった場合 ( $A = 0$ ) の比較として以下で定義される。

$$TE = E[Y(1) - Y(0)] = E[Y(1, M(1)) - Y(0, M(0))]$$

総合効果は、自然な直接効果 (natural direct effect; NDE) と自然な間接効果 (natural indirect effect; NIE) に分解される (Robins and Greenland, 1992; Pearl, 2001)。自然な直接効果は中間変数を  $M(0)$  に固定した場合の曝露効果として以下で定義される。

$$NDE = E[Y(1, M(0)) - Y(0, M(0))]$$

NDE は中間変数の値を曝露なし ( $A = 0$ ) の条件下での“自然な”値で評価しているため、自然な直接効果と呼ばれる (Pearl, 2001, 2012a)。同様に、自然な間接効果は曝露を  $A = 1$  に固定したもとで、中間変数が  $M(0)$  から  $M(1)$  に変化した場合の比較として以下で定義される。

$$NIE = E[Y(1, M(1)) - Y(1, M(0))]$$

NIE は総合効果のうち、曝露を受けることによって中間変数の値が変わることによる結果変数への影響を表している (Hafeman and Schwartz, 2009; Hafeman and VanderWeele, 2011)。この時、以下のような総合効果の分解が成立する。

$$TE = NIE + NDE$$

効果の分解は個人レベルでも成立する。すなわち、対象者  $i (i = 1, \dots, n)$  に対して総合効果、自然な直接効果・間接効果はそれぞれ  $TE_i = Y_i(1) - Y_i(0)$ ,  $NDE_i = Y_i(1, M_i(0)) - Y_i(0, M_i(0))$ ,  $NIE_i = Y_i(1, M_i(1)) - Y_i(1, M_i(0))$  で定義され、 $TE_i = NIE_i + NDE_i$  が成立する。なお、以上の自然な直接効果・間接効果の定義においては、 $(A, M, Y)$  間の関連に対するモデルの関数形が線形であることや、交互作用が存在しないことは何ら仮定されていない点は重要である。また、7 節で詳細を述べるが、 $TE_i = \{Y_i(1, M_i(1)) - Y_i(0, M_i(1))\} + \{Y_i(0, M_i(1)) - Y_i(0, M_i(0))\}$  という異なる効果の分解も可能である。

Pearl (2001) は直接効果のもう 1 つの定義として、以下の制御された直接効果 (controlled direct effect; CDE) を与えている。制御された直接効果は、中間変数  $M$  を  $m$  という水準に固定した場合の曝露効果として以下で定義される。

$$CDE(m) = E[Y(1, m) - Y(0, m)]$$

図1において、曝露変数  $A$  を喫煙に対する介入 ( $A = 1$ : 喫煙,  $A = 0$ : 禁煙), 中間変数  $M$  を LDL コレステロール値, 結果変数  $Y$  を心筋梗塞の発症有無 ( $Y = 1$ : あり,  $Y = 0$ : なし) とする. 自然な直接効果はコレステロール値を禁煙した場合の値 ( $M(0)$ ) に保った時の, 喫煙を続けた場合 ( $A = 1$ ) と禁煙した場合 ( $A = 0$ ) の比較であり, 自然な間接効果は曝露を喫煙あり ( $A = 1$ ) に固定したもとの, コレステロール値が禁煙した場合の値 ( $M(0)$ ) から禁煙しなかった場合 ( $M(1)$ ) の値に変化した場合の比較となる. 一方, 制御された直接効果は, コレステロール値をある値  $m$  に集団で一様に保った時の, 喫煙を続けた場合 ( $A = 1$ ) と禁煙した場合 ( $A = 0$ ) の比較である.

一方, 間接効果を中間変数を制御する形で定義することは難しい. なぜなら, 曝露変数の結果変数に対する直接的な影響を遮る形で, 何らかの変数がある値に固定することは一般には不可能であるためである (Pearl, 2012a). 特に, 総合効果から制御された直接効果を引いたものは, 間接効果の指標として適切ではない (VanderWeele, 2009b). なぜなら, 総合効果から制御された直接効果を引いたものは, 間接効果の指標として明らかに必要な条件である「曝露が中間変数に影響を与えない場合, すなわち全ての対象者について  $M(1) = M(0)$  が成立する場合, 間接効果は 0 にならない」という条件を必ずしも満たさないためである. VanderWeele (2011) は曝露変数の結果変数に対する効果が複数の中間変数によって完全に媒介されているという強い仮定のもとで, 制御された間接効果の定義を与えている.

それに対し, 既に述べたように, 自然な直接効果・間接効果の概念では, 総合効果から自然な直接効果を引くことによって, 自然な間接効果を導くことができる. しかしながら, 自然な直接効果の定義に含まれる  $Y(1, M(0))$  は, 現実的に観察するのが難しいという問題点が存在する.  $Y(1, M(0))$  を実際に観察する方法はあるであろうか. 先ほどの喫煙と心筋梗塞の例では,  $Y(1, M(0))$  を観察するためには①禁煙とは異なる脂質への介入法がある (例: 高脂血症治療薬) こと, ②喫煙の脂質への効果のみをブロックできる方法がある (例: 脂質上昇作用のみをなくした煙草) ことが要求される (Robins and Greenland, 1992). 一方で, Pearl (2011) は, 効果の分解を行って曝露効果がどのような経路でもたらされているかのメカニズムの検討を行うためには, 中間変数を  $M(0)$  に固定する方法が実際に存在することは要求されない旨を述べている.

#### 4. 識別

前節で定義した自然な直接効果および自然な間接効果において, 集団全体での  $Y(1, M(1))$ ,  $Y(1, M(0))$ , および  $Y(0, M(0))$  の期待値を推定するためには, データからは確認できない複数の仮定が必要となる. 自然な直接効果と間接効果の識別に対する 1 つの十分条件は, 以下の 4 つの仮定が成立することである (Pearl, 2001).

