

観察研究におけるバイアスの感度解析

竹内 文乃¹・野間 久史²

(受付 2013 年 10 月 30 日；改訂 2014 年 3 月 26 日；採択 3 月 27 日)

要 旨

疫学研究をはじめとする観察研究の結果には、ランダム誤差だけでなく各種バイアスの影響による不確実性が含まれる。具体的には対象者の選択や参加、追跡の偏り、測定誤差や未測定の変数因子の存在などが、状況によっては結果に重大な影響を与えるバイアスになるとされる。ただ、研究対象者や参加者が偏っているか、対象者の脱落がランダムに起きているかどうか、測定誤差はどの程度か、重要だが測定されていない変数因子がないかどうかは、研究の解析データを用いるだけでは評価することはできない。そこで、これらの影響を評価するには何らかの情報に基づいて感度解析を実施することが必要になるが、もしこのようなバイアスがあれば…というシナリオを複数設定しての感度解析は結果の解釈が困難となる場合が多いこともあり、必要性は十分に知られながらも、通常これらを考慮した解析が行われるケースは決して多くない。本稿では、近年、疫学研究で大きく発展したモンテカルロ法を用いる確率的感度解析に特に焦点を当て、バイアス解析に関する一連の方法および適用事例を総括する。

キーワード：観察研究，疫学研究，不確実性，バイアス，確率的感度解析。

1. はじめに

1995 年、Science 誌に ‘Epidemiology Faces Its Limits’ (Taubes and Mann, 1995) と題された論文が発表され、疫学に関わる研究者に大きな衝撃を与えた。この論文では、当時の疫学研究で報告された多くの健康に関するエビデンスに対して「まったく同じ曝露・疾病に関する研究を行っておきながら、結果が食い違ったり、結論が異なったりすることがある」と指摘しており、いくつかの具体的な事例を示しながら、「疫学はもはや信じるに足りないのではないか」という厳しい論評を下している。

慢性疾患の疫学研究の歴史では、早期の頃は、喫煙と肺がんの関連など、相対リスクが何十倍にも高まるリスク要因の解明に対して大きな成果を上げてきた。これらのリスク要因の評価では、そもそもの関連が非常に強い場合、高度な統計手法などを利用しなくても、関連を検出することができた。しかし、現代の疫学研究では、相対リスクがただかたかた 2 倍にもならない比較的軽微なリスク要因の評価が中心となっており、これらの因果関係は、観察研究では避けることができないバイアスや変数因子などによるさまざまな系統的な誤差の影響によって、たやすく捻じ曲げられてしまう。Taubes and Mann (1995) では、当代を代表する疫学者・生物統計学者へのインタビューによる解説もまとめられており、先述の健康に関するエビデンスが食い違ってしまう問題の最も大きな要因のひとつとして、この系統的な誤差の影響を挙げている。

¹ 国立環境研究所 環境健康研究センター：〒 305-8506 茨城県つくば市小野川 16-2; takeuchi.ayano@nies.go.jp

² 統計数理研究所：〒 190-8562 東京都立川市緑町 10-3; noma@ism.ac.jp

これにより、曝露効果の評価に用いられる生物統計手法も、それらの理論体系における解釈の妥当性が足元から崩れ去ることになる。Washington大学のNorman Breslow教授は、以下のように述べている。‘What is the fundamental quality of the data, and to what extent are there biases in the data that cannot be controlled by statistical analysis? One of the dangers of having all these fancy mathematical techniques is people will think they have been able to control for things that are inherently not controllable. (Taubes and Mann, 1995; p. 168)’ 20世紀後半からの統計理論のめざましい発展によって、複雑で高度な数理モデリングや統計解析手法の利用が可能となったが、これらによって得られる P 値・信頼区間はそもそも「系統的な誤差がまったくない」という前提のもとで、ランダムな誤差のみを評価したものである。当然ながら、この仮定が成り立たない場合、これらの手法は妥当な解釈ができない(Greenland, 1990; Rothman et al., 2008)。

バイアスの影響は、仮にその存在がわかったとしても、それがどの程度のものであるのかは知る術はなく、統計解析で調整することはできない。したがって、次善の策として、結果への影響を定量的に評価するために、感度解析(sensitivity analysis)の方法が、古くから研究されてきた(Greenland, 1996; Rothman et al., 2008; Lash et al., 2009)。しかしながら、従来の方法では、限られたシナリオと限定的な設定のもとでの評価しか可能ではなく、バイアスについての不確実性を考慮したり、複数のバイアスが複合的に影響したりする場合などには容易に適用することができないという強い制約があった。近年になって、これらの問題を解決するために、観察研究におけるバイアスの確率的感度解析(probabilistic sensitivity analysis)の方法論が飛躍的に発展した。この方法は、バイアスの不確実性を表すバイアスパラメータ(bias parameter)を統計モデルに導入し、その不確実性の程度を、バイアスパラメータの確率分布として組み込んで、曝露効果の推定の感度解析を行うという方法である(Hoffman and Hammonds, 1994; Lash and Silliman, 2000; Greenland, 2001, 2003, 2005; Lash and Fink, 2003; Phillips, 2003)。特に、近年では、モンテカルロシミュレーションによって、曝露効果の推定量の不確実性を総合的に評価する、数値的な感度解析の方法論が発展しており、モンテカルロ感度解析(Monte-Carlo sensitivity analysis)やモンテカルロリスク評価(Monte-Carlo risk assessment)という名称で、広まりつつある。その数理的な枠組みは、モンテカルロシミュレーションに基づく比較的簡便なものであり、複数のバイアスの影響を同時に評価する場合にも、複合バイアスマデリング(multiple bias modeling; Greenland, 2005)によるベイズ流の枠組みのもとで、共通のモンテカルロ感度解析の方法を利用した評価が可能である。

本稿では、疫学の観察研究で、新たに確立しつつある、確率的感度解析や複合バイアスマデリングを、モンテカルロ感度解析の枠組みで扱う方法論について、事例を交えた平易な解説を行う。本稿では、最も基本的な 2×2 表の解析における解説を行うこととして、Greenland et al. (1994)による職業性の樹脂曝露と肺がんについてのケースコントロール研究を事例とした解説を行う。2節で、この事例を用いた、未測定交絡因子、測定バイアス、選択バイアスについてのモンテカルロ感度解析の具体的な方法と、解析事例の紹介を行う。3節で、このモンテカルロ感度解析の理論的な解説を行い、この方法が本質的にベイズ流の解析手法として正当化されることを示す。最後に4節で、本稿のまとめを行う。

2. 確率的感度解析：Greenland et al.(1994)を事例に

ここでは、Greenland(1996)、Greenland and Lash(2008)による、職業性の樹脂(resin)曝露と肺がんについてのケースコントロール研究(Greenland et al., 1994)を事例として、 2×2 表のオッズ比の推定におけるモンテカルロ感度解析の事例を紹介する。表1に、このケースコントロー

表 1. 樹脂曝露と肺がんに関するケースコントロール研究 (Greenland et al., 1994).

