

観察研究における感度分析の勧め 実践編 第2部

・・・観察研究に基づく意思決定に関わる全ての方へ・・・

日本製薬工業協会

データサイエンス部会

2016年タスクフォース2

2019年4月

## 第 2 部

---

4. 不確実性を表せる感度分析手法 .....	8
4.1. Probabilistic sensitivity analysis (PSA) .....	8
4.1.1. 概要と特徴 .....	9
4.1.2. シミュレーションによる実際の方法：Monte Carlo simulation .....	11
4.1.3. パラメータの分布 .....	12
4.1.4. シミュレーション回数 .....	18
4.1.5. パラメータに誤差を加味する .....	19
4.1.6. 複数の独立なパラメータ .....	20
4.1.7. 関連したパラメータ分布 .....	22
4.1.8. 各種バイアスにおける PSA の設定と結果の例 .....	23
4.2. Multiple bias (Probabilistic) modelling .....	36
5. 感度分析の準備から結果の報告まで .....	41
5.1. 感度分析の必要性を評価する .....	41
5.2. 感度分析の目的と手法との関係 .....	42
5.3. 研究デザイン、データの収集方法、得られている情報を整理する .....	43
5.4. 感度分析すべきバイアスを選択する .....	48
5.4.1. バイアスの原因を整理して分類する .....	50
5.4.2. 感度分析のモデルに使用するバイアスを特定する（バイアスの候補が複数ある場合） .....	51
5.4.3. 感度分析のモデルに使用するバイアスを特定する（バイアスの候補が複数ある場合） .....	52
5.5. 感度分析の方法を選択する .....	65
5.5.1. 手法の選択：情報量とシンプルさとのバランス、結果説明の容易さ .....	65
5.5.2. 利用可能な感度分析の方法の情報源・ツールと新しいモデルの作成 .....	66
5.6. 感度分析モデルにパラメータを設定する .....	67
5.6.1. 感度分析パラメータに関する情報源と収集 .....	67
5.6.2. 利用可能な情報の種類と計算方法 .....	80
5.6.3. 感度分析パラメータに値と分布を割り当てる .....	81

5.6.4. パラメータ特性の影響評価（シナリオの反映の確認） .....	89
5.6.5. モデル, パラメータと補正結果の吟味 .....	93
5.7. 感度分析の結果を提示し解釈する .....	97
5.7.1. 感度分析の方法を提示する .....	98
5.7.2. 感度分析の結果を提示する .....	102
5.7.3. 感度分析結果の解釈 .....	108
5.8. 教科書的ではないが現実的で簡便な手順による感度分析 .....	110
6. 感度分析の実施例 .....	112
6.1. 例 1：主に Simple sensitivity analysis を用いたもの .....	112
6.1.1. Chang ら（2016）の論文 .....	112
6.1.2. 感度分析の必要性 .....	113
6.1.3. どのようなバイアスを感度分析の対象とするか？ .....	113
6.1.4. 乳酸アシドーシスの誤分類：乳酸アシドーシスを尤もらしく特定可能するには？ .....	115
6.1.5. 誤分類の感度分析 .....	117
6.2. 例 2：主に Probabilistic sensitivity analysis（PSA：確率的感度分析）を用いたもの .....	121
6.2.1. S Youngquist ら（2008）の論文 .....	121
6.2.2. 本事例における感度分析の必要性 .....	122
6.2.3. Probabilistic sensitivity analysis の実施 .....	122
6.2.4. 結果のまとめと考察 .....	125
6.2.5. 読者・意思決定者による追加的考察 .....	126
終わりに .....	129
REFERENCE .....	131
Appendix1. 感度分析の必要性 .....	136
1) 研究の質による評価：計画～解析における精度とバイアスから .....	136
2) 研究の重要性による評価：結果の社会的影響から .....	137
1. 感度分析が必要不可欠である場合 .....	137
2. 感度分析が推奨される場合 .....	138
3. 感度分析が不可欠ではない場合 .....	139
Appendix2. 分布に基づく乱数の生成 .....	141

1) 台形分布 (Trapezoidal), 三角分布 (Triangular) .....	141
2) ベータ分布 .....	143
3) 対数正規分布.....	144
4) ロジット・ロジスティック (logit-logistic) 分布.....	144
5) 相関のある分布.....	146
6) 分布に関する注意事項.....	149
Appendix3. 感度分析を行う Tool, macro .....	153
1) Excel workbook .....	153
2) SAS macro .....	163
3) STATA.....	164
4) R Package.....	164



## 第2部の初めに

観察研究においてバイアス<sup>注1</sup>への対処は重要であり、研究のデザイン並びに解析において種々の方法が用いられています<sup>1), 2)</sup>。本書は解析方法を対象としますが、層別解析やモデルを用いて共変量の影響を調整した上で関心がある推定量を求める“調整解析”がテーマではありません。むしろこうした解析では調整が難しい、あるいは不可能なバイアスの影響を評価するために行われる定量的バイアス分析 (quantitative bias analysis) ≒ 感度分析 (sensitivity analysis)<sup>注2,3)</sup>をテーマとします。

実践編第1部では感度分析のうち、特に曝露やアウトカムの有無など2値の変数を対象に未測定 (未知) 交絡要因、測定誤差 (誤分類<sup>3)</sup>、選択バイアスの影響を検討することを中心として未調整の集計表に対して行える古典的な感度分析手法の基本となる任意の固定した値を用いる Simple な手法を扱い、計算に用いる数式を導く過程や教科書的なバイアスの構造を説明しました。本書では、感度分析の条件 (パラメータ) に分布 (範囲) を持たせるシミュレーションによる確率的感度解析を扱います。例えば、新薬 X の潜在的リスクについて製造販売後データベース調査として case-control 研究を行い Table 4-1 の結果が得られたとします。使用したデータベースではアウトカムの陽性的中度のみバリデーション調査が行われており、全体で 80%、入院外来の別、受診頻度などにより 60%から 90%のばらつきが報告されていたとします。この状況では疑いながらも保守的にリスクとして特定されたとされるかもしれません。しかし感度分析では、例えば、新薬 X 使用患者の陽性的中度は非使用患者より平均 10% 低く (2 例程度が誤って case と分類される)、報告がない陰性的中度は曝露に関わらず 100%に近いと仮定した設定と補正<sup>注4</sup>した結果は Table 4-2 となります。この場合の Systematic error のオッズ比は 3.3%が 1 以下となります<sup>注5</sup>。このような情報があれば重大なバイアスの可能性のある結果でも合理的な意思決定につなげられるのではないのでしょうか。また同時にデータベースのバリデーションを考える上でも感度分析の考え方、モデルは欠かせないものと考えますがいかがでしょうか。

Table 4-1 架空の製造販売後データベース調査 (case-control 研究) の結果

	曝露*あり	曝露*なし	計	オッズ比=2 95%信頼区間 1.15~3.49 *: 曝露は新薬 X
cases	25	50	75	
controls	60	240	300	

注1: 「バイアス」という単語は本書において広く研究の結果を誤らせ得る要因 (交絡要因, 測定誤差, 脱落などを含む) 及び結果の誤りを表現する単語として用いています。

注2: MEMO: 本書では、「感度分析」を用います。疫学辞典 第5版 Miquel Porta (編), 日本疫学会 (訳)に「sensitivity analysis」はありますが、「bias analysis」はありません。また、Handbook of Epidemiology 第2版 (2014) の p 685-706 に「Sensitivity Analysis and Bias Analysis」として Greenland S. が説明していますが、本書では結果へのバイアスの影響に重点を置いた頑健性の評価を中心に扱っておりまた、偏りがなくとも結果に影響が生じる場合があることから「感度分析」としました。なお、臨床試験においても robustness (ICH E9 日本語訳は「安定性」) を評価する方法として「感度分析」という用語が用いられます。こちらは「試験の結果が間違っていないこと」を評価するのに対して、観察研究は未測定交絡因子など様々な要因によりバイアスが生じている可能性がある前提ですので「どの位間違っているか」その程度の評価を目的とします。

注3 薬剤使用等に関する曝露情報や、アウトカムの発生に関する情報に誤りがあるとバイアスが生じることがある。分類上の誤り、という意味で 誤分類 と呼ぶ。実質的には、曝露やアウトカムの有無、といったカテゴリー・データの測定や扱いが誤っているせいで生じるバイアスなので「測定誤差」の一種とも考えられる。

注4: 調整解析と区別するため、本書では感度分析の結果を“補正後”, “補正された”などと表記します。

注5: シミュレーションによって得られる分布から読み取っている。設定は異なるが方法は 5.6.3.1 を参照。

Table 4-2 アウトカムの誤分類によるバイアスの感度分析の設定と結果

仮定	陽性的中度	新薬 X あり:0.65~0.8 新薬 X なし:0.75~0.9 <sup>注6</sup>	
	陰性的中度	0.999~1 (偽陰性が case の最大 5 割程度)	
	non-case の数	60,000	
補正後オッズ比		Systematic error	1.6 (0.98~2.81)
中央値 (95%頻度区間)		Total error	1.62 (0.75~3.58)

観察研究におけるバイアスは様々ですので、本書ではその一部を紹介したに過ぎません。可能な限り事例並びに手法を紹介するリファレンスを付けましたので実施に当たってはそちらもご参照ください。更に感度分析について学ぼうとされる場合、総説としては Lash T.L. (2014) <sup>4)注7</sup>, Uddin M.J. (2016) <sup>5)</sup>, 基本的手法については Greenland S. (1996) <sup>6)</sup>が、定量的手法は Greenland S. (2005) <sup>7)</sup>, 全般的教科書として Lash T.L. (2009) <sup>8)</sup>, Rothman K.J. (2008) <sup>2)</sup>などもありますので感度分析について体系的に学びたい場合は、まずこれらを参照するのが良いでしょう。

また、本書では入門編で用いた Lash T.L.の教科書 <sup>8)</sup>で用いられている Excel<sup>Appendix3.1) 2</sup>を参考に新たに Excel で作成した Tool (本書中で ESATJ と呼ぶことにします) <sup>Appendix3.1) 1</sup>を使用しています。ESATJ は本書で扱う感度分析の手法のうち、簡単な紹介に止めているレコードレベルのデータを用いる手法以外、つまり、複数のバイアスを統合する Multiple bias analysis を含め、四分表に対して本書で扱う全ての方法に対応しています。ESATJ は本書を公開している日本製薬工業協会の WEB 上に公開しています。

## ➤ 本書（実践編第2部）であつかう範囲

入門編（2017年10月）では、観察研究における感度分析を行う意味を理解いただくことに重点を置き、理論や数式を用いず、手順と基本的で簡便な手法を紹介しました。実践編 第1部ではモデルを用いたバイアスの説明と、感度分析のモデル（数式）及び単純な感度分析の方法について説明しています。本書（実践編 第2部）では、まず第4章で原因（曝露）及び結果（アウトカム）が各々二値（あり/なし）の場合のシミュレーションを用いる手法の原理と計算方法を含めて感度分析のモデルを紹介し、第5章以降で感度分析を行う際の考え方、手順並びに注意事項に加えて、事例並びに実践に伴う課題解決の手がかりを紹介します。とりあえず ESATJ を用いて感度分析を行いたい、あるいは読み始めて「難しい」と感じられましたら、先に入門編を一読して本書の第4章の「★」印の段落（網掛け）を

注6：層によって 60%から 90%のばらつきがあるとしているため、新薬 X 有無間で 10%ずらしています。ただし、65%以下ではエラーが生じるため下限を 65%としています。

注7：MEMO：本書では引用文献は原則、引用文献番号のみを表示しますが、特に参考となると考える文献等については筆頭著者名と発行年を表示しています。

一読して実際に定量的感度分析を行う際の解説である第5章に進んでいただくことも可能です。また、計算式については実践編第1部をご一読ください。

なお、引用しているWEBページの最終アクセス日は全て2019年1月16日です。

➤ 対象とする読者 . . . . 観察研究に基づく意思決定に関わる全ての方へ . . . .

実践編では、製造販売後調査等も含め、実務としては観察研究に従事している、あるいは安全監視として日常的に研究論文を評価していながら、しかし、こうした量的な感度分析にはあまり馴染みがなかった、という読者を対象としています。実践編 第1部ではこうした分析を実務として行うであろう統計解析担当者も対象に含めバイアスと感度分析のモデル、数式を解説しました。そして本書（実践編 第2部）では、実際に感度分析を行えるように、また統計解析担当者の側から計画者に対してより具体的な提案ができるように、感度分析の具体的な手法にも踏み込んで解説を行い、課題解決の手がかりとなる論文の紹介を行いました。

特に診療情報データベースを用いた薬剤疫学研究を行う場合のみでなく、そういった薬剤疫学研究の結果を日常的に評価する方<sup>注8</sup>も是非とも本書をご一読ください。

---

注8：市販後安全監視業務の一環として文献を評価し、研究報告を行う場合など。5.8 教科書的ではないが現実的で簡便な手順による感度分析 参照

#### 4. 不確実性を表せる感度分析手法

本章は技術的な説明を多く含むので、Tool を用いてとりあえず定量的な感度分析を行いたい場合は読み飛ばして第 5 章に進むことが可能である。その場合でも[入門編](#)ならびに本章の「★」のドロップキャップ（1 文字目を 2 行で表示）で始まる網掛けのある段落は第 5 章の前にご一読いただきたい。

第 1 部ではバイアスと感度分析のモデルの説明のため、コホート研究である WTCHR を数値例に用いたが、この章では確率を用いた感度分析の事例として Greenland が教科書<sup>2), 9)</sup>などで数値例として用いているレジン樹脂の曝露と肺癌の発生の関係を報告した case-control 研究の結果 (Table 4-1) を用いる。

Table 4-1 職業によるレジン樹脂曝露 (E) と肺癌発生の case-control 研究

	Crude data		
	E=1	E=0	Total
D+	A <sub>1</sub> =45	A <sub>0</sub> =94	M <sub>A</sub> =139
D-	B <sub>1</sub> =257	B <sub>0</sub> =945	M <sub>B</sub> =1202

E=1: レジン樹脂曝露あり, E=0: レジン樹脂曝露なし, D+: 肺癌あり (Disease あり), D-: 肺癌なし

この研究では、年齢、性別などを調整した結果が報告された。調整のない D+ (case) と D- (control) 各群のレジン樹脂の曝露状況である「Crude data」は得られているが、重要な交絡要因である喫煙状況は得られていない。

##### 4.1. Probabilistic sensitivity analysis (PSA)



第 1 部の手法では個々のパラメータの値があらかじめ定められているかのように考えてきた。しかし、パラメータの真の値がいくつかの値かは実際にはわからない。無論、内部検証研究あるいは外部検証研究<sup>注9</sup>等の結果からある程度見通しをつけられる場合もある。ところがこうした検証研究の結果もまたランダム誤差を含む値であるのみでなく、系統誤差（バイアス）と無縁ではない。このため検証研究から得られた値をそのまま用いてよいのか、何らかの判断や対処は必要である。こうしてパラメータの値をどのように定めるか、つまり問題となるバイアスについてどのような仮定を置くかについても解析者の判断に委ねられ、説明が求められる。これまでの手法では個々のパラメータの値の重み（取り得る確率）を同様に扱ってきた。特に Multidimensional analysis（多次元感度分析）では解析者はパラメータがある範囲の何れの値も同じ確率で取り得ると考え、任意の範囲の重みが等しい（「一様分布」という）分布からいくつかの代表値を取り出した、と考えることができる。この考えを少し広めると任意の少数個の値を取り出すのではなく、任意の範囲の一様分布からより多くの（数千、数万個の）値を取り出して扱うことが考えられる。そしてさらに発展させて、ある範囲の

注9 内部検証研究は 5.6.1.1 内部検証データ (internal validation study)、外部検証研究は 5.6.1.2 外部検証データ (external validation study) を参照。

中でより多く生じ得る値（検証研究の点推定値など）とあまり生じない値（検証研究の推定値の信頼区間の両端など）の分布を想定することが考えられる。このことは、解析者が仮定するパラメータの特性のみでなく、その影響をも“分布”として、視覚的および統計的性質として表し、第三者（報告の読者、意思決定者、反論者など）と共有でき、修正もできるという重要な機能を提供する。

パラメータに分布を与える場合でも、感度や特異度などのパラメータを用いて補正した値を得るモデル（計算式）自体は Simple sensitivity analysis と同じであるが、パラメータに分布に従った様々な値を代入するシミュレーションを用いることで、分布として補正した値を得られる PSA の用途として、

- ・ パラメータに信頼区間など精度を含む明確な根拠がある、validation study から妥当な推定値が得られているなど。
- ・ パラメータに曖昧さがある場合や明確な根拠がない。
- ・ パラメータがある範囲で変化した場合の結果への影響を示す、ある程度大きなバイアスがある場合の影響を定量的に示すなど。
- ・ 補正後の値が任意の値（オッズ比が 1 以下あるいは 3 以上など）になる条件（パラメータ）を探索する。
- ・ 補正後の値が任意の値以下、以上となる確率を得る。

などの場合に用いられる。高度な手法でオプションも多いため、項目立てて説明する。

#### 4.1.1. 概要と特徴

★ PSA（確率的感度分析）では、あらかじめ個々のパラメータが従う分布を定める。この分布は交絡や誤分類、選択バイアス等、懸念されるバイアスの程度に対する検証研究の結果やその他、何らかの解析者の想定を適切に表現するように設定する。また、この分布からランダムに抽出した値を用いて Simple sensitivity analysis と同じ補正式を用いた計算を行い、その結果を分布として得る方法である。こうした手順によりその解析者の想定のもとでの「真の値？」<sup>注10</sup>の分布を求めることができる。このシミュレーションの手順を、あらためて示すと以下ようになる。

1. あるバイアスの補正に必要なパラメータごとに分布を設定し、その分布からランダムに値を抽出する。
2. 単純な感度分析で用いた補正式に、1 で得た値を代入して補正值を求める。また、得られた補正值をデータとして保存する。
3. 1 と 2 の手順を十分な回数だけ繰り返し、補正值の度数分布を求める。また、分布の中央値や 95% 区間（2.5% 点、97.5% 点）等を求める。

注10： 実際にバイアスのない真の値が得られるのではないため本書では「真の値?」としている。



これの大まかなイメージは Figure 4-1 のようになる。用いるパラメータをランダムに選ぶため、Multidimensional analysis のように特定の値が必ず使用されるということはない。その代わりに、パラメータの特定の値（の周辺又は範囲）の分布密度を高くする（サンプリングされる確率を大きくする）ことで、パラメータおよび補正後の値に確率を反映させることができる。

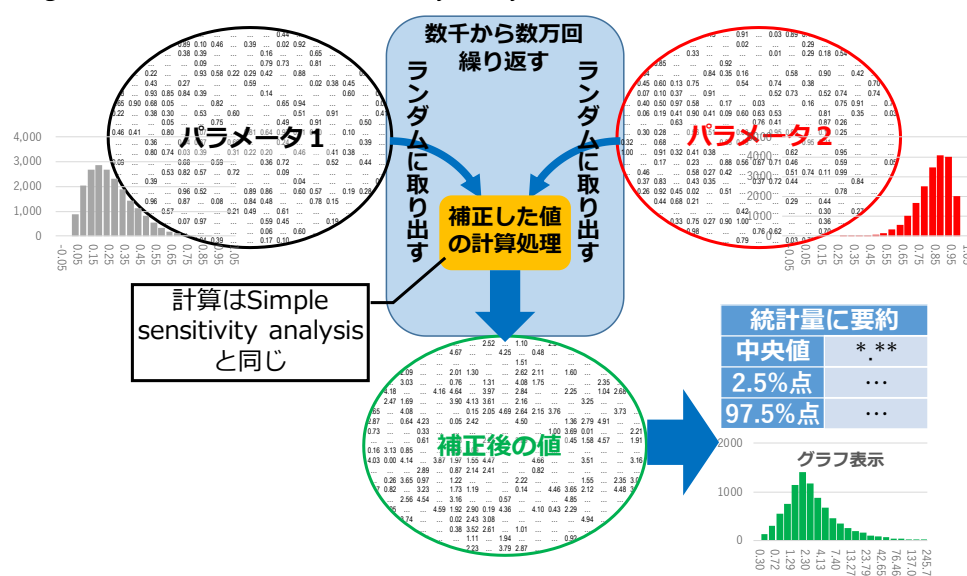
上記の 3 で求めた 95% 区間（95% interval）は頻度区間、シミュレーション区間（simulation interval）と呼ばれることもあるがいわゆる信頼区間ではない。計算の過程から伺えるようにこの区間は以下の 3 つの要素から成り立っている。

- ・ 観察されたデータ
- ・ 観察されたデータと系統誤差（バイアス）を結びつける補正式
- ・ 補正式のパラメータに代入する値（の分布）

これらのうち 2 番目の要素は解析者がそのバイアスに対して抱いている構造的な想定を補正式として表現したものである。1 番目と 2 番目は Simple sensitivity analysis と同じであるが、3 番目の要素は、解析者がそのバイアスに対して抱いている量的な想定を分布として表現したものである。つまり、この補正後の値の区間は解析者がバイアスに対して抱いている構造的かつ量的な想定を効果指標の「真の値？」が取りうる範囲として表現したもので、と解釈することができる。

PSA は手法としては 1 つの名前であるが、様々な分布の定義を用いることができ、さらにパラメータ間に相関関係を設定するなど、多くのバリエーションがあるので内容を分けて説明する。

Figure 4-1 Probabilistic sensitivity analysis のイメージ



#### 4.1.2. シミュレーションによる実際の方法：Monte Carlo simulation

PSA を行う方法として提案され、汎用されている Monte Carlo simulation（モンテカルロシミュレーション）を用いる方法<sup>10)</sup>は複雑な因果の階層のモデル化は不要で複数の要因を同時に感度分析する（補正する）ことも容易となる<sup>11),14),15),16),17)</sup>。Monte Carlo simulation を用いた感度解析は Monte Carlo sensitivity analysis（MCSA）と呼ばれる。<sup>18),19), 4), 7)注11</sup>

★ MCSA を簡単に説明すると、与えた確率密度関数（分布）に従ってパラメータの値を発生（サンプリング）させて Simple sensitivity analysis の方法で計算した結果（オッズ比等）をプールし、プールした結果の分布情報（中央値と頻度区間、ヒストグラムなど）を求める手法である。パラメータ数や範囲によるが、通常 10,000 回以上計算を繰り返す<sup>注12</sup>。

MCSA の例として Table 4-1 の数値例の未測定交絡要因 C（喫煙有無）の感度分析について見てみる。感度分析パラメータとして喫煙と肺癌のオッズ比（ $OR_{DC}$ ）は 2～10、レジン曝露ありの層の喫煙率（ $P_{C1}$ ）、曝露なしの喫煙率（ $P_{C0}$ ）は何れも 0.1～0.6 と想定したとする。しかし何れも範囲の両端は生じ難く、 $P_{C1}$  と  $P_{C0}$  の間に特別な関係（全く同じ、相関があるなど）はないと仮定してシナリオを  $OR_{DC}$  は平均  $\ln(5)$ 、標準偏差 0.4 の対数正規分布、 $P_{C1}$ 、 $P_{C0}$  は平均 0.35、標準偏差 0.13 の正規分布<sup>注13</sup>とし、 $OR_{DC}$ 、 $P_{C1}$ 、 $P_{C0}$  は互いに独立（ $P_{C1} > P_{C0}$  の制約なし）とする。本書のために作成した Excel で感度分析を行う Tool「ESATJ」では Figure 4-2 のように設定する<sup>注14</sup>。このようなパラメータの分布からランダムに取り出した値の組み合わせ毎に、第 1 部第 3 章で説明した各種バイアス補正の計算式を用いて補正した値（この例ではオッズ比とリスク比）及びこれにランダム誤差を加えた値を計算した結果を Table 4-2 のようにテーブルとして出力し（つまり、この各行が Simple Sensitivity Analysis である）、補正後の値について中央値、95%頻度区間などの任意の統計量やグラフを得る。Figure 4-3 はこのシナリオにより 2 万回のサンプリングを行った ESATJ の結果である（ $OR_{DC}$ 、 $P_{C1}$ 、 $P_{C0}$  が負の値及び  $P_{C1}$ 、 $P_{C0}$  が 1 を超える場合は除外されている：truncated normal probability distribution）。出力については後程説明するが、 $OR_{DC}$  の中央値は 1.76、95%頻度区間は 0.93～3.32 であること、1 以下となる場合があるが、頻度は多くないことなどの情報が得られる。Table 4-2 から例えば補正後のオッズ比が 2 以上になる割合や、その場合の各パラメータの範囲なども容易に知ることができる。

注11 MCSA は PSA の代表的な手法であるが、ベイズアンなど他の方法も含まれる。本書では特に MCSA のシミュレーションについて説明する場合は MCSA とし、MCSA を含む確率的感度分析全般については PSA と表記する。

注12： 本書で引用した論文の多くは 1 万～2 万回、多いもので 10 万回であった。結果への影響は 4.1.4 を参照

注13： 正規分布は分布の上限と下限を指定できないため、「喫煙率」のように 0 から 1 という範囲が定まった分布ではないが、最も一般的でイメージしやすく、多くのソフトで乱数を得やすいため例示として用いている。

注14： 水色のセルが入力欄、ESATJ の使い方は Appendix3.1) 1 を参照

Table 4-2 PSA のシミュレーション結果の例

Bias Parameters			Syst Error adjusted		+ Rand Error	
p(C+ E+)	p(C+ E-)	OR(C D)	RR	OR	RR	OR
0.33701	0.44707	4.30175	1.83822	2.06307	1.77325	1.52597
0.34770	0.19243	2.84642	1.37469	1.45295	1.71544	1.30764
0.50762	0.44886	7.69773	1.47462	1.60284	1.56964	1.56916
0.27977	0.19973	5.98347	1.35919	1.46701	1.22152	1.57602
0.38241	0.12104	3.86154	1.10815	1.13164	1.02421	1.14114
0.31992	0.24888	9.88996	1.34730	1.47109	1.56417	1.20437
0.32671	0.31363	2.64492	1.60625	1.73566	1.42689	2.40144
.	.	.	.	.	.	.
.	.	.	.	.	.	.
0.29958	0.45180	11.06982	2.02129	2.43201	1.61758	2.37122

シミュレーション  
の回数分の行数

Figure 4-2 ESATJ のパラメータ設定

変数の名称

負のセルなどエラーの表示

パラメータの分布

シミュレーション回数

単位が「人」「人年」の指定

観測された四分表の度数変数

確率が0~1以外、負のセルとなった場合の扱いのオプション

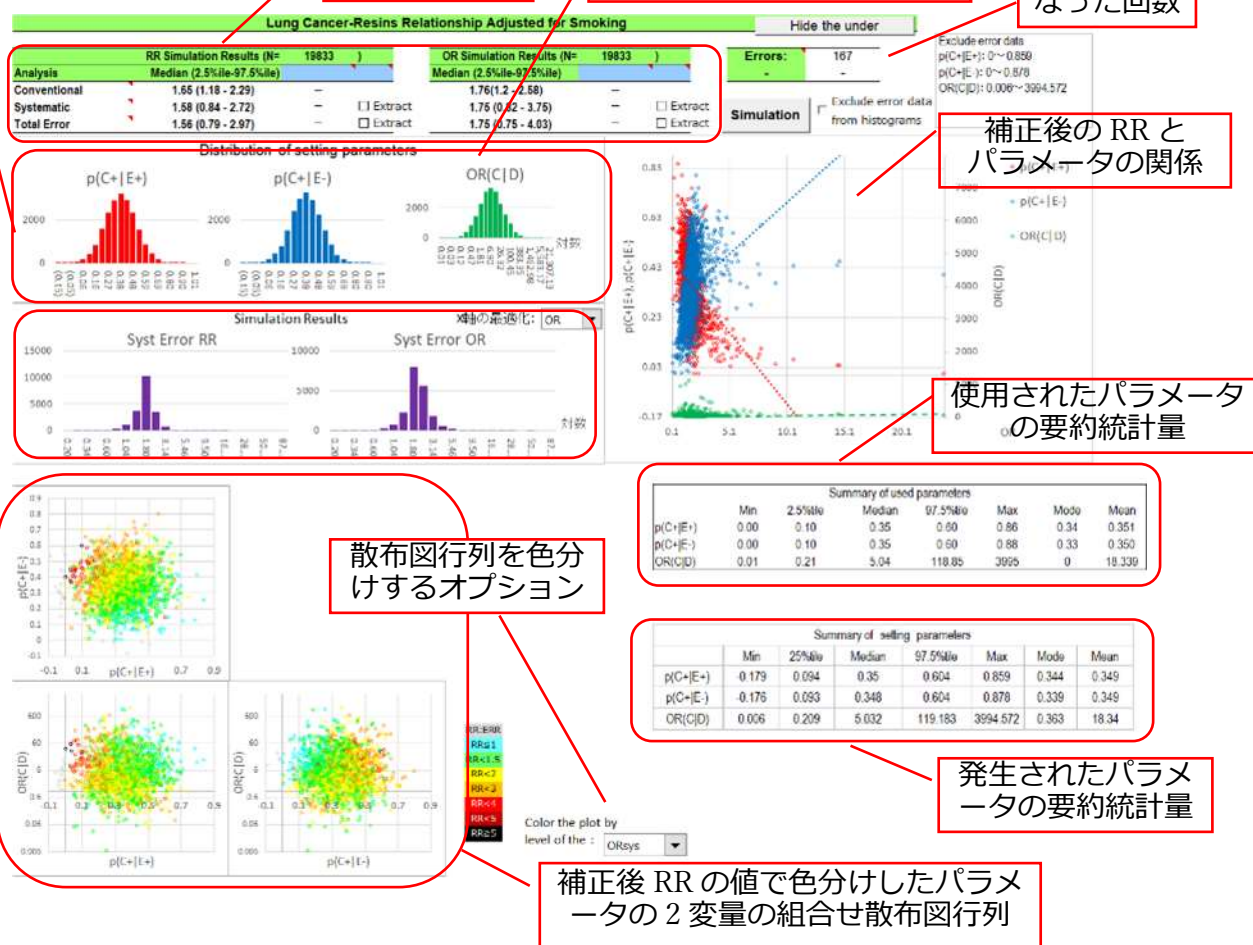
確率を表すパラメータが0~1以外となった場合の扱いのオプション

Figure 4-2 shows the ESATJ parameter settings interface. It includes a 'Variable Names' section with 'Outcome' (Lung Cancer), 'Exposure' (Resins), and 'Confounder' (Smoking). The 'Bias Parameters' section shows 'p (Smoking+|Resins+)' (0.35), 'p (Smoking+|Resins-)' (0.35), and 'OR (Smoking-Lung Cancer)' (1.61). The 'Distribution' section shows 'SD' (0.13), 'α' (0.13), and 'β' (0.4). The 'Simulation' section shows '#Simulation' (20000) and 're-sampling' (FALSE). The 'Data' section shows 'Crude Resins-Lung Cancer Data' and 'Corrected Data' tables. The 'Crude Data' table shows counts for 'Lung Cancer +', 'Lung Cancer -', and 'Total' across 'Resins +' and 'Resins -'. The 'Corrected Data' table shows counts for 'Smoking +', 'Smoking -', and 'Total' across 'Resins +' and 'Resins -'. The 'Crude Data' table has a note 'units are in persons' and the 'Corrected Data' table has a note 'units are in person-years'.

#### 4.1.3. パラメータの分布

★ 感度分析のパラメータに固定値を用いる Simple sensitivity analysis では想定する系統誤差 (systematic error) に不確かさを表現することはできないが、PSA ではパラメータに任意の分布 (平均値、左右の分布の広がり方、偏りなど) を用いることで、解析者が想定する不確実性を反映させることができる。パラメータは、多くの場合、分布を表す累積密度関数の逆関数から得られる。通常、パラメータの分布は最も取りやすいと思われる値を中心に左右に裾が広がる形が想定される。分布において「最も取りやすい値」とは最頻値 (mode) である。左右非対称な分布では最頻値と平均値および中央値は一致せず、一般的な分布関数は最頻値を引数としないため、想定する分布を得ることは容易でない場合がある。比較的単純な方法として、“下限-上限”のみの一様分布、β分布、“下限-最頻値 (点推定値) -上限”の三角形の分布あるいは“下限-最頻区間下限-最頻区間上限-上限”の台形の分布をシナリオとして与える次のような方法が提案されている。(18),(20),(21),(22))



Figure 4-3 MCSA による結果 (ESATJ<sup>1</sup> の出力) OR<sub>DX</sub> の分布補正後の RR, OR  
のヒストグラム補正後の RR,  
OR の要約パラメータのヒストグラム  
OR(C|D)の X 軸は対数エラーと  
なった回数✓ 教科書<sup>8)</sup>にあるシンプルな分布

分布の最小を  $a$ , 最大を  $b$  とし, その間に最も可能性が高い値 (最頻値: mode) を Mode とすると,

- 一様分布 (Uniform Distribution):  $a$  と  $b$  との間に一様に等しい確率密度の分布。
- 三角分布 (Triangular Distribution): 確率が 0 の最小  $a$ , 最大  $b$  と確率が最大の最頻値 Mode の 3 点を直線で結んだ確率密度の分布。
- 台形分布 (Trapezoidal Distribution): 確率が 0 の最小  $a$ , 最大  $b$  と最大の確率を取る最頻値に Mode<sub>1</sub>~Mode<sub>2</sub> の幅 (一様分布) がある確率密度分布。
- ベータ分布は 0~1 の値をとる分布でパラメータ  $\alpha$  と  $\beta$  (何れも正の実数) によって様々な分布となるが,  $\alpha=\beta$  の場合は左右対称,  $\alpha<\beta$  左寄り,  $\alpha>\beta$  右寄りとなる。また最頻値は  $\text{Mode}=(\alpha-1)/(\alpha+\beta-2)$  (但し  $\beta>1$ ) によって得られる。

以上 i)~iii) 3 種類の例を Figure 4-4 に, ベータ分布の例を Figure 4-5 (各グラフの凡例の () 内は  $\alpha$ ,  $\beta$ ) に示す。

Figure 4-4 幾何的確率分布の例

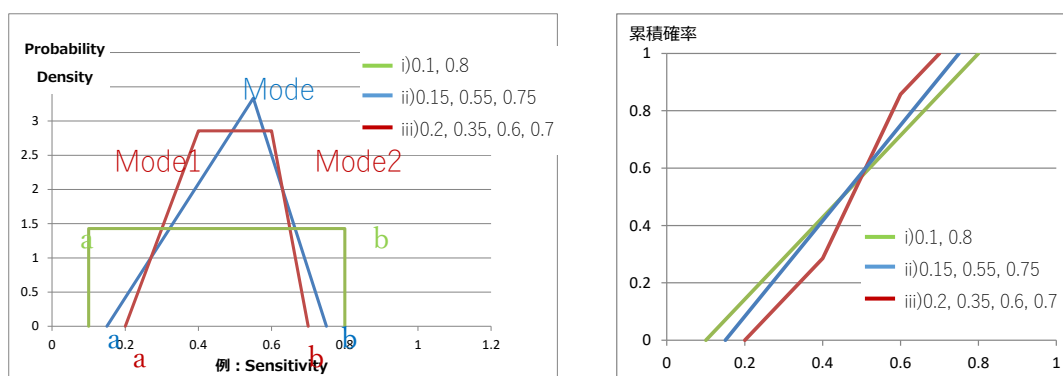
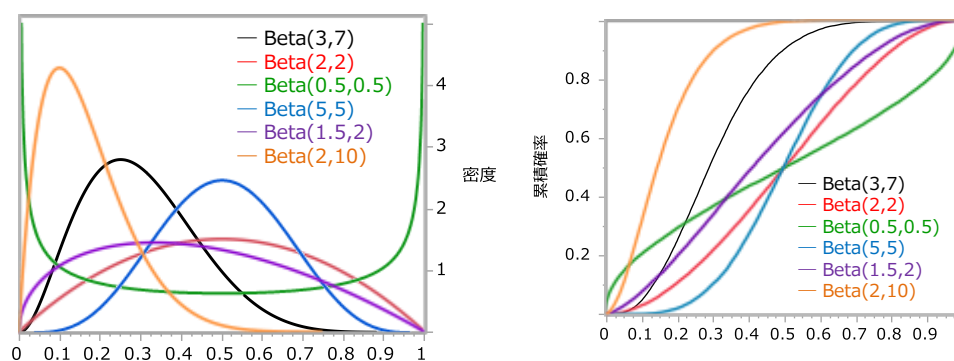


Figure 4-5 Beta 分布の例



一様分布は定義のとおり、解析者が最小値から最大値の間で同じくらい確からしいが、最小値以下および最大値以上は突然生じ得ない（確率が0になる）と仮定していることになり、感度分析のパラメータに“確からしさ”を考慮する確率的感度分析には非現実的な分布である<sup>注15</sup>。対して三角分布並びに台形分布は“最も確か”と考えられる最頻値の左右に直線的に減衰する分布であり、パラメータの情報源から得られる条件による設定が容易であり一様分布より現実的であるが、滑らかではない。そして台形分布は三角分布の“最も確か”と考える値に幅を持たせた分布であり、一様分布の外側に確からしさが減少する区間を表していることになる。

ベータ分布はパラメータ ( $\alpha$ ,  $\beta$ ) により様々な滑らかな分布を表すことができるため便利である。 $\alpha$ と $\beta$ の性質は（分布の最小値=0, 最大値=1とする） $\alpha=\beta>1$ の場合、最頻値=最大値=0.5の左右対称の分布、 $\alpha>\beta>1$ は右に、 $1<\alpha<\beta$ は左に偏った分布、 $\alpha, \beta$ は大きいほど尖った分布である（ベータ分布の詳細は Appendix2.2）参照）。また、感度や特異度などの二項分布に従う出現確率はベータ分布により表すことができる点からも使いやすいであろう。

注15： 5.6.3.1target-adjustment sensitivity analysis における限界値の探索には適している。

本書のために作成した ESATJ は様々な分布 (Appendix3.1) -1-D)参照) をシミュレーションで確かめる機能があるが、その分布を近似するベータ分布のパラメータをシミュレーションにより得ることができる。例えば、感度や特異度などに検証データから得られた平均値と標準偏差を用いて正規分布を用いると 0 以下や 1 以上を生じて不都合な場合があるが、 $\beta$ 分布であれば、範囲を指定することが可能である。(Appendix2-Appendix3.1) -1-E)参照)

一様分布は表計算ソフトを含め、多くのソフトが少なくとも 0~1 の一様乱数を生成する関数が利用可能であるので、任意の最小値 (min)、最大値 (max) の一様乱数  $u$  は、一様乱数  $0 \leq p \leq 1$  から

$$u = \min + (\max - \min) \cdot p$$

により容易に得られ、他の分布も一様分布を基に計算される。主な分布の詳細な説明は Appendix2 を参照。

#### ✓ パラメータの分布形状の重要性



パラメータの分布の最頻値、中央値が近い場合、異なる分布形状を用いても得られる補正後の値の中央値は似通ったものとなるため、主に補正後の中央値に興味がある場合は分布の形状に過度に慎重になる必要はない。しかし、中央値あるいは平均値がほぼ同じでも、分布形状は様々なものがある。例えば Figure 4-6 は中央値がほぼ 0.3 となる正規分布、対数正規分布、ベータ分布であるが、その平均値は 0.30, 0.31, 0.34 であり、最頻値は 0.3, 0.28, 0.08 である。また、5.6.3.1 で説明するが、補正後の値がある基準値以上/以下となる割合を用いる場合などは分布の形状の影響が大きくなる。

通常、感度分析の目的はバイアスを補正した真の値を得ることではなく、バイアスの影響を大まかに調べることにあつて、分布形状の指定に大きな労力を割く意義は薄いとも言えるかもしれない。確認のため、Table 4-1 (レジン樹脂の曝露) の数値例について

- ・ 未測定交絡要因の感度分析をレジン曝露ありの喫煙者の割合  $p(\text{Smoking} + |\text{Resins} +)$  を最小 10% から最大 50% の範囲、
- ・ それ以外は曝露なしの喫煙者の割合を 30%、喫煙による肺癌のオッズ比は 5 の固定
- ・  $p(\text{Smoking} + |\text{Resins} +)$  に一様分布、三角分布、台形分布およびベータ分布を一様分布以外は中央値がほぼ 0.35 となる

このようにパラメータとそれを用いた場合の補正後のオッズ比 (OR<sub>adj</sub>) の中央値と 95% 頻度区間を Table 4-3 に、ヒストグラムを Figure 4-7 に示す<sup>注16</sup>。Table 4-3 の数値はパラメータの分布の中央値が他と異なる一様分布以外の 3 つの OR<sub>Raja</sub> の中央値はほぼ同じ値となっている。しかし、頻度区間の広さは分布の裾の形状の特性から  $\beta$  分布が狭くなっている。

注16： 意図的に各々特徴のある形状の分布としている

さらに Figure 4-7 を見ると、パラメータの分布の特徴は全く異なる印象を与え、補正後のオッズ比の分布も最頻値（ピーク）が明らかに異なるものとなっている。前に「分布形状の指定に大きな労力を割く意義は薄いとも言えるかもしれない」としたが、興味があるのが補正後の値の中央値ではなく、頻度区間や、補正後の値がある値以上、または以下になる確率など、分布形状の影響が大きい場合は、ある程度は慎重になる必要がある。とは言え、分布の詳細な形状に明確な情報が得られることはほとんどないので、解析者の想定する分布の範囲、歪度（最頻度の位置）、尖度（分布の尖り具合、裾の重さ）が概ね表現できる分布を選ばざるを得ない。そのためには得られたパラメータの分布を視覚的に確認し、問題がありそうなら複数の分布を設定し、結果を比較することが薦められる。（ESATJ では様々な分布を近似するベータ分布のパラメータを容易に得ることができる。Appendix2-Appendix3.1） - 1-E)参照)