最初の 2 つの仮定は, 以下の条件付き独立性で与えられる.

$$(4.1) \quad M(a) \perp\!\!\!\perp A|C, \quad \forall m$$

$$(4.2) \quad Y(a, m) \perp\!\!\!\perp A|C, \quad \forall a, m$$

これらの仮定は, 共変量  $C$  を与えれば曝露変数-中間変数間, 曝露変数-結果変数間に未測定の変絡がないことを意味する. 曝露に関するランダム化を行った研究では, これらの仮定は満たされる. 識別に関する 3 つ目の仮定は以下で与えられる.

$$(4.3) \quad Y(a, m) \perp\!\!\!\perp M|A, C \quad \forall a, m$$

この仮定は,  $(A, C)$  を与えれば中間変数-結果変数間に未測定の変絡がないことを意味する. 最後に 4 つ目の仮定は以下で与えられる.

$$(4.4) \quad Y(a, m) \perp\!\!\!\perp M(0)|C \quad \forall a, m$$

この仮定は解釈が難しいが、本質的には曝露変数の影響を受ける中間変数-結果変数間の交絡因子  $L$  (図3参照)が存在しないことを意味する (VanderWeele, 2009c)。

図2は、識別のための仮定を満たす状況を表す因果グラフである。 $C_1$ は曝露変数-結果変数間、曝露変数-中間変数間の交絡因子を表す。 $C_2$ は中間変数間結果変数間の交絡因子を表す。図2において $C_1$ から $C_2$ への影響(矢線)がない場合、あるいは $C_2$ から $C_1$ への影響(矢線)がある場合でも、以下に述べる識別に関する結果は同様に成立する。直接効果・間接効果の推定を行うためには、研究者は交絡因子 $C = (C_1, C_2)$ を注意深く考え測定する必要があるが、本節で考えているように $C$ が曝露変数の影響を受けないことが分かっている場合には、全ての交絡因子を用いて調整することを基本として考えれば良く、必ずしも $C_1$ と $C_2$ を明確に区別する必要はない。なお、仮定(4.1)–(4.3)は、変数間の時間的な順序性を前提としていることに注意が必要である (Valeri and VanderWeele, 2013)。すなわち、曝露変数-中間変数間、および曝露変数-結果変数間に未測定の交絡がないという仮定(4.1)、(4.2)においては、曝露変数が中間変数と結果変数に対して、時間的に先行していることが暗に仮定されている。同様に、中間変数-結果変数間に未測定の交絡がないという仮定(4.3)は、中間変数が結果変数に時間的に先行していることを暗に仮定している。これらの時間的な順序性が満たされない状況では、解析結果を因果的に解釈することは出来ない。

(4.1)–(4.4)の仮定のもとで、 $E[Y(1, M(0))]$ が識別可能であることを以下に示す。まず、一致性の仮定を用いて以下が成立する。

$$(4.5) \quad \begin{aligned} E[Y(1, M(0))] &= \sum_c \sum_m E[Y(1, M(0)) | M(0) = m, c] \Pr[M(0) = m | c] p(c) \\ &= \sum_c \sum_m E[Y(1, m) | M(0) = m, c] \Pr[M(0) = m | c] p(c) \quad \because \text{一致性} \end{aligned}$$

ここで(4.5)最右辺中の $E[Y(1, m) | M(0) = m, c]$ 、 $\Pr[M(0) = m | c]$ についてそれぞれ

$$(4.6) \quad \begin{aligned} E[Y(1, m) | M(0) = m, c] &= E[Y(1, m) | c] && \because \text{仮定(4.4)} \\ &= E[Y(1, m) | A = 1, c] && \because \text{仮定(4.2)} \\ &= E[Y(1, m) | A = 1, M = m, c] && \because \text{仮定(4.3)} \\ &= E[Y | A = 1, M = m, c] && \because \text{一致性} \end{aligned}$$

$$(4.7) \quad \begin{aligned} \Pr[M(0) = m | c] &= \Pr[M(0) = m | A = 0, c] && \because \text{仮定(4.1)} \\ &= \Pr[M = m | A = 0, c] && \because \text{一致性} \end{aligned}$$

が成立する。(4.6)および(4.7)最右辺は観察データから推定可能な量であり、それらを(4.5)に代入することで

$$(4.8) \quad E[Y(1, M(0))] = \sum_c \sum_m E[Y | A = 1, M = m, c] \Pr[M = m | A = 0, c] p(c)$$

を得る。同様にして、 $a = 0, 1$ について

$$(4.9) \quad E[Y(a, M(a))] = \sum_c \sum_m E[Y | A = a, M = m, c] \Pr[M = m | A = a, c] p(c)$$

を得ることができる。したがって、(4.8)、(4.9)より自然な直接効果および自然な間接効果は以下のように表現することができる (Pearl, 2001)。

$$(4.10) \quad \begin{aligned} \text{NDE} &= \sum_c \sum_m \{E[Y | A = 1, M = m, c] - E[Y | A = 0, M = m, c]\} \\ &\quad \times \Pr[M = m | A = 0, c] p(c) \end{aligned}$$

$$(4.11) \quad \text{NIE} = \sum_c \sum_m E[Y|A = 1, M = m, c] \\ \times \{\text{Pr}[M = m|A = 1, c] - \text{Pr}[M = m|A = 0, c]\}p(c)$$

(4.10), (4.11)が妥当であるための十分条件は他にも複数提案されており (Robins, 2003; Petersen et al., 2006; Imai et al., 2010b; Hafeman and VanderWeele, 2011), Imai et al. (2010b)に分かりやすくまとめられている. 一方, 制御された直接効果を推定するためには仮定(4.1), (4.2)の2つが満たされていれば十分である. これらの仮定のもとで, 制御された直接効果は以下のように表現することができる (Pearl, 2001).

$$(4.12) \quad \text{CDE}(m) = \sum_c \sum_m \{E[Y|A = 1, M = m, c] - E[Y|A = 0, M = m, c]\}p(c)$$

(4.10), (4.12)より, ( $C = c$ で条件付けた)自然な直接効果は, 条件付き確率  $\text{Pr}[M = m|A = 0, c]$ を用いて制御された直接効果  $\text{CDE}(m)$ の重み付き平均をとったものであることが理解できる.