	樹脂曝露あり ($X = 1$)	樹脂曝露なし ($X = 0$)	合計
Cases ($D = 1$)	$A_{1+} = 45$	$A_{0+} = 94$	$M_{1+} = 139$
Controls ($D = 0$)	$B_{1+} = 257$	$B_{0+} = 945$	$M_{0+} = 1202$

ル研究の結果を示している。曝露変数を $X (= 0, 1)$ ，結果変数を $D (= 0, 1)$ とする。この研究では、あらかじめ年齢と死亡年での調整を行っており、無調整のオッズ比は 1.77 (95%信頼区間：1.18, 2.64) となっている。

2.1 未測定の交絡因子によるバイアス

この事例では、肺がんの重要なリスク要因である喫煙についてのデータは得られておらず、喫煙による調整は行われていない。当然ながら、喫煙と樹脂曝露との間に相関があるようであれば、このオッズ比の推定値には交絡によるバイアスが生じる可能性がある。ここでは、喫煙を未測定の交絡因子と考え、未測定の交絡因子を $Z (= 0, 1)$ として、 Z の影響を調整したオッズ比の感度解析を行うことを考える。当然ながら、この感度解析を行うためには、(1) 樹脂曝露群、非曝露群ごとの喫煙-肺がんのオッズ比、(2) コントロール群における樹脂曝露群、非曝露群ごとの喫煙者の割合、(3) コントロール群における樹脂曝露者の割合、の 3 つの要因が規定されなくてはならない。(3) については、表 1 のデータから既知であるが、(1)、(2) については、そもそも未知量であるため、なんらかのシナリオを想定したもとの感度解析を行うしかない。さて、ここでは、樹脂曝露群、非曝露群ごとの喫煙-肺がんのオッズ比を OR_{DZ} (層ごとにオッズ比は共通であるという仮定を置くが、共通オッズ比の仮定が成り立たない場合への拡張は容易である) とし、樹脂曝露群と非曝露群における喫煙者の割合を P_{Z1}, P_{Z0} とする。この 3 つのパラメータを、感度解析におけるバイアスの大きさを規定するパラメータとして、バイアスパラメータ (bias parameter) と呼ぶ。2 値の交絡因子 Z による調整を行う場合の層ごとの期待セル度数を、表 2 のように表記することとする。コントロール群が、源泉集団を代表するものであると仮定すると $B_{11} = P_{Z1}B_{1+}, B_{01} = P_{Z0}B_{0+}$ が成り立つ。層ごとの OR_{DZ} が共通であるという仮定から、

$$OR_{DZ} = \frac{A_{11}B_{10}}{A_{10}B_{11}}, \quad OR_{DZ} = \frac{A_{01}B_{00}}{A_{00}B_{01}}$$

という 2 つの方程式を解くと、

$$(2.1) \quad \begin{aligned} A_{11} &= OR_{DZ}A_{1+}B_{11}/(OR_{DZ}B_{11} + B_{1+} - B_{11}) \\ A_{01} &= OR_{DZ}A_{0+}B_{01}/(OR_{DZ}B_{01} + B_{0+} - B_{01}) \end{aligned}$$

という関係式が成り立つ。これによって得られる $A_{11}, A_{01}, B_{11}, B_{01}$ を表 2 に代入することにより、 Z を調整したオッズ比 OR_{DX} の推定値を得ることができる。このように、2 値変数の交絡因子の外的な調整による感度解析の基本的な考え方は、はじめ Cornfield et al. (1959) によって与えられ、Bross (1966, 1967), Schlesselman (1978), Yanagawa (1984), Axelson and Steenland (1988), Gail et al. (1988), Simon (1980) などによって発展した。3 水準以上の場合の方法論も、Greenland (1987), Flanders and Khoury (1990), Rosenbaum (1995) などによって研究されている。これらの古典的な感度解析では、バイアスパラメータ P_{Z1}, P_{Z0}, OR_{DZ} になんらかの値を入れたもとのバイアス調整推定量 OR_{DX} の値を求める。例として、Greenland (1996)

表 2. 感度解析のデータのレイアウト(2 値の交絡因子 Z による調整を行う場合;セルごとの期待度数).

	Z = 1			Z = 0		
	X = 1	X = 0	合計	X = 1	X = 0	合計
Cases						
D = 1	A_{11}	A_{01}	M_{11}	A_{10}	A_{00}	M_{10}
Controls						
D = 0	B_{11}	B_{01}	M_{01}	B_{10}	B_{00}	M_{00}

表 3. 未測定交絡因子のバイアスの感度解析の結果(Greenland, 1996, p.1109 より引用).
 P_{Z1} , P_{Z0} , OR_{DZ} の値ごとのオッズ比 OR_{DX} の推定値.

P_{Z1}	P_{Z0}	OR_{DZ}	OR_{DX}		
			5	10	15
0.40	0.30	1.56	1.49	1.42	1.39
0.55	0.45	1.49	1.54	1.49	1.48
0.70	0.60	1.56	1.57	1.54	1.53
0.45	0.25	2.45	1.26	1.13	1.09
0.60	0.40	2.25	1.35	1.27	1.24
0.75	0.55	2.45	1.41	1.35	1.33

による,いくつかのシナリオを仮定したもとの OR_{DX} を表 3 に示す.ほとんどすべてのシナリオで,喫煙による交絡は, OR_{DX} の推定に影響を及ぼしていることがわかるが,その程度の大きさはシナリオによって異なり,樹脂曝露が肺がんの発症に与える影響は概ね否定できないという結論にはなりそうである.

ただし,当然ながら,これらの解析には,いくつかの重要な問題があり,例えば,バイアスパラメータは,単一もしくは有限個の値に固定され,そのそれぞれで曝露効果の推定値が得られるのみであり,「バイアスパラメータの値の尤もらしさ(plausibility)」といった「シナリオの重み」を考慮することができない.また,従来信頼区間のように,推定の不確実性を定量的に評価することもできない.これらは,感度解析の結果の解釈そのものに影響を及ぼすものである.近年になって,疫学の観察研究で発展した確率的感度解析の方法は,このバイアスについてのシナリオの不確実性を考慮した総合的な評価を可能とする方法である.具体的には,ベイズ統計学と同じように,不確実性の程度を,バイアスパラメータの確率分布として定量的にモデル化し,それを曝露効果の推定に対する不確実性として組み込むという方法である.このバイアスパラメータを組み込んだモデルを,バイアスマデル(bias model)という.ここでは,Greenland and Lash(2008)と同じく,以下のバイアスマデルを仮定する.