Figure 4-6 中央値がほぼ 0.3 となる分布

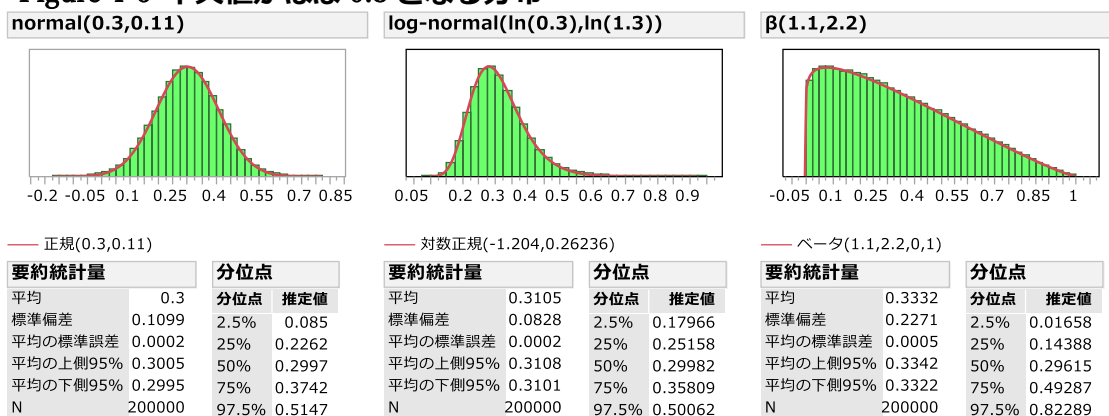
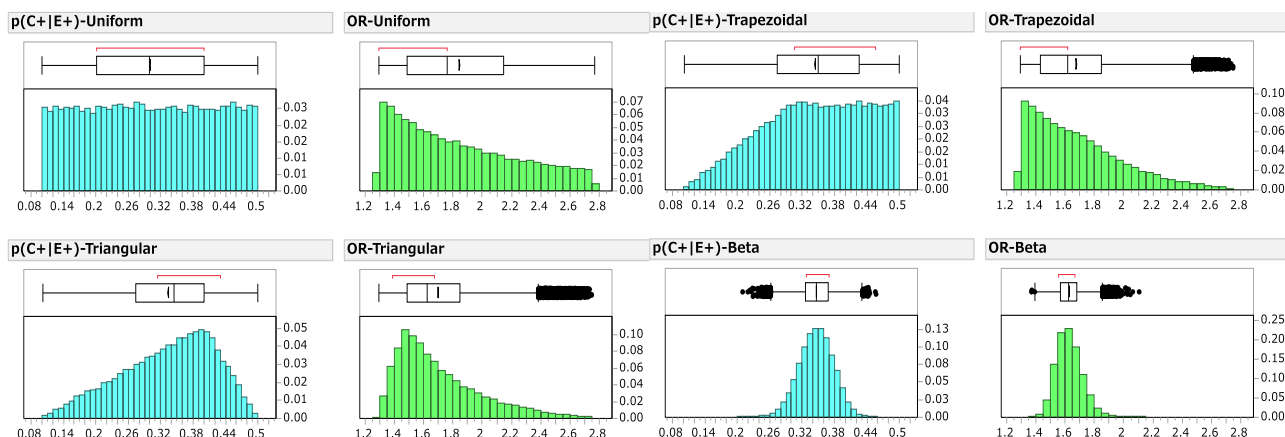


Table 4-3 各種分布のパラメータと補正後オッズ比の頻度区間

分布型	Min	Mode1	Mode2	Max	$\alpha$	$\beta$	Mean	Median	Mode	ORaja Median(0.025-0.975)
Uniform	0.1	—	—	0.5	—	—	0.298	0.298	—	1.762(1.308-2.689)
Triangular	0.1	0.4	—	0.5	—	—	0.333	0.344	0.394	1.626(1.349-2.389)
Trapezoidal	0.1	0.3	0.5	0.5	—	—	0.345	0.351	0.358	1.616(1.303-2.393)
Beta	0.1	—	—	0.5	24	15	0.347	0.347	0.346	1.622(1.482-1.813)

Figure 4-7 パラメータの各種分布と補正後リスク比





## ✓ 一様分布を用いる場合の注意

一様分布は、範囲内のすべての値を取る確率が等しく、分布の端までくまなく密度をもつので、上で見たように効果指標の補正值の分布も広い範囲で大きな密度をもち、95%区間が広くなりがちである。このため、あるパラメータの分布として、わたしたちが通常いづく想定とはかけ離れているかもしれない。例えば、選択バイアスのパラメータは四分表の各セルの抽出率（4変数）が基本であるが選択オッズ比<sup>注17</sup>を用いることで、パラメータの数を1つにすることができる。しかし、各セルの抽出率が一様分布の場合の選択オッズ比は対数正規分布の形となるため（4.1.8.1 参照）、選択オッズ比に一様分布を用いることは、各セルの抽出率の分布に両端が多く、中央が少ない不自然な分布を仮定することになる。このように複合的なパラメータには注意が必要である。また、範囲外の値を取る確率が突然0になる、といった状況が考えにくい場合もある。

## ✓ 矛盾する（エラーとなる）パラメータ

感度分析パラメータの下限と上限は例えば感度、特異度は0～1の間の値<sup>注18</sup>しか取り得ないように理論的な範囲を有する場合以外は何処まで広くするかに議論を伴う。闇雲に広く設定することはエラー（例えば誤分類の感度・特異度では四分表のセルの値が負となる）が発生する場合があり、また過剰な偏りの補正結果を提示することになる<sup>2), 29), 21)</sup>。例えば、Table 4-1の数値例に曝露の誤分類パラメータとして感度=0.8、特異度=0.9（アウトカムの有無の層に共通）とした場合、補正後の $A_0 = -14.1$ と負の値になる（Table 4-4 参照）。誤分類であれば、分類できず除かれた範囲に限定することもある<sup>注19</sup>。また、複数のパラメータがある場合に矛盾を生じることが多くなり得る（例えば感度×特異度）が矛盾に気付くことは困難である。パラメータそのものの及び分布の両端は補正された結果にほとんど影響しないものもある。このようなパラメータは使用しなくても（切り捨てる又は固定値とする）殆ど結果に影響しないとされている。<sup>2), 29), 21), 7)</sup> 特に MCSA のようなシミュレーションのサマリーを結果として用いる場合は事前にパラメータと補正後の値の散布図や Multidimensional analysis 等によって感度分析パラメータの特性を確認することを推奨する。

基本的な MCSA は用いるパラメータのサンプリングをパラメータ毎に独立して行うため、例えば、パラメータ1とパラメータ2があり、パラメータ1>パラメータ2であることを仮定して、パラメータ1を0.5～1の一様分布、パラメータ2に0.3～0.8の一様分布を用いても、各々の分布からランダムに抜き出したパラメータの組み合わせが用いられるため Figure 4-8 のように、仮定に反して20%の組み合わせでパラメータ1<パラメータ2となる。このような仮定への対処は4.1.7. 関連したパラメータ分布 で説明する。

なお、分布の作成については Appendix1、分布を選ぶことについては 5.6.3.3, 5.6.4, 5.6.5 でさらに説明を加えた。

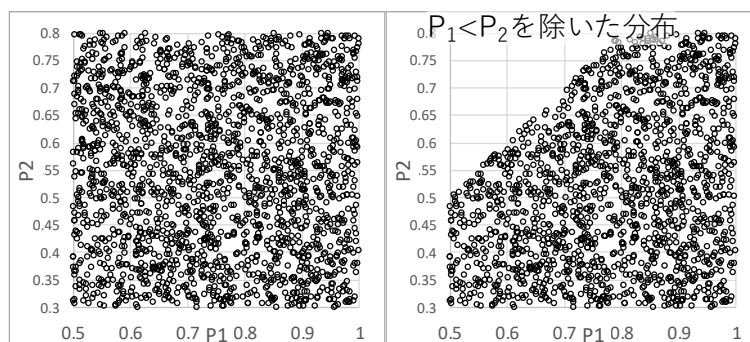
注17 選択オッズ比は第1部 3.6.7 および 4.1.8.1 参照

注18：実際は「無情報」の場合で0.5であるので、0.5～1である。

注19：5.6.3.2 参照。

Table 4-4 数値例に曝露の誤分類として感度=0.8, 特異度=0.9 とした場合

	Crude data			Adjusted data	
	E=1	E=0	Total	E=1	E=0
D+	A <sub>1</sub> =45	A <sub>0</sub> =94	M <sub>A</sub> =139	21.1	-14.1
D-	B <sub>1</sub> =257	B <sub>0</sub> =945	M <sub>B</sub> =1202	280.9	1053.1

Figure 4-8 MCSA におけるパラメータ 1 (P<sub>1</sub>) とパラメータ 2 (P<sub>2</sub>) の分

## 4.1.4. シミュレーション回数

★ PSA (MCSA) は用いるパラメータ以外に、シミュレーション（サンプリング）回数にも注意が必要である。モデル、パラメータの数、パラメータの範囲によって異なるが、例えば誤分類では感度と特異度で最大 4 つ（曝露の誤分類がアウトカムの有無別で異なる値を取る場合など）のパラメータが用いられ、各々 10 個の値を組み合わせた場合で  $10^4 = 10,000$  組となる。試行として行う場合は計算時間との兼ね合いで少なめの回数を用いても、最終結果には十分なサンプリング回数（エラーを除いた有効な回数）を設定する必要がある。Table 4-1 について、喫煙を未測定交絡要因として Table 4-5 のパラメータを用いた PSA のサンプリング回数を 500 回、5,000 回、10,000 回および 20,000 回で各々 500 回繰り返した補正後オッズ比の分布を Figure 4-9 に示す。500 回でばらつきは大きく、中央値の最小から最大幅は 0.041、95% 頻度区間の下側 0.025<sup>th</sup> は 0.048、上側 0.975<sup>th</sup> は 0.061 で各頻度点の分布の最大値÷最小値は Table 4-6 のとおりで、5000 回でも中央値の幅は 5% 以内に収まっているが、0.025<sup>th</sup> は 10% を超えている。例えば用いる全てのパラメータの分布区間が狭い場合やパラメータの数が少ない場合、あるいはオッズ比やリスク比の場合は頻度区間の下限が 1 に極めて近い場合以外など、10,000 回でも十分な場合がある。本書を作成するにあたって参照した PSA (MCSA) が行われている文献は 1 万回から 10 万回程度が多いようであった。

Table 4-5 PSA のパラメータ

Bias Parameters	Min	Max	dhistribution
p(喫煙+ レジン樹脂曝露+)	0.2	0.6	uniform
p(喫煙+ レジン樹脂曝露-)	0.1	0.3	uniform
RR(喫煙-肺癌)	2	10	Log- uniform

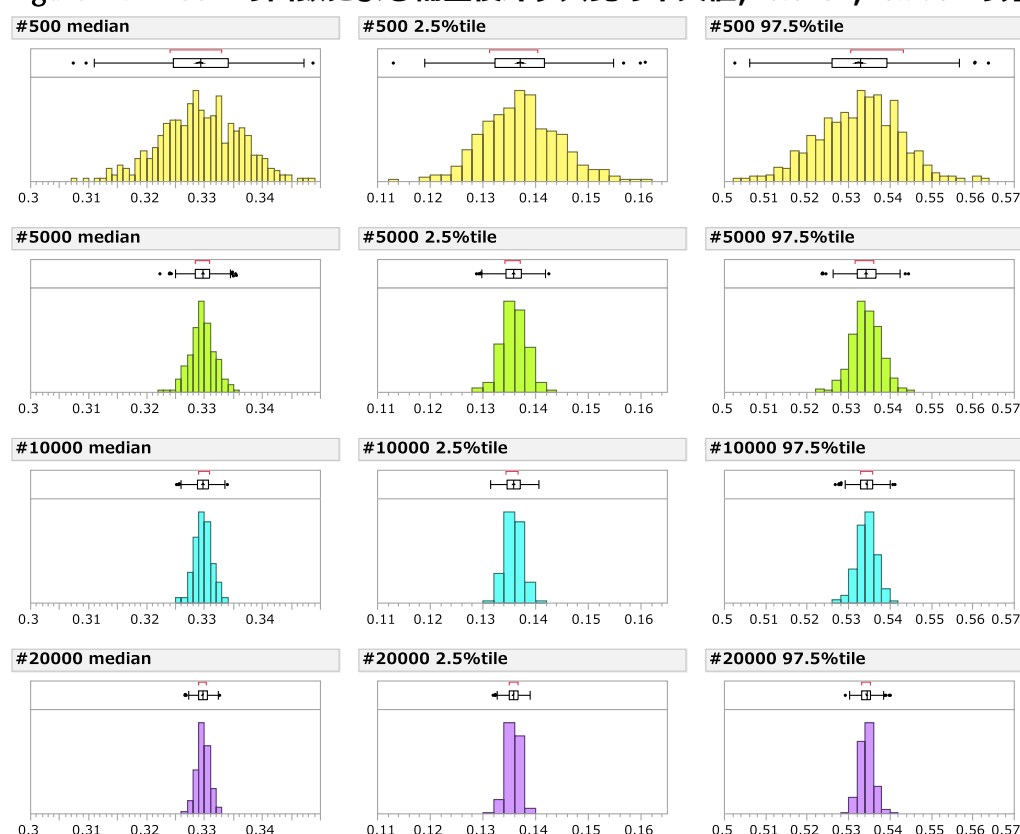
Figure 4-9 MCSA の回数による補正後オッズ比の中央値, 0.025<sup>th</sup>, 0.975<sup>th</sup>の分布

Table 4-6 補正後オッズ比の頻度点の最大値÷最小値

	中央値	0.025th	0.975th
500 回	1.13	1.42	1.12
5,000 回	1.04	1.11	1.04
10,000 回	1.03	1.07	1.03
20,000 回	1.02	1.05	1.02

#### 4.1.5. パラメータに誤差を加味する

あらためて確認すると、4.1.3.パラメータの分布で説明した累積密度関数の逆関数から得られるのは、系統誤差のみを考慮した（あるいはそれを除去した）補正值である。バイアスに関するある想定のもとでその想定が正しいとしたら真の値はこうに違いない、という推定値（真の値？）をあるいは分布として求めたに過ぎない。

しかし、実際に同様な研究を繰り返し行ったとするとこうした値にさらに偶然誤差（random error）を加えた推定値が得られるはずである。つまり、このままではパラメータに“誤差”としてのばらつきがなく信頼区間が狭くなるため、random error を組合せる必要があることが指摘されている<sup>2),21),22)</sup>。

計算方法は系統誤差のみを考慮した補正值からその研究の効果指標のばらつき（標準誤差等）を参照して発生させた乱数を偶然誤差とみなして引く。つまり

$$(\text{偶然誤差も考慮した補正值}) = (\text{系統誤差のみを考慮した補正值}) - (\text{偶然誤差})$$

という計算を行う。もともと「系統誤差のみを考慮した補正值」は、いったん偶然誤差を無視して（推定値）-（系統誤差）に相当する量を近似的に求めたものである。以上の計算により偶然誤差も考慮した「真の値？」を求めたことになる。つまり

$$(\text{真の値?}) = (\text{系統誤差のみを考慮した補正值}) - (\text{偶然誤差}) = (\text{推定値}) - (\text{系統誤差}) - (\text{偶然誤差})$$

偶然誤差は、Observed（観察された結果）の標準偏差を持つ正規分布の乱数が用いられている。例えば、リスク比について Table 4-1 の表記と数値を用いると、オッズ比の標準偏差（SD）は

$$SD = (1/A_1 - 1/B_1 + 1/A_0 - 1/B_0)^{1/2} = (1/45 - 1/257 + 1/94 - 1/1202)^{1/2} = 0.1677$$

である。標準正規分布  $N(0,1)$  の乱数を  $n_i$  とすると、誤差を加えたオッズ比は以下となる。（つまり、平均 0、標準偏差 SD の正規分布の乱数となる）

$$OR_{DEi} = \text{EXP}[\ln(OR_{DEi}) - n_i SD]$$

★ 仮にバイアスが存在しない（系統誤差が 0 の）研究を行ったとしても、真の値は不正確にしかわからない。多くの場合、信頼区間の範囲のどこに真の値があったとしても得られたデータと矛盾しない、などと考える。したがって、系統誤差と偶然誤差をともに考慮したとき、かつバイアスに関する想定が正しかったとしたら、おそらく真の値はこのくらいだろうという推定値（真の値？）をこうした感度分析によって調べることができる。誤差を加えるか否かは、感度分析の目的、パラメータ分布の設定により使い分ける。例えば、パラメータに一樣分布を与え、指定した範囲の補正された値（オッズ比など）の分布とパラメータの関係を確認する場合などは不要と考えられる。

#### 4.1.6. 複数の独立なパラメータ

すでに説明しているが、感度分析に用いる補正式は通常、複数のパラメータが必要であり、PSA では各々が異なる分布を持つ。複数のパラメータのうち、一つに任意の幅のある分布を与え、他を固定値とした場合と、全てに任意の幅のある分布を与えた場合では補正後の値のばらつきは後者で広くなることは想像がつくであろう。

Table 4-1 の数値例について未測定交絡要因（喫煙）の感度分析に用いるパラメータは

- ・ レジン曝露ありの場合の喫煙者の割合  $p(C+|E+)$
- ・ レジン曝露なしの場合の喫煙者の割合  $p(C+|E-)$
- ・ 喫煙による肺癌のオッズ比  $OR(C|D)$

の 3 つである。このパラメータに Table 4-7 の設定を用いて 5 万回のシミュレーションを行った場合の補正オッズ比（偶然誤差無）の中央値と 95% 頻度区間は 1.25 (0.78 - 1.74) である。

では、補正後の値として標準的に用いられる中央値はどのようなになるかを例示する。



## ✓ パラメータの点推定値と Simple sensitivity analysis (SSA) との関係

パラメータに一樣分布を用いる場合はすでに説明したように Multidimensional analysis の単純な拡張で、パラメータの組み合わせは全て同じ確率であるため、PSA の補正後の値の中央値はパラメータの中央値（＝平均値）のみを用いた SSA の結果と変わらない。

★ しかし、PSA のパラメータは多くの場合、最も想定される点推定値（代表値）の両側に頻度が漸減する分布が想定される。そこで、パラメータの各種点推定値（Mean, Median, Mode）を用いて SSA を行い、PSA の結果と比較すると、その結果は Table 4-8 のとおりでパラメータの Median を用いた場合は PSA の結果（Median）に近いが、一致はしていない。用いたパラメータの分布は Figure 4-10 である。視覚的にも解析者が想定するパラメータの点推定値は Mode（最頻値）であり、SSA を行う場合はこの値が用いられるであろう。シミュレーションでは分布の密度に応じてパラメータが使用されるが、同時に中央値から上側と下側の使用される確率はともに 50%であるため、中央値に近い値となる。このように偏り（左右非対称）のある分布をパラメータに用いる場合は点推定値の定義に注意が必要である。

Table 4-7 未測定交絡要因のパラメータ設定

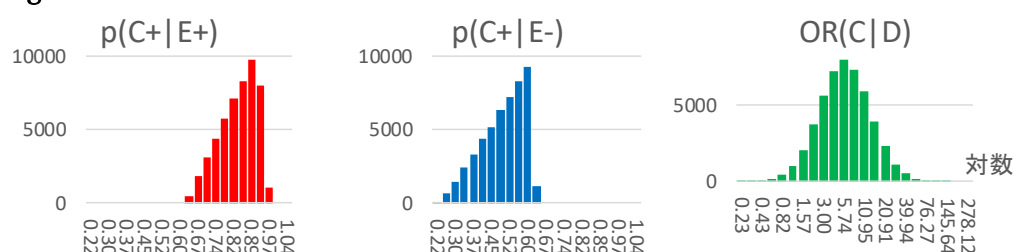
パラメータ	分布関数	Mean <sup>#</sup>	Median <sup>#</sup>	Mode <sup>#</sup>
p(C+ E+)	三角分布：Min0.6, Max0.95, Mode0.9	0.82	0.83	0.88
p(C+ E-)	ベータ分布 Min0.2, Max0.6, $\alpha$ 2.2, $\beta$ 1.02	0.47	0.49	0.58
OR(C D)	対数正規分布：Median5, SD0.8	6.88	4.98	3.0

#：5万回のシミュレーションの結果

Table 4-8 PSA とパラメータの各種点推定値による補正後オッズ比

	PSA			SSA		
	Mean	Median	Mode	Mean	Median	Mode
補正後オッズ比	1.25	1.25	1.30	1.14	1.21	1.38

Figure 4-10 用いたパラメータの分布



## 4.1.7. 関連したパラメータ分布



単純な PSA は例えば誤分類の感度と特異度を独立の分布として扱うが、関連のある分布とすることで例えば

- ・ 通常、感度を高めれば、特異度は低下するため特異度を感度の分布と負の相関関係のある分布とする。
- ・ 共変量を有する割合が常に曝露群に多い（又は少ない）。
- ・ 一方の感度が高（あるいは低）ければ、他方も高い（あるいは低い）場合が多く分布の下限と上限の組み合わせは生じにくい。
- ・ 曝露の有無によって測定（検査）の実施率が異なり、アウトカム（結果）の誤分類に偏りが生じる場合は、感度には差異誤分類を用いるが、特異度は共通と考えられる。このような状況は、特異度パラメータ間の相関係数=1 と同じである。

などの場合も PSA でパラメータの分布の発生（発生）の仕方を工夫することで扱いが可能である。相関のある分布（乱数）の生成方法は Appendix2.5）を参照。

相関を用いない場合と用いた場合の違いを Table 4-1 の例について例示する。Figure 4-12 は相関を用いず、特異度に一様分布を設定している。特異度に感度との相関を用いる設定（相関係数  $r$  を -0.7 としている）を Figure 4-11 に示す。結果の Figure 4-13 と Figure 4-14 を比較すると、まず特異度（Sp 右側の赤と青のヒストグラム）パラメータの分布は相関を用いない場合と用いる場合でほぼ同じとなっているが、下の相関行列は Sp(D+) と Sp(D-) の間に相関関係が表れている。そして結果である補正後のオッズ比は相関を用いない場合は 1.71 (2.5<sup>th</sup> ~ 97.5<sup>th</sup> : 1.07 ~ 2.83), 相関を用いた場合は 1.71 (0.94 ~ 3.23) とわずかであるが頻度区間は広く、緑色のヒストグラムでも裾が厚くなっている。この例では相関係数に負の値、つまり逆相関を設定したため補正後の頻度区間は広がったが、正の値である  $\rho=0.7$  とした場合の補正後オッズ比は 1.71 (1.27 ~ 2.38) と狭い頻度区間が得られる。（正の相関の出力は提示していない）

Figure 4-12 曝露の誤分類のパラメータ。相関を用いない場合の設定

Input Bias Parameters

The parameters are **Se / Sp** Select Method **Setting of Probabilistic** Submit

Variable Names  
Outcome 肺癌  
Exposure レジン樹脂  
Error Check No errors found

Bias Parameters Min  
Se (肺癌+) 0.80  
Se (肺癌-) 0.70  
Sp (肺癌+) 0.90  
Sp (肺癌-) 0.90  
☐ Corrected Variance?

Mod1 Mod2  
Max /  $\alpha$ ,  $p$  /  $\beta$  Distribution  
1 1 Uniform  
0.9 1 Uniform  
1 1 Uniform  
1 1 Uniform  
20000 ← #Simulation

Figure 4-11 相関を用いる場合の設定。

Input Bias Parameters

The parameters are **Se / Sp** Select Method **Setting of Probabilistic** Submit

Variable Names  
Outcome 肺癌  
Exposure レジン樹脂  
Error Check No errors found

Bias Parameters Min  
Se (肺癌+) 0.80  
Se (肺癌-) 0.70  
Sp (肺癌+) 0.90  
Sp (肺癌-) 0.90  
☐ Corrected Variance?

Mod1 Mod2  
Max /  $\alpha$ ,  $p$  /  $\beta$  Distribution  
1 1 Uniform  
0.9 1 Uniform  
1 -0.7 Correlate 1  
1 -0.7 Correlate 2  
20000 ← #Simulation

Figure 4-13 相関を用いない場合の結果

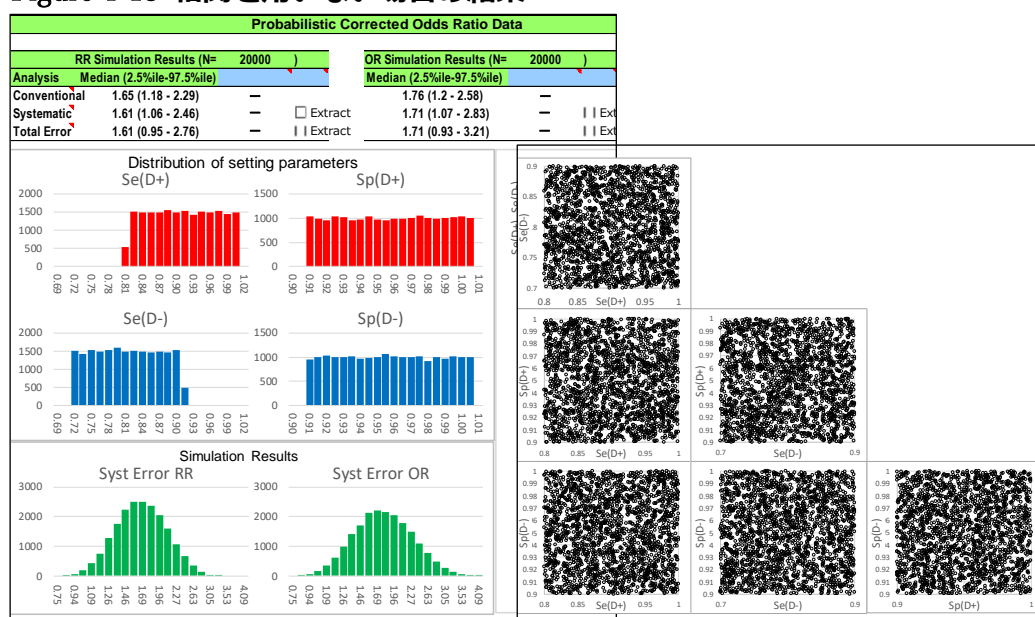
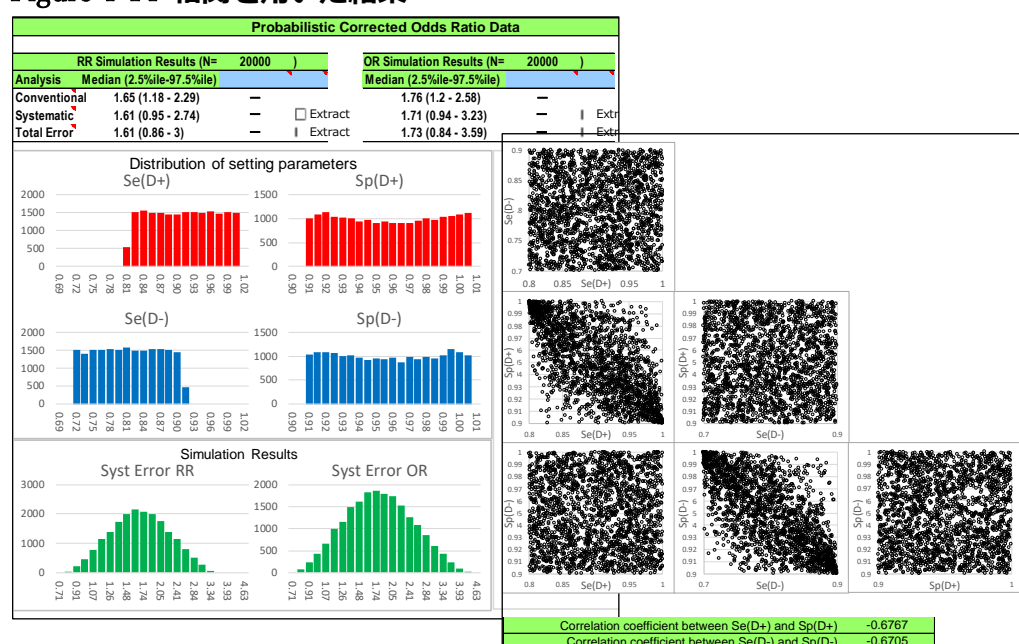


Figure 4-14 相関を用いた結果



## 4.1.8. 各種バイアスにおける PSA の設定と結果の例

ここまで、PSA の個々の技術、仕様について説明した。以下では、レジ樹脂曝露と肺癌の関係を評価した case-control 研究である Table 4-1 の数値例を用いて ESATJ (Excel により感度分析を行える Tool) により実施した結果を紹介する<sup>注20</sup>。例示のためにパラメータを設定しているため、パラメータは明確な根拠に基づいていない。また、教科書<sup>2)</sup>から変更しているものもある。

注20 : ESATJ の使い方は Appendix3.1) 1 を参照

## 4.1.8.1. パラメータが1つの場合（選択バイアスに選択オッズ比を用いる場合）

選択バイアスのパラメータ設定は複数の方法があり、選択オッズ比（第1部 3.6.6 参照）のみを用いてバイアスを補正した結果を得ることができる。選択オッズ比（ $OR_{select}$ ）は、四分表の各セルの抽出率（ $S_{A0}$ ,  $S_{A1}$ ,  $S_{B0}$ ,  $S_{B1}$ <sup>注21</sup>）を用いて

$$OR_{select} = \frac{S_{A1}S_{B0}}{S_{A0}S_{B1}} \quad \text{数式 4-1}$$

である。そして選択バイアスを補正したオッズ比（ $OR_{adj}$ ）は

$$OR_{adj} = OR_{obs} \cdot OR_{select}^{-1} \quad \text{数式 4-2}$$

が成り立つ。なお、各セルの抽出率の分布が一様分布の場合の選択オッズ比（ $OR_{select}$ ）の分布は対数正規分布となる。

確認のため、 $S_{A0}$ ,  $S_{A1}$ を共に0.5~0.7,  $S_{B0}$ を0.06~0.09,  $S_{B1}$ を0.04~0.08の一様分布とした場合（曝露のある場合に control の抽出率が高い仮定）の選択オッズ比（ $OR_{select}$ ）の分布を Figure 4-15 に示す。ここから  $OR_{select}$  を平均 $-0.233=\ln(0.792)$ 、標準偏差 0.268 の対数正規分布として ESATJ を用いて PSA を行ったの結果を Figure 4-16 に示す。パラメータが1つのみであるため、出力はシンプルで、 $OR_{select}$  と補正後オッズ比（OR）の関係も右下のグラフのように1本の曲線で表せる。補正後オッズ比（ $OR_{adj}$ ）は滑らかな対数正規分布となっている。しかし、 $OR_{select}$  に0.32~1.87の一様分布を用いた場合は Figure 4-17 のように Logit 型の分布となり、中央値も信頼区間も異なる結果となる。数式 4-1 は分布のある1つの変数による式であるが、他の複数のパラメータを用いる感度分析のモデルでも、1つのパラメータに分布を用い、他を固定値（定数）とした場合は同様にシンプルな結果になる。

Figure 4-15 各セルの抽出率に一様分布を用いた  $OR_{select}$  の分布と正規分布のあてはめ



注21：添え字の「A」はアウトカムあり、「B」はアウトカムなし、「0」は曝露なし、「1」は曝露ありを表す



Figure 4-16 ORselect に対数正規分布を用いた選択バイアスの補正結果



Figure 4-17 ORselect に一様分布を用いた選択バイアスの補正結果



## 4.1.8.2. 各層に共通のパラメータを用いる場合（曝露の非差異誤分類の例）

曝露の誤分類ではアウトカムの有無の層毎に曝露の測定について感度・特異度（ $Se_A$ ,  $Se_B$ ,  $Sp_A$ ,  $Sp_B$ <sup>注22</sup>）または陽性的中度・陰性的中度（ $Ppv_A$ ,  $Ppv_B$ ,  $Npv_A$ ,  $Npv_B$ ）をパラメータに用いるが、Simple sensitivity analysis 及び Multidimensional analysis ではアウトカムの層により誤分類に差異がない場合は共通の値（例  $Se_A = Se_B = 0.9$  など）を用いる。これをパラメータに分布を用いる PSA において表現する場合は分布から取り出した 1 つの値を両方（例えば  $Se_A$ ,  $Se_B$ ）に与える方法と、同じ分布から取り出した 2 つの値を各々与える方法（例えば一様分布から取り出した  $u_A \neq u_B$  を用いて  $Se_A = u_A$ ,  $Se_B = u_B$  とする）がある。前者は 4.1.7 及び 4.1.8.4 のパラメータの相関係数を「1」とした場合であり、後者は相関係数を「0」（無相関）とした場合である。後者のパラメータは平均（分布）は等しいが、個々の組み合わせでは異なる値の組み合わせとなる。このため感度分析の結果（補正後のオッズ比など）の中央値は等しいが、頻度区間は後者で広くなる。Table 4-1 の数値例にアウトカムの有無で共通のパラメータ（Table 4-9）を用いる場合（Figure 4-18）と、同じ分布を用いてアウトカムの有無各々でパラメータ（Table 4-10）をサンプリングした結果（Figure 4-19）を示す。前者の補正後オッズ比は中央値 1.94、95%頻度区間は 1.78~4.09 であるのに対して、後者は 1.92 と 1.02~7.14 と中央値はほぼ等しいが、区間は広がっている。パラメータの分布は感度の 95%下限が異なるように見えるが、正確には 0.8666 と 0.8648 であり、シミュレーションの誤差である。

この設定はどちらも状況、仮定によってあり得るものであり、教科書<sup>2)</sup>では両方を実施したうえで後者が選択されている<sup>注23</sup>。

Table 4-9 アウトカム有無の層に同じ曝露の感度・特異度を用いるパラメータ

Bias Parameters	Min	Max	$\alpha$	$\beta$	Distribution
Se(Lung Cancer±)	0.8	1	0.9	0.8	expit(logit)
Sp(Lung Cancer±)	0.8	1	0.77	0.8	expit(logit)
$u = \text{Uniform}(0 \sim 1)$ $g = \ln(\alpha / (1 - \alpha)) + \beta * \ln(u / (1 - u))$ $\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * (\exp(g) / (\exp(g) + 1))$					

Table 4-10 アウトカム有無の層に同じ分布から個々に感度・特異度を取り出すパラメータ

Bias Parameters	Min	Max	$\alpha$	$\beta$	Distribution
Se(Lung Cancer+)	0.8	1	0.9	0.8	expit(logit)
Se(Lung Cancer-)	0.8	1	0.9	0.8	expit(logit)
Sp(Lung Cancer+)	0.8	1	0.77	0.8	expit(logit)
Sp(Lung Cancer-)	0.8	1	0.77	0.8	expit(logit)

注22：添え字の「A」はアウトカムあり、「B」はアウトカムなしを表す

注23：パラメータには分布があり、標本を抽出するごとにばらつくと考えられる場合は層毎にもばらつきが生じると考えるであろう。この考えを反映すれば後者を用いるのが自然である。

Figure 4-18 Table 4-9のパラメータを用いた結果

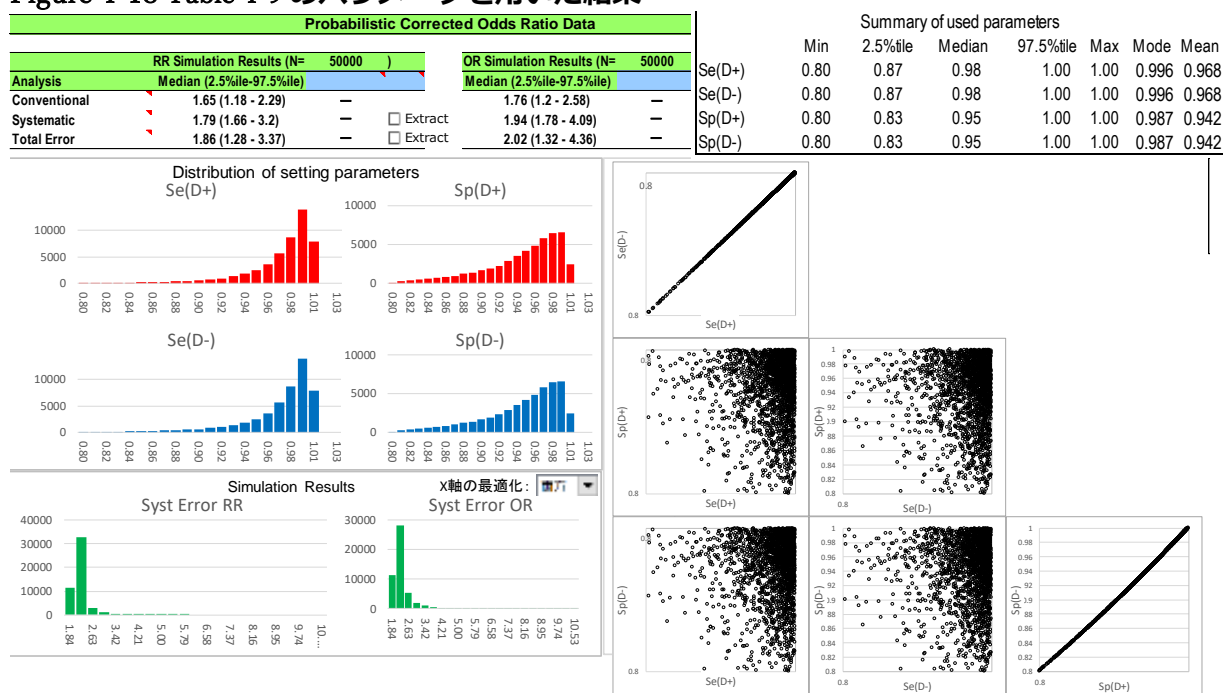


Figure 4-19 アウトカム有無の層毎に同じ分布のパラメータを用いた結果

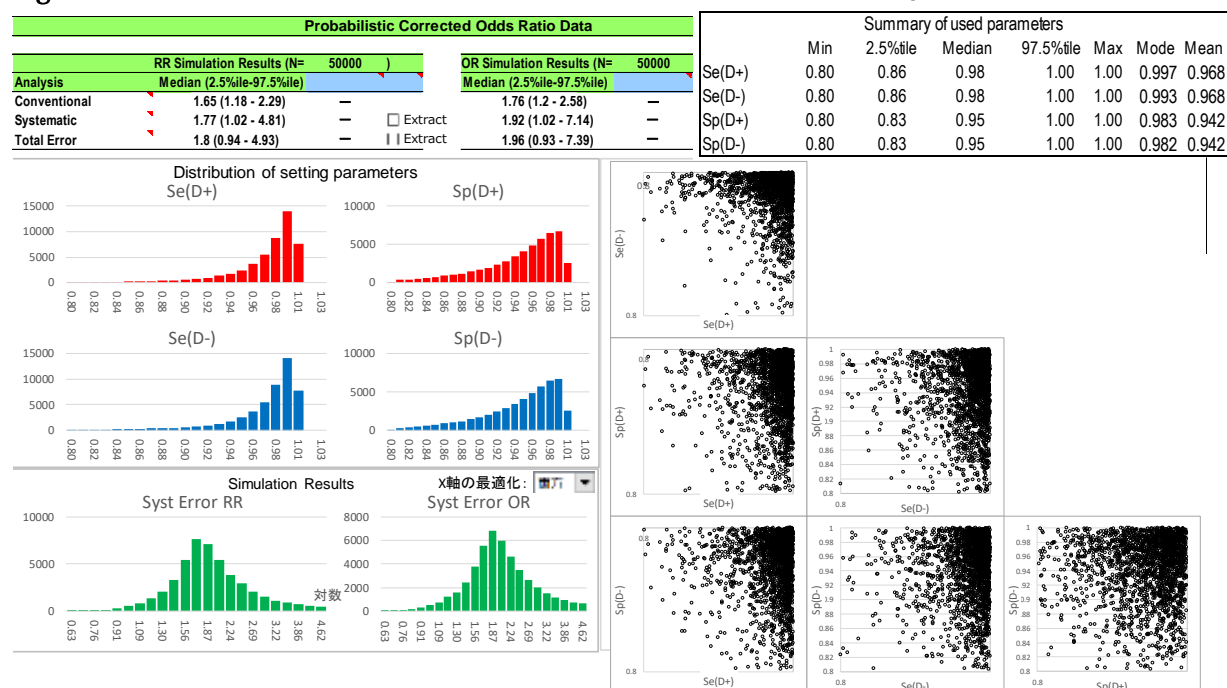
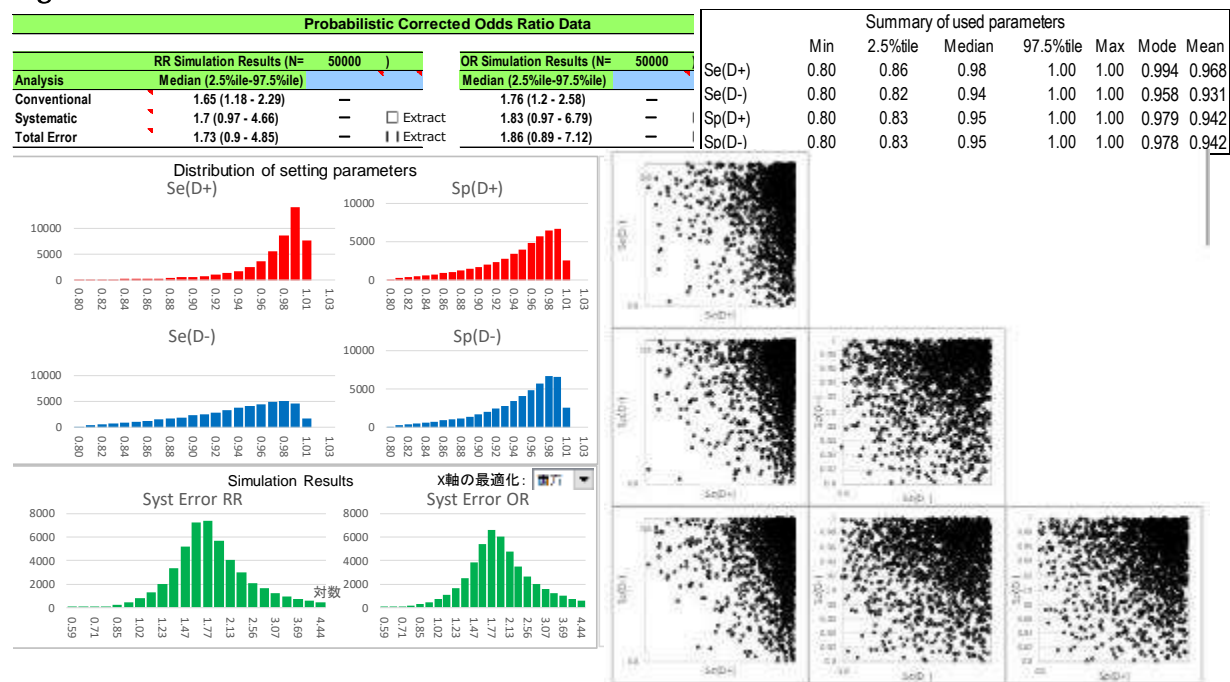


Figure 4-20 感度に差異があるパラメータの結果



## 参考 1 ESATJ における設定方法

Table 4-9 のパラメータ設定を ESATJ で行う場合は、Se(D+)と Se(D-)および Sp(D+)と Sp(D-)の間に相関係数 = 1 として設定することで行える。Se(D+)と Sp(D+)に分布の関数の種類を設定し、Se(D-)と Sp(D-)に各々同じパラメータを入力して、分布の関数の種類に Se(D+)と相関することを指定する Correlate 1 と Sp(D+)と相関することを指定する Correlate 3 を選択して相関係数「ρ」を設定する。

The parameters are **Se / Sp** Select Method **Setting of Probabilistic** **Submit** **UN PROTECT**

Variable Names		Bias Parameters		Mod1 Mod2		Distributio		ρ	
			Min	Max	α	β			
Outcome	Lung Cancer	Se (Lung Cancer+)	0.80	1	0.9	0.8	expit(logit)		
Exposure	Resins	Se (Lung Cancer-)	0.80	1	0.7	0.8	Correlate 1		1
		Sp (Lung Cancer+)	0.80	1	0.77	0.8	expit(logit)		
		Sp (Lung Cancer-)	0.80	1	0.77	0.8	Correlate 3		1

50000 ← #Simulation ☐ re-sampling

Handling of Negative Cells **ERR**

Corrected Variance?