$Y, M$ それぞれに対して以下の線形モデルが成立する場合,

$$(4.13) \quad E[Y|A = a, M = m, C = c] = \beta_0 + \beta_1 a + \beta_2 m + \beta_3 am + \beta_4 c, \\ E[M|A = a, C = c] = \gamma_0 + \gamma_1 a + \gamma_2 c$$

自然な直接効果・間接効果および制御された直接効果はそれぞれ

$$\text{NDE} = \beta_1 + \beta_3(\gamma_0 + \gamma_2 E[C]), \\ \text{NIE} = (\beta_2 + \beta_3)\gamma_1, \\ \text{CDE}(m) = \beta_1 + \beta_3 m$$

と表現することができる. ここでモデル(4.13)において  $A$ と  $M$ の間に交互作用が存在しない場合, すなわち  $\beta_3 = 0$ の場合, 自然な直接効果および制御された直接効果は  $\beta_1$ と等しく, 自然な間接効果は  $\beta_2\gamma_1$ と等しい. したがってこの時, Baron and Kenny (1986)による古典的なアプローチの結果に帰着する. このことから, 自然な直接効果・間接効果の概念は, 古典的なアプローチを拡張したものとなっていることが分かる. 自然な間接効果・直接効果および制御された直接効果の推定量の分散推定については, デルタ法 (VanderWeele and Vansteelandt, 2009)あるいはブートストラップ法 (Efron and Tibshirani, 1993)を適用することが考えられる. 現実のデータ解析においては, 結果変数あるいは中間変数が2値の場合には, それぞれの条件付き期待値のモデル化においてロジスティック回帰モデルが用いられることが多い. Valeri and VanderWeele (2014)は, 結果変数のモデル化にロジスティック回帰モデルを用いた場合の, 稀な疾病の仮定のもとでの自然な直接効果・間接効果の回帰パラメータによる表現を与えている.

2値の結果変数に関しては, ここまで述べてきたようなリスク差のスケールではなく, リスク比(あるいはオッズ比)のスケールで効果を分解することが提案されている (VanderWeele and Vansteelandt, 2010). この場合,  $C = c$ を与えたもとでの条件付きリスク比(総合効果)は, 以下のように自然な間接効果と直接効果の積で表すことが出来る.

$$\frac{\text{Pr}[Y(1) = 1|c]}{\text{Pr}[Y(0) = 1|c]} = \frac{\text{Pr}[Y(1, M(1)) = 1|c]}{\text{Pr}[Y(0, M(0)) = 1|c]} = \frac{\text{Pr}[Y(1, M(1)) = 1|c]}{\text{Pr}[Y(1, M(0)) = 1|c]} \times \frac{\text{Pr}[Y(1, M(0)) = 1|c]}{\text{Pr}[Y(0, M(0)) = 1|c]}$$

同様に, 制御された直接効果は  $E[Y(1, m)|c]/E[Y(0, m)|c]$ で与えられる. オッズ比スケールでは, 総合効果は以下のように分解される.

$$\frac{\text{Pr}[Y(1) = 1|c]/\{1 - \text{Pr}[Y(1) = 1|c]\}}{\text{Pr}[Y(0) = 1|c]/\{1 - \text{Pr}[Y(0) = 1|c]\}} = \frac{\text{Pr}[Y(1, M(1)) = 1|c]/\{1 - \text{Pr}[Y(1, M(1)) = 1|c]\}}{\text{Pr}[Y(1, M(0)) = 1|c]/\{1 - \text{Pr}[Y(1, M(0)) = 1|c]\}}$$

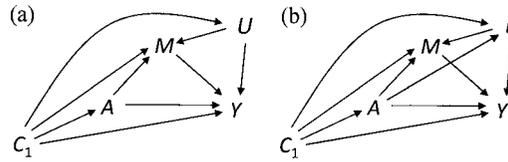


図 3. 識別のための仮定を満たさない因果グラフ. (a) 中間変数-結果変数間の未測定の交絡因子  $U$  が存在する場合, (b) 曝露変数の影響を受ける交絡因子  $L$  が存在する場合.

$$\times \frac{\Pr[Y(1, M(0)) = 1|c]/\{1 - \Pr[Y(1, M(0)) = 1|c]\}}{\Pr[Y(0, M(0)) = 1|c]/\{1 - \Pr[Y(0, M(0)) = 1|c]\}}$$

本節で述べた方法の大部分は, Imai et al. (2010a) の開発した R パッケージ, あるいは Valeri and VanderWeele (2013) の作成した SAS および SPSS マクロで実装可能である. 直接効果・間接効果の推定方法には, 他にも周辺構造モデルに基づく IPW (inverse probability weighting) 法など, 観察データをモデル化する部分の異なる複数の方法が提案されている (Goetgheul et al., 2008; VanderWeele, 2009c; Tchetgen Tchetgen and Shpitser, 2012).

## 5. 感度解析

自然な直接効果・間接効果を推定するための仮定のうち, 仮定 (4.3) および (4.4) は, 曝露変数のランダム割付を行うランダム化比較試験においても成立が保証されない. また, コホート研究では, 曝露変数に関わる既知の交絡因子に関しては十分に測定されている場合が多いが, 中間変数-結果変数間の交絡因子については, デザイン段階で詳細に検討されていない可能性がある. さらに, 仮定 (4.4) は, 中間変数および結果変数が曝露の測定から時間的にかなり経過してから測定される場合, 大変厳しい仮定である. 中間変数-結果変数間の未測定の交絡や, 曝露に影響を受ける交絡因子が想定される場合には, 感度解析を行い, それらの仮定の崩れがどの程度結果に影響を与えるかを見積もっておく必要がある.