1. P_{Z1} , P_{Z0} には,それぞれ独立に連続型一様分布 $U(0.40, 0.70)$ を仮定する.この一様分布の上限・下限は,当時の米国の喫煙者の割合が半数以上であることを反映して定められている(Greenland and Lash, 2008).

2. OR_{DZ} には, 95%確率区間が 5, 15 となるような対数正規分布を仮定する. すなわち, $\log(OR_{DZ}) \sim N(2.159, 0.280^2)$ とする.

モンテカルロ感度解析では, このバイアスモデルからのバイアスパラメータのサンプリングを行い, 「バイアスパラメータの値の尤もらしさ」を反映した, バイアス調整オッズ比の分布を求める. このバイアス調整オッズ比の分布も, ベイズ統計学と同じ要領で「バイアスの不確実性を反映した OR_{DX} の確率分布(後に述べるように, これはベイズ推定を近似しているものと解釈することができるので, そのまま「 OR_{DX} の事後分布」と解釈できる)」という解釈になる. 具体的には, 以下のようなアルゴリズムで OR_{DX} の分布を求める.

1. 先に仮定したバイアスモデルから P_{Z1}, P_{Z0}, OR_{DZ} のサンプリングを行う.
2. これらのバイアスパラメータの値を用いて, (2.1)式から $A_{11}, A_{01}, B_{11}, B_{01}$ を求め, バイアス調整オッズ比の推定値 $OR_{DX,k}$ を求める.
3. 十分多くの回数(数万~数十万回), 1, 2のステップを繰り返し $OR_{DX,k}$ の分布を求める.

ここでは, 50,000回のサンプリングを行い, $OR_{DX,k}$ ($k = 1, 2, \dots, 50,000$) の分布を求めた. 図 1(b)に, この分布を示している. 分布の中央値は, 1.76 であり, 無調整オッズ比とほとんど同じ値になっている. 2.5, 97.5 パーセンタイルは 1.25, 2.48 となっており, ここで仮定したバイアスモデルの不確実性は, データそのものが持つランダムな誤差による不確実性と概ね同程度であるということがわかる. 比較のために, 図 1(a)に, 無調整オッズ比の漸近分布 $N(\log(1.77), 0.1944^2)$ からの 50,000 の乱数のヒストグラムを示している. ここでは, 単純なバイアスモデルを仮定したが, バイアスパラメータ間の相関を考慮したり, より複雑な分布を利用することも可能である. 詳細については, Greenland and Lash (2008) をご参照いただきたい.

一方で, ここで得られたバイアス調整オッズ比 $OR_{DX,k}$ の 95%区間 (1.25, 2.48) は, バイアスの不確実性のみを考慮したものであり, 従来の統計手法が考慮している, 統計的なランダムな誤差を無視している. これに, ランダムな誤差による不確実性を組み込んだものとするためには, 上記のモンテカルロシミュレーションで得られる OR_{DX} の分布にランダムな誤差による不確実性を上乗せする必要がある. これについては, OR_{DX} の最尤推定量の漸近分布から, 無調整対数オッズ比の標準誤差 $SE = (1/24 + 1/94 + 1/257 + 1/945)^{1/2} = 0.1944$ をもとにした誤差項を, 上記のアルゴリズムで得られる $OR_{DX,k}$ に加えてやればよい.

1. 想定したバイアスモデルから P_{Z1}, P_{Z0}, OR_{DZ} のサンプリングを行う.
2. これらのバイアスパラメータの値を用いて, (2.1)式から $A_{11}, A_{01}, B_{11}, B_{01}$ を求め, バイアス調整をしたオッズ比の推定値 $OR_{DX,k}$ を求める.
3. ランダムな誤差を反映する変数 e_k を $N(0, 0.1944^2)$ からサンプリングし, $OR_{DX,k}$ に加え, $OR_{DX,k}^* = \exp[\log(OR_{DX,k}) + e_k]$ を求める.
4. 十分多くの回数(数万~数十万回), 1~3のステップを繰り返し, $OR_{DX,k}^*$ の分布を求める.

50,000回のサンプリングを行った $OR_{DX,k}^*$ ($k = 1, 2, \dots, 50,000$) の分布を, 図 1(c)に示した. この分布は「バイアスとランダムな誤差の両方の不確実性を反映したオッズ比 OR_{DX} の確率分布」となる. 分布の中央値は 1.76, 95%区間は (1.05, 2.96) となる. 無調整オッズ比の 95%信頼区間に比べると, 1.30 倍ほど広がっており, これが, バイアスモデルによって加わった不確実性となる. 3のステップでは, 古典的な対数オッズ比の漸近正規性によるラフな近似を用いているが, より高い精度が期待できるリサンプリングや Bootstrap 法を用いたアプローチも可能である (Greenland and Lash, 2008).