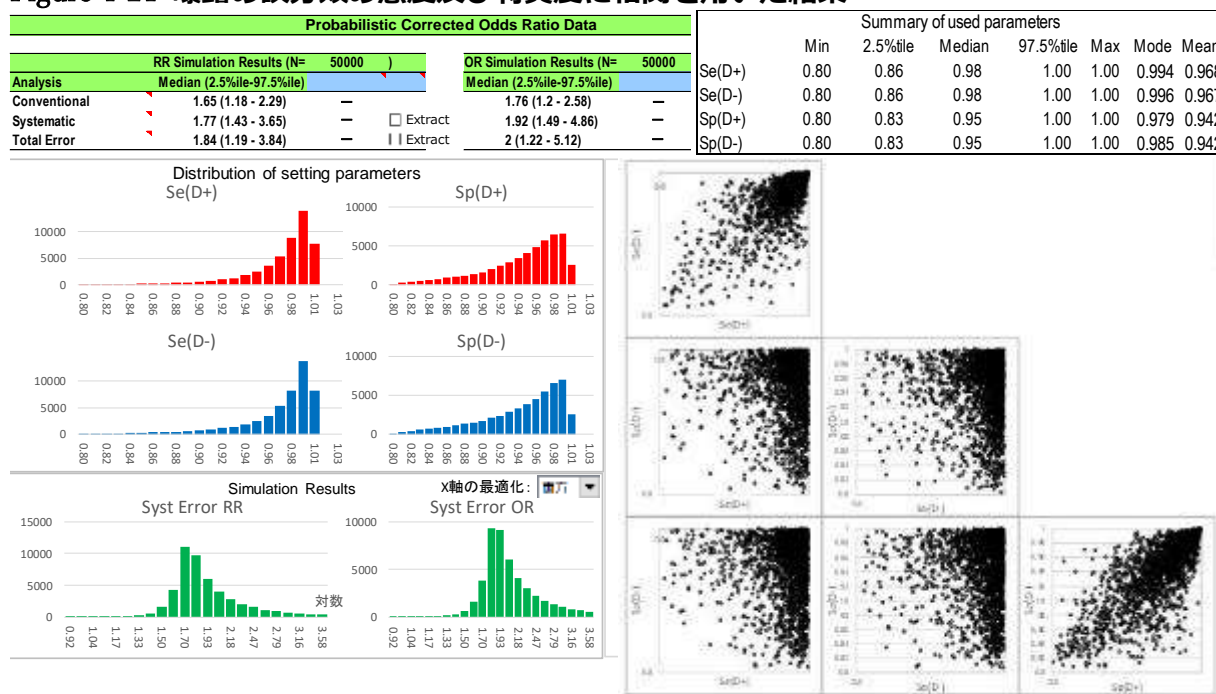
## 4.1.8.3. 層により異なる分布のパラメータを用いる場合（曝露の差異誤分類の例）

曝露又はアウトカムの層によって分布（範囲、中央値など）の異なるパラメータを用いる場合がある。誤分類では差異（的）誤分類と呼ばれる。例えば、4.1.8.2 ではアウトカムの有無間で曝露の誤分類パラメータは同じ分布であると仮定したが、この事例は case-control 研究であり、一般的に指摘されるように想起バイアスにより controls 群で曝露の報告漏れが多い（曝露ありの識別の感度が低い）ことを想定して感度に範囲は同じであるが、中央値がアウトカムありは 0.98、なしは 0.94 となるように Table 4-11 のパラメータを用いると、Figure 4-20 の結果が得られる。補正後のアウトカムなしの群の曝露の頻度が高くなる方向の設定であるため、差異がない設定（Figure 4-19）に比べて補正後オッズ比の中央値は小さいが、補正前より小さくはなっていない。

Table 4-11 感度に差異のあるパラメータ

Bias Parameters	Min	Max	$\alpha$	$\beta$	Distribution
Se(Lung Cancer+)	0.8	1	0.9	0.8	expit(logit)
Se(Lung Cancer-)	0.8	1	0.7	0.8	expit(logit)
Sp(Lung Cancer+)	0.8	1	0.77	0.8	expit(logit)
Sp(Lung Cancer-)	0.8	1	0.77	0.8	expit(logit)
$u = \text{Uniform}(0 \sim 1), \quad g = \ln(\alpha / (1 - \alpha)) + \beta * \ln(u / (1 - u))$ $\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * (\exp(g) / (\exp(g) + 1))$					

Figure 4-21 曝露の誤分類の感度及び特異度に相関を用いた結果



## 4.1.8.4. 相関のあるパラメータを用いる場合（曝露の誤分類の例）

## ✓ 同じ分布（非差異誤分類）の場合

★ 曝露の有無毎のアウトカムの感度（Table 4-11）のように同じ分布で、異なる値を取るが、分布の左右反対側との組み合わせは生じにくいと考えられる場合がある。例えば、過去に曝露の誤分類について複数の報告があり全体での感度の大小と各種の層毎の感度に相関を認める場合や、誤分類の生じる過程から、一方が絶対値として高い場合に他方が低くなると考えにくい場合がある。このような場合に分布間に相関を設定する方法が考えられる。

Figure 4-22 に 4.1.8.3 のアウトカム有の感度（Se(D+)）を 0.8 から 0.02 刻みで層別にアウトカムなしの感度（Se(D-)）の分布を示す。Se(D+)が 0.98 以上の場合でも、0.82 未満の場合でも Se(D-)の分布は同じである。同じ分布に相関（Se(D+), Se(D-)間および Sp(D+), Sp(D-)間ともに相関係数=0.8）を設定した Table 4-12 の場合の分布は Figure 4-23, 結果は Figure 4-21 である。相関を用いない Figure 4-19 とパラメータのヒストグラムおよび要約統計量は同じであるが、散布図行列は異なるものとなっている。また、補正後のオッズ比の中央値は同じであるが、頻度区間は相関を用いた場合に狭くなっている。

Table 4-12 分布が等しく相関のあるパラメータ

Bias Parameters	Min	Max	$\alpha$	$\beta$	Distribution	
Se(Lung Cancer+)	0.8	1	0.9	0.8	expit(logit)	相関係数=0.8
Se(Lung Cancer-)	0.8	1	0.9	0.8	expit(logit)	
Sp(Lung Cancer+)	0.8	1	0.77	0.8	expit(logit)	相関係数=0.8
Sp(Lung Cancer-)	0.8	1	0.77	0.8	expit(logit)	

Figure 4-22 アウトカムありの曝露の感度区分ごとのアウトカムなし群感度の分布

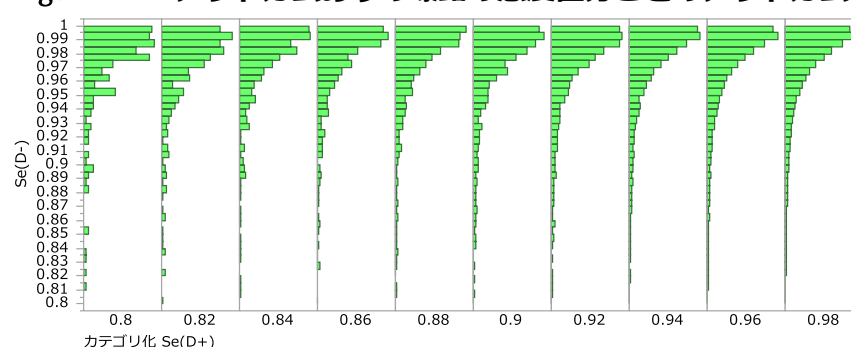
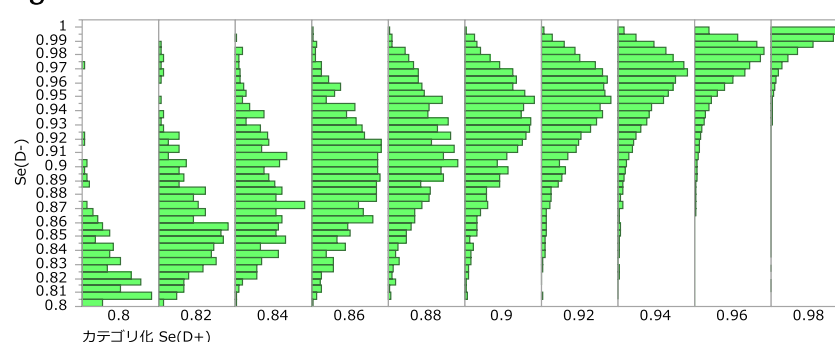


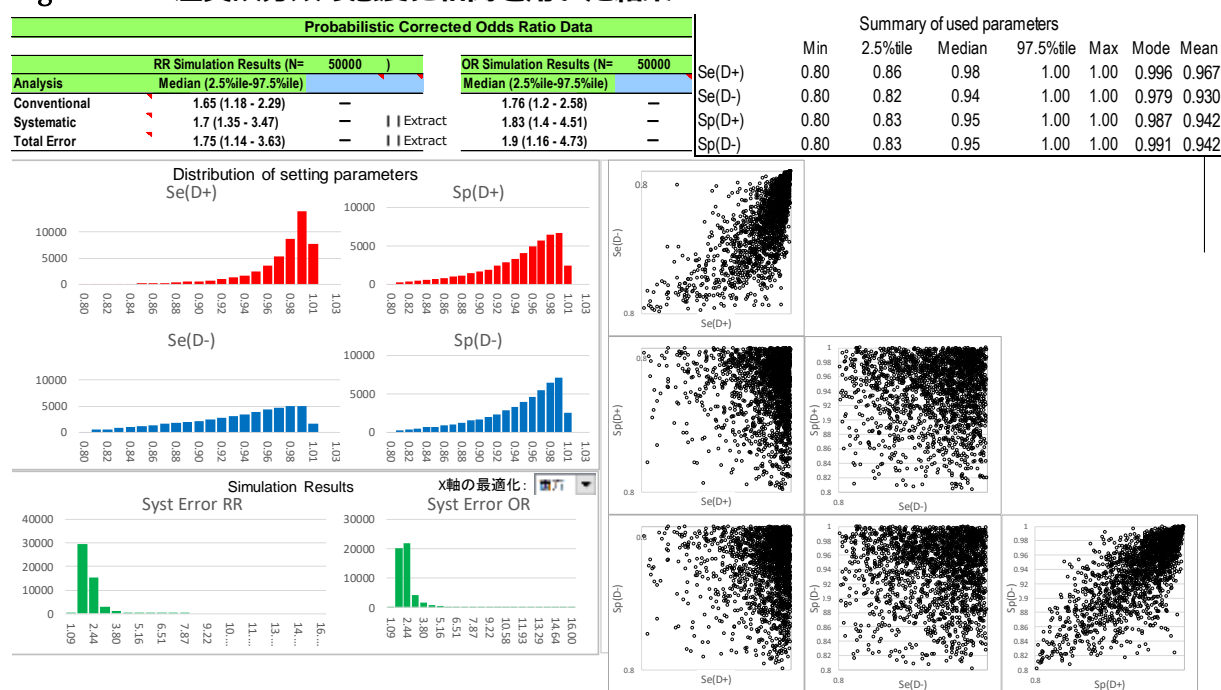
Figure 4-23 感度に相関係数=0.8 の相関のある場合の分布



### ✓ 異なる分布（差異誤分類）の場合

★ 前の例は曝露の誤分類の感度、特異度がアウトカムの有無によって差異がない場合であるが、差異がある場合でも同様のことが可能である。Figure 4-24 は Table 4-11 の曝露の誤分類の感度に差異がある場合について、Table 4-12 と同様に相関（相関係数 = 0.8）を設定した結果である。相関を用いない結果（Figure 4-20）とパラメータの要約統計量及びパラメータ毎の分布（ヒストグラム）は同じであるが、散布図行列は異なった分布を示している。このように、異なる分布間にも相関関係を持たせることが可能である。曝露の誤分類を補正したオッズ比は、前述の分布に差異がない場合と同様、中央値は等しいが、頻度区間は狭くなっている。

Figure 4-24 差異誤分類の感度に相関を用いた結果



#### 4.1.8.5. 他のパラメータに依存するパラメータを用いる場合（未測定交絡要因の例）

##### ✓ 独立の場合（比較のための説明）

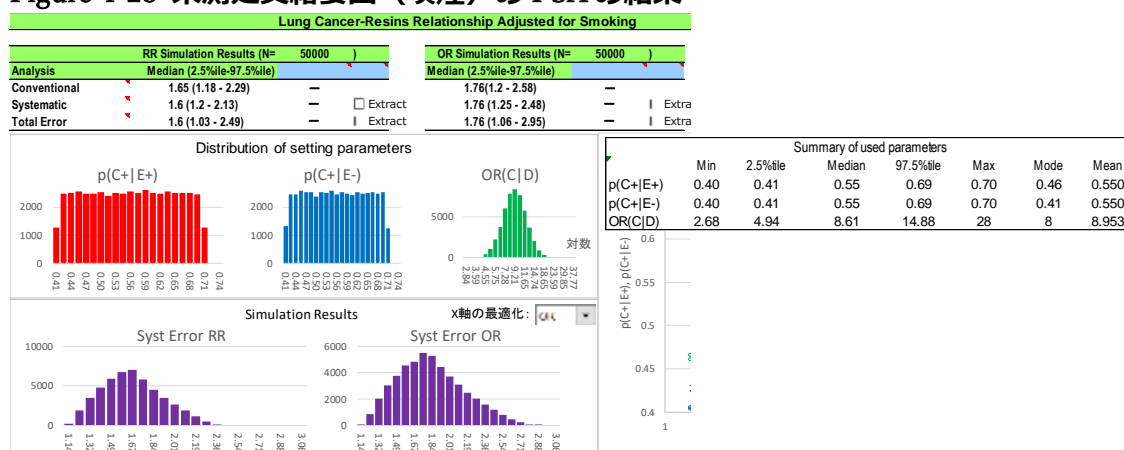
レジン樹脂曝露と肺癌の研究では、肺がんのリスクを高める喫煙状況について情報が得られていなかった。当時の米国の喫煙率は50%を超えており、レジン樹脂曝露によって喫煙率が異なる根拠がないため、喫煙者の割合としては40%から70%程度が現実的だと考えられること、また樹脂を扱う職場であれそれ以外の職場であれ、喫煙者の占める割合が極端に異なることはない（差がある根拠がない）と考えられるとされているため<sup>9)</sup>、レジン樹脂曝露ありの喫煙者の割合  $p(\text{Smoking+}|\text{Resins+})$  となしの割合  $p(\text{Smoking+}|\text{Resins-})$  の分布に差異はない0.4~0.7の一様分布、喫煙による肺癌リスクのオッズ比 ( $OR_{DC}$ ) を95%信頼区間が5~15となる対数正規分布をシナリオ (Table 4-13) とする。

Figure 4-25 に ESATJ を用いてシミュレーション回数 5 万回の PSA を行った結果を示す。下部のパラメータの要約統計量を確認すると、喫煙のオッズ比の中央値は 8.61, 95% 頻度区間は 4.94~14.88 とシナリオを反映していることが確認できる。補正前のオッズ比 OR (Resins-Lung Cancer) は 1.76 (95% 信頼区間 1.2~2.58) に対して喫煙を補正した結果,  $OR_{\text{systematic}}$  の中央値は 1.76 (95% 頻度区間 1.25~2.48), ランダムエラーを加えた結果は  $OR_{\text{totalerror}}$  の中央値は 1.76 (95% 頻度区間 1.06~2.95) であり, 中央値は補正の前後で変化はく, 頻度区間は広がっているが 1 を含んでいないため, 喫煙を考慮してもレジン樹脂と肺癌リスクにはやはり関係がありそうである。

Table 4-13 未測定交絡要因（喫煙）のバイアスパラメータ

Bias Parameters	Min	Max	Mean	SD	Distribution
p(Smoking+ Resins+)	0.40	0.70	—	—	Uniform
p(Smoking+ Resins-)	0.40	0.70	—	—	Uniform
OR(Smoking-Lung Cancer)	—	—	2.15	0.28	Log-normal

Figure 4-25 未測定交絡要因（喫煙）の PSA の結果



#### ✓ パラメータに依存するパラメータ設定

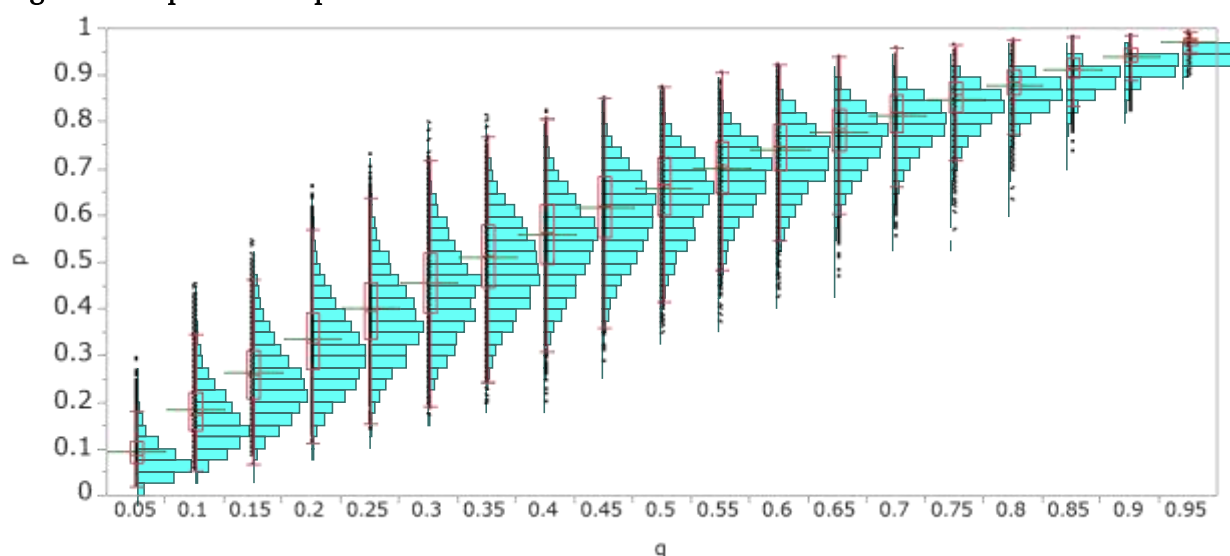
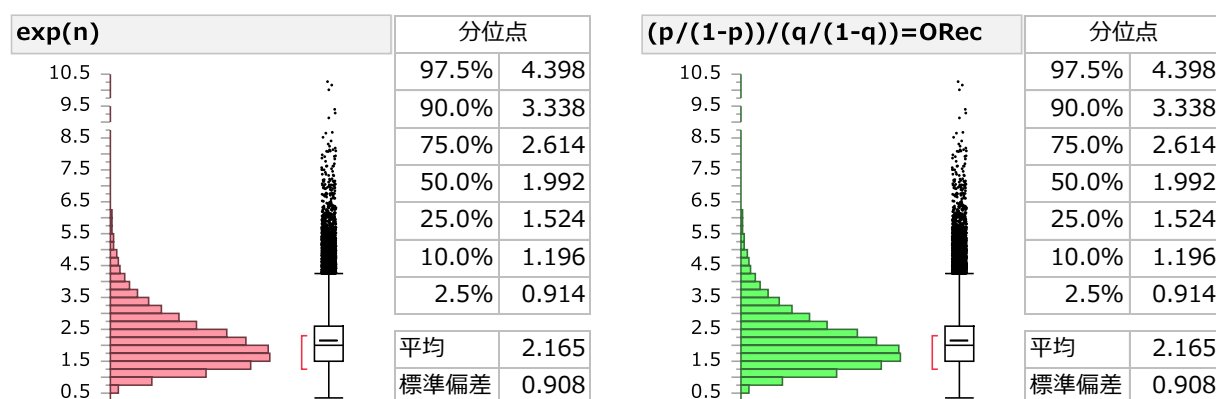
上記ではレジン樹脂曝露の層間で喫煙率の分布は同じであるが, その分布の中で全く独立に喫煙が生じるモデルを用いたが, 一方を基準にその周囲でばらついていると考えることも可能で, 第 1 部 参考 3-5 で説明したが, 利用可能なバイアスに関する情報も相対的な値の場合がある。曝露 (ex レジン樹脂) の有無と要因 (ex 喫煙) のオッズ比を  $OR_{EC}$ ,  $p=p(\text{Smoking+}|\text{Resins+})$ ,  $q=p(\text{Smoking+}|\text{Resins-})$ ,  $n=\text{normal}(\ln(\text{mean}), \text{SD})$  とすると

$$OR_{EC} = \frac{p/(1-p)}{q/(1-q)} = \exp(n) \quad \text{から}$$

$$p = \frac{1}{\frac{(1-q)}{q \cdot \exp(n)} + 1}$$

数式 4-3

により,  $p(\text{Smoking+|Resins-})$ と相対的な  $p(\text{Smoking+|Resins+})$ が得られる.  $\text{mean} = 2$ ,  $\text{SD} = \sqrt{0.4}$  の場合の  $p$  と  $q$  の関係 ( $q$  は 0.05, 0.1・・・0.95) は Figure 4-26 のとおりで, 0 及び 1 に近い場合は偏った対数正規分布型となっている. 与えた分布  $\exp(\ln(\text{mean}), \text{SD})$ と得られた  $p$ ,  $q$  から求めら  $OR_{EC}$ の分布は Figure 4-27 のとおり等しいことが確認できる.

Figure 4-26  $q$  に対する  $p$  の分布Figure 4-27 与えた分布  $\exp(n)$  と得られた  $OR_{EC}$  が同じである確認

レジン樹脂曝露ありの喫煙率をレジン樹脂曝露なしの喫煙率との相対値 ( $OR_{EC}$ ) を曝露なしの喫煙率を中心に対数正規分布する数式 4-3 を用いて Table 4-14 のパラメータを用いた結果は,  $p(\text{Smoking+|Resins+})$  となしの割合  $p(\text{Smoking+|Resins-})$  の中央値及び平均値はほぼ等しく,  $OR_{systematic}$  の中央値は 1.76 (95%頻度区間 1.64~1.9) と Figure 4-25 とは等しいが, 頻度区間は異

なっている。これは Figure 4-29 のとおり、喫煙率の組み合わせに相関関係が生じてパラメータの組み合わせのばらつきが小さくなったためである。

Table 4-14 相対的な喫煙率の設定

Bias Parameters	Min	Max	Mean	SD	Distribution
$p(\text{Smoking+} \text{Resins+})$	—	—	1	0.4	$OR_{EC}$
$p(\text{Smoking+} \text{Resins-})$	0.40	0.70	—	—	Uniform
$OR(\text{Smoking-Lung Cancer})$	—	—	2.15	0.28	Log-normal

Figure 4-28 相対的な喫煙率を用いた結果

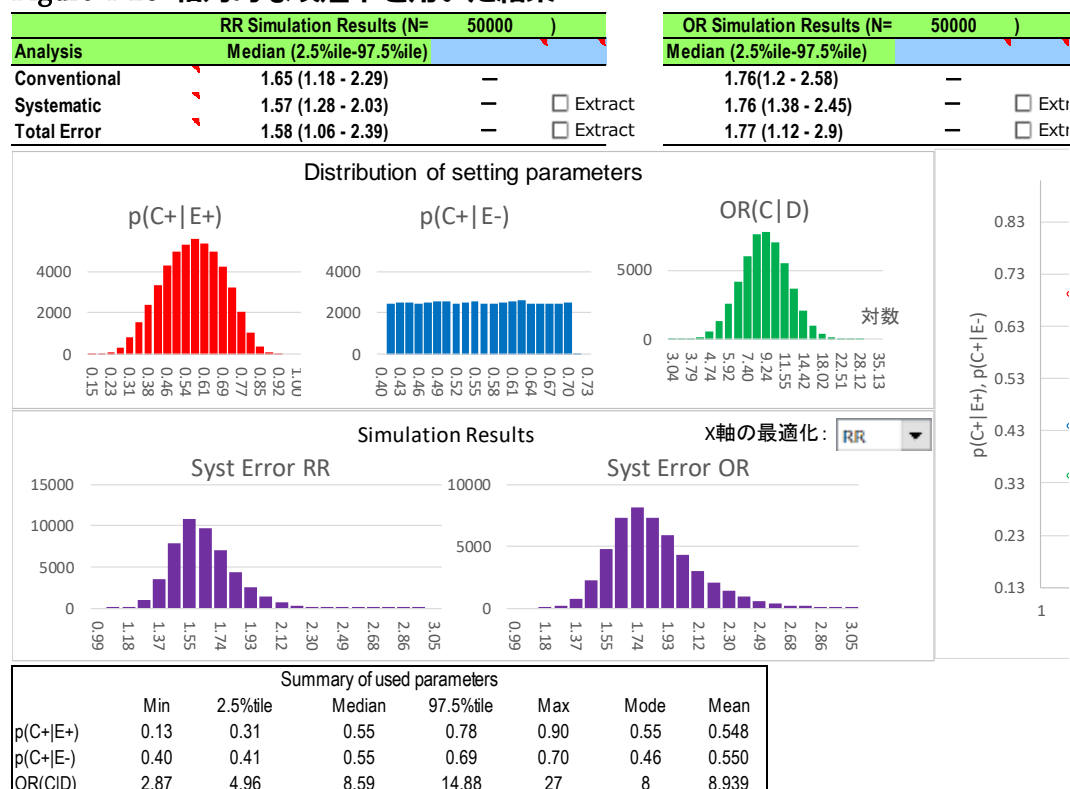
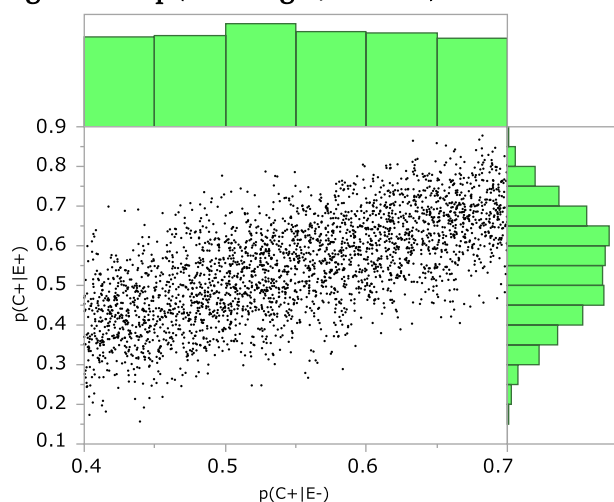


Figure 4-29  $p(\text{Smoking+}|\text{Resins+})$  となしの割合  $p(\text{Smoking+}|\text{Resins-})$  の関係





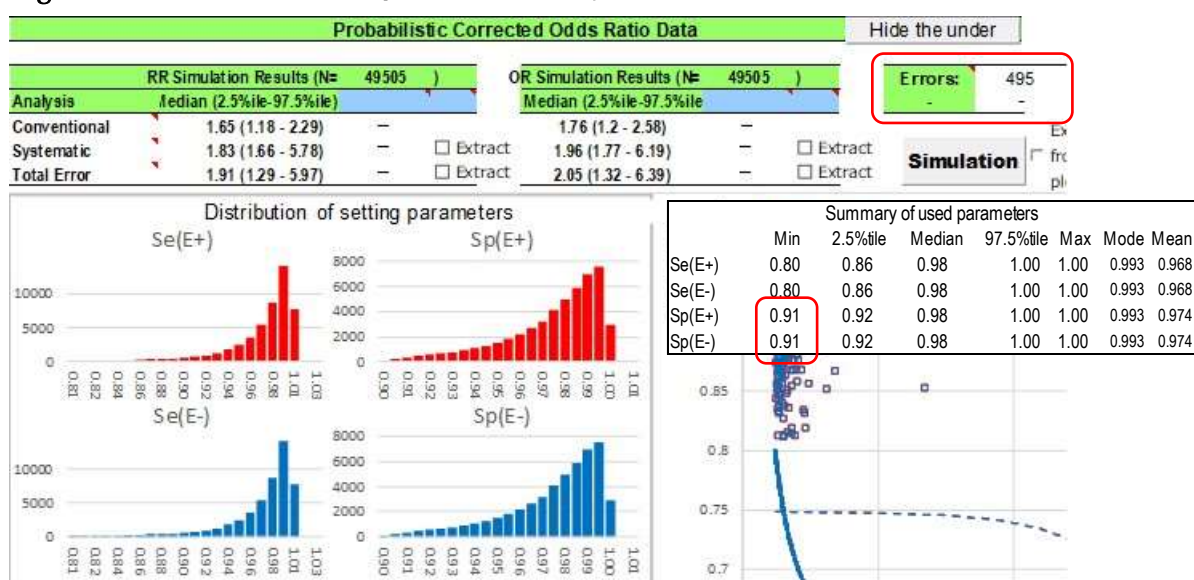
## 4.1.8.6. エラーの生じるパラメータ（アウトカムの誤分類の例）

★ 4.1.3 で説明しているが、パラメータの分布に正規分布など、最小値、最大値を設定できない分布関数を用いる場合、パラメータとしてあり得ない値（感度、特異度が1を超える又は負の値など）の場合は、バイアスの補正の計算の前にパラメータのみで判断可能であるが、誤分類ではパラメータとしてはあり得ても、計算の結果、四分表のセルに負の値が生じ、エラーとなる場合がある。ここでは、Table 4-1 の数値例を用い、アウトカムの誤分類<sup>注24</sup>として Table 4-15 の設定（曝露の有無間のパラメータは一致している）を用いてシミュレーション回数 50,000 回の PSA を行くと、約 1%のエラーが生じた（Figure 4-30 参照）。この場合は、特異度が 0.90~0.9095 の場合にエラーが生じておりパラメータの分布は特異度でわずかに狭くなるが、感度は変わらない（Figure 4-30）。パラメータの組み合わせを Figure 4-31 により確認すると、感度の値に関わらずエラーが生じることがわかる。この例の場合はエラーとなるのは分布の裾で低頻度であるので、エラーが生じていることを無視（報告しない）してもほぼ影響は生じないが、エラーの有無と内容の確認を行い、エラーが生じないようにパラメータを修正しない場合はエラーが生じている（その結果、パラメータの設定と使用された値にずれがある）ことは報告すべきである。

Table 4-15 曝露の誤分類の感度分析パラメータ

Bias Parameters	Min	Max	$\alpha$	$\beta$	Distribution
Se(Resins±)	0.8	1	0.9	0.8	expit(logit)
Sp(Resins ±)	0.9	1	0.8	0.8	expit(logit)
$u = \text{Uniform}(0 \sim 1), g = \ln(\alpha / (1 - \alpha)) + \beta * \ln(u / (1 - u))$ $\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * (\exp(g) / (\exp(g) + 1))$					

Figure 4-30 曝露の誤分類の結果（エラーあり）



注24： Table 4-1 の数値例は case-control 研究であるので、cases と controls は抽出率が異なるため、単純な四分表に対する感度分析の曝露の誤分類とはならないが、計算例として用いている。仮に抽出率を cases は 100%, controls は 1%とした場合、Table 4-1 の数値例は B<sub>1</sub> セルが 25700, B<sub>2</sub> セルが 94500 となるので、この設定ではシミュレーションの約 40%がエラーとなる。

Figure 4-31 感度と特異度の組み合わせ（赤はエラーの組み合わせ）

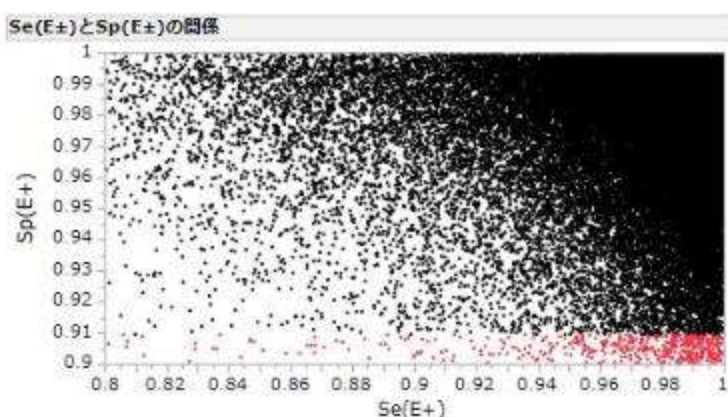
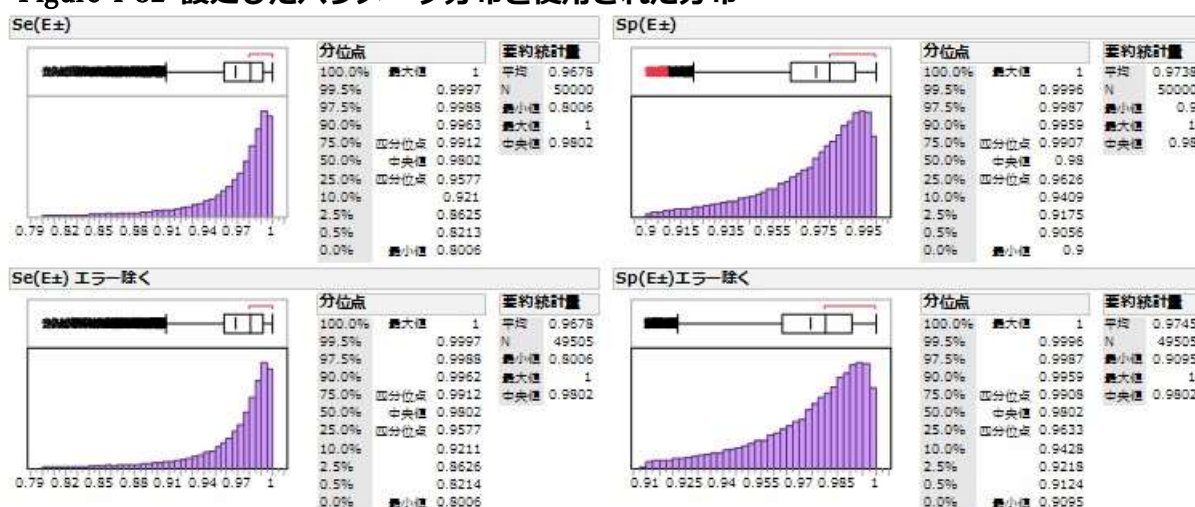


Figure 4-32 設定したパラメータ分布と使用された分布



## 4.2. Multiple bias (Probabilistic) modelling

観察研究において重要なバイアスの原因は1つとは限らず、曝露とアウトカム両方の誤分類があるなど、複数存在する場合は少なくない。このような複数のバイアス各々に対するPSAの結果を併合することで、複数のバイアスが同時に働いた場合の影響を提示する手法である。用途としては複数のバイアス（モデル）があり、

- ・ 結果への影響の方向に一貫性がなく、統合した評価が必要な場合
- ・ 個々の影響度は大きくないが、結果への影響に一貫性があり、全体では大きな影響が考えられる場合。
- ・ 複数の独立した未測定交絡変数が想定され、各々にパラメータの根拠となるデータがあり1つの未測定交絡要因のモデルに統合できない場合。



などである。MCSA による Probabilistic sensitivity analysis を用いる場合は複数要因への拡張は要因間が独立という条件の下で任意の確率密度関数による感度分析パラメータを用いて繰り返し計算した結果の分布を用いることができる (Combined analysis) <sup>7),2),8)</sup>。

具体的な手順は、

- ・ 個々のバイアスの原因について PSA (MCSA) を行い、バイアスの原因各々のパラメータを決め、その補正結果を確認する。
- ・ そして組み合わせる要因を選び、順序を設定する。
- ・ 改めて MCSA を順に行うが、その際に 1 つ前の MCSA の各サンプリングで計算された補正後の四分表 (Corrected Data) を用いる<sup>注25</sup>

ことを繰り返すことで併合した値 (combined results) を得る (誤分類、選択バイアスは各サンプリングの補正後オッズ比を積算するのではない)。

4.1.4 で PSA におけるシミュレーション回数は、パラメータの数とその範囲に応じて十分な回数を設定する必要があることを説明したが、Multiple bias analyses は複数の PSA を併合するため、結果的に用いられるパラメータの数が多くなる。従って、PSA の場合よりも多いシミュレーション回数が必要となることに注意する。なお、ESATJ では個々の PSA は最大 10 万回であるが、Multiple bias analyses では 20 万回を最大としている。

教科書<sup>2)</sup>にレジジン樹脂の曝露と肺癌の発生の関係を報告した case-control 研究のデータ (Table 4-1) を用いてパラメータの設定から Multiple bias analyses の実施までが説明されている。教科書とは若干異なる部分があるが、4.1.8 で例示した各種バイアスの PSA による感度分析から、個々のバイアスの検討では、未測定交絡 (偏りなし Figure 4-25, あり Figure 4-28)、選択バイアス (Figure 4-16)、曝露の誤分類 (Figure 4-21)、アウトカムの誤分類 (Figure 4-30) の 5 つを検討して、アウトカムの誤分類、選択バイアス、未測定交絡 (偏りあり) の 3 種類を Multiple bias analyses に用いることとする。Table 4-16 はそのパラメータの要約である。未測定交絡要因 (喫煙) の設定は 2 種類あるが、このうち 2 つ目 (曝露なしに対する曝露ありの喫煙率のオッズ) を用いる。青いセルの番号は、併合する順番である。Figure 4-33, Figure 4-34 はこの例の 20 万回のシミュレーションの結果である<sup>注26</sup>。

複数のバイアスを組合せる Multiple bias modelling は、感度分析モデルに含める要因間に原因結果関係が生じ、計算する順序を考慮する必要が生じる場合がある<sup>7),6)</sup>。順番の考え方は、計算をバイアスの影響を受けた結果の表から始めるため、バイアスの発生順を巻き戻す (研究

注25 : 未測定交絡要因および選択バイアスは四分表を用いずに補正前のオッズ比を用いて計算可能であるが、その場合、補正後の四分表は一意に定まらないため、次のステップ四分表が必要とならないようにする必要がある。また、共変量の誤分類を扱う場合は、共変量で層別した四分表が必要であるため、最初の Step で行う、あるいは途中で扱う場合は、観測された共変量による 2 つの四分表各々のセルへの分割割合をパラメータとして扱うなど工夫が必要となる。

注26 : 教科書では選択バイアスのパラメータは台形分布の対数変換であるのを、対数正規分布に、アウトカムの特異度の最小値が 0.8 となる式であるが 0.9 未満では負のセルが生じるため 0.9 に修正している。なお、上側の Forest plots は教科書にはない。

プロセスでバイアスが生じる逆順) ように行う。例えばアウトカムの誤分類と選択バイアスがある場合、アウトカムの誤分類の補正後に選択バイアスを補正する。なお、用いる四分表のセルの度数が変わるためエラー（負のセルを生じる）となるパラメータも変化する可能性がある。

Figure 4-33 を異なる順序（曝露の誤分類→選択バイアス→未測定交絡要因）で行った場合の比較を Figure 4-35 に示す。教科書で行われている順番の補正後オッズ比が 2.49（95%頻度区間 1.17～7.16）であるのに対して、選択バイアスとアウトカムの誤分類の順序を入れ替えて行った結果は 2.28（1.18～4.71）になっている。これは Table 4-17 のとおり、各ステップで補正された結果の 4 分表のセルの度数が異なるためである。

Table 4-16 各 PSA のパラメータ設定の要約

↓ SELECT COMBINED ORDER		Bias Parameters	Min	Mode1/ $\alpha$	Mode2/ $\beta$	Max	$\rho$	Distribution
NO	UNMEASURED	p(Smoking+ Resins+)	0.40			0.70		Uniform
	CONFOUNDING without effect modification	p(Smoking+ Resins-)	0.40			0.70		Uniform
		OR(Smoking-Lung Cancer)	8.60			0.28		Log-normal
3	UNMEASURED	OR(Smoking-Resins)	1			0.639		ORec
	CONFOUNDING without effect modification	p(Smoking+ Resins-)	0.4			0.7		Uniform
		OR(Smoking-Lung Cancer)	8.66			0.28		Log-normal
2	SELECTION BIAS	Use Selection OR	-0.233			0.268		Log-normal
1	EXPOSURE MISCLASSIFICATION	Se (Lung Cancer+)	0.8	0	0.8	1		expit(logit)
		Se (Lung Cancer-)	0.8	1	0.8	1	0.8	Correlate with Se(D+)
		Sp (Lung Cancer+)	0.8	1	0.8	1		expit(logit)
		Sp (Lung Cancer-)	0.8	1	0.8	1	0.8	Correlate with Sp(D+)
NO	OUTCOME MISCLASSIFICATION	Se(Resins+)	0.8	0	0.8	1		expit(logit)
		Se(Resins-)	0.8	1		1	1	Correlate with SeE
		Sp(Resins+)	0.9	1	0.8	1		expit(logit)
		Sp(Resins-)	0.9	1		1	1	Correlate with SpE