本節では, 仮定 (4.3), (4.4) のどちらか(あるいは両方)が崩れた場合に用いることのできる感度解析手法として, VanderWeele and Chiba (2014) の方法を紹介する. 図 3 に, 仮定 (4.3), (4.4) が成立しない因果グラフの例を示す. 図 3(a) は中間変数-結果変数間の未測定の交絡因子  $U$  が存在する場合, 図 3(b) は曝露変数の影響を受ける交絡因子  $L$  が存在する場合である.

データから推定できない感度パラメータ  $\delta(m, c)$  を以下のように定義する.

$$\delta(m, c) = E[Y(1, m)|A = 1, M = m, c] - E[Y(1, m)|A = 0, M = m, c]$$

2.1 節で述べた周産期疫学の例では, 感度パラメータ  $\delta(1, c) = E[Y(1, 1)|A = 1, M = 1, c] - E[Y(1, 1)|A = 0, M = 1, c]$  は, 母親が適切なケアを受けて正期産であった集団 ( $A = 1, M = 1$ ) と, 不適切なケアを受けて正期産であった集団 ( $A = 0, M = 1$ ) という, 2つの異なるサブグループにおける, 「(仮想的に) “適切なケアを受けて正期産であった場合の” 乳児死亡割合の差」である. 恐らく, 適切なケアを受けて正期産であった集団 ( $A = 1, M = 1$ ) は, 不適切なケアを受けても正期産であった集団 ( $A = 0, M = 1$ ) よりも総じて不健康であると思われるので, 乳児死亡割合は前者の方が高いと考えられる. この場合,  $\delta(1, c) > 0$  となる. 一方, 感度パラメータ  $\delta(0, c) = E[Y(1, 0)|A = 1, M = 0, c] - E[Y(1, 0)|A = 0, M = 0, c]$  は, 母親が適切なケアを受けて早産であった集団 ( $A = 1, M = 0$ ) と, 不適切なケアを受けて早産であった集団 ( $A = 0, M = 0$ ) という, 2つの異なるサブグループにおける, 「(仮想的に) “適切なケアを受けて早産であった

場合の「乳児死亡割合の差」である。ここでは恐らく、適切なケアを受けても早産であった集団 ( $A = 1, M = 0$ ) は、不適切なケアを受けて早産であった集団 ( $A = 0, M = 0$ ) よりも総じて不健康的であろうから、乳児死亡割合は前者の方が高いと考えられる。この場合も、 $\delta(0, c) > 0$  となる。

ここで(4.10), (4.11)の右辺を  $Q_{NDE}, Q_{NIE}$  と置くと、仮定(4.1), (4.2)のみのもとで以下が成立する (VanderWeele and Chiba, 2014)。

$$(5.1) \quad \begin{aligned} NDE &= Q_{NDE} - \sum_c \sum_m \delta(m, c) \Pr[M = m | A = 0, c] p(c), \\ NIE &= Q_{NIE} + \sum_c \sum_m \delta(m, c) \Pr[M = m | A = 0, c] p(c) \end{aligned}$$

もし感度パラメータ  $\delta(m, c)$  の値が既知であれば、(5.1)を用いて自然な直接効果・間接効果は漸近的にバイアスなく推定可能である。現実には  $\delta(m, c)$  の真値を知ることはできない。しかしながら、 $\delta(m, c)$  の値を尤もらしいと考えられる範囲で動かすことによって、自然な直接効果・間接効果の推定値がどのように変化するかを検討することができる。 $\delta(m, c)$  の値の範囲の決め方については、前段落と同様の議論に基づいて  $\delta(m, c)$  の符号を問題に応じて決めることができれば、NDE の下限(あるいは上限)が  $Q_{NDE}$  で与えられることとなる。Chiba and Suzuki (2012) は感度パラメータを  $\delta(m) = E[Y(1, m) | A = 1, M = m] - E[Y(1, m) | A = 0, M = m]$  と  $c$  に依存しない形で与える、より簡便な感度解析法を提案しており、先験的な知識に基づく(データからは確認できない)仮定のもとで  $\delta(m)$  の値の範囲を決める他の方法をいくつか提案している。これらの方法は、本節で紹介した VanderWeele and Chiba (2014) の方法においても応用可能であると考えられる。先験的な知識に基づいて  $\delta(m, c)$  の範囲を決めるのが難しい場合、曝露変数、中間変数、結果変数が全て2値の場合には、曝露変数に関する未測定の変動の仮定のもとで、自然な直接効果・間接効果のバウンドをデータから推定することができる (Sjölander, 2009)。Chiba and Taguri (2013) は結果変数および中間変数が2値以外の場合のバウンドについても論じている。

(5.1)を用いてバイアス補正を行った場合の推定誤差の評価は、(5.1)右辺第2項に含まれる  $\Pr[M = m | A = 0, c]$  および  $p(c)$  の推定を考慮に入れる必要があるが、VanderWeele and Chiba (2014) はブートストラップ法を用いることを推奨している。同様の条件のもとで適用できる感度解析法は Imai and Yamamoto (2013) や Vansteelandt and VanderWeele (2012) で提案されている。これらの方法では、曝露変数に影響を受けうる中間変数-結果変数間の交絡因子  $L$  が測定されていることが前提となっている。本節冒頭で述べたとおり、中間変数-結果変数間の交絡因子については十分多くの因子が測定されていない可能性が高いことを考えると、 $L$  の測定が不要である点は VanderWeele and Chiba (2014) の方法の利点である。

なお、制御された直接効果については、 $L$  が存在していたとしても(4.2)および、(4.3)を拡張した以下の仮定のもとで

$$Y(a, m) \perp\!\!\!\perp M | A, C, L \quad \forall a, m$$

識別が可能である (Robins and Greenland, 1992; Robins, 2003)。

## 6. 解析事例

4節, 5節で紹介した方法の適用事例として、2.1節で述べた NCHS の提供している出生証明書と乳児死亡のリンケージデータ (2003年度) に関するデータ解析事例を紹介する。曝露変数 ( $A$ ) である妊婦の出生前ケアについては、Adequacy of Prenatal Care Utilization Index (APCUI; Kotelchuck, 1994) にならい、出生前ケア(妊娠中の定期的な健診およびスクリーニング検査)の開始時期と密度(単位時間あたりの受診回数)に基づいてカテゴリー化を行った。正期産 ( $M$ ) は在胎週数が37週以降での出産とし、乳児死亡 ( $Y$ ) は出生後1年以内の死亡とした。交絡因子 ( $C$ )

表 1. 出生前ケアおよび早産の有無別の交絡因子の分布(数値は%).