4 節で, 詳細についての解説は行うが, このモンテカルロ感度解析のアルゴリズムは, バイ

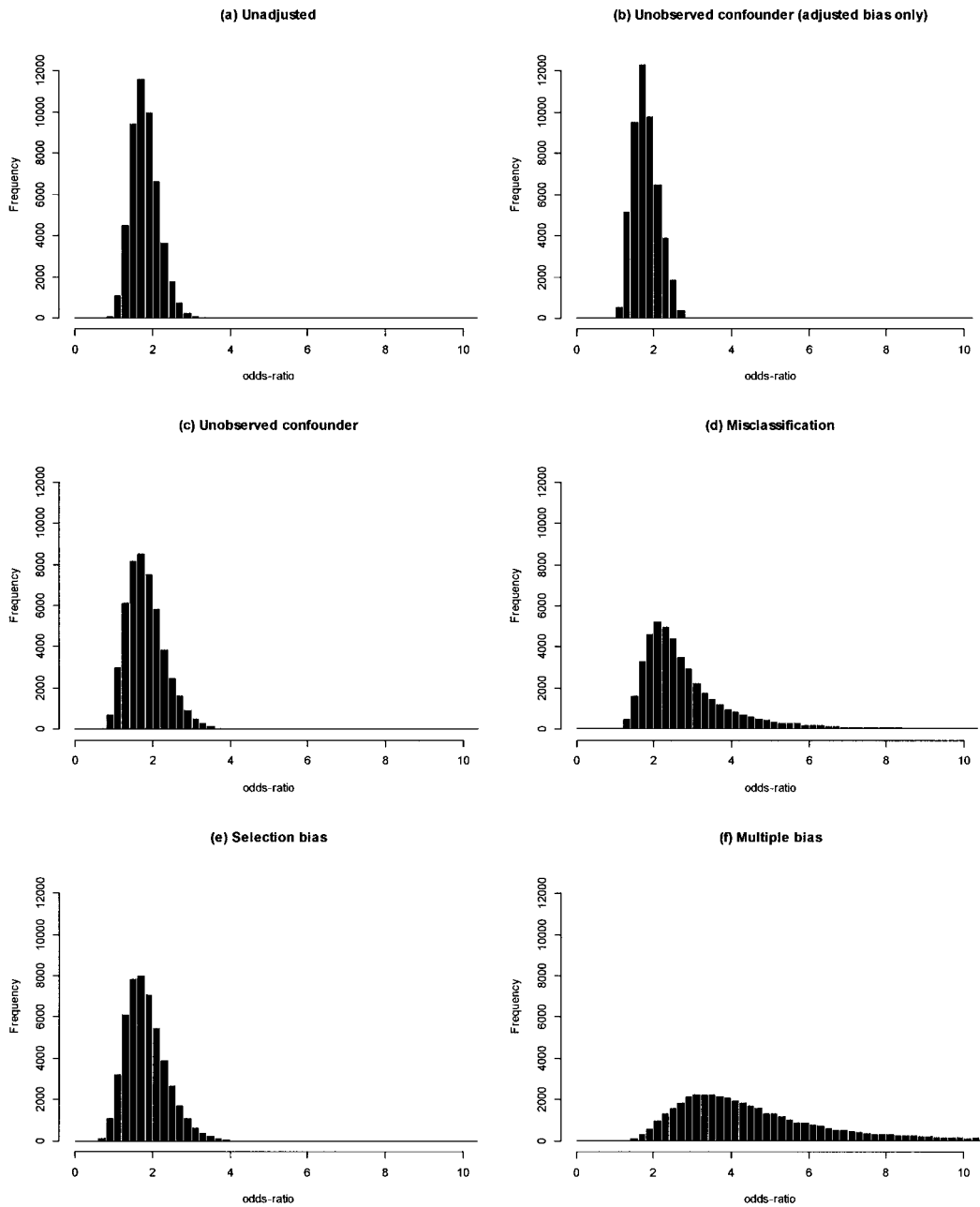


図 1. モンテカルロ感度解析における OR_{DX} の分布：(a)無調整の OR_{DX} の漸近分布からのサンプル，(b)未測定交絡因子によるバイアス(バイアスのみの不確実性を反映した調整オッズ比)，(c)未測定交絡因子によるバイアス(バイアスとランダムな誤差の不確実性を反映した調整オッズ比)，(d)誤分類バイアス，(e) 選択バイアス，(f) 複合バイアスモデル。

アスパラメータを組み込んだベイズモデルのもとでの OR_{DX} の周辺事後分布からのサンプリングを近似しており、上記のアルゴリズムは、その近似的なベイズ推定量と 95%信用区間を求めていることに他ならない。

2.2 測定誤差によるバイアス

ほとんどすべての疫学研究で、なんらかの形での測定バイアス (measurement bias) は存在する。一般的に、離散変数では、誤分類 (classification error, misclassification) ともいわれる。測定誤差の影響の評価方法についての研究の歴史は古く、簡便な解析的方法 (Copeland et al., 1977; Greenland, 1982; Kleinbaum et al., 1984) から、高度な数値計算を用いた方法 (Barron, 1977; Greenland and Kleinbaum, 1983; Carroll et al., 1984, 1995) まで、多くの方法が開発されている。ここでは、単純に表 1 のケースコントロール研究の分割表の解析での曝露変数の誤分類の問題のみを考える。ケース群、コントロール群での誤分類の程度に差がない場合を non-differential な誤分類といい、差がある場合を differential な誤分類という。例えば、ケースコントロール研究での思い出しバイアス (recall bias) は、differential な誤分類にあたる。

記法として、真の曝露変数 $X (= 0, 1)$ と、測定誤差を伴って実際に観測 (分類) される変数 $X^* (= 0, 1)$ を定義する。分類の正確性を表す以下の 4 つのバイアス パラメータを定義する。

$$\begin{aligned} Se &= \text{Sensitivity} \\ &= p(X^* = 1 | X = 1) \\ Fn &= \text{False-negative probability} \\ &= p(X^* = 0 | X = 1) = 1 - Se \\ Sp &= \text{Specificity} \\ &= p(X^* = 0 | X = 0) \\ Fp &= \text{False-positive probability} \\ &= p(X^* = 1 | X = 0) = 1 - Sp \end{aligned}$$

ここで、コントロール群において、真に曝露を受けた対象者 ($X = 1$) の数を B_1 、真に曝露を受けていない対象者 ($X = 0$) の数を B_0 とすると、「曝露を受けた」と分類される対象者の期待数は、

$$(2.2) \quad B_1^* = SeB_1 + FpB_0$$

同じく、「曝露を受けていない」と分類される対象者の期待数は、

$$(2.3) \quad B_0^* = FnB_1 + SpB_0$$

となる。 $Se + Fn = Sp + Fp = 1$ なので B_0^*, B_1^* の合計は、誤分類によって不変であり、

$$\begin{aligned} M_0 &= B_1 + B_0 = (Se + Fn)B_1 + (Sp + Fp)B_0 \\ &= SeB_1 + FpB_0 + FnB_1 + SpB_0 = B_1^* + B_0^* \end{aligned}$$

となる。実際のデータからは B_0^*, B_1^* しか観測することはできない。式 (2.3) から、

$$B_0 = (B_0^* - FnB_1) / Sp$$

であり、式 (2.2) にこれを代入して、 B_1 について解くと、

$$(2.4) \quad B_1 = (SpB_1^* - FpB_0^*) / (SeSp - FnFp)$$

表4. 誤分類バイアスの感度解析の結果(Greenland, 1996, p.1111より引用). Se , Sp の値ごとのオッズ比 OR_{DX} の推定値.

Cases		Controls				
Se	Sp	Se :	0.90	0.80	0.90	0.80
		Sp :	0.90	0.90	0.80	0.80
0.90	0.90		2.34*	2.00	19.3	16.5
0.80	0.90		2.83	2.42*	23.3	19.9
0.90	0.80		1.29	1.11	10.7*	9.1
0.80	0.80		1.57	1.34	12.9	11.0*

* Non-differential な誤分類.

となる. $B_0 = M_0 - B_1$ であるため, これにより, (B_0, B_1) を計算することができる. こうして得られる B_0, B_1 は, 感度・特異度についての仮定を置いたもとで得られる推定値(期待値)である. $SeSp = FnFp$ である場合, これらの解は不定となり, $SeSp < FnFp$ の場合, 負値をとる. 後者の場合, 曝露変数の分類が, でたため(無作為)な分類よりも悪いということを意味している.