Figure 4-33 Multiple bias analysis の結果の例 (Forest plots)

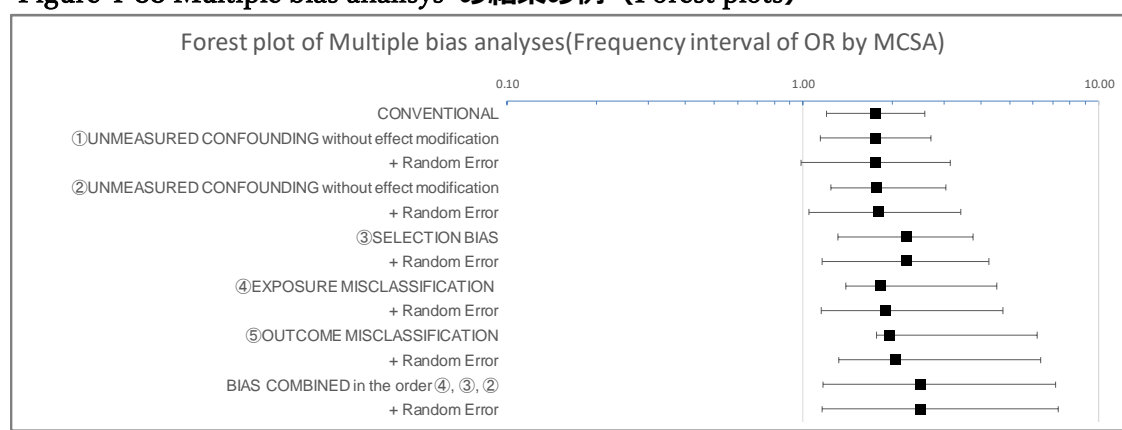


Figure 4-34 Multiple bias analyses の結果例（要約）

	Without Incorporating Random Error			With Incorporating Random Error		
	Median	Percentiles 2.5th-97.5th	Ratio of Limits	Median	Percentiles 2.5th-97.5th	Ratio of Limits
CONVENTIONAL	1.76	— —	—	1.76	1.20 ~ 2.58	2.14
① UNMEASURED CONFOUNDING without effect modification p(C+ E+)~Uniform(0.4,0.7), p(C+ E-)~Uniform(0.4,0.7), OR(C-E)~Log-normal(8.6,0.28),	1.76	1.14 ~ 2.71	2.37	1.76	0.99 ~ 3.15	3.19
② UNMEASURED CONFOUNDING without effect modification OR(C-E)~ORec(1,p(C+ E+),0.64), p(C+ E+)~Uniform(0.4,0.7), OR(C-D)~Log-normal(8.66,0.28),	1.76	1.25 ~ 3.05	2.45	1.79	1.05 ~ 3.42	3.26
BIAS COMBINED	2.49	1.17 ~ 7.17	6.10			
③ SELECTION BIAS Selection OR~Log-normal(-0.23,0.27)	2.23	1.31 ~ 3.77	2.87	2.22	1.16 ~ 4.26	3.67
BIAS COMBINED	2.42	1.29 ~ 6.31	4.91			
④ EXPOSURE MISCLASSIFICATION Se(D+)~expit(logit)(0.8,0,0.8,1) Se(D-)~Correlate with Se(D+)(0.8,1,0.8,1,0.8) Sp(D+)~expit(logit)(0.8,1,0.8,1) Sp(D-)~Correlate with Sp(D+)(0.8,1,0.8,1,0.8)	1.83	1.40 ~ 4.51	3.23	1.90	1.16 ~ 4.73	4.10
BIAS COMBINED	1.83	1.40 ~ 4.51	3.23			
⑤ OUTCOME MISCLASSIFICATION SeE~expit(logit)(0.8,0,0.8,1), SpE~expit(logit)(0.9,1,0.8,1),	1.96	1.77 ~ 6.19	3.49	2.05	1.32 ~ 6.39	4.85
BIAS COMBINED in the order ④, ③, ②	2.49	1.17 ~ 7.17	6.13	2.49	1.15 ~ 7.28	6.34
Uniform(min,max), Log-normal(mean,SD), expit(logit)(min,a,β,max), ORec(min,p(C+ E+),SD), Correlate with Parameter(Pr1,Pr2,Pr3,Pr4,p)						

Figure 4-35 Multiple bias analyses の順番の影響

	Without Incorporating Random Error			With Incorporating Random Error		
	Median	Percentiles 2.5th-97.5th	Ratio of Limits	Median	Percentiles 2.5th-97.5th	Ratio of Limits
CONVENTIONAL	1.76	— —	—	1.76	1.20 ~ 2.58	2.14
① UNMEASURED CONFOUNDING without effect modification p(C+ E+)~Uniform(0.4,0.7), p(C+ E-)~Uniform(0.4,0.7), OR(C-E)~Log-normal(8.6,0.28),	1.76	1.14 ~ 2.71	2.37	1.76	0.99 ~ 3.15	3.19
② UNMEASURED CONFOUNDING without effect modification OR(C-E)~ORec(1,p(C+ E+),0.64), p(C+ E+)~Uniform(0.4,0.7), OR(C-D)~Log-normal(8.66,0.28),	1.76	1.25 ~ 3.05	2.45	1.79	1.05 ~ 3.42	3.26
BIAS COMBINED	2.26	1.18 ~ 4.71	3.99			
③ SELECTION BIAS Selection OR~Log-normal(-0.23,0.27)	2.23	1.31 ~ 3.77	2.87	2.22	1.16 ~ 4.26	3.67
BIAS COMBINED	2.23	1.31 ~ 3.77	2.87			
④ EXPOSURE MISCLASSIFICATION Se(D+)~expit(logit)(0.8,0,0.8,1) Se(D-)~Correlate with Se(D+)(0.8,1,0.8,1,0.8) Sp(D+)~expit(logit)(0.8,1,0.8,1) Sp(D-)~Correlate with Sp(D+)(0.8,1,0.8,1,0.8)	1.83	1.40 ~ 4.51	3.23	1.90	1.16 ~ 4.73	4.10
BIAS COMBINED	2.23	1.31 ~ 3.78	2.88			
⑤ OUTCOME MISCLASSIFICATION SeE~expit(logit)(0.8,0,0.8,1), SpE~expit(logit)(0.9,1,0.8,1),	1.96	1.77 ~ 6.19	3.49	2.05	1.32 ~ 6.39	4.85
BIAS COMBINED in the order ③, ④, ②	2.27	1.18 ~ 4.72	4.00	2.27	1.16 ~ 4.80	4.15
Uniform(min,max), Log-normal(mean,SD), expit(logit)(min,a,β,max), ORec(min,p(C+ E+),SD), Correlate with Parameter(Pr1,Pr2,Pr3,Pr4,p)						

Table 4-17 Multiple bias modelling の各ステップのセルの度数

EXPOSURE MISCLASSIFICATION						→ SELECTION BIAS□				
Cells	a	b	c	d	OR <sub>adj</sub>	a	b	c	d	OR <sub>adj</sub>
2.5%tile	26.8	90.5	73.1	912.0	1.40	27.3	89.1	72.6	898.6	1.31
50%tile	41.7	97.3	226.9	975.1	1.83	50.6	96.4	224.7	966.7	2.23
97.5%tile	48.5	112.2	290.0	1128.9	4.51	87.4	111.6	287.2	1122.2	3.77

SELECTION BIAS□						→ EXPOSURE MISCLASSIFICATION				
Cells	a	b	c	d	OR <sub>adj</sub>	a	b	c	d	OR <sub>adj</sub>
2.5%tile	33.8	90.5	247.5	910.2	1.77	33.8	90.5	247.5	910.2	1.40
50%tile	56.4	93.2	254.7	936.7	1.96	56.4	93.2	254.7	936.7	1.83
97.5%tile	92.7	94.8	259.2	953.2	6.19	92.7	94.8	259.2	953.2	4.51

事例としては、インフルエンザワクチンとナルコレプシーの関係を評価した Bollaerts K の論文<sup>23)</sup>は、未測定交絡、選択バイアス及び曝露とアウトカムの誤分類を MCSA による計算を Step1 曝露の誤分類、Step2 アウトカムの誤分類、Step3 選択バイアス、Step4 未測定交絡要因、Step5 ランダムエラーの順で行っている。また、Gerdtz C<sup>24)</sup>は中絶関連死亡率の測定誤差を評価し、死因の誤分類、選択バイアスの順で扱っているが、両論文にはパラメータの設定および各段階の計算の数式、分布を含めたパラメータの詳細および Supporting Information, Additional files として R のコードが添付されており、Bollaerts K の論文とともに教科書<sup>7),2),8)</sup>よりも分かりやすい。また、日本動脈硬化症研究コホートを用いて動脈硬化性疾患に対する血圧の影響を血圧の誤測定と未測定交絡要因である糖尿病の罹患について MCSA の評価を行った Takeuchi A の論文<sup>25)</sup>は高血圧の程度の分類の誤分類と未測定交絡要因について感度分析を行い、MCSA およびモデルの数式、理論および、用いるパラメータの情報源の詳細な考察が行われており、MCSA 一般におけるパラメータの考察の参考となる。Momoli F の携帯電話の使用とグリオーマ、聴神経腫などとの関係を評価し、選択バイアスと想起バイアスを MCSA による補正を行っている論文は<sup>26)</sup>詳細な計算方法などは省略されているが、一般的な報告における Multiple bias analysis の報告の仕方の参考になる。

このようなバイアスを個別に考慮し、その結果を次のバイアスで用いる単純な方法ではなく、例えば各バイアス間に相関がある、あるいはバイアス間に相互作用（修飾）がある場合などは本書では扱わないが、Bayesian sensitivity analysis を用いることになる<sup>11),12),13)</sup>。

## 5. 感度分析の準備から結果の報告まで

ここまで、感度分析の基礎知識として種々のバイアスの説明、感度分析の理論並びに手法の説明を行ってきた。第1章冒頭で「感度分析では、方法の選択から結果の解釈まで、その多くが解析者の手に委ねられている」と説明したように、感度分析は用いようによっては研究の結果を恣意的に強調することも否定することも可能である。このため、感度分析に用いた仮定の説明性、透明性が重要であるため、本章では感度分析の必要性の判断から、結果の解釈までを説明する。

感度分析は大きく分けてその研究を行った研究者自身が行う場合とそれ以外の第三者（例えば論文の読者、製薬企業・規制当局の安全性担当者など）が行う場合があるが、両者では研究に対する知識量、感度分析の目的に大きな違いが生じる。さらに前者はあらかじめ研究の計画に組み入れられる場合と事後的に行われる場合がある。あらかじめ感度分析を計画する場合は感度分析の必要性の判断に研究の結果を用いることができないが、その研究の中で感度分析に必要なデータを収集することで感度分析の精度を高められる利点がある。また研究者自身であれば、症例単位の詳細なデータを用いる高度な感度分析の手法を用いることが可能であるのに対して、第三者が行う場合は研究について調べ、背景・周辺情報を収集してバイアスを特定する必要があること、通常は症例単位の詳細なデータにアクセスできないなどの違いがある。

本章では教科書に習いながら、医療情報データベースの二次利用及び第三者が感度分析を行う場合を含めて、各作業に分けて事例を含めて説明するが、「5.1 感度分析の必要性を評価する」から「5.4 感度分析すべきバイアスを選択する」は明確な順番があるのではなく、並行する部分が多い。簡潔な概要は[入門編](#)を参照いただきたい。

本章では多くの計算事例を提示している。グラフを伴う感度分析は全て本書のために作成した Excel book の Tool ESATJ<sup>Appendix3.1)</sup><sup>1)</sup>を用いて行いその結果を表示している。追加で用いる確率分布のグラフ表示などは SAS 社の JMP®13 を用いている。

### 5.1. 感度分析の必要性を評価する

感度分析は常に必要なものではなく、行う必要がない場合、行うメリットがないあるいは行うにしても簡便な方法で十分な場合もある<sup>注27</sup>。研究結果と政策的な意思決定との関係、社会的な影響度、緊急性、情報量や効果の大きさといった“研究の重要性”と、その研究そのものの内的妥当性、外的妥当性といった“研究の質”（データの質を含む）の2つの軸の組み合わせによって感度分析の必要性、実用性および求められる精度は異なる。従って、感度分析を行うか

---

注27：我々製薬企業の市販後安全性評価者などが医薬品の安全性に関する論文を評価するなどのため行う感度分析は、容易に入手できる情報を用いるのが通常であろう。このような場合、本書で主に扱う簡便な手法には労力を要しないため、感度分析の要否を深く検討する必要はなく、恣意的（「リスクがあるとは言えない」といった結論を前提とした）にならないように注意することで十分であろう。



否かを定める前に研究の主要な結果から導き出される因果関係の重要性と結果が用いられる場面<sup>注28</sup>を明らかにして、結果及び判断が誤っていた場合の損失・リスクを考える必要がある。

以上のようなことが教科書<sup>2), 4), 6), 7), 8)</sup>には説明されている。しかし、医療情報データベースなど、行おうとする研究の目的外で収集されたデータを二次的に利用する場合は誤分類により結果が誤っている可能性が大きいいため標準的に感度分析が行われるべきであろう。また、製薬企業において実施する観察研究や研究報告のための論文の評価は措置の判断につながるので標準的に頑健性の評価を目的に感度分析を実施する（そのため要否の判断は不要である）ことも考えられるため、詳細は Appendix1 に教科書に習い研究の質と研究の意義、重要性の視点で説明している。validation study を実施する場合の判断も含まれるので必要な場合は参照されたい。

## 5.2. 感度分析の目的と手法との関係

感度分析は研究や調査から得られていない情報を用いるため、解析者の意図次第で様々な結果を導くことが可能であるという危険な側面を有する。感度分析はバイアスを特定、定量化してバイアスがない場合の値を推定するというより、仮定したバイアスを補正した結果を用いて、研究から得られた推定値の確からしさや信頼性の評価が主な目的である。つまり、感度分析の目的と位置づけはバイアスの仮定の仕方と研究結果の確からしさが反映される。例えば、次の3種類に大別するとしよう。

- ① 現実に想定されるバイアスの影響を確認する。  
過去の研究、経験からある程度具体的に情報が得られているバイアスの影響を確認する。一般的な研究結果の報告に記載される“本研究の限界”の定量的な評価を行う。「現実にあるような」バイアスの量を用いるため、その根拠となるデータの収集、説明が重要となる。
- ② 補正後が任意の値となるバイアスの量を示す。  
補正後のオッズ比が1以下や2以上となるなど、意思決定への影響などから設定した目標（閾値）の値となる場合を求める。target-adjustment sensitivity analysis<sup>4)</sup>とよばれる手法で、結果が覆るようなバイアスの程度を定量的に示す目的などで用いられる。一見、簡単なようであるが、1つのバイアスは、通常、2つ以上のパラメータにより定義されるため、一意に定まらない場合が多い。
- ③ 現実的ではない大きなバイアスを想定した結果を示す。  
その研究自体（解析で調整に用いた要因）や、過去の研究、経験などから想定されるよりも大きなバイアスを仮定した結果を求めることで、想定される範囲を超える大きなバイアスによっても結果が覆らない（頑健である）ことを示す。

---

注28：我々製薬企業にとっては、医薬品の安全確保措置の要否の判断など

①は validation study やこれまでの研究から感度分析に用いる情報（アウトカムの誤分類であれば、曝露の層毎の感度、特異度など）がある場合であり、感度分析を行う研究の結果（対象が benefit/risk の何れであるか、曝露効果の大小）に影響を受けない。しかし、常にバイアスについて妥当と思われる根拠があると限らない。そういった場合、結果の頑健性を示す目的で②のように結果が覆る状況を示したり、③のように頑健性を示すことでバイアスの量とその結果への影響を見える化することは研究結果の客観的な評価に役立つであろう。また、バイアスの細かな発生過程の特定や検討ではなく「そのような大きさのバイアスの可能性」に論点を集約できる。このような場合、結果が benefit の存在を示している場合は結果を過大にするバイアス、risk が小さい（ない）ことを示している場合は結果を過小にするバイアスを用いる、つまり「結果の言い過ぎ」を評価する目的とするのが適切であろう。

また、これらの分類は感度分析の方法の選択とも関係する。③は解析者が「極端である」と考えるパラメータを用いた結果を示せば良いため、1組～数組の固定したパラメータを用いて Simple sensitivity analysis (SSA) 又は Multidimensional analysis (MA) を行うことで目的は達せられるが、①は利用可能なバイアスに関する情報に応じて1組～数組の固定したパラメータを用いて SSA 又は MA を行うことも、分布を用いて Probabilistic sensitivity analysis (PSA) を行うこともある。何れを用いるかはバイアスの情報の確からしさや感度分析の必要性などを考慮して選ばれるであろう。対して②はパラメータを一意に特定できないため、分布を用いる PSA の結果、あるいは多くのパラメータの組み合わせによる MA の結果から、条件を満たすパラメータの分布や組み合わせを示すことになる。

### 5.3. 研究デザイン、データの収集方法、得られている情報を整理する

研究者以外の第三者が感度分析を行う場合は必須の作業であるが、研究者自身が行う場合でも本書で扱う感度分析は未知を含む測定されていない、つまりデータからは分からないバイアス<sup>注1</sup>の補正を試みることであるため、その前に分かっていること、つまり研究デザインにおける母集団との関係（抽出）、選択/除外基準、曝露、アウトカムの定義、Time Window の設定、データの収集、コーディング・変換、測定されている要因、主たる解析で調整解析を行っている場合は調整に用いた変数の因果効果（関係）を整理して確認することは有用であろう。多くは通常の研究の計画、実施、集計解析、結果の報告（論文）で作成されているが、文章のみで説明されている場合は図表に整理すると見通しが良くなる。更に、感度分析の対象となり得る解析に用いた仮定、曝露、アウトカムの測定結果と真の値の関係、測定できなかった交絡要因を加える。整理には DAG <sup>注29</sup>を用いて視覚的に行うことが推奨されている。複雑な連鎖がある場合などは DAG による整理が必要かもしれないが、本書で扱う集計結果（四分表）に対して行う感度分析においては要因の複雑な連鎖は扱わないため、簡便に図表として整理することで十分と考える。

注29： 有向非巡回グラフ Directed Acyclic Graph. 第1部 参考 3-6 参照

## ✓ 医療情報データベースを二次利用する研究

医療情報データベースを利用する場合はデータの質と曝露・アウトカムの定義が問題となる。本来、研究は目的に合わせて集団を特定し、必要なデータを収集する。そのデータには一定の標準化、計画的な測定、欠測の防止を含む質の管理が行われ、研究の妥当性を形作られる。しかし、医療情報データベースを利用する場合、そのデータは本来の目的に合わせて発生/収集されており、例外を除いて、研究に必要な質の管理、曝露・アウトカムの測定は考慮されていない。これらの問題はデータ欠損（包括、傷病名の件数の制限などもある）、データ粒度（例えば標準病名マスタは約2万語あるのに対してDPCで用いるICD-10の4桁コードは約1万語）、誤分類、脱落、無症候や未確定の診断、評価・成績情報が含まれないなどいくつかの原因がある<sup>27)</sup>。更にこれらは例えば曝露などによって研究のアウトカムとして扱う疾患の診断につながる検査の実施率が異なる場合もある（5.4.3.3 参照）。保険償還のためのデータ（レセプトデータ）であれば、保険償還という目的との相互作用を生じ、償還に必要な範囲、情報の粒度で記録され、不要な情報、例えば医療資源を要しない合併症や有害事象は意図的に記録されないという強いバイアス（偽陰性）を生じさせる<sup>注30</sup>。このような系統誤差は曝露、アウトカム、交絡要因の系統的な誤分類（例えば通常は傷病名が付与されない場合でも高額な医薬品を使用する場合には選択的に傷病名が付与される）となり、研究に深刻な影響を生じ得る。こういったことへの対処として、データがどのように発生、記録されているのか、どのようにコード化されているのかを確認することは可能性のあるバイアスを確認する第1歩である。

例えば、Figure 5-1 はヘルスケアデータベースを疫学研究に用いるレビュー論文<sup>28)27)</sup>で用いられた、情報がデータベースに取り込まれる過程とバイアスを図にまとめたものでミスコーディングや請求漏れなどをバイアスの要因として整理されている。データを二次利用する研究では曝露やアウトカムの定義による誤分類のみでなく、データの発生、登録過程にもバイアスの原因が潜んでいる。このようなデータベース化のプロセスは研究者自身がかかわらない部分であり見落とされやすい。用いるデータベースが提供しているフロー図などを利用してこのように整理することは難しくなく、バイアスの検討に役立つであろう。

この図を日本の場合に置き換えて病名について「Diagnosis」から「Filing and adjudication of final claims」の間の具体的な説明を追加すると、一般的に、診断の内容は「診察記事（経過記録）」に文章として記録され、必要に応じて傷病名の項目に傷病名のリストまたはテキスト入力により記録され、検査を行うために付与した場合は「疑い」の修飾語が付与される。この「傷病名の項目」は施設のシステムによって様々で、1つの電子カルテに診療録、保険請求（診療報酬明細書）用、DPC用の3種類がある場合や、各々のシステムが分かれている場合、電子カルテの統合された項目からシステムが各々の目的で振り分ける場合など、様々である。「Coding of claims」は電子カルテを用いている場合でも標準病名マスタへの変換、検査や処置に対応する病名がない場合の追加、DPCレセプトは一旦DPCデータ（様式1）を経由するため、傷病名の丸め（標準病名マスタからICD-10に）、病名数の切り捨て（入院後発症傷

注30： データベースを用いる多くの場合で曝露、アウトカム、併存疾患などが「ない」ことはデータが記録されていないとして定義される。このため、「〇〇がない」にはデータの欠損が多く含まれることとなる。

病名は最大 4 件まで) など加工が行われる (Table 5-1 参照)<sup>注31</sup>。紙のカルテの場合は医療事務の職員などが読み取ってコーディング・入力が行われる。この際、保険請求上必要なレベルで病名の丸め (例えば「二型糖尿病」が単に「糖尿病」とされる) が行われることが少なくない。病名の変更、追加、削除があれば「Filing of complete claims」の前に主治医に確認が行われる。さらに、支払い審査機関から疑義照会があれば必要な修正が行われる。このように、データベースに情報が格納されるまでに様々なプロセスがあり、電子カルテから直接データを取り出す場合でも欠損を含め誤りの可能性に注意が必要である。

Figure 5-1 health care utilization databases のデータ発生プロセスとバイアスの図示の例<sup>27)</sup>

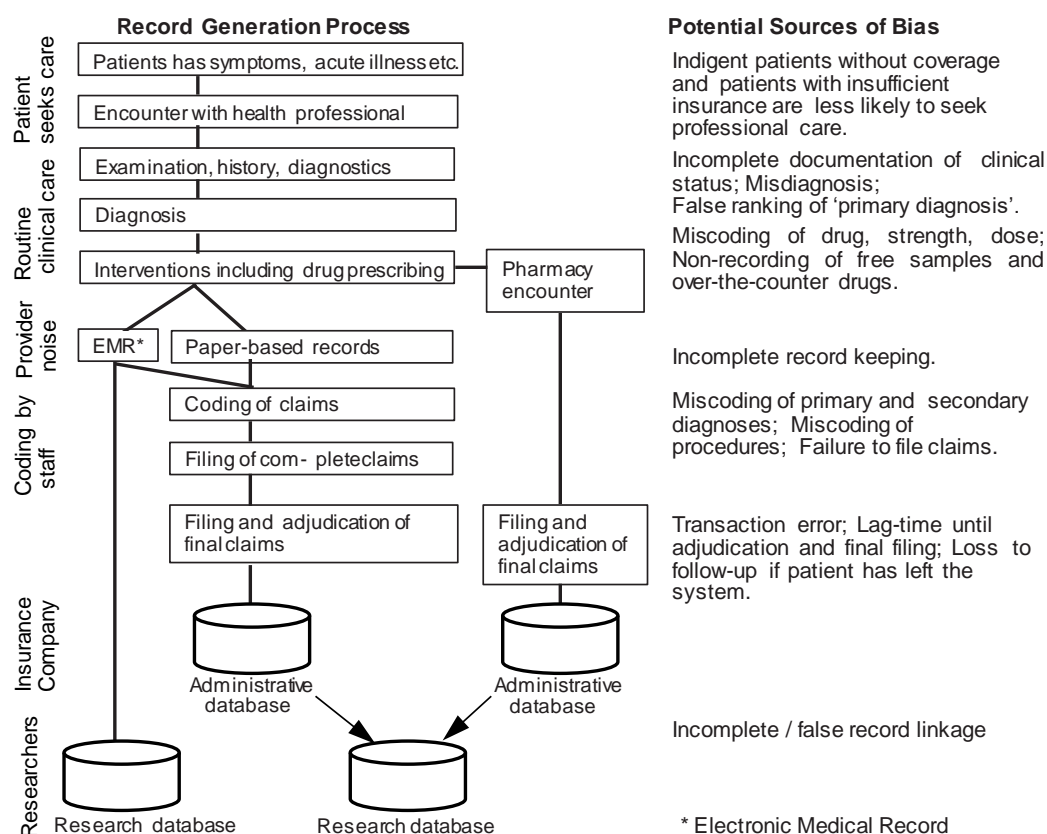


Fig. 2. The generation of health care utilization databases and potential sources of errors/bias.

Figure 5-2 は同じ論文の曝露の Time window の図であるが、真の曝露が調剤請求データに反映される様子と 15 日以内の処方切れ目を埋めて曝露として扱われていることなどが整理されている。曝露期間の扱いはアウトカムとの時間関係を特定する基本である。長期の曝露によるがんのリスクの評価など、正確な曝露 (期間) の特定が不要な場合もあるが、短期間の曝露あるいは曝露中に生じるアウトカムを扱う場合などはこのような図示により、研究で曝露中のアウトカムとして扱うものの中に非曝露のものが混在する可能性を判断できる。Figure 5-3 は

注31 2014 年の電子カルテ普及率は一般病院で 34%, 診療所で 35%である (「未来投資会議構造改革徹底推進会合 第 1 回「健康・医療・介護」会合」資料  
<https://www.kantei.go.jp/jp/singi/keizaisaisei/miraitoshikaigi/suishinkaigo2018/health/dai1/siryou2.pdf> )



社会不安障害（SAD）がその後のうつ病に及ぼす影響を評価した初期発達段階精神医学研究（EDSP）について感度分析を行った報告で作成された図である<sup>31)</sup>。対象の抽出、分類のフロー図にバイアスが考えられる箇所が書き込まれている。

Figure 5-2 曝露の time window の図示の例<sup>27)</sup>

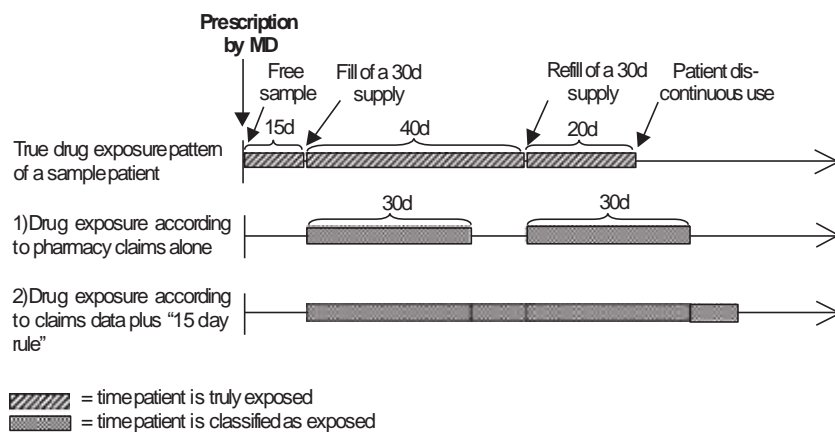


Fig. 3. Typical causes for drug exposure misclassification in longitudinal claims database studies.

Figure 5-3 フローチャートによる整理の例

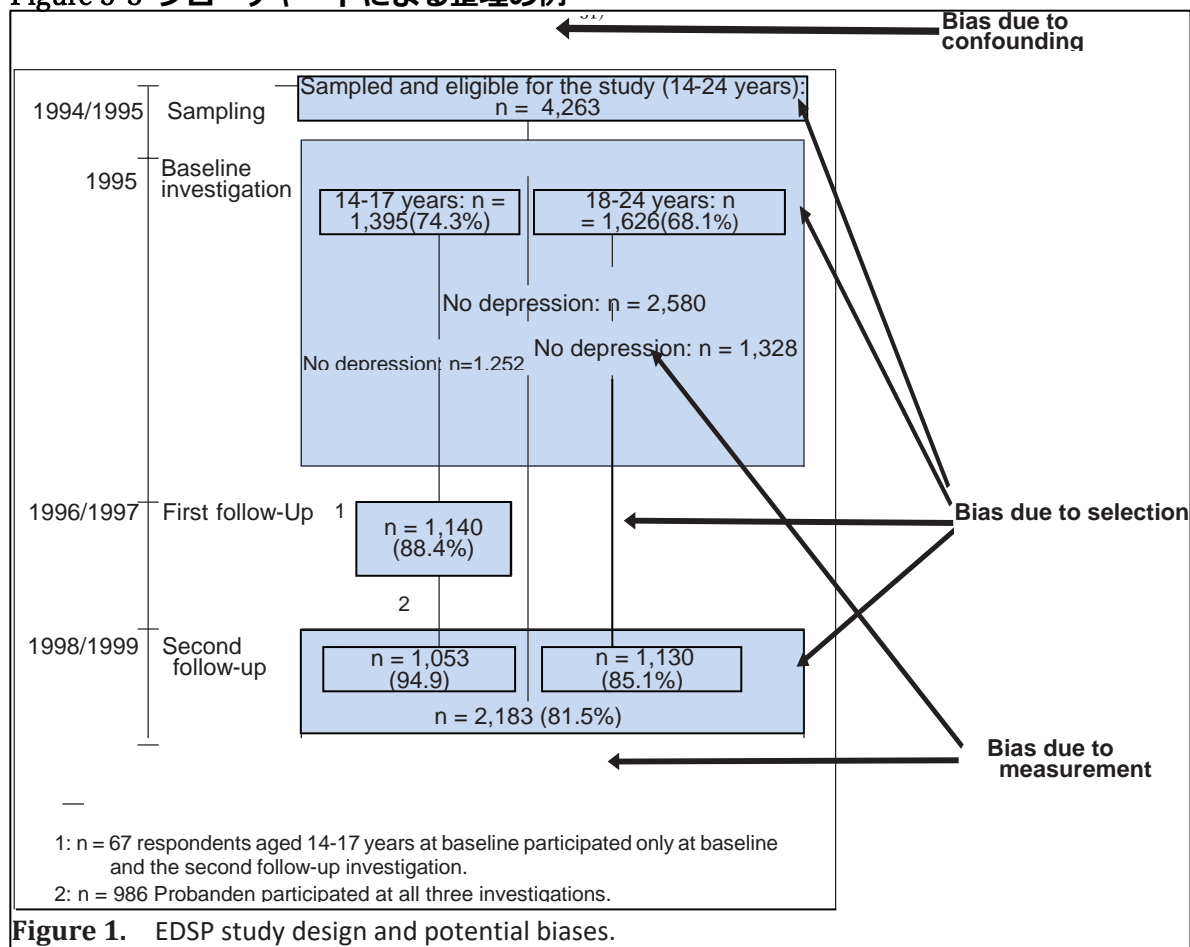


Figure 1. EDSP study design and potential biases.



これらからわかるように、例えばアウトカムの誤分類といっても複数の原因、結果の現れ方が混在（選択に影響する場合と測定に影響する場合）していることがわかり、複雑に見えるため感度分析は困難に感じるかもしれない。しかし、例えばベースラインで過去のアウトカムの記録漏れで本来、除外されるべき対象者が選択された場合、一見アウトカムの誤分類に思われるが、第1部 3.7.5「選択前の集団でのアウトカムの誤分類」で説明しているように選択バイアスであるので、アウトカムの誤分類としては扱わない。観察期間中のアウトカムの誤分類が、ある患者は診断もれ、他の患者はコーディング漏れ、ある患者は転記漏れなど、異なる原因であるかもしれないが、全体を総合してアウトカムの誤分類の感度、特異度（又は陽性的中度、陰性的中度<sup>注32</sup>）の2つの値（パラメータ）で表現できる。複数の未測定交絡要因がある場合でも、感度分析では個々の交絡要因の評価は目的では無いため「包含した交絡要因」として一まとめに扱うことが可能である。だからと言ってこのような詳細な整理、分析が不要なのではない。この情報はこれらのパラメータを見積もること、行った感度分析の説明に重要である。また、測定され、主要な解析で調整された交絡要因個々についての吟味は必要ではないが、5.6.3で紹介するように測定された要因の影響の大きさは未測定交絡要因の感度分析のパラメータを考える基準の一つとして意味がある。<sup>注33</sup>

Table 5-1 DPC における傷病名の扱い

図表 5：DPC レセプトの作成に必要な傷病名の一覧

記載欄	留意事項
①「傷病名」欄	・ 医療資源を最も投入した傷病名を選択する。 ・ 入院中の主要な傷病名・病態に基づき決定する。
②「副傷病名」欄	（診断群分類点数表に定義されている副傷病名がある場合は記載する。）
③「傷病情報」欄	
「主傷病名」	医師が医学的判断に基づき決定した傷病名を記載する。（医療資源の投入量の多寡によらず、医師の判断で決定する）
「入院の契機となった傷病名」	今回入院し治療する必要があると判断する根拠となった傷病名を1つ記載する。
「医療資源を2番目に投入した傷病名」	医療資源を2番目に投入した傷病名を1つ記載する。
「入院時併存傷病名」（最大4つ）	・ 診断群分類の決定に影響を与えない場合であっても、診療上、重要な傷病名は記載する必要がある。 ・ 入院時に併存している傷病名について、重要なものから最大4つまで記載する。
「入院後発症傷病名」（最大4つ）	・ 診断群分類の決定に影響を与えない場合であっても、診療上、重要な傷病名は記載する必要がある。 ・ 入院後に発症した傷病名について、重要なものから最大4つまで記載する。

図表 6：DPC データの様式1の作成に必要な傷病名の一覧

調査項目	留意事項
「主傷病名」	退院時サマリーの主傷病欄に記入された傷病名を入力する。
「入院の契機となった傷病名」	入院の契機となった傷病名を入力する。

注32：陽性的中度、陰性的中度とは、測定結果が陽性又は陰性のうち、正しい場合の割合。5.4.3.3 参照

注33：多くの共変量が観察されている場合、主解析において観察されたデータから機械的にモデルを作成する変数選択法を用いられていることがある。この方法は相関に基づく手法であり、因果の向きは無視される、疑似相関が生じる、中間変数が除かれないなど問題が多く、因果効果の推定では用いるべきでないことが指摘されている<sup>4),32)</sup>。データの得られていない要因、バイアスを扱い、単にデータの相関に基づくのではなく背景にある機序、意味を想定する感度分析では、このような機械的な結果の吟味も必要である。

「医療資源を最も投入した傷病名」	入院期間中、複数の病態が存在する場合は医療資源を最も投入した傷病名で、請求した手術等の診療行為と一致する傷病名を入力する。
「医療資源を2番目に投入した傷病名」	医療資源を2番目に投入した傷病名は、「入院時併存症名」もしくは「入院後発症疾患名」のいずれかに必ず入力する。
「入院時併存症名」(最大10)	<ul style="list-style-type: none"> <li>・医療資源の投入量に影響を及ぼしたと判断される入院時併存症がある場合には必ず入力する。</li> <li>・以下に該当するものがある場合は入力すること。 <ol style="list-style-type: none"> <li>1. 診断群分類点数表に定義された副傷病名</li> <li>2. 慢性腎不全</li> <li>3. 血友病・HIV感染症</li> <li>4. 併存精神疾患</li> </ol> </li> </ul>
「入院後発症疾患名」(最大10)	<ul style="list-style-type: none"> <li>・医療資源の投入量に影響を及ぼしたと判断される入院後発症疾患がある場合には必ず入力する。</li> <li>・以下に該当するものがある場合は入力すること。 <ol style="list-style-type: none"> <li>1. 診断群分類点数表に定義された副傷病名</li> <li>2. 術後合併症</li> </ol> </li> </ul>

### Ⅲ. コーディングの基本的な考え方

#### 1. 診療録の記載および診療報酬の請求における傷病名の選択

##### ○重要なポイント

診療報酬の請求は診療録の記載に基づいて行われる必要があり、DPCの決定の際にも、診療録の記載に基づき適切に行わなければならない。

○ 医師法第24条において、「医師は、診療をしたときは、遅滞なく診療に関する事項を診療録に記載しなければならない。」と規定されており、その記載事項については医師法施行規則第23条に規定されている。

○ また、療養担当規則第8条（診療録の記載及び整備）及び第22条（診療録の記載）に診療録に係る規定があり、診療録の記載は診療報酬請求の根拠となるものであるため、レセプトに記載された事項は、診療録に記載されていなければならない。

#### 2. コーディングの基本と傷病名選択の定義

##### ○重要なポイント

- ・DPCコーディングの基本は医療資源に基づく「医療資源病名」の選択にある。
- ・対象となる期間は、DPC算定病床に入院していた期間である。

（2）最も医療資源を投入した傷病名は、1入院期間を対象として退院時に1つを決定する

（4）原則として医療資源病名と実施した手術、処置には乖離がないこと

（DPC/PDPS 傷病名コーディングテキスト <https://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000202577.pdf>）

## 5.4. 感度分析すべきバイアスを選択する

感度分析を行う場合は対象とするバイアス（要因）を決めることになる。5.3で「分かっていること」を整理しながら、バイアスの生じそうな箇所を確認した。観察研究の殆どの場合で何らかのバイアスの影響を受けているが、研究実施者自身が感度分析を行う場合は、研究の計画段階及び報告書の作成において考慮すべきバイアスは十分に認識されているはずである。しかし、第三者が感度分析を行う場合は背景情報を収集、確認しながら、論文などの報告書から読み取る必要がある。加えて、得られていない情報を対象とする感度分析では見落としを生じやすいため、全体を眺めて順を追って検討する。教科書的には専門家の参加やDAGの利用などを推奨されているが（参考2）、本書が主な対象として扱う結果を四分表による要約が可能な研究については、簡便な図表による整理を行い、疑わしいバイアスは簡便に感度分析を行えるExcelのToolを用いてSimple sensitivity analysisを行い、凡その影響度を確認することは効

率が良いと考える。簡便な図表による整理は、例えば観察研究で通常作成される Figure 5-1, Figure 5-3 のようにデータ収集、標本の抽出（Enrolment of patients）のフローチャートを用いて、組み入れ/除外の基準、曝露、アウトカムの判別、打ち切りなどについてバイアスが生じる過程を確認するが、前項 5.3 で丁寧な整理を行ってれば、バイアスの候補は具体的に見えているであろう。それでも、例えばアウトカムの測定に関する原因としては

- ・ 発現時期の誤り（ベースラインですでに生じていたものを曝露後に測定など）、
- ・ 打ち切り、
- ・ 測定漏れ、誤った測定、
- ・ 記録・コーディングのミス、除外又は対象外
- ・ 識別定義（アルゴリズム）の精度

など複数が考えられるが、これらをバイアスに分類すると、

- ・ 研究対象者の抽出に影響する場合（1 番目と 2 番目）は選択バイアスモデル
- ・ 研究の対象とすべきアウトカムの有無が誤る場合はアウトカムの誤分類

であるなど、用いるモデルが異なる。また、最後の 1 つはデータの扱いによるものであるため測定されたデータから情報が得られるが、それ以外はデータの品質の問題であるため測定されたデータからは精度に関する情報は得られないという違いがあるが、3 番目から 5 番目は 1 つの誤分類モデルにより扱うことができる。

感度分析は 1 つのモデルのみを扱う単純な方法で十分な場合もあれば複数のモデルを用いた確率的で定量的な分析が必要な場合もあるため、全体を俯瞰して扱うバイアスを取捨選択する必要がある。しかし、具体的にバイアスを生じている個所が特定されても、その影響度、重要性に見当をつけることが容易ではない場合、簡便に感度分析を行える Simple sensitivity analysis によりバイアスの影響度を確認することは重要なバイアスの候補を検討するのに有効であろう。

## 参考 2 教科書的なバイアスの特定

第三者が感度分析を行う場合、一人で行えることもあるが、Cochrane の観察研究におけるバイアス評価 tool ROBINS-I の論文<sup>33)</sup>では対象分野の臨床の専門家と疫学の専門家の参加が重要であるとしている。また、直接測定されない、目に見えない要因について考えるとき、図表による視覚化が有効である。国際医薬経済・アウトカム研究学会（The International Society for Pharmacoeconomics and Outcomes Research : ISPOR）はデータの二次利用による非ランダム化有効性研究のバイアスを軽減させるガイダンス<sup>34)</sup>で主解析の前に因果モデルを有向非巡回グラフ [Directed Acyclic Graph : DAG]<sup>注34)</sup>により明確にすること、異なる交絡について感度分析を実施する場合は、それぞれの感度分析の前提を表す追加の DAG を報告することを推奨しており、ENCePP の観察研究の技術的ガイド<sup>35)</sup>でも潜在的変数を含めた因果モデ