	全体 n=1,307,767	ケア適切 n=1,058,306	ケア不適切 n=249,461	正期産 n=1,232,203	早産 n=75,564
母親の年齢					
19歳以下	10.7	8.6	19.6	10.3	17.3
20歳-34歳	78.9	80.4	72.9	79.3	72.2
35歳以上	10.4	11.0	7.6	10.4	10.5
母親の人種					
白人	77.7	80.2	67.3	78.5	65.1
黒人	15.1	12.7	25.4	14.3	28.3
その他	7.2	7.1	7.3	7.2	6.6
母親の教育歴					
12年未満	21.7	17.4	40.2	21.0	32.8
12年以上	78.3	82.6	59.8	79.0	67.2
母親の婚姻状態					
既婚	64.2	69.9	59.8	65.3	45.1
未婚	35.8	30.1	40.2	34.7	54.9
多胎妊娠					
なし	98.7	98.8	98.0	99.1	91.2
あり(双子以上)	1.3	1.2	2.0	0.9	8.8

表 2. 出生前ケア別の正期産割合.

正期産	ケア適切	ケア不適切
なし	37,980	27,584
あり	1,020,326	211,877
正期産割合 (%)	96.4	84.9

としては、母親の年齢(19歳以下/20歳-34歳/35歳以上)、母親の人種(白人/黒人/その他)、母親の教育歴(12年未満/12年以上)、母親の婚姻状態(既婚/未婚)、多胎妊娠の有無を考えた。解析対象はAPCUIにおいて適切(adequate)、あるいは不適切(inadequate)に分類され、解析に用いるデータに欠測のない1,307,767例である。

表1に出生前ケア(適切/不適切)および出産時期(正期産/早産)ごとの交絡因子の分布を示す。表2に出生前ケア別の正期産割合、表3に出生前ケアおよび出産時期の組合わせごとの乳児死亡割合を示す。表2より、正期産に対する未調整の出生前ケアの曝露効果はリスク差で  $96.4 - 84.9 = 11.5\%$  (95%信頼区間 [CI]: 11.33 to 11.62) となった。また、表3より、乳児死亡割合に関する未調整の曝露の総合効果は、リスク差で  $0.52 - 1.57 = -1.05\%$  (95% CI: -1.10 to -1.00) となった。また、中間変数で層別した場合、リスク差は正期産で  $0.30 - 0.75 = -0.45\%$  (95% CI: -0.49 to -0.41)、早産で  $6.57 - 6.19 = 0.38\%$  (95% CI: 0.03 to 0.73) となった。これらの結果は未調整の解析結果ではあるものの、①正期産(M)に対する出生前ケア(A)の曝露効果が大きいこと(表2)、および②乳児死亡(Y)に対する出生前ケア(A)の効果は、正期産(M)の有無で層別すると絶対値で減少すること(表3)、の2点を考えると、乳児死亡割合に対する出生前ケアの効果は、正期産の増加を介した間接効果で大部分を説明できる可能性がある。

交絡を調整した解析においては、乳児死亡Yに関しては、A, M, Cの主効果およびA-M間の交互作用項を含むロジスティック回帰モデルを仮定した。同様に、正期産Mに関しても、

表 3. 出生前ケアおよび正期産有無の組合わせごとの乳児死亡割合.

乳児死亡	正期産		早産		全体	
	ケア適切	ケア不適切	ケア適切	ケア不適切	ケア適切	ケア不適切
なし	1,017,268	210,285	35,483	35,257	1,052,751	245,542
あり	3,058	1,592	2,497	2,327	5,555	3,919
死亡割合 (%)	0.30	0.75	6.57	6.19	0.52	1.57

表 4. 自然な直接効果・間接効果の推定結果. 括弧内はブートストラップ法による 95%信頼限界.

	リスク差 (%)			NIE/TE (%)
	NDE	NIE	TE	
交絡調整	-0.202 (-0.256, -0.138)	-0.601 (-0.634, -0.585)	-0.803 (-0.858, -0.753)	74.8 (69.9, 81.9)
感度解析1	-0.487 (-0.548, -0.431)	-0.316 (-0.334, -0.292)	-0.803 (-0.858, -0.753)	39.4 (35.3, 43.6)
感度解析2	-0.585 (-0.646, -0.529)	-0.219 (-0.242, -0.194)	-0.803 (-0.858, -0.753)	27.2 (23.5, 31.0)

A, C の主効果を線形項として含むロジスティック回帰モデルを仮定した。これらのモデルでは交絡因子 C 同士, C-A 間あるいは C-M 間の交互作用は仮定しなかった。これらの交互作用を仮定したモデルを適用することも当然可能であるが、もし実際に交互作用が存在していたとしても、推定の際に平均化されるため、直接効果・間接効果の推定に与える影響は A-M 間の交互作用を誤って無視してしまった場合と比較すると大きくないと考えられる。

感度解析については、曝露変数の影響を受けうる交絡因子 L として妊娠中の生活習慣(喫煙や飲酒など)や妊娠高血圧症などの合併症が考えられる。出生前ケアの影響で妊娠中の喫煙や飲酒、合併症は減少すると考えられるが、同時に不適切な生活習慣や合併症の発症は、早産および乳児死亡を増加させる中間変数-結果変数間の交絡因子である可能性がある。ここでは、感度パラメータ  $\delta(1, c)$  および  $\delta(0, c)$  について以下のように見積もった。まず、 $\delta(1, c)$  については、 $E[Y(1, 1)|A = 1, M = 1] = E[Y|A = 1, M = 1]$  の推定値である 0.3%の約 4 分の 1 である場合として 0.075% (シナリオ 1), 約 3 分の 1 である場合として 0.10% (シナリオ 2) を与えた。次に、 $\delta(0, c)$  については、 $E[Y(1, 0)|A = 1, M = 0] = E[Y|A = 1, M = 0]$  の推定値である 6.6%の約 4 分の 1 である場合として 1.65% (シナリオ 1), 約 3 分の 1 である場合として 2.2% (シナリオ 2) を与えた。なお、ここでは簡単のため感度パラメータは交絡因子には依存しないものと仮定した。