一般的には, 曝露変数の分類は, 無作為な分類よりも正確であるという前提で感度解析を行う. 式(2.4)と同様に, ケース群における, 真に曝露を受けた人数を A_1 , 真に曝露を受けていない人数を A_0 とすると,

$$(2.5) \quad A_1 = (SpA_1^* - FpA_0^*) / (SeSp - FnFp)$$

となる. また, ケースの総数 M_1 に対して, $A_0 = M_1 - A_1$ となる. 古典的な誤分類の感度解析は, (Se, Sp) に適当な値を与えたもとの, (2.4) (2.5)式からセル度数の期待値を求め, バイアス調整を行ったオッズ比 OR_{DX} を求める. Greenland(1996)による, いくつかのシナリオのもとの感度解析の結果を, 表4に示した. よく知られているように, Non-differential な誤分類のもとでは, 誤分類によるオッズ比のバイアスは必ず帰無仮説の(保守的な)方向に入る. これについては, 曝露変数が3水準以上であったり, 他の変数に誤分類が入ったりする場合には, その限りではない(Dosemeci et al., 1990). また, 表4の結果からは, ケース群がコントロール群よりも「より曝露を受けたと誤分類される傾向が大きい」という場合でも, バイアス調整オッズ比が, 無調整オッズ比よりも大きく出るというシナリオも出ている. これは, いわゆるケースコントロール研究での「思い出しバイアス」の設定であり, 思い出しバイアスが必ずしも曝露効果を過大に推定するバイアスを生じさせるわけではないことを示している. また, この事例では, Se よりも Sp のほうが推定値のバイアスに与える影響が大きく出ている. これは, そもそも曝露を受けた対象者の割合が小さいためである.

この設定でも, 前節と同様に Se, Sp をバイアスパラメータとして, 同じ手順で, モンテカルロ感度解析を実行することができる. ここでは, Non-differential な誤分類の仮定を置き, ケース群・コントロール群の Se, Sp が共通で, 台形分布 Trapezoidal (0.75, 0.85, 0.95, 1.00) に従うという仮定のもとでのモンテカルロ感度解析を行う. ランダムな誤差も組み込んだ OR_{DX} の分布を, 図1(d)に示す. 誤分類バイアスのみを調整したオッズ比の分布の中央値は2.45, 95%区間は(1.87, 14.25)であり, ランダムな誤差も含めた調整オッズ比の中央値は2.56, 95%区間は(1.49, 14.68)であった. バイアス調整オッズ比の95%区間は, 無調整オッズ比の95%信頼区間

よりもかなり広く、誤分類バイアスによる不確実性は、この事例に関しては、ランダムな誤差よりもかなり強いと考えられる。

同様の誤分類の感度解析についての研究は、古くから、疾病の誤分類 (Greenland, 1996)、交絡因子の誤分類 (Savitz and Baron, 1989)、複数の変数についての誤分類 (Kristensen, 1992; Chavance et al., 1992)、バリデーション研究のデータを用いた評価方法 (Wacholder et al., 1993; Tennenbein, 1970; Green, 1983; Marshall, 1990; Brenner and Gefeller, 1993) など、多くの方法論が開発されている。モンテカルロ感度解析は、これらの方法にも、同様に適用することができる。

2.3 選択バイアス

選択バイアス (selection bias) は、結論を適用したいターゲット集団からの対象者の選択におけるバイアスであり、ケースコントロール研究では、コントロールの選択におけるバイアスの問題が古くから問題とされてきた (Rothman et al., 2008)。例えば、Science 誌での Taubes and Mann (1995) でも、Boston 大学の Charles Poole 教授が、1980 年代の低レベル電磁場の曝露と白血病・脳腫瘍のケースコントロール研究を行った際の Random digit dialing によるコントロールの選択 (ランダムに構成した電話番号で、コントロールを選び出す方法) でのバイアスの実例を紹介しており、この研究では、社会経済的に低い階層の人々が、コントロール群に選ばれにくいというバイアスがあったと述べている。貧しい階層のひとは、そもそもこのような研究に参加したがないという傾向があり、また、日中、自宅に居ることも少なく、当時の留守番電話機能のある電話を持っていないひが多かったという (Taubes and Mann, 1995)。この選択バイアスにより、この研究では、対象者集団の偏りはもちろん、統計解析の結果でも、貧しい階層に顕著な母乳育児、母親の喫煙などが、リスク要因として上がってきたという。

選択バイアスの感度解析におけるケースコントロール研究のオッズ比の分解方法はよく知られたものがある (Kleinbaum et al., 1984; Rothman et al., 2008)。 S_{A1}, S_{B1} を、それぞれ曝露を受けたケース・非ケースが、ケース群・コントロール群へサンプリングされる確率であるとする。このとき、ケースコントロール研究でのそれぞれのサンプリング数の期待値は、 $A_1/S_{A1}, B_1/S_{B1}$ となる。同様に、 S_{A0}, S_{B0} を、それぞれ曝露を受けていないケース・非ケースのサンプリング確率とすると、それぞれのサンプリング数の期待値は、 $A_0/S_{A0}, B_0/S_{B0}$ となる。このとき、選択バイアスを調整したオッズ比は、

$$\frac{(A_1/S_{A1})(B_0/S_{B0})}{(A_0/S_{A0})(B_1/S_{B1})} = \left(\frac{A_1 B_0}{A_0 B_1} \right) \left(\frac{S_{A1} S_{B0}}{S_{A0} S_{B1}} \right)^{-1}$$

と書ける。すなわち、選択バイアスの調整オッズ比は、単純に無調整オッズ比を、選択バイアスの調整因子 $S^* = S_{A1} S_{B0} / S_{A0} S_{B1}$ で割ったものとして表すことができる。当然ながら、 $S^* = 1$ のとき、選択バイアスの影響はなくなり、ケースコントロールオッズ比で、偏りのない推定が可能である。これは、例えば、疾病の有無・曝露の有無と対象者の選択が独立である場合に成り立つ (Greenland, 1996; Greenland and Lash, 2008)。

モンテカルロ感度解析は、 $S_{A1}, S_{B1}, S_{A0}, S_{B0}$ をバイアスパラメータとしても実行可能であるが、ここではバイアス調整因子 S^* を直接バイアスパラメータと見なして、感度解析を行う。バイアスモデルは、95% 確率区間が (0.70, 1.50) となるような対数正規分布を仮定する。すなわち、 $\log(S^*) \sim N(0, 0.21^2)$ とする。ランダムな誤差も組み込んだ OR_{DX} の分布は、図 1(e) となる。選択バイアスを調整したオッズ比の分布の中央値は 1.76、95% 区間は (1.17, 2.65)、ランダムな誤差を考慮した調整オッズ比の分布の中央値は 1.76、95% 区間は (1.01, 3.10) であった。

このバイアスモデルは、単純なケースコントロール研究に限らず拡張可能である。たとえば Density sampling を行ったケースコントロール研究の場合には、 $S_{A1}, S_{B1}, S_{A0}, S_{B0}$ が人時間

表 5. Greenland et al. (1994)の事例におけるモンテカルロ感度解析の結果.