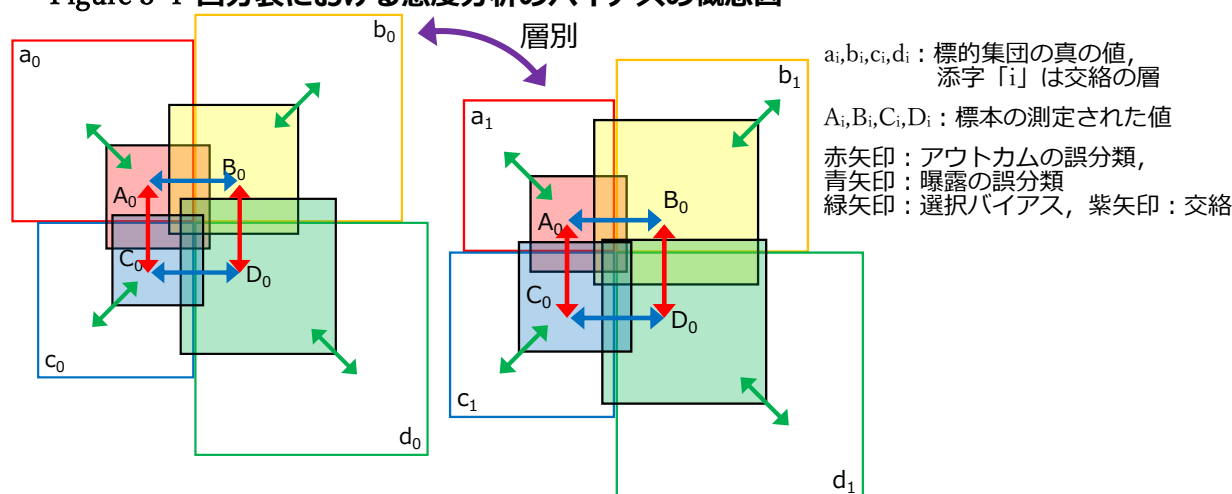
注34 : DAG については第 1 部 参考 3-6 参照

ルの検討には因果関係を構成する原因と結果にバイアスの候補を加えた DAG の作成が有効であるとされている。感度分析では、測定可能な変数を用いる必要はないため、通常は測定できない潜在的変数も扱う事が可能で効率的な場合がある。

#### 5.4.1. バイアスの原因を整理して分類する

前の「5.3 研究デザイン、データの収集方法、得られている情報を整理する」でバイアスが生じそうな箇所を整理した結果、観察研究におけるバイアスは選択バイアス、交絡、情報バイアスの3種類に概ね分類されるが、その中に様々な細かな分類、名称（65個に分類している文献もある<sup>36)</sup>）があるように、実際の研究で生じているであろうバイアスを正しく整理するのは簡単ではない場合がある。しかし、Lash T.L. (2014)<sup>4)</sup>は「考慮すべきバイアスの潜在的な原因」として調整不能な交絡、誤った調整、選択バイアス、測定エラーに分けて説明しているように、四分表に対する感度分析は Figure 5-4 のように、種々のバイアスがあっても、結果的に「そのようなバイアスが生じたとして、四分表のセルはどのように変わるだろうか」を考え、バイアスの補正を測定されたセル間の移動（誤分類モデル）、標的集団からの抽出率の補正（選択バイアスモデル）、異なる層の分割（交絡モデル）の何れかに落とし込み、補正した四分表を得ることになるため、感度分析においてはバイアスの細かな分類、名称の特定に労力を割く必要はないであろう。問題となるバイアスは個々の研究で異なるが、既に報告されている交絡要因、曝露・アウトカムなどの測定精度を念頭に置きながら、Figure 5-1, Figure 5-2, Figure 5-3 のようにバイアスを生じる過程から検討する。Figure 5-1, Figure 5-2 は米国の医療情報データベースを用いる場合には一般的なものではあるが<sup>注35</sup>、保険制度や医療環境、データベースの構築方法による違いが生じるので用いるデータベースについて確認が必要である。それでも、バイアスを感度分析の補正式に正しく分類することが難しい場合は第1部第1章や、総説として Rothman K.J. (2008)<sup>2)</sup>の CHAPTER9, 12, 景山 茂 (2016)<sup>37)</sup>の第5章-3~5等を参照いただきたい。

Figure 5-4 四分表における感度分析のバイアスの概念図



注35： Figure 5-1 の医療用医薬品のサンプル使用は日本ではほとんど考えなくていいであろう。



#### 5.4.2. 感度分析のモデルに使用するバイアスを特定する（バイアスの候補が複数ある場合）

最終的に感度分析の対象とするバイアスを選ぶステップである。繰り返しになるが、研究の様々なプロセスでいくつものバイアスの原因が見つかる場合でも、本書の対象である四分表に要約できる研究では、研究の各段階で確認された様々なバイアスの原因は未測定交絡要因、曝露の誤分類、アウトカムの誤分類、選択バイアスといった感度分析で大別されるモデルに分類、集約されるため、個々の細かな原因ごとに感度分析の対象とする必要はない。この利点として、モデル内の主な要因などからそのモデルのパラメータを見積もり、モデル全体で結果に与え得る影響のマグネチュードが大きなものを特定することで、細かな個々の原因ごとの検討を少なくとも減らすことができる。しかし、感度分析の対象とするバイアスについてはモデルに関する量（併存率、オッズ比、感度、特異度など）の情報を集めるために細かな原因レベルで取舍選択も必要となる場合もある。また、“バイアス”は比較する層間で何らかの偏り（差異）が生じていることを連想させるが、感度分析で扱う曝露やアウトカムの誤分類は偏りがなくても（各層の感度、特異度が等しくても）観察されるオッズ比やリスク比が過少になる場合があるため、感度分析の対象となる場合がある<sup>注36</sup>。加えて、実際に計算してみると、ほとんどの観察研究で課題とされる未測定交絡要因が結果に大きな影響を及ぼす条件は限定的であるのに対して、あまり課題として扱われない誤分類、選択バイアスは容易に結果に大きな影響を与え得るため、感度分析が必要な場合が少なくない。

では、その研究ではどのバイアス分類の影響が大きいのであろうか？ 対象が四分表にまとめられた結果であれば“百聞は一見に如かず”のことわざのとおり、パラメータに任意の固定値を用いるシンプルな感度分析を行える Excel 等の Tool を用いて適当な（容易に想定される）パラメータを用いた結果を確認するのが感度分析の要否を判断する最も簡便で確実な方法であると考える。

とは言え、そのようなパラメータの選択にはパラメータの作用のイメージが必要かもしれない。バイアスのモデルごとにどのような状況が研究の結果への影響が大きくなり得る（感度分析による補正結果と研究結果の乖離が大きくなる）かのイメージをつかんでいただくため、主なモデルについて、どういう場合に影響が大きく、あるいはその影響はどの方向（例えばオッズ比が小さくなる/大きくなる）かを説明する。

なお、各バイアスの一般論的メカニズムは第 1 部第 1 章に紹介しており、バイアスを補正するモデルの計算は第 1 部第 3 章で説明している。

---

注36 第 1 部 3.2.2 参照



### 5.4.3. 感度分析のモデルに使用するバイアスを特定する（バイアスの候補が複数ある場合）

#### 5.4.3.1. 未測定交絡要因モデルの場合

研究の結果に影響を与える曝露，アウトカムの両方に関連する要因であるが，測定できていない又は未知の要因である．研究対象のアウトカムを多く生じさせることが報告されている要因（肺癌における喫煙など）が未測定であっても，比較する群間でバランスが取れている又は，必ずしもバランスは取れていないかもしれないが，その要因の保有率が両群とも極めて小さいと推定される場合など結果への実質的影響は小さくなるため感度分析の対象とはならない．そのようなリスク要因が大きな割合で保有され，曝露の有無間で偏りが生じていると考えられる場合は感度分析の対象となり得る．具体的には，曝露群に多く偏っている場合はリスク比やオッズ比を大きくする方向に働き，対照群に多く偏っている場合は小さくなるほうに働く．感度分析に用いる補正前の四分表のオッズ比を  $OR_{obs}$ ，補正後のオッズ比を  $OR_{adj}$ 、未測定交絡要因（C）とアウトカム（D）のオッズ比を  $OR_{DC}$ ，曝露なしの C の保有率を  $p(C+|E-)=P_{C0}$ ，曝露ありの場合を  $p(C+|E+)=P_{C1}$  と表すと，その比は

$$\frac{OR_{adj}}{OR_{obs}} = \frac{(OR_{DC} - 1)P_{C0} + 1}{(OR_{DC} - 1)P_{C1} + 1} \quad \text{数式 5-1}$$

により表せる（第 1 部 3.1.6, 参考 3-3 参照）．通常  $OR_{DC}$  は 1 より大きいので  $P_{C0} > P_{C1}$  の場合に補正後のオッズ比は大きく， $P_{C0} < P_{C1}$  の場合に小さくなることがわかる．つまり，

- ・  $OR_{adj} > OR_{obs}$  となるバイアスを想定する場合は  $P_{C0} > P_{C1}$
- ・  $OR_{adj} < OR_{obs}$  となるバイアスを想定する場合は  $P_{C0} < P_{C1}$

となるパラメータを用いる．さらにこの式から  $(OR_{DC} - 1)P_{C0} = g$ ， $P_{C1}/P_{C0} = q$  とすると

$$\frac{OR_{adj}}{OR_{obs}} = \frac{g + 1}{gq + 1} = \frac{g + 1}{\left(g + \frac{1}{q}\right)q} \quad \text{数式 5-2}$$

$q > 1$  の場合，

$$1 > \frac{g + 1}{gq + 1}, \quad \frac{g + 1}{\left(g + \frac{1}{q}\right)} > 1 \quad \text{であるので} \quad 1 > \frac{OR_{adj}}{OR_{obs}} > \frac{1}{q}$$

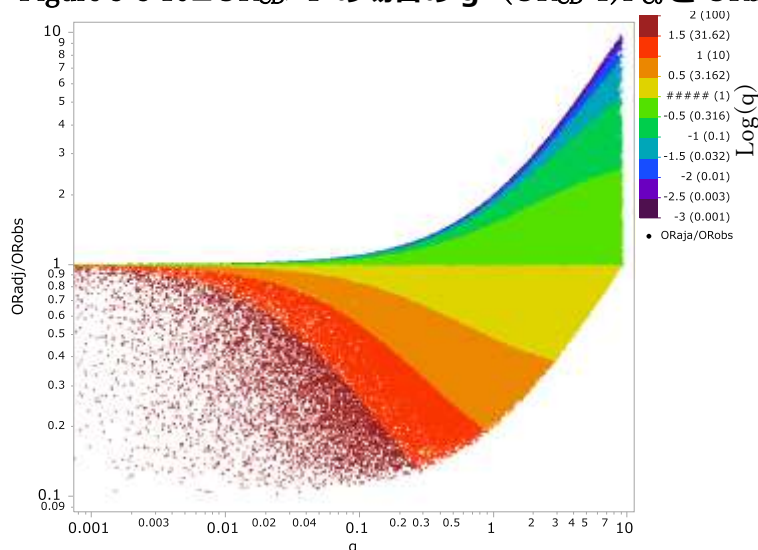
同様に  $0 < q < 1$  の場合

$$\frac{g + 1}{gq + 1} > 1, \quad 1 > \frac{g + 1}{\left(g + \frac{1}{q}\right)} > 0 \quad \text{であるので} \quad \frac{1}{q} > \frac{OR_{adj}}{OR_{obs}} > 1$$

となる．これから， $OR_{DC}$  の大きさに関わらず， $q > 1$  の場合の  $OR_{adj}$  の最大は  $OR_{obs} \cdot q$ ， $q < 1$  の場合  $OR_{adj}$  の最小は  $OR_{obs} \cdot q$  であることが確認できるので， $P_{C0}$  と  $P_{C1}$  の比を想定することで感度分析が不要な場合の判断が可能である．加えて，例えば，研究結果がリスクを評価するもので，オッズ比が小さい（例えば 1.2 程度）場合，曝露群に多く偏っている（ $P_{C1} > P_{C0}$ ）と考えられる要因は感度分析の対象としないが，対照群に偏っている場合は過小評価が考えられるため対象とするといった判断も行える．Figure 5-5 に数式 5-2 を用いて， $10 \geq$

$OR_{DC} > 1$ ,  $P_{C0}$ ,  $P_{C1}$  とともに 0~1 の一様分布によりシミュレーションした結果の  $g$  と  $OR_{Adj}/OR_{Obs}$  の関係を示す。プロットは  $q$  の値により色分けをしている。プロットの色が混在していないことから  $g = (OR_{DC} - 1)P_{C0}$ ,  $q = P_{C1}/P_{C0}$  を仮定することでバイアスの影響度が評価可能なことが確認できる。

Figure 5-5  $10 \geq OR_{CD} > 1$  の場合の  $g = (OR_{CD} - 1)P_{C0}$  と  $OR_{Adj}/OR_{Obs}$



#### 5.4.3.2. 曝露の誤分類モデルの場合

誤分類は有無の二値変数であれば最終的にあり、なしに集約されるが、後で説明するアウトカムが時間軸の中で“点”である場合が多いのに対して、曝露は多くの場合、期間という“線”として扱われ、加えて期間の長さや曝露量により 3 区分以上に分類されるため、複雑になりやすい。ここでは本書で対象としている二値区分に絞って扱う。また原理は同じであるのでアウトカムの誤分類も含めて説明するので適宜読み替えていただきたい。

曝露とアウトカムの四分表 Table 5-2 において、曝露又はアウトカムの誤分類はこの四分表の各セルの値が誤っており、感度分析ではこれを正しいと思われる値に補正する操作が行われ、曝露の誤分類の場合は Table 5-2 の青の矢印、アウトカムの誤分類は赤の矢印間で値が移動する<sup>注37</sup>。Table 5-2 の記号を用いると、オッズ比は  $A_1 \times B_0 \div (A_0 \times B_1)$  であるので、対角線上のセルの組 ( $A_1$ ,  $B_0$  又は  $A_0$ ,  $B_1$ ) の一方が大きく、他方が小さく、またはその反対になる場合に、オッズ比への影響が大きくなる。また、誤分類はセルの値の大きさに比例して生じるため、Table 5-2 の矢印間で値の大きいほうで誤りが生じる場合に他方の値が大きくなる。一般に case-control 研究では“曝露あり”より“なし”が多いため曝露の特異度 (“なし”を正しくなしと測定できる割合) が、また、cohort 研究では“アウトカムあり”より“なし”が多いため、アウトカムの特異度の影響が大きくなりやすい。

注37：四分表内で入れ替わるのではなく、表の外と入れ替わる場合（つまり表の対象の抽出に誤りがある）は誤分類であるが選択バイアスとして扱う。

Table 5-2 曝露有無，アウトカム有無の四分表

	曝露+ (E=1)	曝露- (E=0)	計
アウトカム+	A <sub>1</sub> = 200	A <sub>0</sub> = 300	M <sub>A</sub> = A <sub>1</sub> + A <sub>0</sub> 500
アウトカム-	B <sub>1</sub> =250	B <sub>0</sub> = 750	M <sub>B</sub> = B <sub>1</sub> + B <sub>0</sub> 1000
計	N <sub>1</sub> = A <sub>1</sub> + B <sub>1</sub> 450	N <sub>0</sub> = A <sub>0</sub> + B <sub>0</sub> 1050	N <sub>-</sub> = 1500

曝露の誤分類を感度 (Se<sub>A</sub>, Se<sub>B</sub>)，特異度 (Sp<sub>A</sub>, Sp<sub>B</sub>) と Table 5-2 のセルの記号を用いて誤分類のある観測データと誤分類を補正した場合を整理すると Table 5-3 となる。

Table 5-3 曝露の誤分類の数式整理

		E=1	E=0	Total
Crude data	D*=1	A <sub>1</sub> *=Se <sub>a</sub> A <sub>1</sub> +Fp <sub>a</sub> A <sub>0</sub>	A <sub>0</sub> *=Sp <sub>a</sub> A <sub>0</sub> +Fn <sub>a</sub> A <sub>1</sub>	M <sub>1</sub>
	D*=0	B <sub>1</sub> *=Se <sub>b</sub> B <sub>1</sub> +Fp <sub>b</sub> B <sub>0</sub>	B <sub>0</sub> *=Sp <sub>b</sub> B <sub>0</sub> +Fn <sub>b</sub> B <sub>1</sub>	M <sub>0</sub>
	Total*	N <sub>1</sub> *=A <sub>1</sub> *+B <sub>1</sub> *	N <sub>0</sub> *=A <sub>0</sub> *+B <sub>0</sub> *	
誤分類のなす data	D=1	A <sub>1</sub> =[Sp <sub>a</sub> A <sub>1</sub> *-(1-Sp <sub>a</sub> )A <sub>0</sub> *]/(Se <sub>a</sub> +Sp <sub>a</sub> -1) =[Sp <sub>a</sub> (A <sub>1</sub> *+A <sub>0</sub> *)-A <sub>0</sub> *]/(Se <sub>a</sub> +Sp <sub>a</sub> -1) =[A <sub>1</sub> *-(1-Sp <sub>a</sub> )M <sub>1</sub> ]/(Se <sub>a</sub> +Sp <sub>a</sub> -1)	A <sub>0</sub> =[Se <sub>a</sub> A <sub>0</sub> *-(1-Se <sub>a</sub> )A <sub>1</sub> *]/(Se <sub>a</sub> +Sp <sub>a</sub> -1) =[Se <sub>a</sub> (A <sub>0</sub> *+A <sub>1</sub> *)-A <sub>1</sub> *]/(Se <sub>a</sub> +Sp <sub>a</sub> -1) =[A <sub>0</sub> *-(1-Se <sub>a</sub> )M <sub>1</sub> ]/(Se <sub>a</sub> +Sp <sub>a</sub> -1)	M <sub>1</sub>
	D=0	B <sub>1</sub> =[Sp <sub>b</sub> B <sub>1</sub> *-(1-Sp <sub>b</sub> )B <sub>0</sub> *]/(Se <sub>b</sub> +Sp <sub>b</sub> -1) =[Sp <sub>b</sub> (B <sub>1</sub> *+B <sub>0</sub> *)-B <sub>0</sub> *]/(Se <sub>b</sub> +Sp <sub>b</sub> -1) =[B <sub>1</sub> *-(1-Sp <sub>b</sub> )M <sub>0</sub> ]/(Se <sub>b</sub> +Sp <sub>b</sub> -1)	B <sub>0</sub> =[Se <sub>b</sub> B <sub>0</sub> *-(1-Se <sub>b</sub> )B <sub>1</sub> *]/(Se <sub>b</sub> +Sp <sub>b</sub> -1) =[Se <sub>b</sub> (B <sub>0</sub> *+B <sub>1</sub> *)-B <sub>1</sub> *]/(Se <sub>b</sub> +Sp <sub>b</sub> -1) =[B <sub>0</sub> *-(1-Se <sub>b</sub> )M <sub>0</sub> ]/(Se <sub>b</sub> +Sp <sub>b</sub> -1)	M <sub>0</sub>
	Total	N <sub>1</sub> =A <sub>1</sub> +B <sub>1</sub>	N <sub>0</sub> =A <sub>0</sub> +B <sub>0</sub>	

ここから，感度分析に用いる補正前のオッズ比 OR<sub>obs</sub> と補正後のオッズ比 OR<sub>adj</sub> は

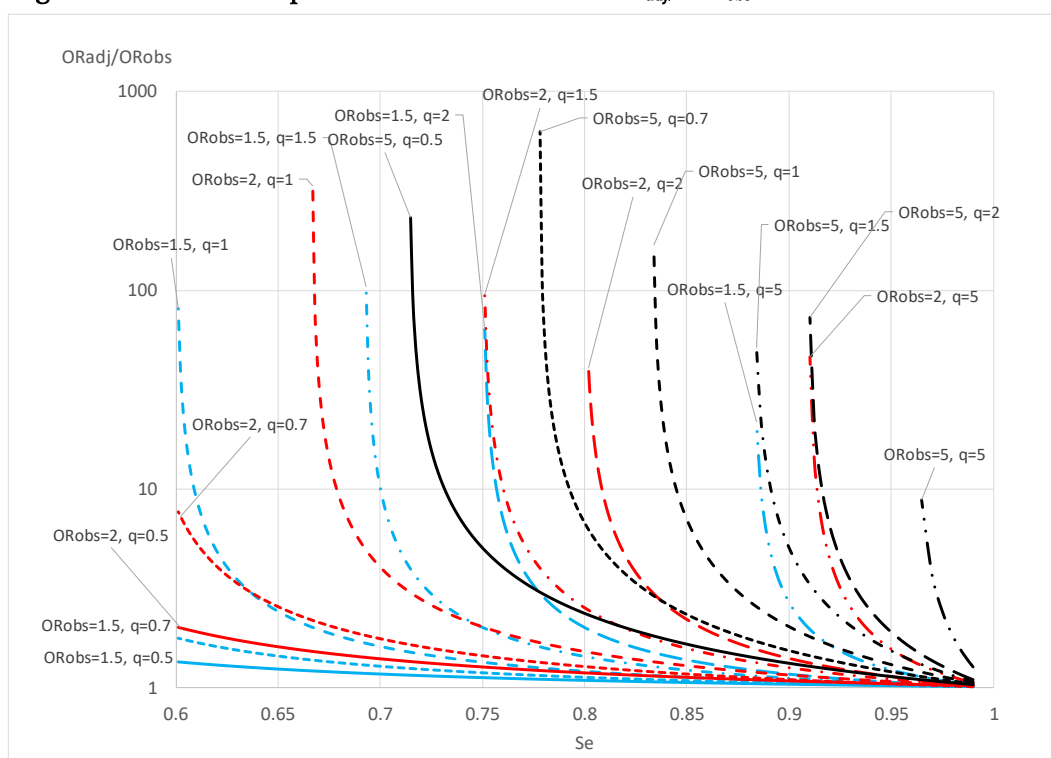
$$OR_{obs} = \frac{A_1 B_0}{A_0 B_1}, \quad OR_{adj} = \frac{[Sp_A(A_1 + A_0) - A_0] \cdot [Se_B(B_1 + B_0) - B_1]}{[Sp_B(B_1 + B_0) - B_0] \cdot [Se_A(A_1 + A_0) - A_1]}$$

OR<sub>adj</sub> と OR<sub>obs</sub> の関係を表すため，OR<sub>adj</sub> の式に OR<sub>obs</sub> を用いる．曝露なしの場合のオッズを p とし，B<sub>1</sub>=B<sub>0</sub>q と表すと，A<sub>0</sub>=pB<sub>0</sub>，A<sub>1</sub>=OR<sub>obs</sub>pB<sub>0</sub>q であるので，

$$OR_{adj} = \frac{[Sp_A(OR_{obs}pB_0q + pB_0) - pB_0] \cdot [Se_B(B_0q + B_0) - B_0q]}{[Sp_B(B_0q + B_0) - B_0] \cdot [Se_A(OR_{obs}pB_0q + pB_0) - OR_{obs}pB_0q]}$$

$$= \frac{[Sp_A(OR_{obs}q + 1) - 1] \cdot [Se_B(q + 1) - q]}{[Sp_B(q + 1) - 1] \cdot [Se_A(OR_{obs}q + 1) - OR_{obs}q]}$$

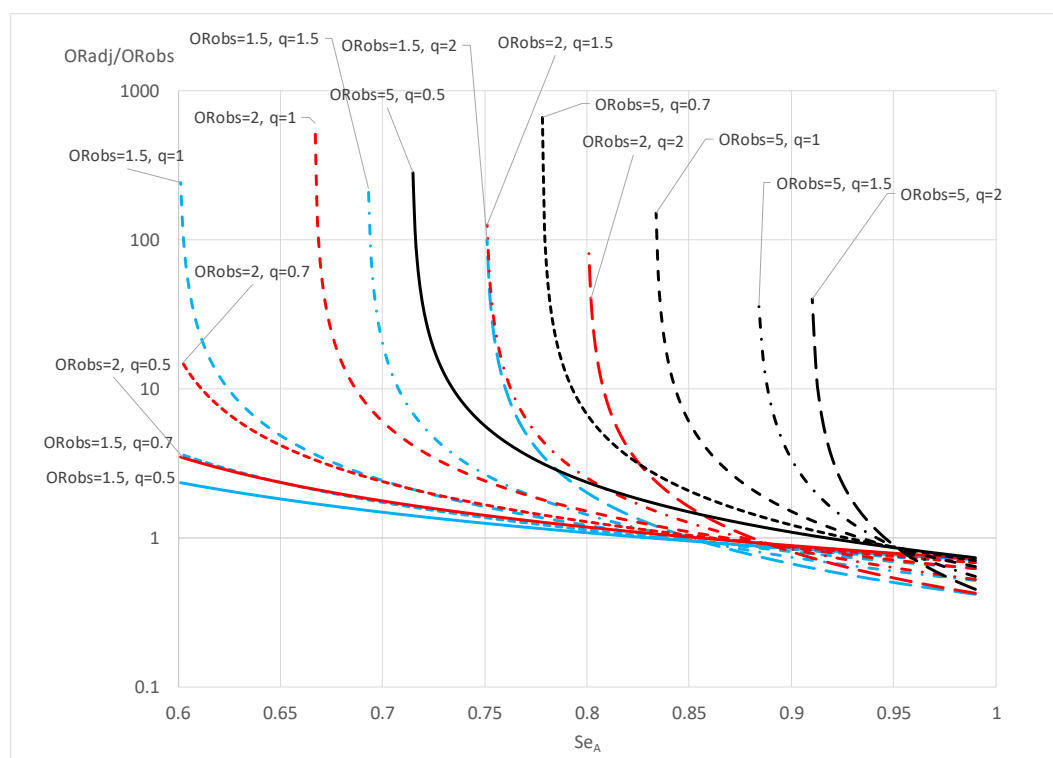
となるが，この式から OR<sub>adj</sub> / OR<sub>obs</sub> をイメージするのは難しい．Sp<sub>A</sub>=Sp<sub>B</sub>=0.99，と固定し，感度に差異がなく (Se<sub>A</sub>=Se<sub>B</sub>=0.99~0.6)，OR<sub>obs</sub> が 1.5, 2, 5, q が 0.5, 0.7, 1, 1.5, 2, 5 の場合の感度と OR<sub>adj</sub>/OR<sub>obs</sub> の関係を Figure 5-6 に示す．ここから，例えば，OR<sub>obs</sub>=1.5, q=5 の場合の OR<sub>adj</sub> / OR<sub>obs</sub>=2.67 であるように，曝露の感度および特異度に差異がなく，かつ感度が 0.9 の場合でも観測されたオッズ比は相当に過少とななっている場合があることを読み取れる．これは同時に感度，特異度がともに非差異の場合は補正後のオッズ比が観測された値より小さくなることはない事も示している．

Figure 5-6 特異度 ( $Sp$ ) = 0.99 とした場合の  $OR_{adj}/OR_{obs}$ 

#### ✓ 差異誤分類は影響が大きくなる

しかし、差異がある場合は、補正後のオッズ比が小さく ( $OR_{adj}/OR_{obs} < 1$ ) なる場合が生じる。このことを確認するため、アウトカムなしの曝露の感度 ( $Se_B$ ) を 0.8 に固定した場合の Figure 5-6 と同様のシミュレーションの結果を Figure 5-7 に示す。この例では  $Se_A > 0.85$  の場合に  $OR_{adj} / OR_{obs} < 1$  となる場合が生じている。つまり、観測されたオッズ比が 1 より大きい場合、曝露の誤分類で感度、特異度が非差異（アウトカムの有無間で共通）であれば誤分類を補正したオッズ比は大きくなる場合はあっても、1 よりも小さくなることはないため、測定されたオッズ比がある程度大きいことによる意思決定には影響がない（感度分析は必要ではない）が、小さいことによる意思決定（意思決定の保留）の場合には感度分析は必要である。更に、差異があり、 $Se_A > Se_B$  の場合は、観測されたオッズ比が大きくなっても、補正後のオッズ比は 1 よりも小さくなる場合があるため、感度分析は必要である。

誤分類はこのように差異がある場合 4 つのパラメータを用い複雑であるので、Excel 等の簡便な Tool でおおよその影響を確認することを推奨する。

Figure 5-7  $Se_B=0.8$  に固定した場合

では、こういった場合に曝露の誤分類は生じるであろうか？ 曝露の有無はアウトカムとの因果関係を考慮して時間関係を含めて定義される。また、アウトカムが短期間の曝露により生じやすい場合は曝露が新たに生じたものである必要があるなど、単に物質（医薬品）の特定ではなく、時点・期間の特定が必要である。改めて曝露の誤分類が生じる原因の事例として、5.3の Figure 5-2 曝露の time window の図示の例を見てみる。曝露は処方医薬品であるが、真の曝露は初めに無料のサンプルが 15 日間使用され、そのあと 30 日分の処方を 2 回受けたのち、2 回目の処方の 20 日後に使用が中止されている（1 回目は非使用日が少なくとも 10 日あるため 40 日となっている）。これは inception cohort（新規使用患者）を対象とした研究であれば無料のサンプルを使用していた患者は本来除外されるべき患者であるため選択バイアスであるが、case-control 研究である場合は誤分類である場合がある。また、無料のサンプルがなかったとしても、研究で使用する情報である調剤請求データは 30 日分の処方が 40 日の間を空けて 2 回生じている。そして研究での扱い（集計）は、調剤毎に調剤日数に服薬コンプライアンスを考慮して 15 日を加えた期間を曝露期間として扱っているため、真の曝露期間に対して開始日が 15 日遅く、終了日が 25 日後で曝露期間は 85 日間となっている。また仮に、アウトカムが生じた日のみを曝露の有無の判別対象とした場合（発現日に曝露があれば曝露あり、なければ曝露なし）、この 85 日のうち、真の曝露は 60 日であるため、 $1 - 60 \div 85 = 29.4\%$  の確率で曝露なしがありと誤って分類（偽陽性）されることになる。これを逆に真の曝露期間から見ると、65 日中 60 日が曝露ありとなるため 92% が正しく曝露あり（この確率を感度と呼ぶ）、8% が誤って曝露なし（偽陰性）となる。これは対象者毎の確認であるが、こういった考えを拡張して感度分析では曝露の有無やアウトカムの有無の群ごとにこのようなことが生じている割合



に加えてスイッチ OTC がある場合は OTC を使用している割合などを扱う。Figure 5-2 は保険償還のデータであるため医療機関をまたいで曝露状況を把握できるが、研究が医療機関の診療情報（電子カルテ）データベースを用いる場合には他の医療機関の診療情報は欠損しており、inception cohort に限定した new user design を用いる場合に、選択バイアスである既使用患者が混在するバイアス（prevalent user bias）がより多く生じやすいので注意が必要である。また、このような服薬コンプライアンスなどによる誤分類は多くの場合、曝露群のみに生じるため差異誤分類となる。また、異なる曝露を比較対象とした場合でも、例えば新薬と従来から使用されている薬剤では 1 回の処方日数に長短が生じるなどにより誤分類の確率が異なる差異誤分類となる。

#### 5.4.3.3. アウトカムの誤分類モデルの場合

アウトカムの誤分類を扱う数式も曝露の誤分類同様の特徴を有している。曝露の誤分類同様に、アウトカムの誤分類を感度（ $Se_1$ ,  $Se_0$ ）、特異度（ $Sp_1$ ,  $Sp_0$ ）と Table 5-2 のセルの記号を用いて誤分類のある観測データと誤分類を補正した場合を整理すると Table 5-4 となる。

Table 5-4 アウトカムの誤分類の数式整理

		E=1	E=0	Total
data Crude	D=1	$A_1 = Se_1 A_1^* + Fp_1 B_1^*$	$A_0 = Se_0 A_0^* + Fp_0 B_0^*$	$M_1$
	D=0	$B_1 = Fn_1 A_1^* + Sp_1 B_1^*$	$B_0 = Fn_0 A_0^* + Sp_0 B_0^*$	$M_0$
data 誤分類のない	D*=1	$A_1^* = [Sp_1 A_1 - (1 - Sp_1) B_1] / (Se_1 + Sp_1 - 1)$ $= [Sp_1 (A_1 + B_1) - B_1] / (Se_1 + Sp_1 - 1)$	$A_0^* = [Sp_0 A_0 - (1 - Sp_0) B_0] / (Se_0 + Sp_0 - 1)$ $= [Sp_0 (A_0 + B_0) - B_0] / (Se_0 + Sp_0 - 1)$	$M_1^*$
	D*=0	$B_1^* = [Se_1 B_1 - (1 - Se_1) A_1] / (Se_1 + Sp_1 - 1)$ $= [Se_1 (B_1 + A_1) - A_1] / (Se_1 + Sp_1 - 1)$	$B_0^* = [Se_0 B_0 - (1 - Se_0) A_0] / (Se_0 + Sp_0 - 1)$ $= [Se_0 (B_0 + A_0) - A_0] / (Se_0 + Sp_0 - 1)$	$M_0^*$
	Total	$N_1 = A_1 + B_1 = A_1^* + B_1^*$	$N_0 = A_0 + B_0 = A_0^* + B_0^*$	

ここから、感度分析に用いる補正前の四分表のオッズ比  $OR_{obs}$  と補正後のオッズ比  $OR_{adj}$  は

$$OR_{obs} = \frac{A_1 B_0}{A_0 B_1}, \quad OR_{adj} = \frac{[Sp_1 (A_1 + B_1) - B_1] \cdot [Se_0 (A_0 + B_0) - A_0]}{[Sp_0 (A_0 + B_0) - B_0] \cdot [Se_1 (A_1 + B_1) - A_1]}$$

$OR_{adj}$  と  $OR_{obs}$  の関係を表すため、 $OR_{adj}$  の式に  $OR_{obs}$  を用いる。曝露なしの場合のオッズを  $p$ ,  $B_1 = B_0 q$  と表すと、 $A_0 = p B_0$ ,  $A_1 = OR_{obs} p B_0 q$  であるので、

$$\begin{aligned}
 OR_{adj} &= \frac{[Sp_1 (OR_{obs} p B_0 q + B_0 q) - B_0 q] \cdot [Se_0 (p B_0 + B_0) - p B_0]}{[Sp_0 (p B_0 + B_0) - B_0] \cdot [Se_1 (OR_{obs} p B_0 q + B_0 q) - OR_{obs} p B_0 q]} \\
 &= \frac{[Sp_1 (OR_{obs} p + 1) - 1] \cdot [Se_0 (p + 1) - p]}{[Sp_0 (p + 1) - 1] \cdot [Se_1 (OR_{obs} p + 1) - OR_{obs} p]}
 \end{aligned}$$

数式 5-3

となる。当然であるが、曝露の層間の比である「q」は消えている。この式からでも  $OR_{adj}/OR_{obs}$  をイメージするのは難しいので、 $Sp_1 = Sp_0 = 0.999$  と固定し、感度に差異がなく

( $Se_1=Se_0=0.99\sim0.6$ ),  $OR_{obs}$  が 1.5, 2, 5,  $p$  が 0.002, 0.005, 0.01, 0.05, 0.1, 0.2 の場合の感度と  $OR_{adj}/OR_{obs}$  の関係を Figure 5-8 に示す。線が込み合い、煩雑であるので、 $Se=0.99$  と  $Se=0.6$  の場合の  $OR_{adj}$  を Table 5-5 に整理した。この設定では  $OR_{adj}$  は  $OR_{obs}$  よりほとんどの場合でわずかに大きな値、つまり観察された  $OR_{obs}$  は僅かであっても過少となっている。もともと  $OR_{obs}$  が大きい場合は意思決定への影響はないかもしれないが、 $OR_{obs}=1.5$  で  $p=0.002$  の場合では  $OR_{adj}$  は約 2 であるので、意思決定への影響が考えられる。また、感度に差異がある場合として、曝露なしの層の感度 ( $Se_0$ ) を 0.8 に固定した場合を

Table 5-6 に示す。この場合は、 $OR_{adj}/OR_{obs} > 1$  と  $OR_{adj}/OR_{obs} < 1$  の両方が生じている。アウトカムの誤分類の影響は状況に依存して複雑であるため、そのパターンをまとめた論文もある<sup>38)</sup>。参考までに、その 1 つの例を Table 5-7 に示しておく。

Figure 5-8 特異度 ( $Sp$ ) =0.999 とした場合の  $OR_{adj}/OR_{obs}$

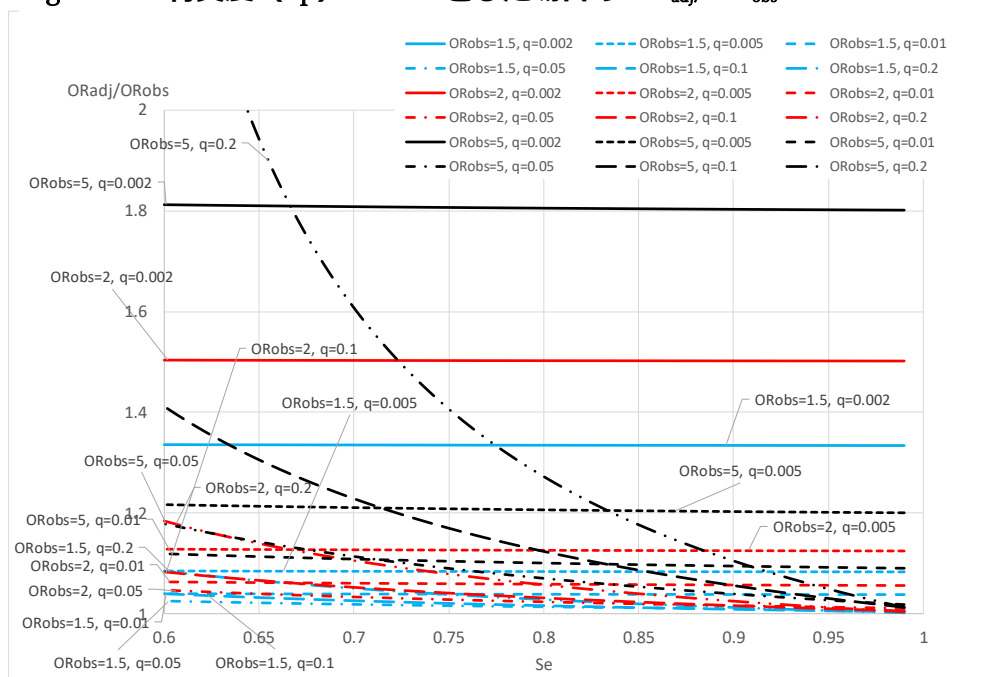


Table 5-5  $Sp_1=Sp_0=0.999$ ,  $Se=0.99, 0.6$  の場合の  $OR_{adj}/OR_{obs}$

$OR_{obs}$	P=曝露なしの場合のオッズ					
	0.002	0.005	0.01	0.05	0.1	0.2
1.5	1.334	1.083	1.037	1.007	1.004	1.003
	1.335	1.085	1.041	1.024	1.041	1.085
2	1.501	1.125	1.056	1.011	1.006	1.005
	1.503	1.129	1.063	1.046	1.082	1.185
5	1.802	1.200	1.089	1.018	1.012	1.012
	1.811	1.217	1.119	1.179	1.411	2.610

上段:  $Se=0.99$ , 下段:  $Se=0.6$

Table 5-6  $Se_0=0.8$ ,  $Se_1=0.99$ , 0.8, 0.6 の場合の  $OR_{adj} / OR_{obs}$ 

$OR_{obs}$	P					
	0.002	0.005	0.01	0.05	0.1	0.2
1.5	1.077	0.804	1.212	0.807	1.455	0.813
	1.334	1.013	1.502	1.023	1.805	1.071
	1.781	1.395	2.006	1.425	2.417	1.606
2	0.874	0.792	0.908	0.793	0.969	0.798
	1.084	1.016	1.127	1.032	1.206	1.123
	1.450	1.449	1.508	1.508	1.625	1.966
5	0.836	0.771	0.851	0.773	0.878	0.779
	1.038	1.029	1.058	1.058	1.100	1.272
	1.393	1.586	1.423	1.732	1.498	3.815

上段:  $Se_1=0.99$ , 中段:  $Se_1=0.8$ , 下段:  $Se_1=0.6$ 

Table 5-7 アウトカムの誤分類が結果のオッズ比に影響するパターン

Categories	$OR_D > 1$	$OR_D < 1$	$OR_D = 1$
A $SE_0 = SE_1 < 1$ , $SP_0 = SP_1 = 1$	Toward the null	Toward the null	Unbiased
B $SP_0 = SP_1 < 1$ , $SE_0 = SE_1 = 1$	Toward the null	Toward the null	Unbiased
G $SE_0 = SE_1 = r < 1$ , $SP_0 = SP_1 = r < 1$	Switchover bias if $r < 0.5$ , toward the null if $r \geq 0.5$	Switchover bias if $r < 0.5$ , toward the null if $r \geq 0.5$	Unbiased
H $SE_0 = SE_1 = r_1$ , $SP_0 = SP_1 = r_2$	Switchover bias if $r_1 + r_2 < 1$ , toward the null if $r_1 + r_2 \geq 1$	Switchover bias if $r_1 + r_2 < 1$ , toward the null if $r_1 + r_2 \geq 1$	Unbiased
C,I $SE_0 < SE_1$ , $SP_0 = SP_1$	Bias in any direction <sup>a</sup>	Switchover bias if $OR_T > 1$ , toward the null if $OR_T \leq 1$	Away from the null with $OR_T > 1$
D,J $SE_0 = SE_1$ , $SP_0 > SP_1$	Bias in any direction <sup>a</sup>	Switchover bias if $OR_T > 1$ , toward the null if $OR_T \leq 1$	Away from the null with $OR_T > 1$
E,K $SE_0 > SE_1$ , $SP_0 = SP_1$	Switchover bias if $OR_T < 1$ , toward the null if $OR_T \geq 1$	Bias in any direction <sup>b</sup>	Away from the null with $OR_T < 1$
F,L $SE_0 = SE_1$ , $SP_0 < SP_1$	Switchover bias if $OR_T < 1$ , toward the null if $OR_T \geq 1$	Bias in any direction <sup>b</sup>	Away from the null with $OR_T < 1$
M $SE_0 > SE_1$ , $SP_0 < SP_1$	Switchover bias if $OR_T < 1$ , toward the null if $OR_T \geq 1$	Bias in any direction <sup>b</sup>	Away from the null with $OR_T < 1$
N $SE_0 < SE_1$ , $SP_0 > SP_1$	Bias in any direction <sup>a</sup>	Switchover bias if $OR_T > 1$ , toward the null if $OR_T \leq 1$	Away from the null with $OR_T > 1$
O $SE_0 < SE_1$ , $SP_0 < SP_1$	Bias in any direction <sup>a</sup>	Bias in any direction <sup>b</sup>	Unbiased or away from the null
P $SE_0 > SE_1$ , $SP_0 > SP_1$	Bias in any direction <sup>a</sup>	Bias in any direction <sup>b</sup>	Unbiased or away from the null

Note.  $OR$  = odds ratio;  $SE_E$  = sensitivity of the test for disease status;  
 $SP_E$  = specificity of the test for disease status. If Exposure=Yes then  $E=1$  else  $E=0$ .  
 $OR_D$ =odds ratio of true disease status without misclassification  
 $OR_T$ = odds ratio of tested disease status with misclassification,  $R_{bias} = OR_T / OR_D$   
a: Switchover bias if  $OR_T < 1$ , toward the null if  $R_{bias} < 1$  and  $OR_T \geq 1$ ,  
unbiased if  $R_{bias} = 1$ , away from the null if  $R_{bias} > 1$ .  
b: Switchover bias if  $OR_T > 1$ , toward the null if  $R_{bias} > 1$  and  $OR_T \leq 1$ ,  
unbiased if  $R_{bias} = 1$ , away from the null if  $R_{bias} < 1$ .