自然な直接効果・間接効果の推定結果および感度解析結果を表 4 に示す。交絡調整後の総合効果、自然な直接効果・間接効果の推定値および 95%信頼区間は、それぞれ -0.80% (ブートストラップ法による 95% CI: -0.858 to -0.753), -0.20% (95% CI: -0.256 to -0.138), -0.60% (95% CI: -0.634 to -0.585) となった。これらの結果から、出生前ケアの乳児死亡に対する総合効果のうちのおよそ 4 分の 3 (75.9%) が正期産の増加を通じた間接効果と考えられることが示唆された。しかしながら、この結果は、解析で考慮していない喫煙や飲酒等の不適切な生活習慣や合併症の発症が正期産群において早産群と比較して少ないこと等による残差交絡が影響している可能性がある。この場合、不適切な生活習慣や合併症の発症は乳児死亡のリスク因子であると考えられるため、これらを考慮していない今回の解析では、中間変数の結果変数に対する効果(した

がって自然な間接効果)は過大評価されていると考えられる。感度解析の結果は、そのようなバイアスを補正したものであると解釈できる。感度解析のシナリオ1では、総合効果に対する間接効果の割合は39.4%、シナリオ2では27.2%となり、これらの場合では間接効果よりもむしろ正期産を介しない直接効果の影響の方が大きく推定された。以上より、本解析事例においては推定のための仮定の崩れによる結果への影響は比較的大きいことが示唆された。なお、NIE/TEの値は[0,1]に収まらない可能性があることに注意が必要である。Wang and Taylor(2002)は同様の指標が[0,1]に収まる十分条件について論じている。また、仮定(4.1)、(4.2)のみに基づく自然な直接効果のバウンド(Chiba and Taguri, 2013)は下限、上限それぞれ-1.57%(95% CI: -1.60 to -1.53), 99.57%(95% CI: 99.54 to 99.61)となった。このような最低限の仮定のもとのバウンドは多くの場合非常に広くなってしまうことが難点ではあるが、感度解析の結果はバウンドの推定結果に反するものであってはならないため、バウンドを推定して結果が矛盾しないことを確認しておくことは有用である。

## 7. 終わりに

本論文では、潜在結果変数モデルに基づく直接効果・間接効果の推定および感度解析について、事例紹介と共に簡単なまとめを行った。今後の課題については、曝露変数の影響を受ける交絡因子  $L$  が存在する場合の識別の問題が挙げられる。近年、Tchetgen Tchetgen and VanderWeele(2012)により識別のための仮定および推定量がいくつか提案されているが、 $L$  を用いた感度解析法や識別可能な新しい推定対象に関する検討(Pearl, 2012b)も含めて、まだまだ研究の余地があると思われる。

本論文では  $Y(1) - Y(0) = \{Y(1, M(1)) - Y(1, M(0))\} + \{Y(1, M(0)) - Y(0, M(0))\}$  という総合効果の分解を考えたが、3節で触れたように  $Y(1) - Y(0) = \{Y(1, M(1)) - Y(0, M(1))\} + \{Y(0, M(1)) - Y(0, M(0))\}$  という分解も可能である。Robins and Greenland(1992)では自然な直接効果  $\{Y(1, M(a)) - Y(0, M(a))\}$  と間接効果  $\{Y(a, M(1)) - Y(a, M(0))\}$  ( $a = 0, 1$ ) のうち、 $a = 0$  の場合を純粋直接効果・間接効果(pure direct effect/indirect effect)、 $a = 1$  の場合を総合直接効果・間接効果(total direct effect/indirect effect)と呼んでいる。総合効果は、本論文のように純粋直接効果と総合間接効果に分解されるか、あるいは総合直接効果と純粋間接効果に分解される。純粋直接効果・間接効果と総合直接効果・間接効果の違いは、曝露変数と中間変数の交互作用効果がどのように分解されるかの違いと考えることができる(Robins and Greenland, 1992; Hafeman and Schwartz, 2009)。このことを確認するために、総合効果の以下の3つの要素への分解を考える(VanderWeele, 2013a)。

$$(7.1) \quad Y(1) - Y(0) = \{Y(1, M(0)) - Y(0, M(0))\} + \{Y(0, M(1)) - Y(0, M(0))\} \\ + \{Y(1, 1) - Y(1, 0) - Y(0, 1) + Y(0, 0)\} \{M(1) - M(0)\}$$

この等式の証明は VanderWeele(2013a)に与えられている。(7.1)の右辺の1番目の要素、2番目の要素はそれぞれ純粋直接効果および純粋間接効果である。3番目の要素は曝露変数と中間変数の交互作用効果  $\{Y(1, 1) - Y(1, 0) - Y(0, 1) + Y(0, 0)\} = \{Y(1, 1) - Y(0, 0)\} - \{Y(1, 0) - Y(0, 0)\} + \{Y(0, 1) - Y(0, 0)\}$  に曝露変数の中間変数への効果  $\{M(1) - M(0)\}$  を乗じたものとなっている。この3番目の要素について、以下の等式が成立する。

$$(7.2) \quad \{Y(1, 1) - Y(1, 0) - Y(0, 1) + Y(0, 0)\} \{M(1) - M(0)\} \\ = \{Y(1, M(1)) - Y(0, M(1))\} - \{Y(1, M(0)) - Y(0, M(0))\} \\ = \{Y(1, M(1)) - Y(1, M(0))\} - \{Y(0, M(1)) - Y(0, M(0))\}$$