バイアスモデル	ランダムな誤差を含めない 場合		ランダムな誤差を含めた 場合	
	Median	2.5th & 97.5th Percentiles	Median	2.5th & 97.5th Percentiles
1. なし (無調整な解析)	1.77	—	1.77	1.18, 2.64
2. 未測定交絡因子によるバイアス	1.76	1.25, 2.48	1.76	1.05, 2.96
3. 誤分類バイアス	2.45	1.87, 14.25	2.56	1.49, 14.68
4. 選択バイアス	1.76	1.17, 2.65	1.76	1.01, 3.10
5. 複合バイアスモデル (2-4を同時にモデル化)	4.26	2.24, 27.66	4.32	2.00, 28.30

あたりの率となる。また、Scharfetein et al. (1999)では、線形回帰モデルへの拡張も行われている。同様のバイアスモデルを用いて、Greenland (2003)やLash and Fink (2003)のように対象者の参加やレスポンス、追跡・脱落に起因するバイアスも調整をすることも可能である。また、メタアナリシスでは、実施された研究の出版における選択プロセスのバイアスは出版バイアス (publication bias)として知られており、多くの感度解析の方法が研究されているが、Bowden et al. (2010)によって、モンテカルロ感度解析を利用した議論も行われている。

2.4 複合バイアスモデリング

ここまでは、それぞれのバイアスを個別に調整する感度解析の方法を説明してきたが、複数のバイアスが同時に懸念される場合、それらを同時に調整した感度解析を行う必要がある。古典的な感度解析の方法では、複数のバイアスを多次元的にモデル化し、調整オッズ比を求めることには、まずは計算上の困難が伴い、解析的な表現が陽に求まらないこともある。仮に、数値的に取り扱うことができたとしても、それぞれのバイアスについて複数のシナリオを考慮する必要がある、そのすべての組み合わせを考慮すると、全体としてのシナリオは膨大なものとなる。自然、解釈も複雑なものとなる。また、バイアス調整の順番が入れ替わると、調整オッズ比の値が変わってしまうという可能性もある (Greenland, 1996; Greenland and Lash, 2008)。この場合にも、モンテカルロ感度解析によるバイアスの同時調整は可能であり、複数のバイアスをモデル化した方法として、複合バイアスモデリング (multiple bias modeling) が提案されている (Greenland, 2005)。次節で解説する通り、これは、ベイズ流の解析としての解釈が可能であり、バイアスパラメータを複数同時にモデル化したベイズモデルにおいて、近似的な事後サンプリングを行うものとなる。バイアス間の関連についても、バイアスパラメータの同時分布における相関のモデル化によって表現することができる。

事例として、2.1節から2.3節で考慮したバイアスを独立にモデル化した複合バイアスモデルにおけるモンテカルロ感度解析を考える。バイアスパラメータは独立であるという仮定になるため、前節までのアルゴリズムの単純な拡張として、未調整のオッズ比に対して、個々のバイアス調整の方法を順に適用していき、それぞれのバイアスの不確実性を加算してだけでよい (Greenland and Lash, 2008)。バイアスの不確実性のみを考慮した調整オッズ比の分布では、中央値は4.26、95%区間は(2.24, 27.66)となり、複数のバイアスを同時に考慮している分、不確実性は大きくなる。ランダムな誤差も考慮すると、中央値は4.32、95%区間は(2.00, 28.30)とな

る。後者の OR_{DX} の分布は、図 1(f) となる。

2 節での感度解析の結果を、表 5 にまとめた。いずれのバイアスモデルのもとでも、95% 区間の下限は、帰無仮説を跨いでいない。特に、すべてのバイアスを複合バイアスモデルで組み込んだ場合でも、この結果は一貫して成り立っている。すなわち、今回の解析で考慮したシナリオのもとでは(概ね Greenland and Lash (2008) によるものと同等)、結論は概ね不変であり、これらの仮定のもとでのバイアスに対するロバストネスが確認できたと言うことはできるだろう。感度解析の結果の解釈・報告の方法についての詳細は、Lash et al. (2009) が参考になるだろう。

3. 確率的感度解析の理論

ここまで Greenland et al. (1994) の事例をもとにして説明を行ってきた確率的感度解析は、もともとは、金融や経済における投機リスク評価、原子炉の安全性評価などリスク評価の分野で用いられてきた方法である (Committee on Risk Assessment of Hazardous Air Pollutants, Commission on Life Sciences, National Research Council, 1994; Vose, 2008)。疫学研究では、近年になって、電磁場への曝露と小児白血病の関連評価への適用 (Phillips, 2003; Greenland, 2003) やシリカ曝露と肺がんの関連評価への適用 (Steenland and Greenland, 2004) をはじめ、Chu et al. (2006)、Eddy et al. (1992)、Fox et al. (2005)、Greenland (2001, 2005)、Hoffman and Hammonds (1994)、Lash and Fink (2003) などによる研究が行われてきた。

モンテカルロ感度解析の数学的定式化として、まず一連の観測データ (D) を用いて興味あるパラメータ $\theta = \theta(\alpha)$ を推定することを考える。従来解析では、系統的バイアスはないという仮定を置き D のみによって θ が識別できるものとして、統計モデル $L(D|\alpha)$ を用いて θ を推定する。これに対してバイアスモデルでは、バイアスパラメータを η としてモデルに導入し、データの分布はバイアスモデルを組み込んだ $L(D|\alpha, \eta)$ によってモデル化されるとする。

バイアスモデルを用いたバイアス調整をベイズ流に扱うことを考えると、 (α, η) の事前分布 $p(\alpha, \eta)$ を与えて $\theta = \theta(\alpha)$ 周辺事後分布を求めることになる (Gelman et al., 2013; Greenland, 2005)。

$$p(\alpha|D) \propto \int L(D|\alpha, \eta)p(\alpha, \eta)d\eta$$

その際は、事前分布 $p(\alpha, \eta)$ を説明するのに、バイアスパラメータ η 自身が既知の共変量と未知のハイパーパラメータ β の関数としてモデル化されることになることもある (Parmigiani, 2002)。このようにして導かれる周辺事後分布における推測では、複雑な階層モデルなどでは特に、マルコフ連鎖モンテカルロ法の実装や収束判断などの計算上の問題が伴うが、モンテカルロ感度解析法では、比較的簡便なモンテカルロサンプリングを組み合わせるだけで実行することができ、それらの問題点を回避することができる。また、後に述べるとおり、複数のバイアスを複合的にモデル化する場合には、階層ベイズモデルでの事後サンプリングのアルゴリズムは、それに比例して複雑なものとなるが、モンテカルロ感度解析は、単純に複数のバイアスに対するモンテカルロシミュレーションを組み合わせるだけで実装できるという利点がある。