出典: Chen Q (2013)<sup>38)</sup>の TABLE 2—Analytic Descriptions of Patterns of Bias Caused by Disease Misclassification

## ✓ アウトカムの陽性的中度によるバイアス補正

医療情報データベースを二次利用する場合に目的とするアウトカムの特定精度をバリデーション研究により評価することが推奨されており<sup>41),42),43)</sup>, アウトカムの感度, 特異度を得ることは望ましいが, 医薬品の安全性研究など, アウトカムの発生頻度が低い場合は特に, 実施の容易さから陽性的中度 (Ppv) のみによる評価が容認されている<sup>41),42)</sup>. これはデータベースを用いる場合, 研究の定義に合致する全ての cases (又はランダムに抽出) を特定することが容易なのに対して十分な数の真の cases を特定することは多大な労力を要するうえ, 十分な高い陽性的中度は擬陽性が少ない, つまり特異度 (Sp) が高いことであり, 仮に Ppv=1 の場合 Sp=1 であるので, 感度が比較する群間で等しい (非差異) 場合には次のように理論上, 正しい相対リスクが得られるためである.

Table 5-2 の記号を用い, 誤分類のない場合のセルの値を各々「\*」付き大文字で表すと

$$A_1^* = [Sp_1 A_1 - (1 - Sp_1) B_1] / (Se_1 + Sp_1 - 1) = [Sp_1 (A_1 + B_1) - B_1] / (Se_1 + Sp_1 - 1) \\ = [A_1 - (1 - Sp_1) N_1] / (Se_1 + Sp_1 - 1)$$

$$B_1^* = [Sp_0 A_0 - (1 - Sp_0) B_0] / (Se_0 + Sp_0 - 1) = [A_0 - (1 - Sp_0) N_0] / (Se_0 + Sp_0 - 1)$$

$$N_1^* = N_1, \quad N_0^* = N_0$$

から感度 (Se), 特異度 (Sp) を用いると, 補正したリスク比 ( $RR_{adj}$ ) は

$$RR_{adj} = \frac{A_1^* / N_0^*}{A_0^* / N_1^*} = \frac{[A_1 - (1 - Sp_1) N_0] / (Se_1 + Sp_1 - 1) N_0}{[A_0 - (1 - Sp_0) N_1] / (Se_0 + Sp_0 - 1) N_1} \quad \text{数式 5-4}$$

$Sp_1 = Sp_0 = 1$  (つまり,  $Ppv_1 = Ppv_0 = 1$ ),  $Se_1 = Se_0 = Se$  の場合, 上記式は

$$RR_{adj} = \frac{A_1^* / N_0^*}{A_0^* / N_1^*} = \frac{A_1 / N_0}{A_0 / N_1} = RR_{obs}$$

となり, 正しい相対リスクを得られることが確認できる. 陽性的中度 (Ppv), 陰性的中度 (Npv) を用いると

$$RR_{adj} = \frac{A_1^* / N_0^*}{A_0^* / N_1^*} = \frac{[Ppv_1 A_1 + (1 - Npv_1) B_1] / N_1}{[Ppv_0 A_0 + (1 - Npv_0) B_0] / N_0} \quad \text{数式 5-5}$$

である. この式では  $Ppv_1 = Ppv_0 = 1$ ,  $Npv_1 = Npv_0$  の場合でも  $RR_{adj} = RR_{obs}$  とはならない. また,

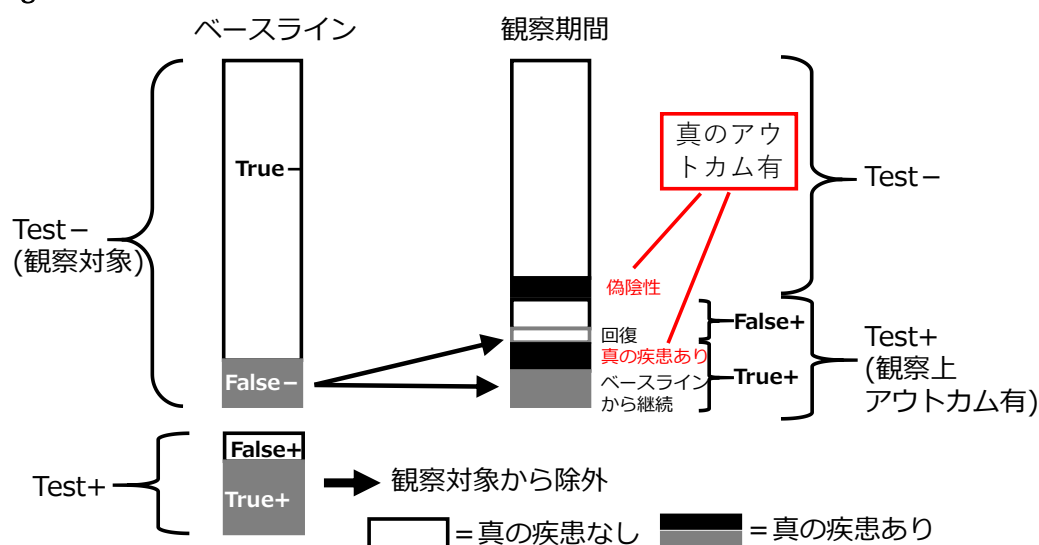
$$Ppv_1 = \frac{Se_1 A_1^*}{A_1}, \quad Ppv_0 = \frac{Se_0 A_0^*}{A_0}, \quad Npv_1 = \frac{Sp_1 B_1^*}{B_1}, \quad Npv_0 = \frac{Sp_0 B_0^*}{B_0}$$

であるので,  $Se_1 = Se_0$ ,  $Sp_1 = Sp_0$  の場合でも  $Ppv_1 = Ppv_0$ ,  $Npv_1 = Npv_0$  とはならない. このように, 陽性的中度は対象集団における有病率の影響を受けるため, 有病率 (アウトカムの発生率) が異なる集団を用いたバリデーション研究から得られた陽性的中度, 陰性的中度を用いる場合は注意が必要である.

以上から、アウトカムの誤分類の感度が曝露の層で差異がある場合は感度分析の必要性が高まるため、医療情報データベースを二次利用する場合に曝露の有無で感度に差異が生じ難いよう、例えば一方には診断名の付与のみにより、他方は加えて検体検査の実施を条件とするなど曝露の有無で異なる定義を用いるというアイデアが考えられるが、FDA の電子健康データを用いた薬剤疫学的安全性研究のガイダンス<sup>43)</sup> (p20) では曝露情報をアウトカムの識別に用いないことが推奨されている（アウトカムの前後の曝露情報の有無を含む）。

具体的に論文での扱いを紹介する。通常、研究は曝露後に新たに生じた事象（疾患の発症、転倒、死亡などのイベント）をアウトカムとして扱うため、ベースラインでは事象を有しない者（case-control 研究では新規のアウトカム）を対象とする。このためアウトカムの有無の分類はベースラインと観察期間の2回行うことになる。ベースラインにおけるアウトカムの特異度の低下（偽陽性）は研究から除外する方向のみに働くため曝露の有無に関係なければ結果に影響することは少ないが、低い感度は研究から除くべき対象を組み込んでしまうため、観察期間中にアウトカムが観察され、新規アウトカムとして扱った場合はアウトカムありの数が過大となる。Pekkanen J (2006) の論文<sup>44)</sup>はこの場合を含めた影響をシミュレーションにより説明している。この論文で説明に用いられている図を改変したものを Figure 5-9 に示す。

Figure 5-9 観察研究におけるアウトカムとする疾患の誤分類の影響<sup>44)</sup>



観察期間の測定（Test）により陽性となり、実際にアウトカムとする疾患があるにもかかわらず誤りが生じる。理解を促すために具体的に説明を加える。例えば、複数の医療機関の電子カルテのデータを用いて新規糖尿病の発症（タイプを問わない）をアウトカムとした cohort 研究を行う場合を考える。患者が糖尿病と診断されており、HbA<sub>1c</sub> の値が継続的に 6.5%以上である場合に「糖尿病患者」とできるとしよう。そして「新規糖尿病の発症」は HbA<sub>1c</sub> の値が継続的に 6.5%未満から「初めて 6.5%以上となる」必要がある。このため継続的に HbA<sub>1c</sub> を測定されていない患者は除く必要がある。このことを含め、Table 5-8 の a, b, c のようなアウトカ



ムの測定に関連したバイアスが考えられる。この例の場合は a, b とともに糖尿病の新規発症としては偽陽性、特異度により扱え、アウトカムの誤分類のように見えるが感度分析のモデルは選択バイアスであり、c は偽陰性としてアウトカムの誤分類である。

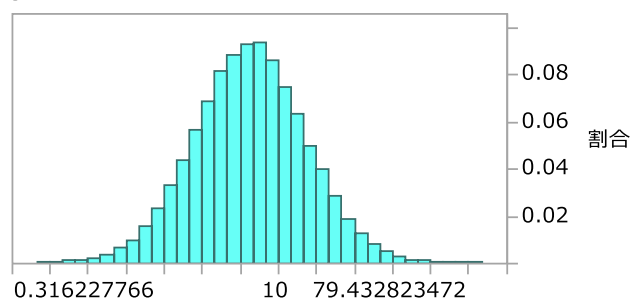
Table 5-8 アウトカムが誤分類される原因の例

- |   |
|---|
| <p>a.仮に「6.5%以上の HbA<sub>1c</sub> の記録がない患者」を未発症とした場合、検査が行われていないだけで実は発症している患者が未発症として扱われ、観察中に検査が実施された場合は誤って新規発症として扱われる。</p> <p>b.他の医療機関で糖尿病治療を受け、HbA<sub>1c</sub> &lt; 6.5%であった患者がコントロール不良となって HbA<sub>1c</sub> ≥ 6.5%となった患者は新規発症として扱われる。</p> <p>c.新規に発症したが、他の医療機関で先に診断され、治療されたため未発症として扱われる。</p> |
|---|

#### ✓ アウトカム測定間隔の差による誤分類のシミュレーション

アウトカムが差異誤分類を生じる場合をシミュレーションにより紹介する。アウトカムの診断や医療情報データベースを二次利用する場合のアウトカム定義に臨床検査の結果が必要な場合がある。新医薬品と既存医薬品の使用患者における臨床検査値異常を指標としたアウトカムの頻度を比較する研究を行ったとする。日常診療では使い慣れた既存医薬品では臨床検査の実施間隔は長く、新薬では相対的に短い間隔で検査される可能性は考えられる。単純化のために全ての患者が処方開始日から新医薬品は3か月間隔、既存医薬品は6か月間隔で検査が行われ、アウトカムは時間に依存せず一定の確率で36か月間に累計10%の患者に生じ、中央値6か月の対数正規分布（Figure 5-10）の期間持続するとすると、生じているアウトカムのうち新医薬品は90%、既存医薬品は76%が検出される（感度に該当する）。つまり、真のリスク比は1であるが、観察される結果は1.18となる。このような情報バイアス（差異誤分類）は施設の種類（病院と診療所）、診療科、併用薬・併存疾患、対象者の脆弱性等によっても生じる。

Figure 5-10 検査異常の持続期間の分布 =EXP(NORM.INV(UNIFORM(0~1),LN(6),1)



#### 5.4.3.4. 選択バイアスモデルの場合

選択バイアスは研究の対象とする条件（組み入れ条件：入院，受診，検査実施，追跡/打ち切りなど，「共通の結果」，DAG では「Collider」とよばれる）を生じる原因が研究の対象であるアウトカム以外に他の要因（曝露，合併症，健康への関心，経済力，保険システムなど，）によって生じる場合に他の要因とアウトカムの間に相関が測定されるバイアスであり，代表性や一般化可能性の問題ではないことは第 1 部 1.1.3 で説明した．アウトカムを生じさせる研究対象の曝露以外の要因の併存割合が組み入れによって曝露の層で差が生じる場合は選択バイアスではなく，交絡である．

選択バイアスにもいろいろな種類があり，Berkson's bias（組み入れ条件である入院を介してアウトカムと無関係な要因（施設特性など）に相関が生じる，特に case-control 研究で問題となり得る），self-selection bias（研究組み入れへの対象者の積極的関与の有無），loss to follow up（ランダムでない脱落・打ち切り），recall bias（アウトカムを生じた症例は曝露の詳細な情報を報告しやすい，特に case-control 研究で問題となり得る）などがある．（詳細は教科書を参照いただきたい）

四分表における選択バイアスの概念は Figure 5-4 の緑の矢印のように，研究の仮説の対象となる集団（標的集団）全体を観察した四分表の各セル（真の値）から研究で用いる標本の曝露の層内の抽出確率が異なる．このため，これまで説明してきた未測定交絡要因は四分表を層に分割し，誤分類は四分表の中で頻度が移動するのみで，四分表全体の数は変わらなかったが，選択バイアスは四分表の各々のセルの研究対象集団からの抽出率の偏りを補正するため，全体の数も変化する．Table 5-2 のセルの記号を用いて各セルに選択される確率を  $S_j$  ( $j$  は  $A_1, A_0, B_1, B_0$ ) とすると，感度分析に用いる補正前の四分表のオッズ比  $OR_{obs}$  と補正後のオッズ比  $OR_{adj}$  の比は

$$\frac{OR_{adj}}{OR_{obs}} = \frac{A_1 \cdot B_0}{A_0 \cdot B_1} \cdot \left( \frac{S_{A_1 S_{B_0}}}{S_{A_0 S_{B_1}}} \right)^{-1} \cdot \frac{A_0 \cdot B_1}{A_1 \cdot B_0} = \frac{S_{A_0 S_{B_1}}}{S_{A_1 S_{B_0}}} = \frac{S_{B_1}}{S_{A_1}} \cdot \frac{S_{A_0}}{S_{B_0}}$$

とできる．数式はとてもシンプルで手計算も容易である． $OR_{obs} > 1$  の場合， $S_{A_0 S_{B_1}} > S_{A_1 S_{B_0}}$  であれば， $OR_{obs}$  は過少， $S_{A_0 S_{B_1}} < S_{A_1 S_{B_0}}$  であれば， $OR_{obs}$  は過大となっていることになるが，曝露の各層のアウトカムの有無毎の抽出率の比に依存し，標本の曝露の層の抽出率には依存しない．つまり，曝露の層内におけるアウトカムの有無の層間の抽出率に違いがある場合にバイアスを生じる．

数式はシンプルであるが，パラメータを推定するのは選択バイアスの性質上難しいかもしれない．このようなことが生じるのは，研究の元となる集団が特定（複数）の医療機関受診者であるホスピタルコホートなど，本来研究の対象とすべき集団（地域住民など）からコホートへの組み入れ（抽出率）が研究で扱う層（曝露，アウトカム）によって偏りがある場合<sup>注38</sup>，アウ

注38：「ある場合」としたが，むしろ組み入れがランダムで偏りがないとできる場合は皆無であり，そもそも地域住民など真の母集団の定義さえできない場合が多いであろう．つまり，このようなコホート（医療情報データベース）を用いる場合は選択バイアスの考慮は必須であると考える．

トカムの診断が自発的受診によるものと検診制度によるものが混在する場合などがあるが、四分表を対象とする単純な方法では、研究の対象とする組込みおよび除外条件の情報の欠測、誤り、追跡不能（打ち切り）など様々なものが感度分析における選択バイアス補正式の対象となる。これにはポストモダンな選択バイアスの定義により、曝露、アウトカム、共変量に関するベースラインの誤分類も含まれる<sup>注39</sup>。

事例として 5.3 で紹介した Figure 5-3 フローチャートによる整理の例は、ミュンヘンとその周辺住民を対象に社会不安障害（SAD）について研究する The Early Developmental Stages of the Psychopathology Study である<sup>31)</sup>。この研究では、住民台帳から 1995 年時点（T0）で 14 から 24 歳の住民をランダムに選び（14～17 歳は 18～24 歳の 2 倍の確率、計 4,263 人）、インタビュー調査を行い、1996 年（T1）と 1998 年（T2）追跡調査が行われている。ベースラインアンケートの回答率は 71%であったため、組入れは 3,021 人であった。14～17 歳は T1, T2 で、18～24 歳は T2 で脱落があり各々 1,053 人、1,130 人となっている。T2 における社会不安障害の発症とうつ（DEP）の関係の未調整のオッズ比 14～17 歳は 0.22, 18～24 歳は 3.26 であるが、脱落率などから抽出率（S）は平均 60%, SAD+DEP+（S(D+E+)) は 63%, SAD-DEP+（S(D-E+)) は 57%とした場合の補正後オッズ比は 14～17 歳は 0.18, 18～24 歳は 2.65 とされている。この研究は初めに無作為抽出（確率的抽出）が行われており、本来、選択バイアスは生じないデザインであるので、アンケートの未回答、脱落による選択バイアスのみが想定された。実際の医薬品の観察研究では地域コミュニティなどポピュレーションベースでランダム抽出を行い、面接にて直接データが収集される選択バイアスを無視可能なデザインは稀であるが、母集団からの抽出率が曝露の有無又はアウトカムの有無で共通（依存しない、独立である）の場合はバイアスを生じない。

再掲：Figure 5-3 フローチャートによる整理の例

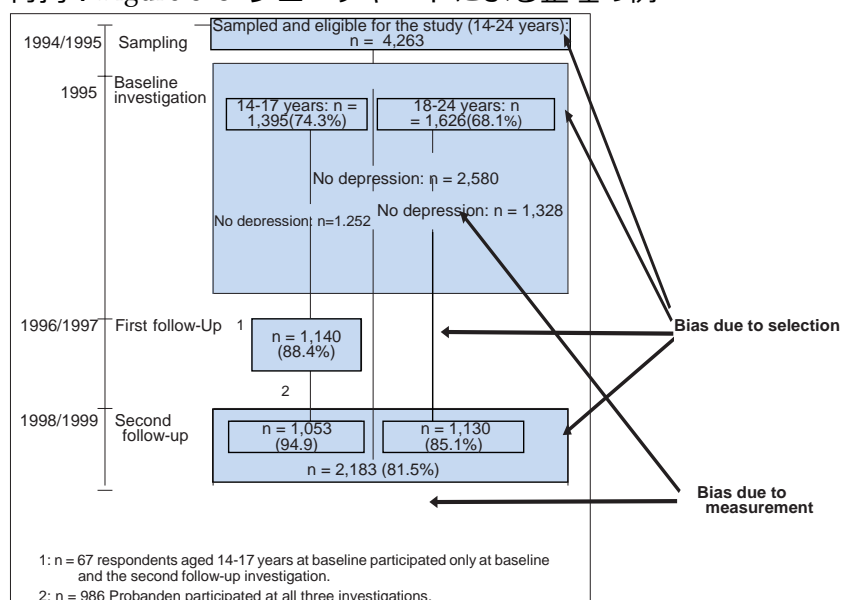


Figure 1. EDSP study design and potential biases.

注39：誤分類として扱う場合、曝露などの分類をあり、なしの2値ではなく、より複雑にする必要があるが、その計算式は複雑であり本書の対象としていない。第3、第4の分類がある場合、便宜的に感度分析の補正を行う計算（数式）に選択バイアスの数式を用いるという分類である。

以上、主なバイアスについて、その概略と結果の統計量がオッズ比の場合について結果への影響が大きくなる場合を紹介した。詳細なバイアスの原因の量（誤分類の感度、特異度など）を調べていなくても経験的あるいは直感的に未測定交絡要因や曝露の誤分類と言ったバイアスに対して「このぐらいあるかもしれない」量を想定することで、研究により報告された結果が感度分析によってどう変わり得るかを見積もることができる。また、繰り返しになるが、簡便な感度分析を行う Tool (ESATJ, Excel<sup>Appendix3.1)</sup> など) を用いて、簡便な手法により想定される場合の影響を確認することは有効な手段であると考える。

## 5.5. 感度分析の方法を選択する

感度分析の手法は単純なものから複雑なものまで種々提案されている。<sup>2),8),6),7),19),12)</sup> ここでは、第1部第3章及び第4章の主な手法について、5.5.1.で手法を選択する際の考え方、5.5.2.で追加の留意事項を説明する。なお、選択したモデルに必要なパラメータが得られない、あるいは利用可能なパラメータに適したモデルがある場合があり、このステップと「5.6.1 感度分析パラメータに関する情報源と収集」のステップは巡回することがある。

### 5.5.1. 手法の選択：情報量とシンプルさとのバランス、結果説明の容易さ

これまで幾度か触れてきたが、確率分布（パラメータのばらつき）を用いない、Simple sensitivity analysis で十分な場合もあれば、複数のバイアスの原因を併合して扱う Multiple bias analysis、さらにレコードレベルで行う高度な手法まで、状況によって必要なあるいは適切な手法は異なる。教科書などにその選択の基準が明記されてはいないが次のような事項が考慮されるであろう。

- a. 研究結果が用いられる場面（意思決定）の重要性。
- b. 感度分析モデルの複雑さ（パラメータの数、分布、バイアスの数）
- c. Validation study が行われている場合。
- d. バイアスの面からはその量（未測定交絡要因の併存率、オッズ比や曝露の誤分類の感度、特異度などのパラメータ）の信頼性/曖昧さ。
- e. 研究結果の面から小さなバイアスでも結果が影響を受ける、反対に大きなバイアスでも影響を受けないことを示す場合。
- f. 感度分析の結果の説明の容易さのバランス

これらのうち、a～d はより高度な手法を選ぶ必要性や実施可能性の要因であり、e, f は簡便な手法を選ぶ要因である。順に見てゆくと、研究結果が重要と思われる意思決定に提供される場合は少なくとも感度分析の仮定の曖昧さの提示は必要であるので、極端な場合でも結果が覆らない、またはわずかなバイアスでも覆ることの提示を目的とする e. のような場合を除いて、確率的な手法である PSA あるいはより高度な手法が適している。感度分析に用いるパラメータを特定の値に固定できる（その根拠となる情報がある）のであれば単純な手法で結果を表現可能であるが、パラメータの数が多く、値のバリエーションが多い、パラメータの値に相



関がある，確率的である場合には PSA を用いる必要が生じる．感度分析の対象となるバイアスの種類が複数ある場合は個々のバイアスの影響の方向が異なる場合，影響の小さい複数のバイアスの累積が危惧される場合は Multiple bias modelling（4.2 参照）が有用な場合がある．

研究内部の validation study によって感度分析の対象が特定され，根拠の妥当な感度分析パラメータ（点推定値）が得られている場合，validation study の信頼性が高くパラメータの精度が高い場合は simple sensitivity analysis で十分な場合もあり得るが，通常は PSA あるいは validation study を最大限に生かすレコードレベルのデータを用いる手法が適切となる．または単に研究の結果が逆転するあるいは常識の範囲では覆らないことを示すのみであれば multidimensional analysis も候補となる<sup>注40</sup>が，仮に限界値を示す場合は多くのシナリオ（感度分析パラメータ）による試行錯誤が必要となる．対して，probabilistic sensitivity analysis（PSA）は結果を信頼区間として提示することが可能で，多くの組み合わせから特定の結果となる場合を取り出すことが可能であり，結果のみの提示を考えれば Multidimensional analysis よりもむしろ簡便である<sup>注41</sup>．感度分析に用いるバイアスのパラメータの精度が低い，曖昧な場合は少なくとも PSA を用いる必要があると考えるが，その場合、用いる分布は例えば特に根拠がない場合は一様分布のような単純な分布を用い<sup>注42</sup>，根拠がある場合はその情報に応じた分布を用いる，あるいは感度分析の目的（結果が逆転するのは例外的な場合であることの提示など）に合わせて選択する必要があると考える．

本書のために多くの事例に当たったが，印象としては PSA（Monte Carlo simulation による）が多く，複数のバイアスの感度分析を行っている場合でも Multiple bias analysis が用いられている例は多くはない．また，Multidimensional analysis も事例は少ないようである．つまり，観察研究における事例としては PSA を中心に，必要に応じてより簡便な手法あるいはより高度な手法が選ばれていると考える．つまり実用上は，最小値と最大値を指定する簡便な一様分布のパラメータを用いた PSA を基準として，分布関数に仮定を反映させるバリエーションがあり，パラメータに相関を設定する複雑なものが用いられている．

5.8 教科書的ではないが現実的で簡便な手順による感度分析で説明しているが，バイアスの影響に見当をつける，補正後オッズ比が 1 になる場合等，特定の値となるパラメータを探すなど，一様分布の PSA の出力を利用することで容易に得られることから，一様分布による PSA は基礎的な手法と言えるであろう．

### 5.5.2. 利用可能な感度分析の方法の情報源・ツールと新しいモデルの作成

感度分析は主要な統計ソフトの標準的なメニュー（プロシージャ，関数など）に含まれていないため補正式，手法のプログラミング（コード）が必要である．これらを研究者自身が作る

注40： 5.6.1.5 結果が無意味または有意になる限界値，5.6.3.1target-adjustment sensitivity analysis<sup>41</sup> 参照

注41： 5.6.3.1target-adjustment sensitivity analysis，5.8 教科書的ではないが現実的で簡便な手順による感度分析 参照

注42： オッズ比など，比のパラメータの場合は注意が必要である．③ 参照



のではなく、文献、書籍、公開されているプログラムなどの情報源、ツールなど、研究に使用され確認されたプログラム（コード）を積極的に用いることが推奨されている。本書で引用している資料にも多くの例示が含まれている。しかし、用いているモデル、設定するパラメータ、特に確率的な手法の分布関数の意味など十分な理解（実際の分布の確認など）が必要である。また、複数の要因を扱う場合など新たなモデルが必要な場合への既存プログラムの適用は困難な場合があり、独自に作成する必要がある。その場合、研究に用いられた実績のあるプログラムと新たに作成したプログラムの結果を比較することはアルゴリズムやプログラムミスの発見に役立つ。「Appendix2 分布に基づく乱数の生成」に WEB に公開されている Excel workbook, SAS Macro, R Package のリストを掲載しているので参考にされたい。

以上、感度分析の手法、特に確率的手法である PSA（MCSA）に重点を置いて、仕組み、特徴、選ぶ際の考え方などを紹介してきた。PSA は与える情報が複数あり、様々な条件を設定可能である。次に感度分析に用いる情報（パラメータ）について説明する。

## 5.6. 感度分析モデルにパラメータを設定する

ここでは、選んだ感度分析のモデルに必要なパラメータの値や分布の情報源、感度分析の目的による使い分け、パラメータと結果の確認について説明する。

感度分析を行う目的、必要性に合わせて投入可能なリソースにより得られる情報の範囲で妥当と考えられる物を選ぶべきであるが、経験的あるいは直感的な値とせざるを得ない、あるいはそれで十分な場合もある。しかし、不適切なパラメータはバイアスを助長させ、確率的な手法によっても克服されないため注意が必要である<sup>46)</sup>。また、利用可能なより妥当性、信頼性の高い情報に合わせて使用するモデルを変更する場合も生じる。

### 5.6.1. 感度分析パラメータに関する情報源と収集

情報源は妥当性から次の 5.6.1.1～5.6.1.4 の順となるが、研究者以外が事後的に研究結果を評価する目的などのため感度分析を行う場合は主に 5.6.1.3 文献、統計情報など、5.6.1.4 専門家の意見、5.6.1.5 結果が無意味または有意味になる限界値を情報源として用いるであろう。

情報は測定されていない共変量に関するものと、測定されているデータの精度（誤り）に関するもの、選択バイアスに関するものに大別できる。共変量は例えば喫煙歴などで、共変量とアウトカムの関係の大きさ（オッズ比など）と曝露の層別の併存割合などが必要な情報である。データの精度は主に曝露およびアウトカムの識別精度を表す情報として感度・特異度あるいは陽性的中度、陰性的中度などが用いられる。前者の共変量は研究対象のポピュレーションに依存しているとみなせる場合は少なくなく、研究対象者から直接データを得られなくてもポピュレーションベースの情報にも妥当性を見いだせるのに対して、後者の原因は研究における測定、データの収集・情報源、扱い、対象者の選択/除外基準、追跡方法などにあり研究固有

である。このため、データ精度及び選択バイアスに関する情報は研究対象から取得することが好ましいが、単に同じ集団であるだけではなく、確率的抽出を行うなど、研究対象の代表性のある標本である必要がある。

内部検証データを用いる場合は検証データを収集するための研究計画と実施が必要となる。内部検証データ以外の場合、対象とする曝露群毎の情報は多くの場合得られない。何れの情報源を用いる場合でもその情報は実施する主研究の測定精度、必要なサンプルサイズに関する重要な情報である。研究結果によってはほんのわずかなバイアス（例えばアウトカムの誤分類の特異度が曝露/非曝露間で0.1%異なる）であっても結果が大きな影響を受ける場合があり、重要な意思決定に用いる研究を行う場合は、研究計画の段階で内部検証データの収集を行うことが望ましい反面、とても大きなバイアスパラメータを用いても結果への影響は限定的な場合があり、常に用いるパラメータに強い根拠は必要ではない。言い換えれば、ありえそうな範囲のバイアス（パラメータ）によって相反する補正結果が得られるような場合に用いるパラメータには、より強い根拠が求められると考える。

Table 5-9 未測定交絡要因の補正のための検証研究の分類と、それらの入手可能性、異なる分析戦略の使用可能性、および複数の関連への可能性に関する一般的な考え方<sup>47)</sup>

	Validation study		
	Internal	External	
Information on disease outcome	Yes	Yes	No
Availability	Rare*	Rare	Frequent
Analytic methods			
Multiple imputation	Yes <sup>†</sup>	Yes <sup>†</sup>	No
2-stage sampling	Yes	No	No
Propensity score calibration	Yes	Yes	Yes
Validation study can be used to adjust multiple associations	Yes (if random sample)	Yes	Yes
Transportability of parameters <sup>‡</sup>	Yes	?	?

\*：可用性は実現可能性と同じではない。内部検証の研究は多くの場合実現困難かもしれないが、特定の条件下では実行可能である。

†：内部妥当性検証研究には、計画的に疾患の予後に関する情報が含まれている。しかし、特定のバリデーションスタディでは、サイズが小さい又は、発生率が低いために、複数の帰属が適用されるため、対象の疾病予後に関する十分な情報が得られない可能性がある。

‡：検証研究からのモデルとそれらに関連するパラメータは、主研究に適用することができる。これは通常、内部検証研究で与えられるが、外部検証研究では仮定する必要がある（つまり、疑わしさがある）。（筆者記）

Table 5-9 はデータベースを用いた薬剤疫学研究における未測定交絡要因の補正に関する論文の検証研究を整理した表である。検証研究は、研究における主研究に用いる集団のより詳細な追加情報を収集する内部検証研究（internal validation study）と、異なる研究として実施される外部検証研究に分けられ、得られるパラメータの性質が異なると説明されている。

#### 5.6.1.1. 内部検証データ (internal validation study)

内部検証研究は言い換えれば nested (or sub) cohort 研究の結果の利用であり、通常、感度分析のための研究の実施が必要である。当然、多くの労力とリソースが必要であり、Table 5-9でも実現可能な場合は稀であるとされているが、有病率や曝露頻度の影響を受ける陽性的中度をバリデーション指標として用いる場合は、陽性的中度の数学的性質上、研究に用いるのと互換性のある対象集団における曝露又はアウトカムの層別の測定が必要であるため、内部検証研究以外から得られた陽性的中度を用いる場合はパラメータに幅を持たせる、仮定について丁寧に説明するなど工夫が必要である。また、後で述べる外部検証データに対してより狭い信頼区間が得られることが報告されている<sup>48)</sup>。なお、内部検証研究を行った場合はその利点を十分に生かすため、レコードレベルのデータを用いて欠損値補完 (multiple imputation of missing data<sup>49)</sup>) の手法が用いられることがある。

内部検証研究の具体的な研究方法としては、無作為または何らかの基準で抽出した一部の集団 (sub cohort) について曝露の区別にアウトカムである診断やアウトカムの区別に曝露である薬剤の使用の精度を評価して感度、特異度 (又は陽性的中度、陰性的中度<sup>注43)</sup>) を測定する、喫煙や既往歴などの重要な交絡要因のアウトカムとの関係 (オッズ比等) や存在割合の情報を収集する。条件は厳しいが、実施する主研究の計画内ではない先行研究など異なる研究計画により同じデータ、同じ曝露、アウトカムについて行われた検証データが利用可能な場合もある (これから行う研究で用いるアウトカムの識別アルゴリズムなど識別条件に用いるデータが整っている必要がある。). (参考 3 参照)

内部検証データの利点は、本研究と同じ分布特性を有していることを保証できれば、確率とランダム誤差 (全体で又は生年齢、曝露などの層毎に) によってバイアスなく当てはめることが可能であり、“仮定”を置く必要がないことである。このためには主研究に用いる集団の最尤推定と見なせる必要がある (つまり内部検証データに対して研究データは MCAR (完全にランダムな欠測) 又は MAR (ランダムな欠測) の欠測データと見なせる)。主研究の対象から無作為抽出された sub cohort であれば MCAR として単純な利用が可能である (参考 4 参照)。

この事例としては、米国の保険請求データ Medicaid Analytic eXtract (MAX) を用いた妊婦の抗うつ薬の使用と子の心臓障害の関係を評価した研究では<sup>50)</sup>、他の障害を含む複数の疾患・障害 (子癇前症、持続性肺高血圧症、先天性心臓奇形、右心室流出路閉塞および他の心臓奇形) の診断について、病院の診療録を確認する大規模な internal validation study が行われた<sup>51)</sup>。この検証研究では Figure 5-11 のように、病院の診療録にアクセス可能な場合に限られ要求したうちレビューできたのは 24~43%であったことなどからランダムなサンプルと見なせないことが限界であるが、属性の確認などから、全体への一般化は可能であると説明され、曝露による差異はなく非差異誤分類であるとされた。結果は薬剤ごとに統計処理されており、アウトカムの誤分類を補正したオッズ比の変化 (補正前を基準として) は 1.3%から 9.7% (変化の最大は補正前オッズ比 1.13 から補正後 1.24) の範囲であった。

注43： 陽性的中度、陰性的中度を用いる場合は注意が必要である 5.4.3.3 参照

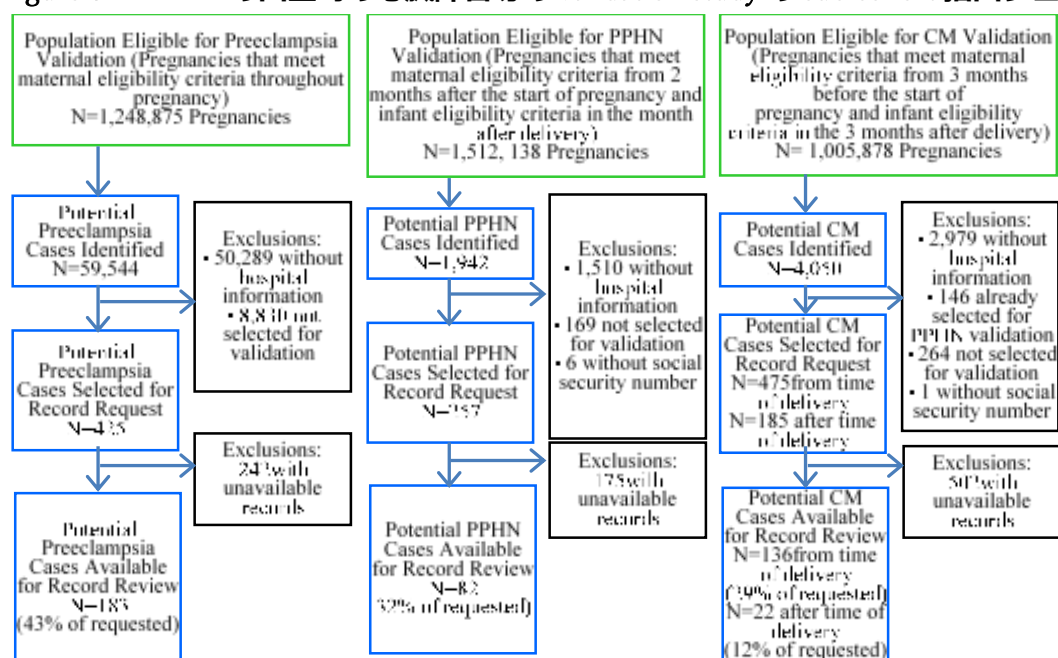
Figure 5-11 MAX の出生時の心臓障害等の validation study の sub cohort 抽出フロー<sup>51)</sup>

Figure 1. Flow chart of pregnancies included in the validation study, Medicaid Analytic eXtract, 2000–2007

### ✓ 可能性のある部分集団を用いる (“all possible cases”)

しかし、研究に用いる医療情報データベースが内部対象研究を実施可能な場合でも、アウトカムの感度と特異度を測定するためには研究対象の集団からランダムな標本を抽出して真のアウトカムなし/ありを評価するため、例えば頻度 1% のアウトカムを 100 例得るには 1 万例の標本が必要になるように、稀なアウトカムを扱う場合に十分な数の真の陽性を得るためには大きな標本が必要となるが、検証研究においてカルテ（の臨床経過の記述）や画像のレビューなどが必要な場合は困難である。こういった場合の工夫として、対象集団のうち、データベースから得られる情報からアウトカムが陽性である可能性のある集団（ケースが多く含まれると期待されるサブグループ）と可能性の極めて低い集団に分けて、前者のみを対象とする方法がある。この前者は“all possible cases” と呼ばれる<sup>52)</sup>。医療情報データベースを用いる場合は傷病名、検査、処置など複数の条件を組み合わせることでアウトカムを定義することが少なくない。こういった条件の何れか 1 つでもある場合を“all possible cases”として扱う。同時に、この方法は複数の選択肢が考えられるアウトカム定義のうち最も効率の良い定義を得ることにもなる。この“all possible cases”を扱う方法は合理的に見え、魅力的であるが、“all possible cases”以外の群に真の cases が（全く、またはほとんど）含まれないという弱い仮定が必要で、この条件が成立しない場合にこの方法で得られる感度は過大なものとなる。実際、「Validation study, all possible cases, diagnosis」といったキーワードで PubMed を検索しても、多くの研究は感度ではなく、陽性的中度が報告されている。<sup>注44</sup>

注44： 陽性的集度と感度から特異度を求めることが可能で（5.6.2 参照）、不確実性を考慮できる感度分析では有用な方法であるとする。



この事例として内部検証研究ではないが、米国の高齢者を対象とする公的医療保険システムである Medicare の保険費償還データベースにおける脳卒中アルゴリズムの検証を紹介する<sup>53)</sup>。この検証は、米国の脳卒中のコホート研究（REasons for Geographic And Racial Differences in Stroke (REGARDS) Study<sup>54),55)</sup> のデータとカルテレビュウによって行われた。REGARDS は、脳卒中の地理的要因と人種によるリスクを評価する目的でランダムに選択した 45 歳以上の白人とアフリカ系黒人に文書による同意を得て行われた約 3 万人のコホート研究である。データは初回に面接で、以降 6 か月ごとに電話によるインタビューによって収集された。つまり、アウトカムである卒中関連イベントは自己申告によるものであった。REGARDS のデータには社会保障番号が含まれているため、REGARDS と Medicare のデータをリンクさせ、両者のデータの何れかで脳卒中と思われる対象者を“all possible cases”として全症例のカルテレビュウによって検証が行われた。結果は脳卒中の感度、特異度、陽性的中度、陰性的中度について Table 5-10 のように属性ごとに報告されている。また、この REGARDS Study を用いて、Medicare と他のコホート研究のデータを突合（リンク）させて研究に用いることを評価する研究が行われ、リンクできた集団の結果の全体への汎化可能性（外挿性）の評価がおこなわれている<sup>56)</sup>。

Table 5-10 REGARDS のカルテレビュウの結果<sup>53)</sup>

Accuracy Measures Stratified by Age, Sex, and Race in Any Stroke Cohort						
Algorithm	Category	No.	PPV (95% CI), %	Sensitivity (95% CI), %	Specificity (95% CI), %	NPV (95% CI), %
AIS	Age					
	65 y	1915	85.4 (70.6-100)	44.2 (29.3-59.1)	99.8 (99.6-100)	98.7 (98.2-99.2)
	65-74 y	9713	90.7 (85.3-96.0)	56.0 (48.8-63.2)	99.9 (99.8-100)	99.2 (99.0-99.3)
	≥ 75 y	3461	87.8 (81.6-93.9)	66.4 (58.7-74.1)	99.6 (99.4-99.8)	98.6 (98.2-99.0)
	Sex					
	Female	7804	92.0 (87.1-96.9)	62.6 (55.4-69.8)	99.9 (99.8-100)	99.2 (98.9-99.4)
	Male	7285	85.8 (87.1-96.9)	55.2 (48.3-62.2)	99.7 (99.6-99.9)	98.8 (98.5-99.0)
	Race					
	Black	5574	93.0 (88.3-97.6)	66.1 (58.8-73.3)	99.8 (99.7-100)	99.0 (98.7-99.2)
	White	9515	86.5 (80.6-92.5)	53.3 (46.5-60.1)	99.8 (99.7-99.9)	99.0 (98.8-99.2)
ICH	Age					
	< 65 y	1915	—	—	—	—
	65-74 y	9713	85.4 (71.3-99.5)	67.3 (50.7-84.0)	100.0 (99.9-100)	99.9 (99.8-100)
	≥ 75 y	3461	93.8 (81.0-100)	69.4 (49.9-88.8)	100.0 (99.9-100)	99.8 (99.7-100)
	Sex					
	Female	7804	86.7 (71.0-100)	77.8 (59.6-96.0)	100.0 (99.9-100)	99.9 (99.9-100)
	Male	7285	91.3 (79.8-100)	61.6 (45.3-78.0)	100.0 (99.9-100)	99.8 (99.7-99.9)
	Race					
	Black	5574	81.4 (64.4-98.5)	82.5 (65.7-99.2)	99.9 (99.9-100)	99.9 (99.9-100)
	White	9515	95.2 (86.1-100)	59.1 (42.5-75.6)	100.0 (100-100)	99.9 (99.8-99.9)
AIS/ICH	Age					
	< 65 y	1915	86.1 (71.9-100)	45.5 (30.7-60.3)	99.8 (99.6-100)	98.7 (98.2-99.2)
	65-74 y	9713	92.2 (87.7-96.8)	57.9 (51.3-64.5)	99.9 (99.8-100)	99.1 (98.9-99.2)
	≥ 75 y	3461	90.1 (84.9-95.4)	67.9 (60.7-75.0)	99.6 (99.4-99.8)	98.4 (98.0-98.8)
	Sex					
	Female	7804	93.1 (88.8-97.4)	64.2 (57.5-71.0)	99.9 (99.8-100)	99.1 (98.9-99.3)
	Male	7285	88.3 (83.1-93.5)	57.3 (50.9-63.7)	99.8 (99.6-99.9)	98.6 (98.4-98.9)
	Race					
	Black	5574	93.1 (88.9-97.4)	68.8 (62.1-75.5)	99.8 (99.7-99.9)	99.0 (98.7-99.2)
	White	9515	89.3 (84.3-94.3)	54.6 (48.3-60.9)	99.8 (99.7-99.9)	98.8 (98.6-99.1)

Abbreviations: AIS, acute ischemic stroke; ICH, intracranial hemorrhage; NPV, negative predictive value; PPV, positive predictive value.