すなわち、交互作用効果に関連する部分である(7.1)の3番目の要素は、「総合直接効果と純粋直接効果の差」あるいは「総合間接効果と純粋間接効果の差」に等しい。以上より、純粋直接効果・間接効果は、中間変数との交互作用効果を含まない“純粋な”効果を意味していること、総合直接効果・間接効果は、中間変数との交互作用効果を含む“総合的な”効果を意味していることが理解できる。なお、Pearl(2001, 2012a)は自然な直接効果・間接効果は純粋直接効果・間接効果のみを指すという立場をとっている。VanderWeele(2013a)は総合効果を(7.1)で与えられる3つの要素に一意的に分解し、それぞれの平均効果を推定することを提案している。この方法が解析結果の解釈においてどの程度有用であるかはまだ不明確であり、今後実データへの適用も含めて議論していく必要があると考えられる。

6節で例示したように、自然な直接効果・間接効果の推定を行った際には、NIE/TEの計算により「総合効果のうちの何割が中間変数を介した間接効果であるか」を検討することは曝露効果のメカニズムの解釈において有用であると考えられる。この指標はProportion Mediated(PM)とも呼ばれており(Imai et al., 2010c; VanderWeele and Vansteelandt, 2010)、統計学的に類似した問題であるサロゲートエンドポイントの評価の文脈でも同様な指標を用いることが提案されている(Wang and Taylor, 2002; Taylor et al., 2005)。Robins and Greenland(1992)やVanderWeele(2013b)はより公衆衛生的に意味のある可能性がある指標として、 $1-CDE(m)/TE$ で定義されるProportion Eliminated(PE)を用いることを提案している。PEは、中間変数を適当な基準値 $m$ に一定に保つことができるような介入が仮に存在する場合、「総合効果のうちの何割を中間変数への介入により除去できるか」と解釈することができ、中間変数への介入による公衆衛生的な効果の1つの指標となりうると考えられる。なお、PEはPMと同様に $[0,1]$ に収まるとは限らない指標である。上記のPEの解釈は少なくとも $[0,1]$ の範囲内に収まることを前提としているので、Wang and Taylor(2002)と同様に、その十分条件について今後検討していく必要があると考えられる。

近年、Daniels et al.(2012)により自然な直接効果や間接効果をベイズ流に推定する方法が提案されている。サロゲートエンドポイントの評価においては、潜在結果変数に基づくものを含めて複数のアプローチが提案されている(Joffe and Greene, 2009; VanderWeele, 2013c)。これら関連した研究や、本節で論じた最新の研究も踏まえて、直接効果・間接効果に関する研究を更に発展させていくことを検討していきたいと考えている。

## 謝 辞

本論文の内容に関して、作成期間を通じて様々なご支援・ご教示を頂きました大分大学の和泉志津恵先生に深く感謝申し上げます。また、草稿段階で有益なコメントを下さいました統計数理研究所の黒木学先生、滋賀医科大学の村上義孝先生、防衛医科大学の佐藤弘樹先生に御礼申し上げます。なお、本論文の一部は、科学研究費補助金(若手研究(B))課題番号24700278の補助を受けて行った。

## 参 考 文 献

- Baron, R. M. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, *Journal of Personality and Social Psychology*, **51**, 1173–1182.
- Chiba, Y. and Suzuki, E. (2012). Causal inference with intermediates: Simple methods for principal strata effects and natural direct effects, *Current Topics in Public Health*(ed. A.J. Rodriguez-Morales), 37–60, InTech, Rijeka.

- Chiba, Y. and Taguri, M. (2013). Alternative monotonicity assumptions for improving bounds on natural direct effects, *International Journal of Biostatistics*, **9**, 235–249.
- Cole, S. R. and Hernán, M. A. (2002). Fallibility in estimating direct effects, *International Journal of Epidemiology*, **31**, 163–165.
- Cole, S. R., Platt R. W., Schisterman, E. F., Chu, H., Westreich, D., Richardson, D. and Poole, C. (2010). Illustrating bias due to conditioning on a collider, *International Journal of Epidemiology*, **39**, 417–420.
- Daniels, M. J., Roy, J. A., Kim, C., Hogan, J. W. and Perri, M. G. (2012). Bayesian inference for the causal effect of mediation, *Biometrics*, **68**, 1028–1036.
- Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*, Chapman and Hall, New York.
- Goetgeluk, S., Vansteelandt, S. and Goetghebeur, E. (2008). Estimation of controlled direct effects, *Journal of the Royal Statistical Society Series B*, **70**, 1049–1066.
- Hafeman, D. M. and Schwartz, S. (2009). Opening the black box: A motivation for the assessment of mediation, *International Journal of Epidemiology*, **38**, 838–845.
- Hafeman, D. M. and VanderWeele, T. J. (2011). Alternative assumptions for the identification of direct and indirect effects, *Epidemiology*, **22**, 753–764.
- Imai, K. and Yamamoto, T. (2013). Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments, *Political Analysis*, **21**, 141–171.
- Imai, K., Keele, L., Tingley, D. and Yamamoto, T. (2010a). Advances in social science research using R, *Causal Mediation Analysis Using R* (ed. H. D. Vinod), 129–154, Springer, New York.
- Imai, K., Keele, L. and Yamamoto, T. (2010b). Identification, inference and sensitivity analysis for causal mediation effects, *Statistical Science*, **25**, 51–71.
- Imai, K., Keele, L. and Tingley, D. (2010c). A general approach to causal mediation analysis, *Psychological Methods*, **15**, 309–334.
- Joffe, M. M. and Greene, T. (2009). Related causal frameworks for surrogate outcomes, *Biometrics*, **65**, 530–538.
- Kotelchuck, M. (1994). The adequacy of prenatal care utilization index: Its US distribution and association with low birthweight, *American Journal of Public Health*, **84**, 1486–1489.
- 黒木 学, 小林史明(2012). 構造的因果モデルについて, 計量生物学, **32**, 119–144.
- MacKinnon, D. (2008). *An Introduction to Statistical Mediation Analysis*, Lawrence Erlbaum Associates, New York.
- Otsuka, K., Taguri, M., Dennis, C. L., Wakutani, K., Awano, M., Yamaguchi, T. and Jimba, M. (2014). Effectiveness of a breastfeeding self-efficacy intervention: Do hospital practices make a difference?, *Maternal and Child Health Journal*, **18**, 296–306.
- Partridge, S., Balayla, J., Holcroft, C. A. and Abenhaim, H. A. (2012). Inadequate prenatal care utilization and risks of infant mortality and poor birth outcome: A retrospective analysis of 28,729,765 U.S. deliveries over 8 years, *American Journal of Perinatology*, **29**, 787–793.
- Pearl, J. (2001). Direct and indirect effects, *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* (eds. J. Breese and D. Koller), 401–420, Morgan Kaufmann, San Francisco.
- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed., Cambridge University Press, New York.
- Pearl, J. (2011). Principal stratification — A goal or a tool?, *International Journal of Biostatistics*, **7**, Issue 1, Article 20.
- Pearl, J. (2012a). The mediation formula — A guide to the assessment of pathways and mechanisms, *Causality: Statistical Perspectives and Applications* (eds. C. Berzuini, P. Dawid and L. Bernardinelli), 151–179, Wiley, New York.