ここまで述べてきたモンテカルロ感度解析法は、前述の事前分布 $p(\alpha, \eta)$ をバイアスパラメータの周辺事前分布 $p(\eta)$ のみの特定で代用したものに相当する。一般に、 α は η の情報なしには部分的にも特定されず、 α についての推定は η の値に完全に依存する (逆に η も α の情報なしには部分的にも特定されない)。モンテカルロ感度解析では、周辺事前分布 $p(\eta)$ からの η の各サンプリング値に対して、 η を固定して推定した興味あるパラメータの推定値を最尤推定量であるとみなす。この場合、 α には一様事前分布 (無情報事前分布) を仮定していることになり、 $p(\eta|D) \cong p(\eta)$ と近似できる。したがって、

$$p(\alpha|D, \eta) \propto \frac{L(D|\alpha, \eta)p(\alpha, \eta)}{p(\eta|D)} \propto L(D|\alpha, \eta)$$

が成り立ち、

$$p(\theta|D, \eta) \propto \int_{\theta(\alpha)=\theta} L(D|\alpha, \eta) d\alpha$$

となる。

大標本の状況では、事後分布 $p(\theta|D, \eta)$ は正規近似 [$\sim N(\hat{\theta}_\eta, \hat{s}_\eta^2)$] が可能であり、 $p(\theta|D, \eta)$ を得るための積分が不要となる。モンテカルロ感度解析では以下の3つの手順を踏むことで以下の事後分布 $p(\theta|D)$ からの擬似的なサンプリングとみなすことができる。

$$p(\theta|D) = \int p(\theta|D, \eta)p(\eta|D) d\eta$$

1. $p(\eta)$ から η の値をランダムサンプリングする
2. 各 η から $\hat{\theta}_\eta$ と標準誤差 \hat{s}_η を算出する
3. $\hat{\theta}_\eta$ を $N(\hat{\theta}_\eta, \hat{s}_\eta^2)$ から再サンプリングするか $\hat{\theta}_\eta$ の分布に $N(0, \hat{s}_\eta^2)$ を加える

このように、モンテカルロ感度解析は、ベイズ流の枠組みで解釈することができる。一方で、さまざまな観点からの考察はされているが、頻度論的な方法としての解釈はできない (Greenland, 2001)。また、 α についての事前分布を導入した方法への拡張は容易である。

複合バイアスマデリングにおけるモンテカルロ感度解析は、特定のバイアスに対して、上記のサンプリングとバイアス調整を実施し、その結果を他のバイアスマデルでの無調整パラメータとして、個別に同様のサンプリングを繰り返すことで、複数のバイアスの同時調整が可能となる。

ここで紹介した解析手法の多くは、Orsini et al. (2008) による STATA モジュール EPISSENS (<http://nicolaorsini.altervista.org/stata/tutorial/e/episens.htm>) で実行することができる。また、Lash et al. (2009) は、疫学における確率的感度解析の教科書であり、背景の理論や計算ツールについてのより詳細な解説がまとめられている。

4. おわりに

観察研究でのバイアスの存在およびその影響については、生物統計の専門家であるとなしと関わらず、広く重要であると認識されている。その一方で定量的な評価はシナリオ依存的な感度解析になることが多く、ほとんどの場合は実施されずに考察部分で限界としての議論がなされるに留まることが多い。バイアスマデルの定式化及び事前分布を仮定してモンテカルロサンプリングを行うことで複数のバイアスを同時考慮することは、このような定性的議論を打開し、観察研究の結果に含まれる不確実性を定量的に評価することにつながる。選択バイアスに関しては、公的な統計データなどを外部データとの比較で事前分布に関する情報を得ることが可能であろう。未測定の変質因子に関しては、すべての対象者で測定されていない共変量を未測定の変質因子とみなし、測定されている一部対象者の情報を事前情報とした欠測補完の手法として利用するという利用方法もある。または、架空の変質因子を想定し、結果を覆さない変質因子はどの程度の影響(未測定の変質因子のオッズ比)をもつものかを逆算することで、そのような変質因子が存在する可能性を議論することも可能である。測定誤差については、このようなバイアスマデルを用いることで、デザイン段階で一部対象者に対して精度の高い測定を併せて実施しておくことで、それを妥当性データとして利用することもできる。既存の観察研究データおよび計画中の観察研究に対してぜひ適用を検討いただきたい。

謝 辞

本稿執筆に際し、度重なる打ち合わせの機会、および助言をいただいた名古屋大学大学院医学系研究科の松井茂之教授、大分大学工学部の和泉志津恵准教授に深く感謝申し上げます。

参 考 文 献

- Axelsson, O. and Steenland, K. (1988). Indirect methods of assessing the effect of tobacco use in occupational studies, *American Journal of Industrial Medicine*, **13**, 105–118.
- Barron, B. A. (1977). The effects of misclassification on the estimation of relative risk, *Biometrics*, **33**, 414–418.
- Bowden, J., Jackson, D. and Thompson, S. D. (2010). Modelling multiple sources of dissemination bias in meta-analysis, *Statistics in Medicine*, **29**, 945–955.
- Brenner, H. and Gefeller, O. (1993). Use of positive predictive value to correct for disease misclassification in epidemiologic studies, *American Journal of Epidemiology*, **138**, 1007–1115.
- Bross, I. D. J. (1966). Spurious effects from an extraneous variable, *Journal of Chronic Diseases*, **19**, 637–647.
- Bross, I. D. J. (1967). Pertinency of an extraneous variable, *Journal of Chronic Diseases*, **20**, 487–495.
- Carroll, R. J., Spiegelman, C. H., Lan, K. K. G., Bailey, K. T. and Abbott, R. D. (1984). On errors-in-variables for binary regression models, *Biometrika*, **71**, 19–25.
- Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*, Chapman and Hall/CRC, London.
- Chavance, M., Dellatolas, G. and Lellouch, J. (1992). Correlated nondifferential misclassifications of disease and exposure, *International Journal of Epidemiology*, **21**, 537–546.
- Chu, H. Z., Wang, S. R. and Greenland, S. (2006). Sensitivity analysis of misclassification: A graphical and a Bayesian approach, *Annals of Epidemiology*, **16**, 834–841.
- Committee on Risk Assessment of Hazardous Air Pollutants, Commission on Life Sciences, National Research Council (1994). *Science and Judgment in Risk Assessment*, National Academy Press, Washington, D.C.
- Copeland, K. T., Checkoway, H., Holbrook, R. H. and McMichael, A. J. (1977). Bias due to misclassification in the estimate of relative risk, *American Journal of Epidemiology*, **105**, 488–495.
- Cornfield, J., Haenszel, W., Hammond, E. C., Lillienfeld, A. M., Shimkin, M. B. and Wynder, E. L. (1959). Smoking and lung cancer: Recent evidence and discussion of some questions, *Journal of the National Cancer Institute*, **22**, 173–203.
- Dosemeci, M., Wacholder, S. and Lubin, J. H. (1990). Does nondifferential misclassification of exposure always bias a true effect toward the null values?, *American Journal of Epidemiology*, **132**, 746–749.
- Eddy, D. M., Hasselblad, V. and Shachter, R. D. (1992). *Meta-analysis by the Confidence Profile Method: The Statistical Synthesis of Evidence*, Academic Press, Boston.
- Flanders, W. D. and Khoury, M. J. (1990). Indirect assessment of confounding: Graphic description and limits on effect of adjusting for covariates, *Epidemiology*, **1**, 239–246.
- Fox, M. P., Lash, T. L. and Greenland, S. (2005). A method to automate probabilistic sensitivity analyses of misclassified binary variables, *International Journal of Epidemiology*, **34**, 1370–1376.
- Gail, M. H., Wacholder, S. and Lubin, J. H. (1988). Indirect corrections for confounding under multiplicative and additive risk models, *American Journal of Industrial Medicine*, **13**, 119–130.
- Gelman, A., Carlin, J. B., Stern, H. S., Funson, D. B., Vehtari, A. and Rubin, D. B. (2013). *Bayesian data analysis*, 3rd ed., Chapman and Hall/CRC, London.