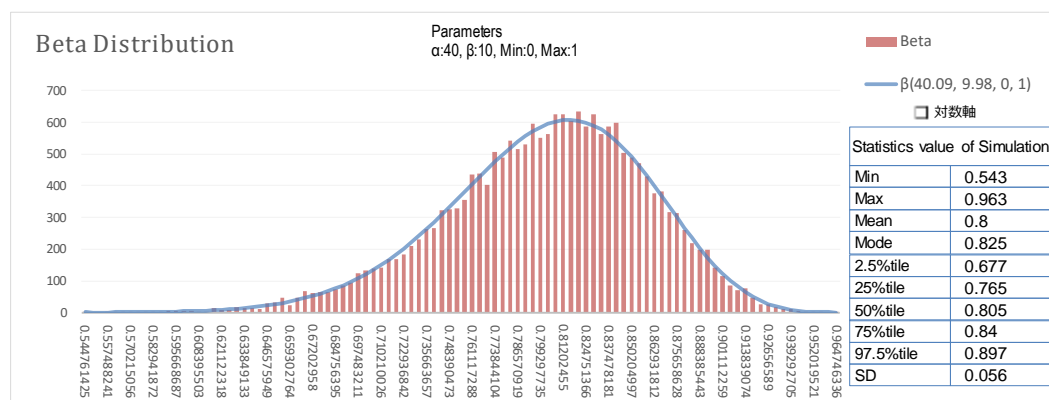


### 参考 3 validation study の実施と結果の利用について

validation study は費用、労力と時間を要し実施が困難な場合が少なくないが、実施する主研究の最良の感度分析パラメータと成り得ると同時に主研究を改善する多くの情報が得られるため、validation study によって研究全体の縮小を必要とする場合であっても、研究のリソースの相当な割合を費やすべきであるとされている。<sup>4)</sup> validation study を行う場合は Strom B.L.

(2012)<sup>57)</sup>の CHAPTER41 に患者の自己申告、保険償還データベース等について方法、事例がレビューされている。その他曝露の誤分類（非差異）についての validation study のデザインと例数設計<sup>58)</sup>の方法、validation study を含めた本研究の設計について<sup>59)</sup>などが報告されている。validation study を行おうとする場合はこれらを参考に感度分析を前提とした無駄のない研究が望まれる。例えば、感度と特異度の何れが重要であるのか、あるいは陽性的中度のみで足りるのかなど<sup>注45</sup>、必要な情報に応じて望ましい validation study は全く異なるものとなる。措置（規制）により曝露を減らした場合のアウトカム減少割合に興味があるのであれば、曝露の誤分類は重要ではない場合もある<sup>60),61)</sup>。case-control 研究におけるアウトカムの特定は感度よりも特異度の影響が大きく、通常の診断技術評価が感度を重視するのとは異なり、特異度を優先する必要があると指摘されている<sup>62)</sup>。アウトカムの特異度は本研究の対象の抽出条件に強い影響を受けると共に稀な事象の場合には僅かな特異度の違いでも影響が大きく、曝露の有無の間に差異がある差異誤分類の場合は更に大きなバイアスとなるが、このような特異度の調査は膨大な労力を要する。

検証データが本研究のランダムな標本であるとみなせる場合は、検証データの点推定値と標準偏差などからパラメータの分布を設定できる。感度や特異度など比率は二項分布に従うが、その信頼区間はベータ分布により求めることができる。一般に、標本数を  $n$ 、出現数を  $k$  とすると、その確率分布は第 1 パラメータ  $k$ 、第 2 パラメータ  $n - k$  のベータ分布に従う。（つまり  $\beta(k, n - k)$ ）例えば、本研究のアウトカムありからランダムに 50 例を抽出して検証した結果、陽性的中度（Ppv）は 0.8（40/50）であった場合、 $Ppv \sim \beta(40, 10)$ （下記分布）として確率的感度分析を行うことができる。



注45：パラメータに陽性的中度を用いる場合は注意が必要である。詳細は 5.4.3.3, 5.6.2 を参照。

#### 参考 4 架空のコホート研究における内部検証研究による感度分析の結果

下記表の架空の1万人のコホート研究について validation study として 10%にあたる 1,000 人をランダムに抽出して測定されたアウトカムの感度と特異度を用いて感度分析を行うと、補正後リスク比の中央値は理論値に近いが、頻度区間はかなり広い。validation study を倍の 2,000 人としても 95%頻度区間は 0.96~8.49 に過ぎず、3,000 人でようやく 1.12~5.72 である。測定が容易な陽性的中度、陰性的中度を用いた場合<sup>注46</sup>、研究の曝露有無、アウトカム有無毎にランダムな 50 人を用いた場合の結果は下部のとおりで、理論値のリスク比 2 に対して 1.47 と大きくずれている。これは Figure 5-12 のとおり、陰性的中度の分布が広くなるため、陰性的中度=1 に固定した場合の補正後 RR は 2.04（95%頻度区間 1.41~3.01）となる。このように稀な生起確率に対するランダムな抽出による validation study は効率が悪いので、効率化のため関心のある部分集団のみを対象とする方法がある。性別、年齢、医療機関等を層としてランダムサンプリングを行うには問題がないが、特に cohort 研究では内部検証研究の対象はベースラインの情報をういて抽出することになるため、曝露やアウトカムを層とする場合は工夫した手法（anamorphic design<sup>63)</sup> または 2-stage design<sup>64)</sup> と呼ばれる<sup>47)</sup>）が提案されている。

		曝露あり	曝露なし
Observed	アウトカムあり	188= A <sub>1</sub>	144= A <sub>0</sub>
	アウトカムなし	4812= B <sub>1</sub>	4856= B <sub>0</sub>
	計	5000= A <sub>1</sub> + B <sub>1</sub> = N <sub>1</sub>	5000= A <sub>0</sub> + B <sub>0</sub> = N <sub>0</sub>
	発現率	3.76%	2.88%
	RR(95%信頼区間)	1.31 (1.05 - 1.62)	
アウトカムの真の発現割合		2%	1%
Adjusted	アウトカムあり	A <sub>1</sub> <sup>*</sup>	A <sub>0</sub> <sup>*</sup>
	アウトカムなし	B <sub>1</sub> <sup>*</sup>	B <sub>0</sub> <sup>*</sup>
	計	A <sub>1</sub> <sup>*</sup> + B <sub>1</sub> <sup>*</sup> = N <sub>1</sub> <sup>*</sup> = N <sub>1</sub>	A <sub>0</sub> <sup>*</sup> + B <sub>0</sub> <sup>*</sup> = N <sub>0</sub> <sup>*</sup> = N <sub>0</sub>
validation study	感度 <sup>#1</sup>	13.5/15=0.9	
	特異度 <sup>#1</sup>	965.3/985=0.98	
感度分析	パラメータ感度 = Se	F 分布 <sup>#2</sup>	
	パラメータ特異度 = Sp	F 分布 <sup>#2</sup>	
	RR <sub>systematic_error</sub> (95%頻度区間)	2.01 (0.66~16.88)	
validation study	陽性的中度 <sup>#3</sup>	24/50=0.48	15/50=0.3
	陰性的中度 <sup>#3</sup>	50/50=1	50/50=1
感度分析	パラメータ陽性的中度 = Pvp	F 分布 <sup>#2</sup>	F 分布 <sup>#2</sup>
	パラメータ陰性的中度 = Pvn	F 分布 <sup>#2</sup>	F 分布 <sup>#2</sup>
	RR <sub>systematic_error</sub> (95%頻度区間)	1.47 (0.34 - 6.29)	
Se : アウトカム分類の感度, Sp : アウトカム分類の特異度			
$A_1^* = \frac{A_1 - (1 - Sp)N_1}{Se + Sp - 1}, B_1^* = \frac{B_1 - (1 - Se)N_1}{Se + Sp - 1}, N_1^* = N_1$			

注46 : 陽性的中度、陰性的中度を用いる場合は注意が必要である。5.4.3.3 を参照

$$A_0^* = \frac{A_0 - (1 - Sp)N_0}{Se + Sp - 1}, B_0 = \frac{B_0 - (1 - Se)N_0}{Se + Sp - 1}, N_1^* = N_1.$$

$Pvp$  : アウトカム分類の陽性的中率,  $Pvn$  : アウトカム分類の陰性的中率

$$A_1^* = PvpA_1 + (1 - Pvn)B_1, A_0^* = PvpA_0 + (1 - Pvn)B_1$$

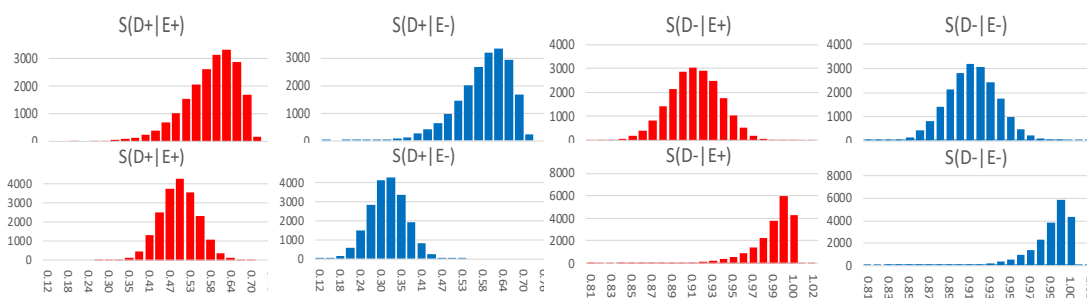
$$B_1^* = (1 - Pvp)A_1 + PvnB_1, B_0^* = (1 - Pvp)A_0 + PvnB_0.$$

#1 : 理論上の感度を 0.9, 特異度を 0.98 としてランダムな標本 1000 人による validation study を行った場合の理論上の頻度

#2 : 比の確率分布は生起数を  $x$ , 試行数を  $n$  とするとき第 1 自由度  $d1=2(n-x+1)$ , 第 2 自由度  $d2=2x$ ,  $u$  を  $0 \sim 1$  の一様分布とすると,  $p=d2/(d1 \cdot F \text{ 分布の逆関数}(u, d1, d2)+d2)$  により得られる. 結果は Figure 5-12 のとおり.

#3 : 理論上の感度を 0.9, 特異度を 0.98 として曝露の有無とアウトカムの有無の組合せ各々からランダムな標本 50 人による validation study を行った結果

Figure 5-12 感度分析パラメータ分布 上段 : 感度/特異度, 下段 : 陽性的中率, 陰性的中



#### 5.6.1.2. 外部検証データ (external validation study)

研究対象集団を含む集団について内部検証以外の方法で測定された検証データである. 研究者自ら外部で実施する検証研究 (external validation study), 先行研究, 同じポピュレーションに対する異なるデータベースの報告など, 様々な情報源が該当する. 測定精度の評価・検証目的で行われたもの (external validation data) と罹患頻度や曝露頻度などの参考情報 (external information) に区別できる.

一般論としては, バイアスの原因に関する量 (誤分類では感度, 特異度など) が複数の報告 (情報源) で一貫していることや, 年齢, 性別, 地域, 疾患, 入院/外来の区分などの層による偏りがなく, 研究対象集団との背景の類似性などを根拠として用いることができ, そのような場合, 外挿の不確かさを考慮したパラメータの分布 (特に広さ) を指定することになると考える

例えば, 多くの疫学研究に用いられている米国の高齢者及び障害者の公的保険給付システム (Medicare) では, Medicare Current Beneficiary Survey としてサンプルした被保険者に多目的のインタビュー調査が行われており, 請求データにはない喫煙などの生活習慣の情報が得られている<sup>59),19)</sup>. また, 内部検証データで紹介した米国の Medicare のデータである MAX の

出生時の障害を調査した研究は MAX を用いて妊産婦、出生児に関する研究の外部検証データとして複数の研究で利用されている

曝露、アウトカムなどの測定精度の場合は、主研究と同じ方法（識別のアルゴリズムなど）のデータが含まれている必要がある。研究対象と同じ集団又はデータベースを用いていても非確率的抽出であるなど主研究に用いる集団の最尤推定として見なせない場合は外部検証データである。本邦においても独立行政法人 医薬品医療機器総合機構の MIHARI Project 以外に 4.1.8 で紹介している Takeuchi A の日本動脈硬化症研究コホートを用いて動脈硬化性疾患について感度分析を行った論文<sup>25)</sup>では、血圧値に基づく高血圧の区分を血圧が 2 回測定された患者のデータを用いてパラメータが設定されている。このように、より正確な測定が行われた部分集団のデータを用いることが考えられるが、その測定の理由（2 回測定したのは 1 回目が高い値であったためかもしれない）が新たなバイアスの原因となりうる。また、山口 拓洋

（2015）<sup>65)</sup>は用いた静脈血栓塞栓症発症、出血性イベントについて、Shiro T（2016）は大腿骨折について<sup>66)</sup> レセプト又は DPC データを用いる研究のために一部の施設を対象に検証研究を行っている。そのほか、大規模ではあるが情報（データ項目等）が限られる医療費償還データベースを用いた研究の外部検証データとして詳細ではあるが小規模な診療録データベースの利用や<sup>14)</sup>同じ地域の医療機関の診療録データベース<sup>11)</sup>が利用可能で有効である場合もある。また、我々製薬企業であれば、製造販売後調査（使用成績調査）のデータから対象とする有害事象の処置に特定の医薬品が使用される頻度など、役立つ情報が得られる場合もある。確率的サンプリングにより統計的に代表性を担保する検証研究を行うことができないため、このような方法は現実的で感度分析に重要な知見を得られる。5.3 で診療録の傷病名の記録（付与）について説明したように、医療機関の診療録データベース及びそこから作成される診療報酬明細書（レセプト）との関係も様々で、施設によりデータ生成過程で差異（情報バイアス）が生じる可能性があること、研究対象の集団とのズレによる選択バイアスを生じる可能性がある。こういった検証研究の結果を感度分析の情報（パラメータ）に用いる場合は検証データ（施設）の多様性を確認することを推奨する。

#### 5.6.1.3. 文献，統計情報など

自ら外部検証研究を行う、あるいは用いるデータについてあらかじめ広範に行われた外部検証研究の結果を用いるのではなく、文献や疾患、公衆衛生に関する調査の統計情報なども情報源となり得、報告（結果）について事後的に感度分析を行う場合は主要な情報源になると考えられる。

文献レビューから測定されていない交絡要因や誤分類による偏りを補正するための情報を収集するが、利用可能な報告が複数あり、その値がばらつく場合は系統的レビューによる統合（meta-analysis）結果を用いることができる<sup>68)</sup>。米国の Sentinel Initiative の活動による医療情報データベースの二次利用による研究で必要となるアウトカムを定義するアルゴリズムのバリデーションの系統的レビューのための文献データベースの検索、レビュー方法<sup>70)</sup>とその結果の報



告<sup>71)</sup>は異なる医療環境、データベースであることに留意は必要であるが文献の検索式の設定、レビュー方法の参考になる。しかし、海外で行われた validation study の結果（アルゴリズム）を医療環境の異なる本邦で用いる場合はその妥当性の評価が必要である<sup>注47</sup>。その他 validation study としてではないが、検査や画像検査による診断精度が報告されていることがある。例えば、入門編の事例として用いた肺炎については、その定義が傷病名の付与のみであったため用いていないが、アルゴリズムとして胸部の画像検査の実施を含めることが考えられる。この場合、胸部画像検査を用いた診断精度がアウトカムの感度・特異度となるが、胸部のレントゲン画像とヘリカル CT 画像を用いた場合の肺炎の診断確率の報告<sup>69)</sup>は感度のパラメータとして利用可能であろう。また、内部検証データ以上に外部検証データ、文献、統計情報には偏り、誤差を含むことを理解して感度分析のパラメータに分布を設定できる PSA を用いるなど不確実性を考慮する必要がある。利用可能な情報を広く探すことは対象とする研究により適した情報を得られる可能性があると同時に、偶然かもしれないが、感度分析実施者にとって都合の良い情報に引き付けられる可能性もある。こういったバイアスを避ける意味でも、先に紹介した系統的レビューの方法と結果の統合は有用であろう。

#### 5.6.1.4. 専門家の意見

内部・外部の検証データ、文献、統計情報の何れも利用できないあるいは文献、統計情報はあるが、一貫性がないなどの理由で不十分な場合があり、専門家<sup>注48</sup>の意見や経験による感度分析パラメータの設定が行われている。各専門家にパラメータの推定値と幅について意見を求める場合、専門家の意見は環境、選択的な経験、好みなどの影響を強く受けやすく、しばしば偏った情報を与えながら偏りは過小評価されやすいため、たった一人の専門家の意見によるのではなく、5.6.1.3 文献、統計情報などの「系統的レビューによる統合」と同様に複数の専門家の意見を統合することが推奨されている<sup>73),74)</sup>。意見を求める場合に専門家自身の経験のみでなく、文献調査により得られた文献の報告を踏まえた意見を求めることで、より一般的な意見が得られると思われる。しかし、専門家自身の経験や考えに沿った報告が重視されるなどバイアスが生じる可能性、文献の公表バイアスの影響に注意しながらであるが、meta-analysis の手法を用いて専門家の意見からパラメータを設定することは可能であると考えられる。

また、大規模な診療情報データベースでは多くの非専門家による診療情報が含まれることが予想されるが、これら非専門家の診断、治療に専門家の意見（ロジック）を当てはめることはバイアスにつながると想像することは容易であろう。診療情報データベースを用いる研究ではアウトカムや曝露の特定を利用可能なデータから識別（類推）するためのアルゴリズムを用いる必要がある。このアルゴリズム自体を専門家に意見を含める場合に専門家と非専門家の診断手順（アルゴリズム）、選好の違いにも留意が必要であると考えられる。

注47： 米国で保険償還データベースにおける abacavir の過敏症反応を Recursive Partitioning を用いたアルゴリズムを作成して validation study が行われているが<sup>72)</sup>、同じ方法を日本でも実施されることが the 30<sup>th</sup> International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Taipei, October 2014 で報告されている。

注48： ここでいう専門家は、研究の対象となる疾患、アウトカムの臨床の専門家に限らず、疫学研究社、用いるデータソースの運用者などが含まれる。



## 5.6.1.5. 結果が無意味または有意になる限界値

5.6.1.1～5.6.1.4 で説明したエビデンスに基づくバイアスの仮定を用いるのではなく、研究の結果の有意性が無くなる（例えばオッズ比が 1）、反対に意味のある大きな差（例えばオッズ比が 2）になる感度分析パラメータを逆算する方法である<sup>19), 75)</sup>。意思決定の基準を限界値とすることであらかじめ設定する感度分析パラメータを少なくできる（事前設定を行わないことも可能であるが結果の提示が難しくなる）。結果は単純な固定値ではなく線形関数または分布となる。未測定交絡要因の場合であれば、数式 5-1 から、曝露の各層における未測定交絡要因を有する割合  $P_{C0} = P_{C1}$  の場合は  $RR_{DC}$  に関係なく  $RR_{adj} = RR_{obs}$  であるが、それ以外は

$$RR_{DC} = 1 - \frac{RR_{adj} - RR_{obs}}{RR_{adj} \cdot P_{C0} - RR_{obs} \cdot P_{C1}} \quad \text{数式 5-6}$$

$$\begin{aligned} P_{C0} &= \frac{RR_{obs}}{RR_{adj}} \left( P_{C1} + \frac{1}{RR_{DC} - 1} \right) - \frac{1}{RR_{DC} - 1} \\ &= \frac{RR_{obs}}{RR_{adj}} \cdot P_{C1} + \left( \frac{RR_{obs}}{RR_{adj}} - 1 \right) \cdot \frac{1}{RR_{DC} - 1} \end{aligned} \quad \text{数式 5-7}$$

を得られる。 $RR_{obs}$  は研究から得られ、 $RR_{adj}$  は限界値として任意に定められるが、固定できない 3 個の変数が残る。数式 5-6 から目標とする  $RR_{adj}$  と  $P_{C0}$ ,  $P_{C1}$  を任意に定めれば、未測定交絡要因のリスク比 ( $RR_{DC}$ ) を得られ、数式 5-7 からは任意の  $RR_{DC}$  における  $P_{C0}$ ,  $P_{C1}$  が得られる。例えば、第 1 部 2.2 で用いた WTCHR の数値例 (Table 5-11) では  $RR_{obs} = 1.732$  であるが、 $RR_{adj} = 1$  となるのは、未測定交絡要因  $RR_{DC}$  が 2, 3, 4, 5 及び 10 の場合で Figure 5-13 のように、 $P_{C0}$ ,  $P_{C1}$  の関係は傾き  $= RR_{obs}/RR_{adj}$  の直線となる。また、未測定交絡要因の場合、バイアスの大きさは①交絡要因とアウトカムの関連の強さ ( $RR_{CD}$ )、②交絡要因と曝露の関連の強さ ( $RR_{EC}$ )、③集団における交絡要因の併存率 ( $P_{C0}$ ,  $P_{C1}$ ) により決まるが、 $P_{C0}$ ,  $P_{C1}$  を定めない“仮定を置かない方法” (Without Assumptions) がある<sup>76)</sup>。この方法では、交絡要因の併存率を問わず数式 5-8 により補正後のリスク比又はオッズ比の下限が計算されるため、最も影響を受ける場合の値が得られるが、同時に保守的な扱いであるため、注意が必要である。

$$\begin{aligned} RR_{CD} &= \frac{(A_{11} + A_{01})(N_{10} + N_{00})}{(A_{10} + A_{00})(N_{11} + N_{01})}, \quad RR_{EC} = \frac{N_{11}N_{00}}{N_{10}N_{01}} \\ RR_{adj} &\geq RR_{obs} / \frac{RR_{EC}RR_{CD}}{RR_{EC} + RR_{CD} - 1} \end{aligned} \quad \text{数式 5-8}$$

アウトカムの誤分類について、実際の検証研究でしばしば行われる陽性的中度のみが測定される場合は、数式 5-5 により

$$Npv_0 = 1 - \frac{[Ppv_1 \cdot A_1 + (1 - Npv_1)B_1] \cdot n}{m \cdot RR_{adj} \cdot B_0} + \frac{Ppv_0 \cdot A_0}{B_0} \quad \text{数式 5-9}$$

とできるが、不定のパラメータが 3 個あるため、やはり一意に定まらない。例えば、WTCHR の数値例において、 $Ppv_1 = Ppv_0 = 1$  の場合に補正後のリスク比  $RR_{adj}$  が 1, 2 及び  $RR_{adj} = RR_{obs}$  となる陰性的中度 ( $Npv_1$ ,  $Npv_0$ ) は Figure 5-14 のようになる。ここで

$RR_{adj} = 2$  の場合,  $Npv_1 = Npv_0$  であれば, 約 0.365 である。これは, 曝露ありで 5896 人, なしで 9865 人が偽の陰性であり, 感度は各々 0.075 と 0.044 であるので, 通常あり得ないであろう。  $RR_{adj} = 1$  の場合は  $Npv_1 = Npv_0 = 0.5346$  であるので感度は 0.099 と 0.059 であり, やはり通常あり得ないであろう。また,  $Npv_1 = Npv_0$  という仮定は感度が曝露の群間で異なる (差異誤分類) ことを意味している。  $RR_{adj} = 1$  の場合において感度に差異が生じない陰性的中度は  $Npv_1 = 0.40$ ,  $Npv_0 = 0.66$  の場合で  $Se_1 = Se_0 = 0.079$  である。

このように, 関数によって評価することも可能であるが, やはり複雑で関数からグラフを作成して視覚的に確認することになるため, パラメータに一様分布を用いた PSA の結果を用いるのが簡便である。その方法などについては 5.6.3.1 target-adjustment sensitivity analysis に具

Figure 5-13 WTCHR の数値例で補正後リスク比

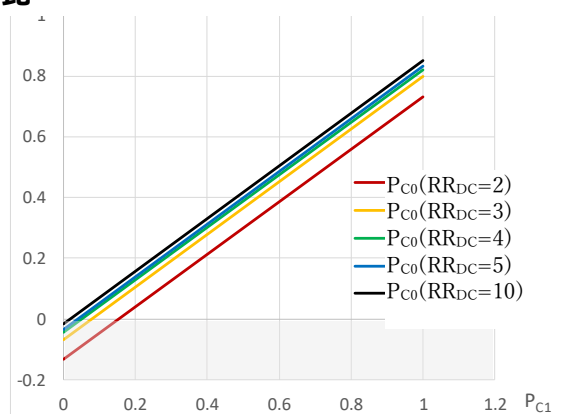
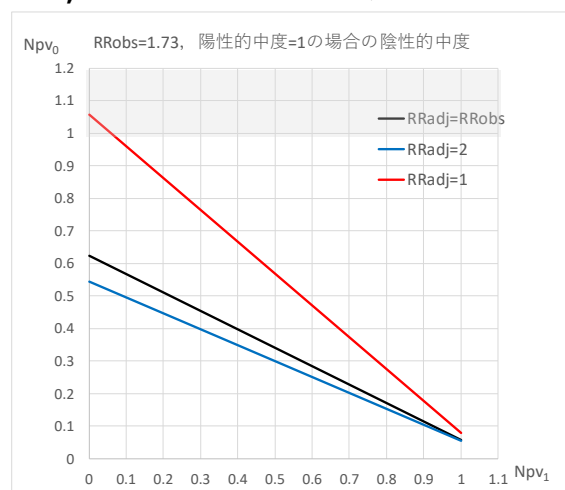


Figure 5-14 WTCHR の数値例で補正後リスク比=1, 2 及び観測値と等しくなる陰性的中度



体的に例示している。

#### 5.6.1.6. 数理的制約

感度, 特異度, 陽性的中度, 陰性的中度は 0~1 の間しか値を取り得ない<sup>注49</sup>, 頻度 (クロス表のセル) は負の値になり得ない, 誤分類の場合は計算式中に感度+特異度-1 による除算があるため, 感度+特異度=1 の場合は計算できないなど, 数理的な制約があることは少なくない。このような制約の範囲はパラメータの分布特性の制約条件となる。補正結果のセルがあり得ない値となるのは事前のセルの値及び感度分析モデルに依存するため予め特定するのではなく, MCSA のシミュレーションの結果から除くことで処理される<sup>7),77),78),注50</sup>。

注49: 定義としては 0~1 であるが, 2 値 (有無, 陽性陰性など) ではランダムな場合で感度および特異度は 0.5 であるので, 現実には 0.5~1 である。

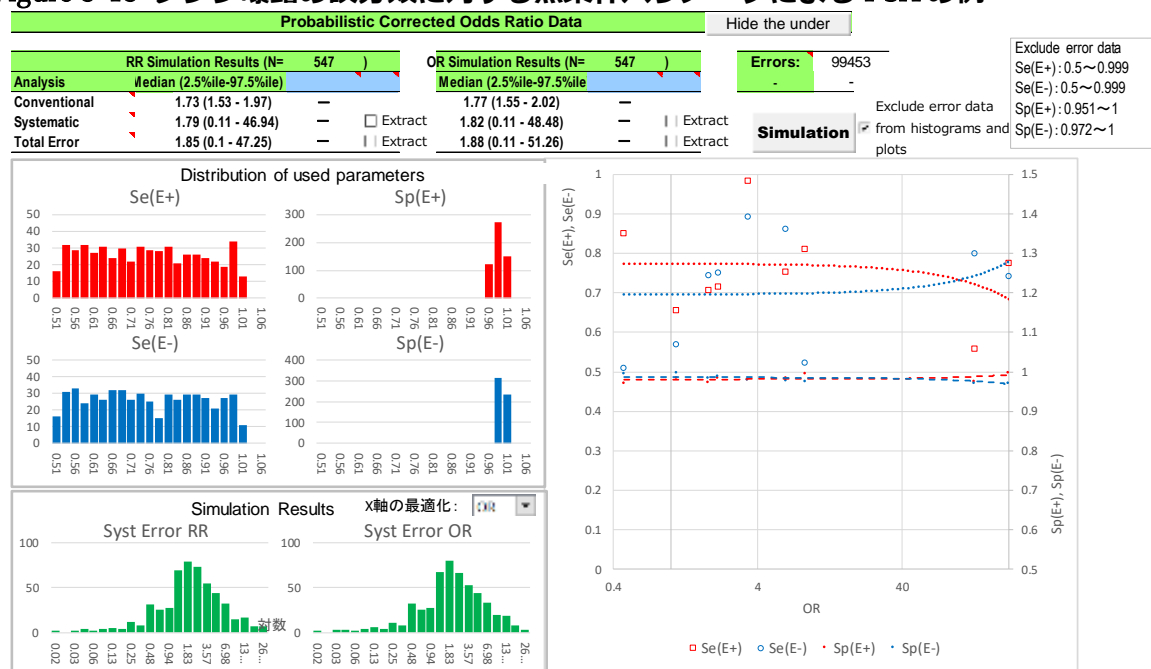
注50: 除かない方法もある。5.6.5 参照

WTCHR の数値 (Table 5-11) を用い、アウトカムの誤分類を例に示す。Figure 5-15 は各感度、特異度に無情報事前分布である  $0.5 \sim 1$  の一様分布を与えて 100,000 回の PSA を行った結果である<sup>注51</sup>。負のセルの発生などにより、100,000 回のサンプリング中、99,453 回のエラーが生じている。実際に用いられた (エラーを除いたもの、右上の「Exclude error data」) パラメータの  $Se$  は与えたものと同じであるが、 $Sp$  は粉塵+で  $0.951 \sim 1$ 、粉塵-で  $0.972 \sim 1$  となっている。この結果から、数理的制約から少なくとも感度 ( $Se$ ) は  $0.5 \sim 1$ 、特異度 ( $Sp$ ) は粉塵+は  $0.95 \sim 1$ 、肺癌-は  $0.97 \sim 1$  の範囲内となる。なお、このような方法では MCSA のサンプリングの多くが無効になる場合があるため、十分なサンプリング回数を設定することに注意する。エラーによりあらかじめ与えたパラメータと用いられたパラメータが異なる場合の結果の提示などについては 5.7.2 感度分析の結果を提示する で説明する。

Table 5-11 WTCHR の数値例を交絡要因の有無で層別した分割表

アウトカム： ぜん息	未知の交絡要因					
	合計 (元のデータ)		$C = 1$		$C = 0$	
	曝露：粉じん		曝露：粉じん		曝露：粉じん	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	$A_1 = 476$	$A_0 = 450$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
$D = 0$	$B_1 = 9,286$	$B_0 = 15,536$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
合計	$N_1 = 9,762$	$N_0 = 15,986$	$N_{11}$	$N_{01}$	$N_{10}$	$N_{00}$

Figure 5-15 レジン曝露の誤分類に対する無条件パラメータによる PSA の例



注51 : アウトカムや曝露の有無を全くランダムに決めても正解確率の期待値は 50%であるため。なお、陽性的中度、陰性的中度にはこの規則は成立しない。

## 5.6.2. 利用可能な情報の種類と計算方法

第1部第3章および第2部5.4.2などで感度分析の各モデルにおける計算方法を紹介してきたが、例えば誤分類であれば、感度・特異度を用いるか、陽性的中度・陰性的中度を用いるか、未測定（未知）交絡要因であれば、併存率を絶対値、相対値（オッズなど）を用いるかなど、同じバイアスについて異なる計算式が存在する。これは、パラメータ（バイアス）に関する利用可能な情報、研究の種類及び指標（リスク比、オッズ比など）により使い分けることとなるが、5.4.3.3のアウトカムの誤分類における陽性的中度・陰性的中度のように、曝露の有無などの層間の偏りの有無の直感的な反映ができない場合がある。実施の容易性から検証研究により陽性的中度（ $Ppv_0 = 0.9$ ,  $Ppv_1 = 0.85$ ）と“all possible cases”（5.6.1.1 参照）により近似的な感度（ $Se_0 = Se_1 = 0.8$ ）が測定されているとする。誤分類を補正する計算式は通常、感度と特異度または、陽性的中度と陰性的中度の組み合わせであるので、差異/非差異を扱いやすい特異度を用いる場合、

$$Se = \frac{A^* Ppv}{A}, Sp = \frac{B^* Npv}{B}, B^* Npv = B^* - \left(\frac{1}{Se} - 1\right) A^* Ppv = B^t, \\ B = B^t + (1 - Ppv) A^* \text{ より}$$

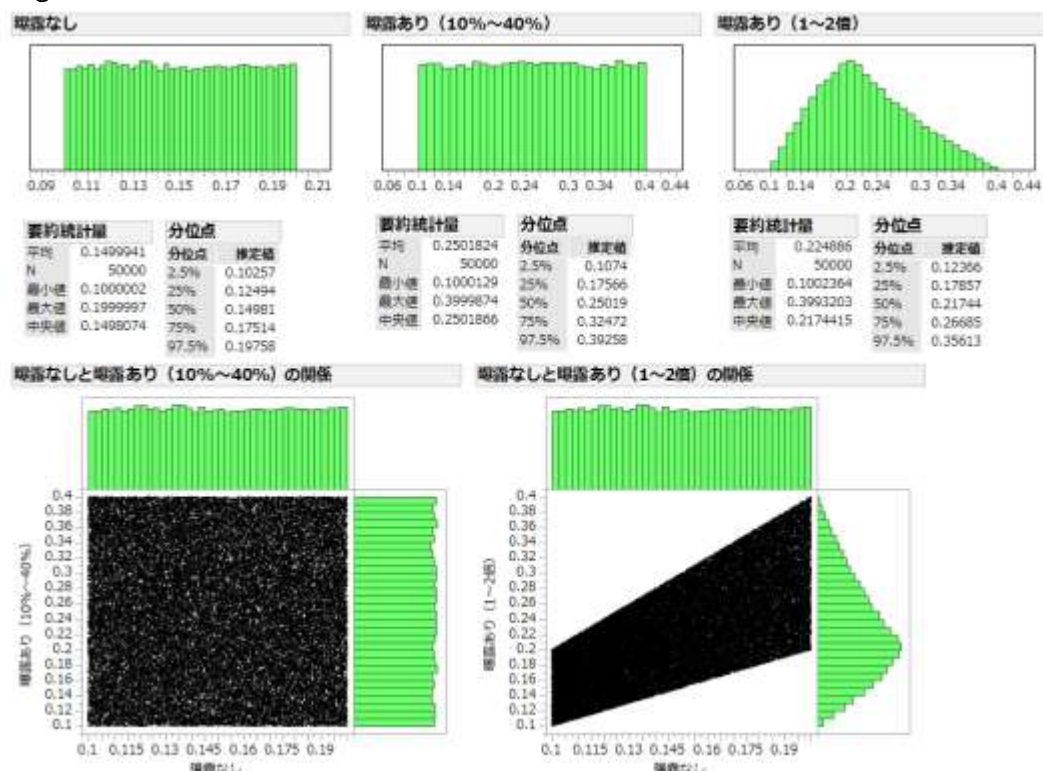
$$Sp = \frac{B^* - (Ppv/Se - 1)A^* - A^* + A^* Ppv}{B^* - (Ppv/Se - 1)A^*} = 1 + \frac{(1 - Ppv)A^*}{B^* - (Ppv/Se - 1)A^*} \quad \text{数式 5-10}$$

として陽性的中度と感度から特異度が求まる。数式 5-10 から第1部の WTCHR の数値例（Table 5-11）におけるアウトカムの誤分類を例にすると  $Sp_1 = 0.992$ ,  $Sp_0 = 0.997$  であり、 $Ppv_0 = Ppv_1 = 0.9$  では  $Sp_1 = 0.992$ ,  $Sp_0 = 0.997$  である。非差異として  $Sp_1 = Sp_0 = 0.997$  となるのは  $Ppv_0 = 0.9$ ,  $Ppv_1 = 0.945$  の場合である。

感度分析の方法に PSA を用いる場合、相対的なパラメータの扱いには工夫と注意が必要である。

例えば、未測定（未知）交絡要因を曝露なしでは 10%～20%（ $P_0$ ）が有しており、曝露ありではその 1 倍～2 倍であるという情報から  $P_0 \times 1 \sim 2$  とする（ $P_1$ ）、あるいは曝露ありを 10%～40%（ $P_1'$ ）として独立に扱う方法も考えられる。これを一様分布とした場合の要因の併存割合パラメータの分布は Figure 5-16 のようになる。 $P_1$  と  $P_1'$  の最小値と最大値はほぼ等しいが、分布の形状は全く異なり、 $P_0$  との関係も異なる。前述の陽性的中度と感度の場合は、陽性的中度に一様分布を用いた場合、特異度も一様分布であるが、 $Ppv_0 = Ppv_1$  は非差異を意味しないため、感度分析の計算式に数式 5-10 を組込むのではなく、 $Se_0$ ,  $Se_1$  に分布を与える。これらのように、パラメータに用いる情報、仮定をより正しく反映できる方法を選ぶ必要がある。

Figure 5-16 未測定交絡要因の併存率の設定方法によるパラメータ分布の比較



### 5.6.3. 感度分析パラメータに値と分布を割り当てる

利用可能な情報源から感度分析モデルのパラメータを設定するが、未測定又は未知の交絡要因について感度分析を行う場合に測定された交絡要因の影響を上回る設定を用いるのは過剰であるという意見がある<sup>79),13)</sup>。測定された交絡要因の調整による変化が僅かな場合に、それより遥かに大きな影響を与える未測定交絡要因があると仮定することは過剰な補正につながる可能性がある。パラメータの設定に感度分析の対象とする要因の知られている影響度の大きさやその確からしさも考慮してパラメータを設定する必要があると考える。

ここでは、パラメータを選ぶための考え方、意味の面から説明する。

#### 5.6.3.1. target-adjustment sensitivity analysis<sup>4)</sup>

補正後のオッズ比が1以下、あるいは2以上になるなど、設定した目標の値となる場合を求める手法である。つまり、5.6.1.5の目標値(限界値)を用いる場合の手法である。この手法は感度分析のモデルと目標値さえ決めれば、事前に感度分析パラメータに具体的な値、分布を決める必要がなく意思決定論的には簡便でかつ理解が容易であるとされている<sup>4)</sup>。5.6.1.5のように、感度分析のモデルの数式を変形して目的の値となるパラメータを解とする関数からその組み合わせを計算することができる。WTCHRの数値例では特徴がないが、Table 5-2の数値例でアウトカムの誤分類に陽性的中度=1の場合の陰性的中度と補正後リスク比( $RR_{adj}$ )の関係はFigure 5-17のようになり、 $RR_{adj} = 1$ の場合は $Npv_1 \leq 1$ 、 $Npv_0 > 1$ であるので、



$RR_{adj} \leq 1$  となることはない。このような特殊な場合を除いて、多くの場合、パラメータは3個以上あるため、一意の値とはならず煩雑であるので<sup>注52</sup>、PSA 又は Multidimensional analysis の結果から条件を満たすパラメータを取り出すほうが容易である。具体的には、“広めの”パラメータを設定した PSA (MCSA) を行い、補正済みオッズ比又はリスク比が限界値以下又は以上となる組み合わせを抽出してパラメータを確認することで、限界値となるパラメータが得られる。あるいは、任意又は無条件（例えばアウトカムの感度、特異度を 0.5~1 の一様分布とする）のパラメータを用いた PSA を行い、補正済みオッズ比又はリスク比が限界値以下又は以上となるサンプリングの頻度を求めることで、限界値の生じやすさを求めることができる。

[5.8 教科書的ではないが現実的で簡便な手順による感度分析 参照]