- Pearl, J. (2012b). Interpretable conditions for identifying direct and indirect effects, Technical Report, R-389, Department of Computer Science, University of California, Los Angeles.
- Petersen, M., Sinisi, S. and van der Laan, M. (2006). Estimation of direct causal effects, *Epidemiology*, **17**, 276–284.
- Robins, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects, *Highly Structured Stochastic Systems* (eds. P. Green, N. L.Hjort and N. Richardson), 70–81, Oxford University Press, New York.
- Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects, *Epidemiology*, **3**, 143–155.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, **66**, 688–701.
- Sjölander, A. (2009). Bounds on natural direct effects in the presence of confounded intermediate variables, *Statistics in Medicine*, **28**, 558–571.
- Taylor, J., Wang, Y. and Thibaut, R. (2005). Counterfactual links to the proportion of treatment effect explained by a surrogate marker, *Biometrics*, **61**, 1102–1111.
- Tchetgen Tchetgen, E. J. and Shpitser, I. (2012). Semiparametric theory for causal mediation analysis: Efficiency bounds, multiple robustness, and sensitivity analysis, *Annals of Statistics*, **40**, 1816–1845.
- Tchetgen Tchetgen, E. J. and VanderWeele, T. J. (2012). On identification of natural direct effects when a confounder of the mediator is directly affected by exposure, Harvard University Biostatistics Working Paper, 148.
- Valeri, L. and VanderWeele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros, *Psychological Methods*, **18**, 137–150.
- VanderWeele, T. J. (2009a). Concerning the consistency assumption in causal inference, *Epidemiology*, **20**, 880–883.
- VanderWeele, T. J. (2009b). Mediation and mechanism, *European Journal of Epidemiology*, **24**, 217–224.
- VanderWeele, T. J. (2009c). Marginal structural models for the estimation of direct and indirect effects, *Epidemiology*, **20**, 18–26.
- VanderWeele, T. J. (2011). Controlled direct and mediated effects: Definition, identification and bounds, *Scandinavian Journal of Statistics*, **38**, 551–563.
- VanderWeele, T. J. (2013a). A three way decomposition of a total effect into direct, indirect, and interactive effects, *Epidemiology*, **24**, 224–232.
- VanderWeele, T. J. (2013b). Policy-relevant proportions for direct effects, *Epidemiology*, **24**, 175–176.
- VanderWeele, T. J. (2013c). Surrogate measures and consistent surrogates (with discussion), *Biometrics*, **69**, 561–581.
- VanderWeele, T. J. and Chiba, Y. (2014). Sensitivity analysis for direct and indirect effects in the presence of exposure-induced mediator-outcome confounders, *Epidemiology, Biostatistics, and Public Health* (in press).
- VanderWeele, T. J. and Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition, *Statistics and Its Interface*, **2**, 457–468.
- VanderWeele, T. J. and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome, *American Journal of Epidemiology*, **172**, 1339–1348.
- VanderWeele, T. J., Asomaning, K., Tchetgen Tchetgen, E. J., Han, Y., Spitz, M. R., Shete, S., Wu, X., Gaborieau, V., Wang, Y., McLaughlin, J., Hung, R. J., Brennan, P., Amos, C. I., Christiani, D. C. and Lin, X. (2012). Genetic variants on 15q25.1, smoking and lung cancer: An assessment

- of mediation and interaction, *American Journal of Epidemiology*, **75**, 1013–1020.
- Vansteelandt, S. and VanderWeele, T. J. (2012). Natural direct and indirect effects on the exposed: Effect decomposition under weaker assumptions, *Biometrics*, **68**, 1019–1027.
- Wang, Y. and Taylor, J. (2002). A measure of the proportion of treatment effect explained by a surrogate marker, *Biometrics*, **58**, 803–812.

## Estimation and Sensitivity Analysis of Direct and Indirect Effects

Masataka Taguri

Department of Epidemiology and Biostatistics, Graduate School of Medicine, Yokohama City University

Estimation of causal effects of an exposure on the disease occurrence is an overarching goal of epidemiological studies. Once it is established that an exposure affects disease risk, questions often arise as to the relative importance of different possible pathways for the effect. This can be done by effect decomposition of a total effect of an exposure into a direct and an indirect effect. This article reviews statistical methods on direct and indirect effects based on a potential outcome model. It discusses the definition, assumptions, and identification formulas for direct and indirect effects. It also discusses a sensitivity analysis method for these effects that is applicable when an assumption of no unmeasured confounding is violated. We apply these methods to the 2003 US birth certificate and infant death data provided by the National Center for Health Statistics.