- Green, M. S. (1983). Use of predictive value to adjust relative risk estimates biased by misclassification of outcome status, *American Journal of Epidemiology*, **117**, 98–105.
- Greenland, S. (1982). The effect of misclassification in matched-pair case-control studies, *American Journal of Epidemiology*, **116**, 402–406.
- Greenland, S. (1987). Quantitative methods in the review of epidemiologic literature, *Epidemiologic Review*, **9**, 1–30.
- Greenland, S. (1990). Randomization, statistics, and causal inference, *Epidemiology*, **1**, 421–429.
- Greenland, S. (1996). Basic methods for sensitivity analysis of biases, *International Journal of Epidemiology*, **25**, 1107–1116.
- Greenland, S. (2001). Sensitivity analysis, Monte Carlo risk analysis, and Bayesian uncertainty assessment, *Risk Analysis*, **21**, 579–583.
- Greenland, S. (2003). The impact of prior distributions for uncontrolled confounding and response bias: A case study of the relation of wire codes and magnetic fields to childhood leukemia, *Journal of the American Statistical Association*, **98**, 47–54.
- Greenland, S. (2005). Multiple-bias modeling for analysis of observational data (with discussion), *Journal of the Royal Statistical Society, Series A*, **168**, 267–306.
- Greenland, S. and Kleinbaum, D. G. (1983). Correcting for misclassification in two-way tables and matched-pair studies, *International Journal of Epidemiology*, **12**, 93–97.
- Greenland, S. and Lash, T. L. (2008). Bias analysis, *Modern Epidemiology* (eds. K. J. Rothman, S. Greenland and T. L. Lash), 345–380, Lippincott-Raven, Philadelphia.
- Greenland, S., Salvan, A., Wegman, D. H., Hallock, M. F. and Smith, T. J. (1994). A case-control study of cancer mortality at a transformer-assembly facility, *International Archives of Occupational and Environmental Health*, **66**, 49–54.
- Hoffman, F. O. and Hammonds, J. S. (1994). Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability, *Risk Analysis*, **14**, 707–712.
- Kleinbaum, D. G., Kupper, L. L. and Morgenstern, H. (1984). *Epidemiologic Research: Principles and Quantitative Methods*, Wiley, Van Nostrand.
- Kristensen, P. (1992). Bias from nondifferential but dependent misclassification of exposure and outcome, *Epidemiology*, **3**, 210–215.
- Lash, T. L. and Fink, A. K. (2003). Semiautomated sensitivity analysis to assess systematic errors in observational data, *Epidemiology*, **14**, 459–466.
- Lash, T. L. and Silliman, R. A. (2000). A sensitivity analysis to separate bias due to confounding from bias due to predicting misclassification by a variable that does both, *Epidemiology*, **11**, 544–549.
- Lash, T. L., Fox, M. P. and Fink, A. K. (2009). *Applying Quantitative Bias Analysis to Epidemiologic Data*, Springer, New York.
- Marshall, R. J. (1990). Validation study methods for estimating exposure proportions and odds ratios with misclassified data, *Journal of Clinical Epidemiology*, **43**, 941–947.
- Orsini, N., Bellocco, R., Bottai, M., Wolk, A. and Greenland, S. (2008). A tool for deterministic and probabilistic sensitivity analysis of epidemiologic studies, *The Stata Journal*, **8**, 29–48.
- Parmigiani, G. (2002). *Modeling in Medical Decision Making — A Bayesian Approach*, Wiley, New York.
- Phillips, C. V. (2003). Quantifying and reporting uncertainty from systematic errors, *Epidemiology*, **14**, 459–466.
- Rosenbaum, P. R. (1995). *Observational Studies*, Springer-Verlag, New York.
- Rothman, K. J., Greenland, S. and Lash, T. L. (2008). *Modern Epidemiology*, 3rd ed., Lippincott

- Williams & Wilkins, Philadelphia.
- Savitz, D. A. and Baron, A. E. (1989). Estimating and correcting for confounder misclassification, *American Journal of Epidemiology*, **129**, 1062–1071.
- Scharfstein, D. O., Rotnitzky, A. and Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models, *Journal of the American Statistical Association*, **94**, 1096–1120.
- Schlesselman, J. J. (1978). Assessing effects of confounding variables, *American Journal of Epidemiology*, **108**, 3–8.
- Simon, R. (1980). Re: ‘Assessing effect of confounding variables’, *American Journal of Epidemiology*, **111**, 127–128.
- Steenland, K. and Greenland, S. (2004). Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer, *American Journal of Epidemiology*, **160**, 384–392.
- Taubes, G. and Mann, C. C. (1995). Epidemiology faces its limits, *Science*, **269**, 164–169.
- Tennenbein, A. (1970). A double sampling scheme for estimating from binomial data with misclassification, *Journal of the American Statistical Association*, **65**, 1350–1361.
- Vose, D. (2008). *Risk Analysis*, John Wiley and Sons, New York.
- Wacholder, S., Armstrong, B. and Hartge, P. (1993). Validation studies using an alloyed gold standard, *American Journal of Epidemiology*, **137**, 1251–1258.
- Yanagawa, T. (1984). Case-control studies: Assessing the effect of a confounding factor, *Biometrika*, **71**, 191–194.

Sensitivity Analysis for Biases in Observational Studies

Ayano Takeuchi¹ and Hisashi Noma²

¹Center for Environmental Health Sciences, National Institute for Environmental Studies

²The Institute of Statistical Mathematics

Most researchers recognize that conventional statistical analysis of observational data require assumptions like no selection bias, no information bias, no unmeasured, missing at random. . . etc. It is almost impossible to assess that these assumptions are met with study data. If these assumptions are unmet, the results from conventional analysis have uncertainty and are biased. In this article, we reviewed bias analysis, focusing on probabilistic sensitivity analysis.