Figure 5-18 は WTCHR の数値例 (Table 5-11) について、未知交絡要因のリスク比 ( $RR(C|D)$ ) を 2~10、曝露群、非曝露群の要因保有率 ( $p(C+|E+)$ ,  $p(C+|E-)$ ) を独立に 1%~100% (全て一様分布) として PSA を行い、補正後リスク比 $\leq 1$  となる場合を抽出 (Figure 5-18 の赤枠のオプションを設定) した結果である。補正後リスク比 $\leq 1$  となるのは計算したシミュレーションのうち 18.7%、パラメータは  $p(C+|E+)$  が 0.124~1,  $p(C+|E-)$  が 0.01~0.521,  $RR(C|D)$  が 2~10 の範囲なのであるが、その組み合わせのうち 21%であること (つまりこの範囲でも 79%は補正後リスク比 $>1$ )、 $p(C+|E+)$  は  $p(C+|E-)$  の概ね 2 倍以上である場合であることがわかる。反対に補正後リスク比 $>1$  となるのは、 $p(C+|E+)$  及び  $p(C+|E-)$  が 0.01 以上の全ての範囲で生じることも確認できる。Figure 5-18 の右下の散布図は補正後リスクの値で色分けしたパラメータの散布図行列である。ここから補正は  $p(C+|E+)$  と  $p(C+|E-)$  の比に依存し、補正後リスクが 1 以下となる組み合わせ中で補正後リスクが 1 を超える場合は僅かであること (水色の中に緑のドットがわずかにある) が確認できる。

また、用いる場合は以下の点にも留意が必要と考える。

- ・ 目標値とする効果として意味のある値の選択は恣意的であり議論の対象となる場合があるため、目標値の意味、根拠を説明できるようにする。
- ・ 限界値を超えるパラメータの組み合わせが複数生じ得るため研究の結果が誤っていることを強調する方向に印象付ける場合がある。特に既に知られている値 (因果関係が確立している曝露・アウトカムのオッズ比など) を目標値とする場合はパラメータに“尤もらしさ” (疑わず鵜呑みにされる) を与え結果の評価に影響する場合がある。

注52：MEMO:第1部 3.1.4 の補正したオッズ比を求める未測定交絡要因の場合の式、3.3.2 曝露の誤分類の場合の式を見てほしい。非差異誤分類であれば、逆算は可能であるが、差異誤分類の場合はこれらの式から補正オッズ比が目標値となるパラメータを逆算するのは容易でない事は明らかであろう。

Figure 5-17 Table 5-2 の数値例で補正後リスク比=1, 2 及び観測値と等しくなる陰性的中度

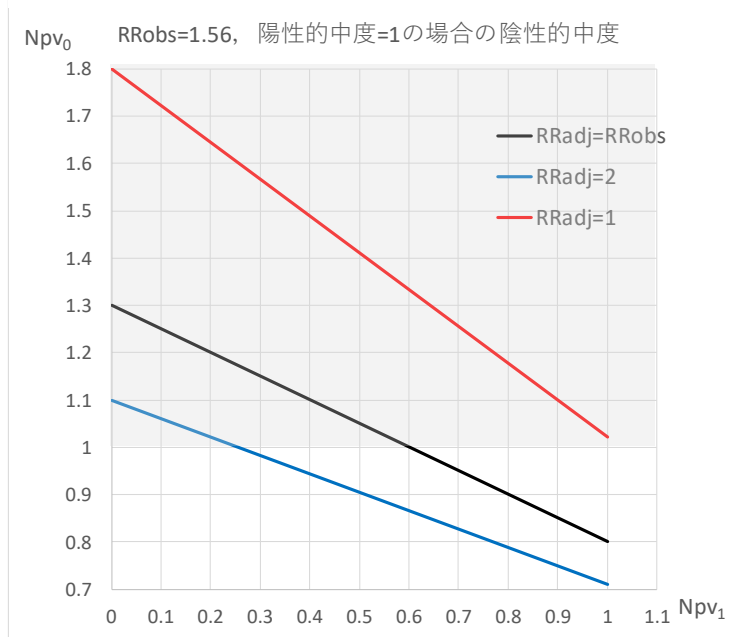
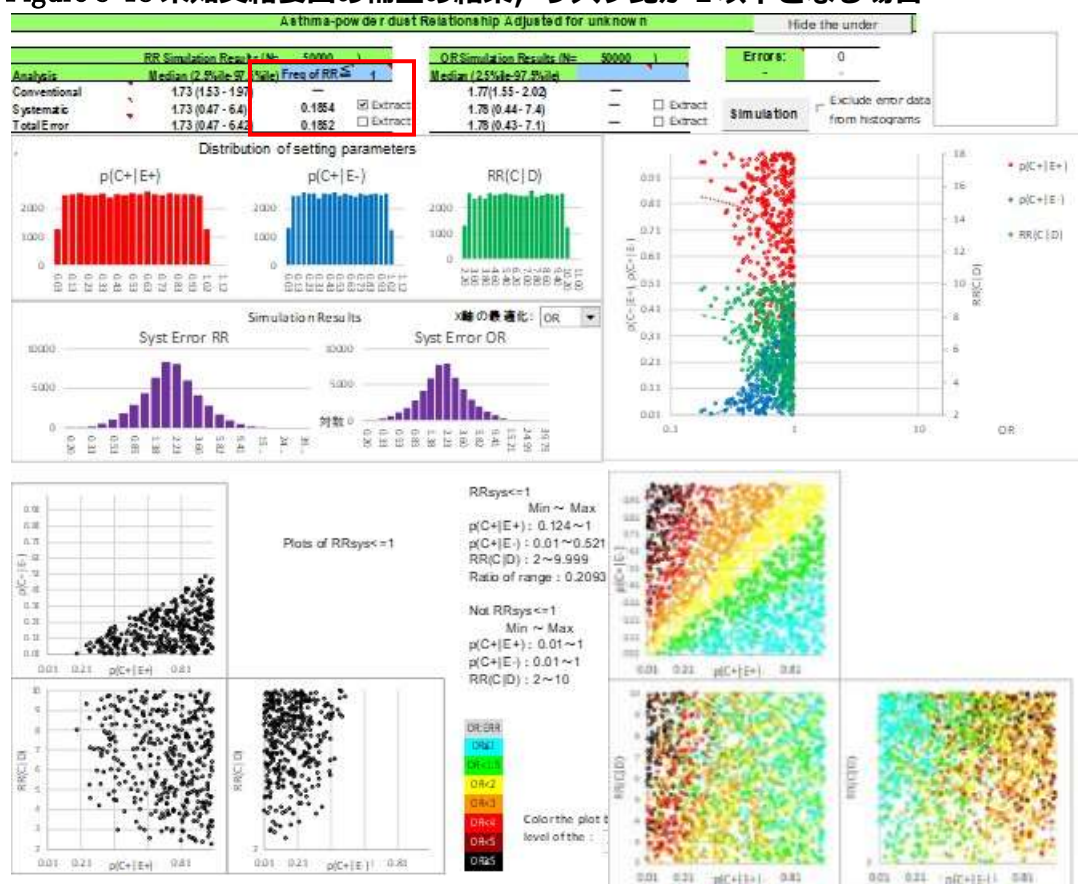


Figure 5-18 未知交絡要因の補正の結果, リスク比が1以下となる場合



### 5.6.3.2. fixed bias-parameter analysis

fixed bias-parameter analysis (FBA)<sup>4)</sup>は感度分析のパラメータに分布を用いず、固定した単一の又はいくつかの組み合わせを用いる方法であり、第1部第3章で説明した Simple sensitivity analysis によって行われる。目標値（限界値）を用いるのではなく内部検証データ、外部検証データ、文献調査あるいは専門家の意見 [5.6.1 参照] から得られた“点”推定値によりパラメータに固定値（の少数の組み合わせ）を設定する。他の手法を含めて、研究者は控えめな値を用いることも反対に不確かさを過小評価することを避けるために妥当な（あり得そうな）範囲内で極端な値を選択することが可能であるが、固定値であるため、その意図の影響が大きくなる。このため特に用いた値を選んだ理由や値の意味の説明が重要である。例えば状況として次のようなものが考えられる。

- ・ 情報源が本研究から○%をランダムに抽出して調査した内部検証データであるので代表性があり、ばらつきを考慮しても（多少パラメータの値を変えても）結果が安定している。
- ・ 外部検証データであるが、本研究と背景が似通っている、あるいはいくつかの要因層別間で差異が少なく外挿可能であると考えた。
- ・ 文献調査の結果、類似の●件の研究で報告されており、その値は□～△と一貫性を認められた。
- ・ パラメータを大きく変えた場合でも結果への影響は小さかった。

結果の提示は単一の値を用いるとしても、研究者（感度分析の実施者）は結果を説明するためパラメータの異なる組み合わせによる補正結果のリストなどから補正された推定値の一貫性、パターンからパラメータの結果への作用、影響度を理解することは意味がある。情報源が内部検証データであっても例えば診断に必要な検査の未実施などにより判定不能なケースが生じるなど何らかの不確かさは存在することに留意し、このことを踏まえた確認も役立つと考える。

例えばアウトカムの特定の内部検証の結果が Table 5-12 であったとする。感度は  $80 \div 100 = 0.8$ 、特異度は  $190 \div 200 = 0.95$  となるが、判定に必要な検査の未実施等により 100 例（25%）に不明が生じている。不明/陰性（validation/研究）の 90 例のうち 20 例が本当は陽性（研究で陰性症例の陽性期待値の 20/210 の約 2 倍）残りは陰性、不明/陽性の 10 例が全て陽性（先の 20 例と合わせて不明 100 例中陽性 30 例であり検査が行われた場合の陽性期待値とほぼ等しい）であるとすれば、感度は  $100 \div 130 = 0.769$  特異度は  $260 \div 270 = 0.96$  である。validation 不明の 100 例全てが陰性の場合、感度は 0.8 のままであるが特異度は  $280 \div 300 = 0.933$  と低下する。

Table 5-12 内部検証データの数値例

		研究			研究は診断，検査の実施，処置の実施から判別．Validation は検査結果と診断と処置を組合せて判別，不明は検査は陽性だが診断も処置もない場合，診断又は処置はあるが検査未実施の場合
		陽性	陰性	計	
validation	陽性	80	20	100	
	陰性	10	190	200	
	不明	10	90	100	
	計	100	300	400	

### 5.6.3.3. probabilistic sensitivity (or bias) analysis

probabilistic sensitivity (or bias) analysis (確率的感度解析：PSA) では，固定値（の組み合わせ）を与える FBA の代わりにパラメータに分布を割り当てる．分布の代表値（中央値や最頻値）と広がり，密度は FBA に用いるのと同じ情報源から設定することもできる．分布の関数として表し方は 4.1 で説明しているので，その設定，選び方を説明する．

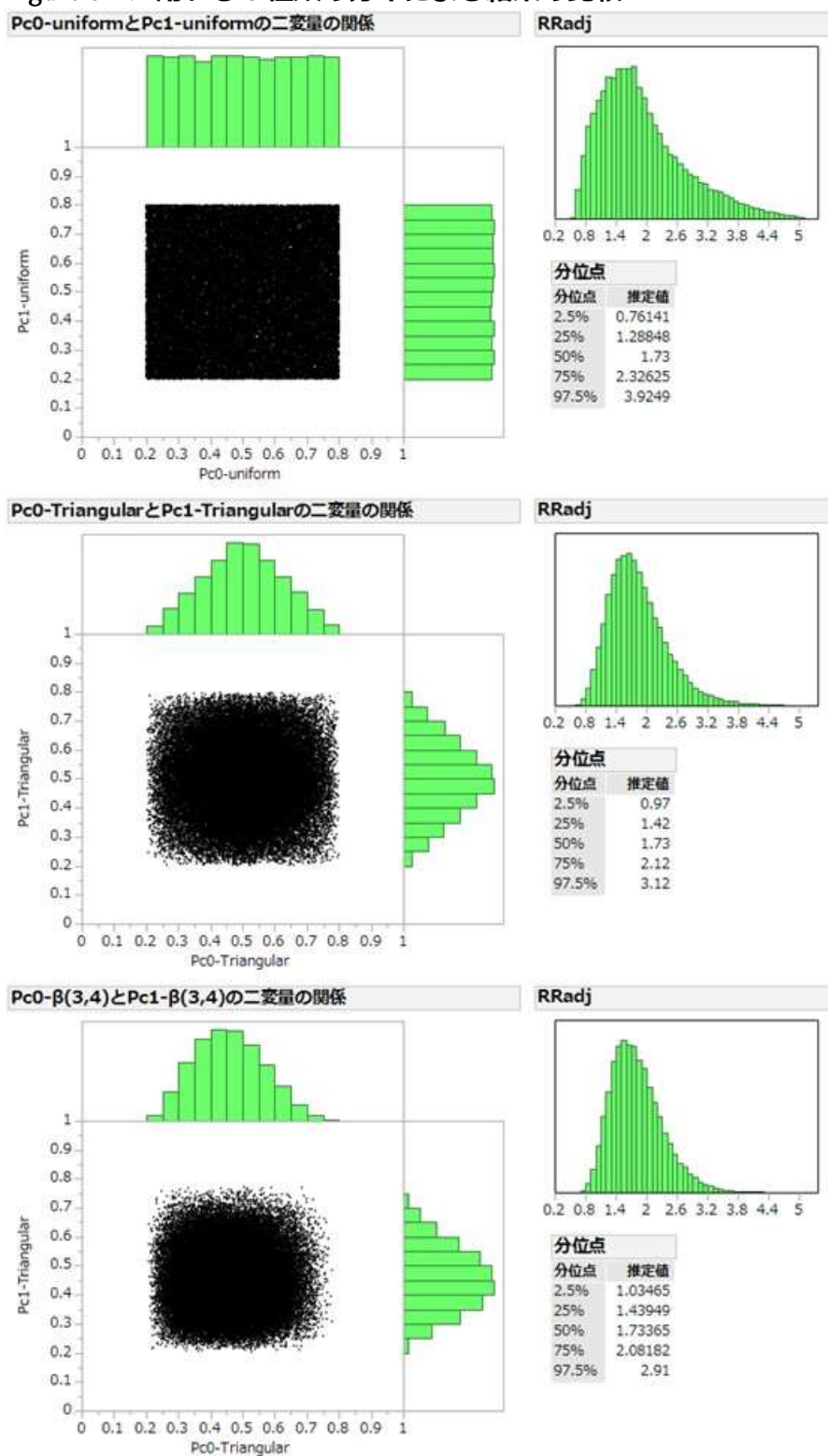
検証データ等からパラメータの値の範囲を得ることはできても分布形状の情報は得られず，任意に決めなければならない場合が多いであろう．検証データから点推定値と 95%信頼区間が得られている場合でも「信頼区間」は 95%の確率で真値が含まれているが，一方で，どこかの辺りに真値が位置するかの情報が得られるのではない．しかし感覚的にはベイズ流の考えから（ベイズでは「信用区間」という）正規分布や二項分布を仮定して両端の頻度が小さくなる分布形状を想定して三角分布やベータ分布を当てはめるのが自然と考えられる（4.1.3 参照）．実際に WTCHR の未知交絡要因（C）の感度分析を未知要因のリスク比  $OR_{DC}=10$  に固定して  $P_{C1}$ ， $P_{C0}$  は独立に最頻値 0.5，最小 0.2，最大 0.8 として一様分布（uniform），三角分布（trigonom），ベータ分布（ $\alpha=3$ ， $\beta=4$ ）を用いた結果は Figure 5-19 に示すとおりである（ $P_{C1}=P_{C0}$  の制約なし）．補正された  $OR_{DE}$  の 95%頻度区間（2.5%～97.5%）は一様分布が最も広がっているが，中央値は補正のないオッズ比（1.76）より小さな値となること，約 1/4 でオッズ比が 1 以下となることは同じである．用いる分布形状により補正結果の不確実性（頻度区間の広さ）に影響を与えるが，分布のパラメータである最頻値（又は中央値），最小値，最大値が同じであれば，つまりこれら分布パラメータに比べて分布形状の影響は多くの場合で限定的である<sup>注53</sup>．また，目的が，精度の良い感度分析を行うことである場合は分布の設定は重要であるが，「オッズ比が 2 以下になるのは」，「大きなバイアスを考慮すると・・・」といった頑健性を示すことが目的の場合には分布の形状は重要でなく，むしろパラメータの影響が明確になる一様分布が適している場合があるであろう．

ただし，未測定交絡要因のリスク比，オッズ比や選択バイアスの Selection OR に一様分布を用いる場合は常用対数とする．詳細は③ 参照．

注53：Beta 分布，正規分布，対数正規分布，その他複雑な分布関数を用いる場合，与えたパラメータは例えば 0.7～1 の範囲を前提としている場合でも，分布の裾が薄く長く，実際に得られる乱数は 0.9～1 のようになる場合がある．このような場合は分布の形状が結果に大きく影響するので注意が必要である．（ESATJ では様々な分布を近似するベータ分布が得られる．Appendix2-Appendix3.1）-1-E)参照）



Figure 5-19 用いる 3 種類の分布による結果の比較





## ✓ 検証データからパラメータを考える

では、validation study の結果からパラメータを設定する場合について考えてみる。

5.6.1.1 内部検証データ（internal validation study）の Table 5-10 のカルテレビューによるアウトカムの精度評価のように、validation study により、性・年齢や何らかの疾患の有無などの層毎に報告されている結果（感度、特異度など感度分析に用いるパラメータ）に大小があり、それが単にバラツキではなく喫煙率の例のように層によって差異があると考えられ、その層の構成が曝露やアウトカムの層で異なる（と考えられる）場合は、これらの情報を用いてシナリオ（バイアスのパラメータ）を設定することが考えられる。

Table 4-1 の数値例の未測定交絡要因である喫煙率について仮定の例示を行う。4.1.8.5 では差異はないとしたが、Table 4-1 の論文に用いられた研究は、変圧器製造施設におけるレジン樹脂を含む 7 種類の化学物質、樹脂の曝露について工場勤務者（退職者を含む）の死亡データとコントロールとして心疾患関連死の死亡データを用いて case-control 研究が行われており、全員が白人男性のデータあった<sup>9)</sup>。また、cases と controls の間では年齢区分はマッチされているが、樹脂の曝露の有無間ではバランスが取れていないため年齢構成に差異が生じ得るとする。仮に Table 4-1 は Table 5-13 の年齢構成であり、住民調査から性・年齢別の喫煙率が測定されていたとする。曝露ありの喫煙率は年齢区分平均と同じ、曝露なしは年齢区分平均×1.1、95%信頼区間は年齢区分平均×0.8～×1.2 と仮定して曝露の層毎の喫煙率をシミュレーションした結果（Figure 5-20）、曝露ありの喫煙率は平均 0.565、標準偏差 0.148、曝露なしは 0.622、0.164 となった。曝露の有無毎の喫煙率の相関係数を 0.7、喫煙と肺癌のオッズ比は 4.1.8.5 と同じ（平均 2.15、標準偏差 0.28）として PSA を行った結果（Figure 5-21）の補正後オッズ比（OR<sub>adj</sub>）は中央値 1.9、95%頻度区間 1.3～2.69 である。非差異の設定の結果（Figure 4-28）は 1.76、1.64～1.9 であったので、補正によりレジン曝露の肺癌リスクは僅かに大きくなる結果となった。

Table 5-10 の例でも、アウトカムの感度は 65 歳未満は 45.5 であるのに対して 75 歳以上は 67.9 と約 1.5 倍になっている。年齢や性別であれば、多くの場合は情報が得られるので、層間のこれらの構成から差異の有無と具体的な感度分析のパラメータ（分布）を得ることは比較的容易であろう。

Table 5-13 年齢区分毎の人数構成と外部調査による想定喫煙率

年齢区分	E=1（率）	E=0（率）	喫煙率
-40	50 (0.17)	80 (0.08)	0.6
-50	60 (0.20)	150 (0.14)	0.7
-60	70 (0.23)	300 (0.29)	0.6
-70	80 (0.26)	200 (0.19)	0.5
70-	42 (0.14)	309 (0.30)	0.4

Figure 5-20 曝露ごと喫煙率のシミュレーション結果

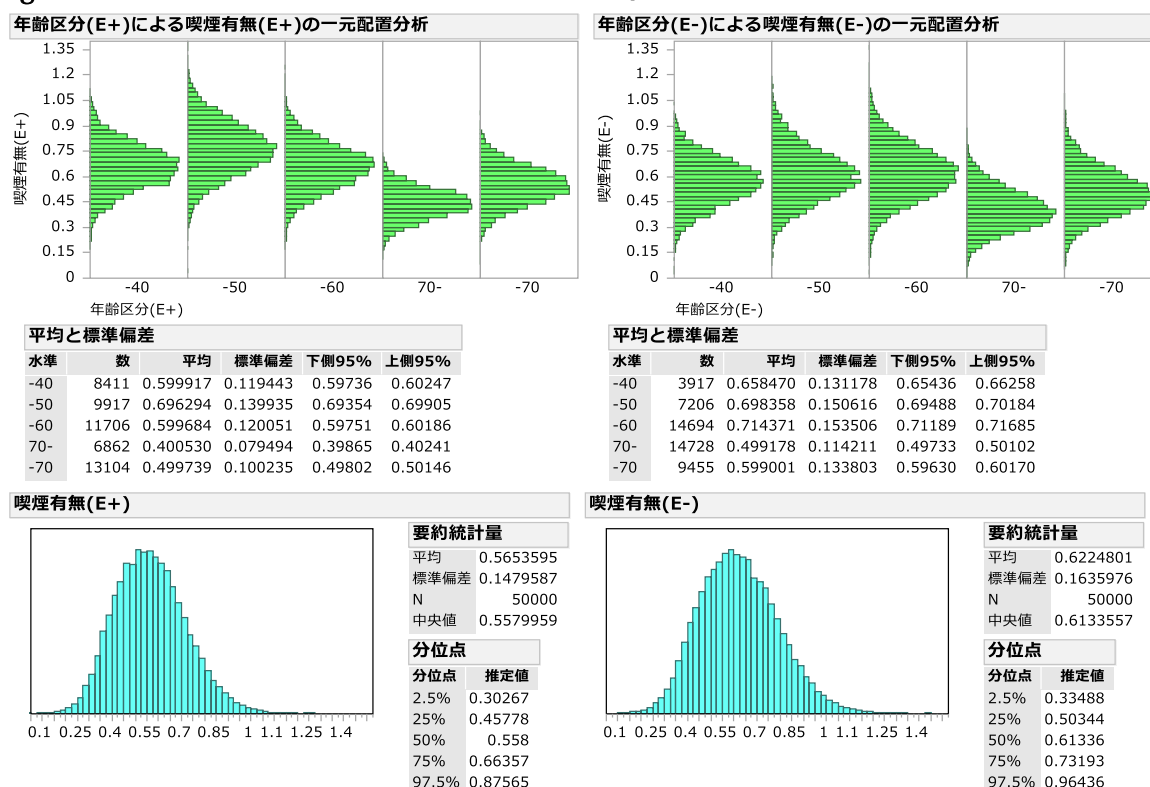


Figure 5-21 喫煙率に差異を用いた結果



#### 5.6.4. パラメータ特性の影響評価（シナリオの反映の確認）

この工程は Simple sensitivity analysis, Multidimensional analysis では不要である。

PSA（確率的感度分析）は用いるパラメータの分布の形状と広がり感度分析の結果に影響を与える。分布の形状は感度分析の結果を中央値と 95% 頻度区間（下側 2.5% と上側 97.5% の値）のみを示すならばあまり影響はないかもしれないが（Figure 5-19 参照）、分布の形状・広がり感度分析の結果を示すことは感度分析の説明に重要であり、少なくとも PSA で用いる MCSA の各サンプリングをテーブルとして出力し、用いたパラメータと補正された値の関係を多角的/視覚的に確認すべきである。加えて、対数正規分布のようにシナリオの最小値、最大値、最頻値を平均値と標準偏差によって表現する必要があるなど、意図したシナリオが表現できている事の確認も必要である。本書が提供する ESATJ では、様々な分布関数（ユーザーが作成した関数を含む）から実際に乱数を生成し、ヒストグラムと統計量を表示して視覚的に確認できるようにしている。また、0 から 1 の範囲以外の感度/特異度、負の頻度など理論上ありえない結果の除外は与えたパラメータの分布を変更することになるのでその影響の評価も示すことを忘れてはならない。（5.6.1.6 数理的制約, 5.7.1 感度分析の方法を提示する 参照）

複数のパラメータを独立に設定（MCSA におけるランダムサンプリング）する場合にパラメータ間の関係から矛盾が生じることがあるため、例えば、MCSA のシミュレーションの組み合わせのリストを出力してまず補正結果のヒストグラム又は box-plot から異常値のないことを確認し、更に各パラメータのヒストグラムを用いて分布を比較してその形状を確認する。もしも補正結果または中間の変数（セルの値など）に異常値がある場合、その異常値の場合のパラメータの組み合わせを確認することでパラメータを適切に修正する。

説明のため、WTCHR（Table 5-11）の未知交絡要因 C のシナリオを

- ・ 未知要因と喘息のリスク比は最頻値 4、範囲を 1 から右は十分に大きい値として 20 程度で分布の両端はほとんど生じない対数正規分布型の分布、
- ・ 未知要因を有する割合は曝露ありは平均 0.5、なしは平均 0.3 でどちらも範囲は 0~1、分布の裾が正規分布型の分布、
- ・ 曝露ありの未知要因を有する割合が曝露なしより大きい（分布が右にシフトしている）。

これを表現するパラメータとして

- ・ 未知要因と喘息のオッズ比  $RR(C|D)$  は平均 4.5 ( $=\ln(1.5)$ )、標準偏差 0.4、
- ・ 粉塵曝露有の未知要因を有する割合  $P(C+|E+)$  は平均 0.5、標準偏差 0.3、
- ・ なしの  $P(C+|E-)$  は平均 0.3、標準偏差 0.2 の正規分布<sup>注54</sup>、

注54：シナリオから  $P(C+|E+)$  と  $P(C+|E-)$  の標準偏差は 0.25 と 0.15 程度が適切であるが、説明のため問題のある分布の例として大きな値としている。

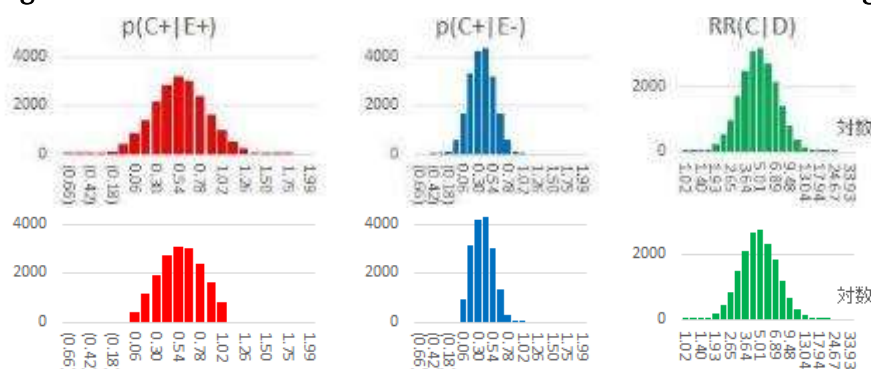
- ・  $P(C+|E+)$ と $P(C+|E-)$ の相関係数に0.7

を用いた MCSA (2 万回) を行った。ESATJ では Figure 5-22 のように、2 万回中、2871 回のエラー（取りえないパラメータの値や負のセル）が生じていることが表示され、発生したパラメータとエラーとならず使用されたパラメータの最小・最大、四分位などが表で表示される。ここから、 $P(C+|E+)$ および $P(C+|E-)$ に0以下と1以上が生じていること $RR(C|D)$ は若干、分布の右側でエラーが多く、平均など分布が左寄りになっていることが確認できる。さらに、「Exclude error data from histograms」をチェックすると、エラーとなる値を除いたヒストグラムが表示される。Figure 5-23 でその前後を見比べると、 $P(C+|E+)$ と $P(C+|E-)$ の分布の両端が切り落とされている。

Figure 5-22 ESATJ の結果



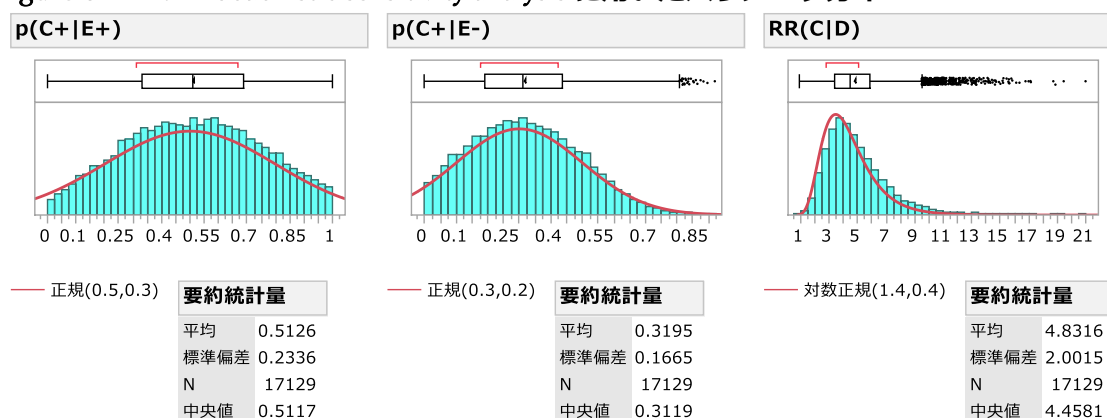
Figure 5-23 パラメータの分布（下は「Exclude error data from histograms」の結果）



では、より詳細にみるため、ESATJ が出力したデータを用いて次の Figure 5-24 にエラーと  
ならず用いられた 3 つのパラメータの実際の分布を示す<sup>注55</sup>。P(C+|E+), P(C+|E-)が 0 以  
下あるいは 1 以上になるなどのエラーにより 2797 件が除かれている。ヒストグラムはエラー  
とならず用いたパラメータのパラメータのヒストグラムに設定した理論分布の（赤色曲線）を重  
ねている。分布を確認すると、

- ・ OR(C|D)は対数正規分布の当てはめのため、右下の平均が 4.87、中央値が 4.48 となっ  
ているが、最頻値は 4.07 である。
- ・ 分布の範囲を指定できない正規分布により生じた異常なパラメータを除外する truncated  
normal probability distribution を用いているため P(C+|E+), P(C+|E-)に 0 が、  
P(C+|E+)に 1 が生じ、平均値もずれている。
- ・ パラメータの設定意図が「0,1 はあり得るが実際にはほぼ生じない」であるのでこのパ  
ラメータは不適切である。

Figure 5-24 4.1 Probabilistic sensitivity analysis に用いたパラメータ分布



対数正規分布であるのでパラメータに最頻値を与えることができず、類推して中央値を与え  
たが、試行錯誤が必要であるので、対数正規分布を用いること自体を見直す必要がありそう  
である。

では、改めて本来、シナリオの「平均値」は分布の最頻値を意図しているとする。先ほど用  
いた分布の最頻値はシナリオを反映しているが、範囲と「正規分布に近い」、つまり裾の頻度  
は極めて小さいという条件は満たせていない<sup>注56</sup>。そこで、次のような logit を用いる分布関数  
(logit-logistic distribution) <sup>注57</sup>

注55：説明のため、 $P_{C1} < P_{C0}$ の仮定を置き、エラーとなる場合が多くなるよう、4.1.2 から変更している。

注56：パラメータの分布関数をそろえる必要はない。ESATJ で分布に相関を設定するのに同じ分布関数であると都  
合がよく、また説明も容易であるためここでは同じ関数を持ちることとしている。

注57：この関数は Greenland S が教科書<sup>2)</sup>で用いた関数で、指定した範囲内の左または右に偏った裾が正規分布型の  
乱数を生成できる。詳細は Appendix 2.4) 参照



$$\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * (\exp(g) / (\exp(g) + 1))$$

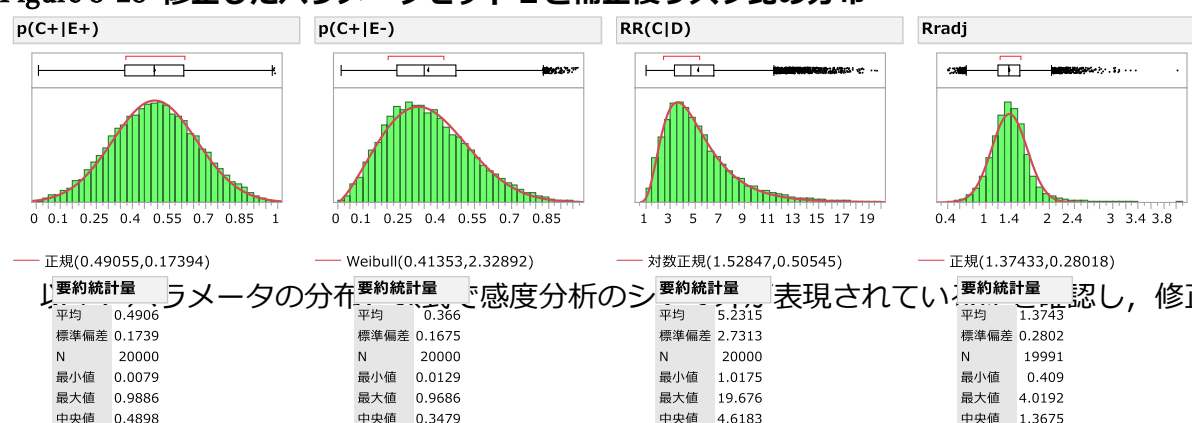
$$u = \text{Uniform}(0 \sim 1), g = \ln(\alpha / (1 - \alpha)) + \beta * \ln(u / (1 - u))$$

を用いて

	Min	Max	$\alpha$	$\beta$
$P_{C1}$	0	1	0.49	0.45
$P_{C0}$	0	1	0.35	0.45
$OR(C D)$	1	20	0.19	0.5

表のパラメータ（これを「パラメータセット2」とし、先のものを「パラメータセット1」とする）を用いると、Figure 5-25 の分布が得られる。ヒストグラムの赤線は得られた分布に近似的に求めた分布で  $P(C+|E+)$  は正規分布、 $P(C+|E-)$  は weibull 分布、 $OR(C|D)$  は対数正規分布である。最頻値、中央値と範囲は  $P(C+|E+)$  は 0.533, 0.489, 0.008~0.989,  $P(C+|E-)$  は 0.300, 0.348, 0.013~0.969,  $OR(C|D)$  は 4.180, 4.618, 1.075~19.676 であり、シナリオの意図がほぼ反映されている。また、このパラメータではエラーとなる  $P(C+|E+)$ ,  $P(C+|E-)$ ,  $OR(C|D)$  は生じていない。

Figure 5-25 修正したパラメータセット2と補正後リスク比の分布



行った。一様分布を用いる以外では、この PSA のシナリオを計算処理で表すためのパラメータ（分布、関数）を決める作業は探索的に行うことになる場合が多い。分布関数は唯一の正解があるのではなく、互換性のあるいくつかの選択肢があるであろう。例として用いた数式はより一般的な  $\beta$  分布によっても比較的近い分布を得ることができる。より影響の大きい事柄として、シナリオの情報に広い意味での「平均値」が得られる場合があるが、それは数学的な単純平均ではなく、最頻値や中央値の意味合いである場合が少なくない。また、PSA(MCSA)の結果は実際に得られた分布であり、結果は頻度として中央値が用いられる。つまり最頻値や中央値を反映しやすい分布を用いる方がパラメータを決める上で効率は良いと考えられる。

ここで感度分析により補正された結果も併せて検討したいところであるが、それでは結果から逆算した感度分析となる可能性がある。感度分析には結果に対して中立で透明性が求められる。

る。このため、感度分析のパラメータの設定はここでを行い、パラメータと補正されたオッズ比の関係の詳細な検討は次の項で扱う。

#### 5.6.5. モデル、パラメータと補正結果の吟味

感度分析においてモデルは重要な要素である。この“モデル”とは個々のパラメータの分布を表す関数とバイアス（例えば曝露、アウトカムの誤分類、未測定交絡要因など）を表現する補正式である。特に群（層）間でパラメータが異なる（差異誤分類の感度、特異度など）場合はリスク比やオッズ比などの補正される量が大きく、慎重さが求められる。PSA の場合はパラメータと補正結果の分布と相関（2 変量の組み合わせ）の確認から問題点を見つけられる場合がある。

#### ✓ パラメータの分布と補正結果の関係を確認する

Figure 5-26 と Figure 5-27 は「5.6.4 パラメータ特性の影響評価（シナリオの反映の確認）」の2つのパラメータセットのパラメータ分布を補正後のリスク比（ $RR_{DE}$ ）で色分けした散布図行列である。 $RR_{DE} \leq 1$  を黄色、 $RR_{DE} \geq 2$  を橙色から黒色で、除外されたパラメータは灰色でプロットしている。下側2つは各パラメータはお互いに独立であり、Figure 5-26 はエラーの灰色のプロットがあるが、これを除いても truncated normal probability distribution による偏った除外は生じていないこと、Figure 5-26 から補正後リスク比（ $RR_{DE}$ ）が小さくなるのは未知の要因を有する者が曝露ありに多く、なしに少ない場合で、未知の要因のリスク比の影響は小さいことが読みとれる。また、Figure 5-27 と Figure 5-26 を見比べると、 $RR(C|D)$ の範囲が狭くなっている以外は、分布の範囲に大きな違いはないように見える。しかし、 $RR(C|D)$ の範囲が狭くなっているため、補正後リスク比（ $RR_{DE}$ ）が2を超えるものは少なくなっている。左上のプロットで  $p(C+|E+) = p(C+|E-)$  のラインより上がパラメータセット2の結果である。Figure 5-27 より多く、 $p(C+|E+) < p(C+|E-)$  がより多く生じていることが確認できる。

Figure 5-26 パラメータセット1の散布図行列

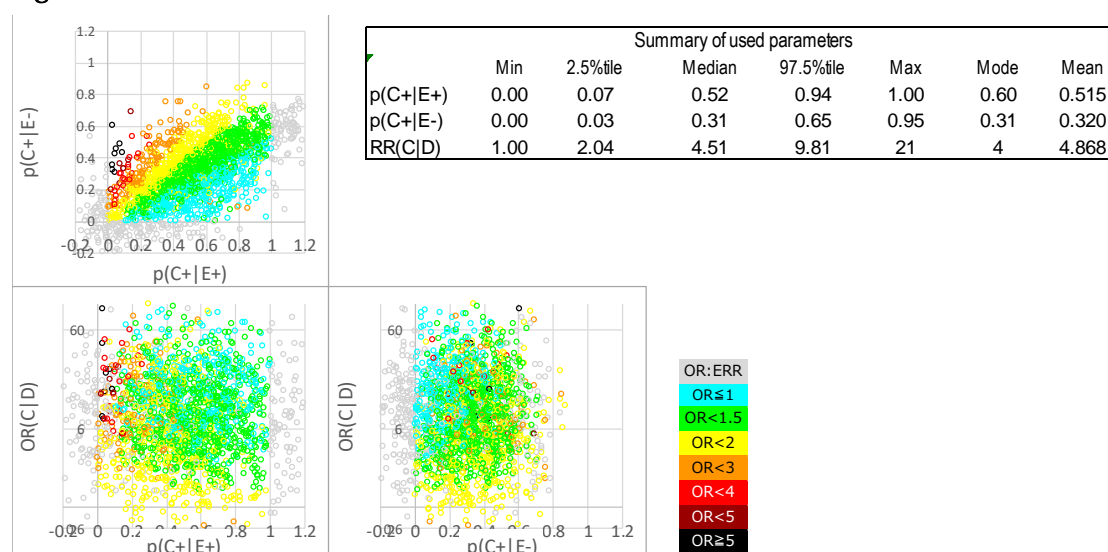
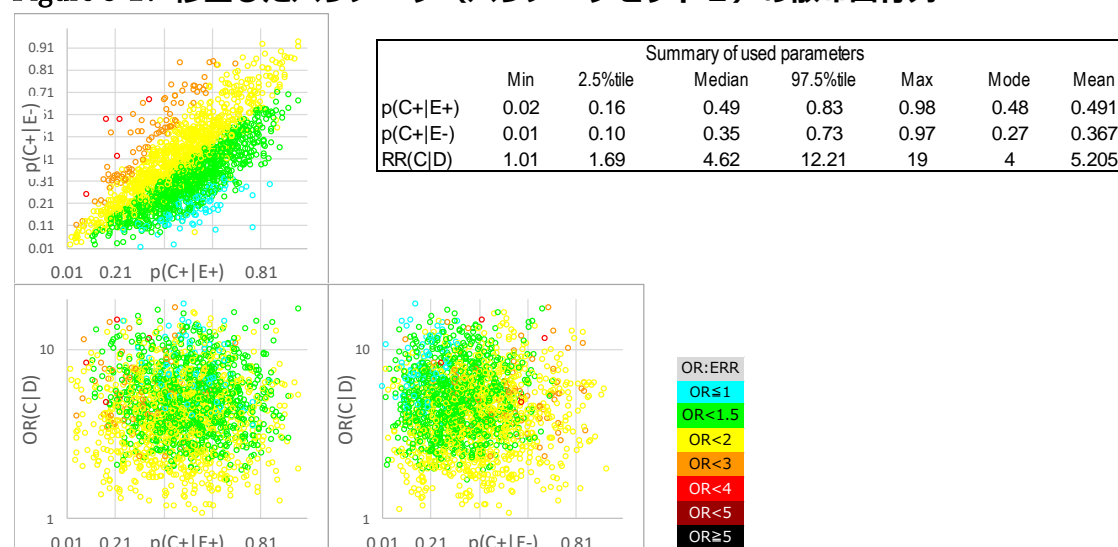


Figure 5-27 修正したパラメータ（パラメータセット2）の散布図行列



この例では、どちらを選ぶべきであろうか、加えて  $p(C+|E+) < p(C+|E-)$  はどうする（含める/除く）べきであろうか。更に Figure 5-27 では未知の要因のリスク比は比較的影響が少ないことが読み取れる。このように影響の少ないパラメータ並びにバイアスの原因は Probabilistic analysis（MCSA）の対象とせず、よりシンプルな方法を用いることが結果の解釈を容易にできる場合がある。次に、補正後のリスク比（ $RR_{DE}$ ）の分布を Figure 5-25 により確認すると、パラメータセット1の95%頻度区間は0.50～2.82、中央値は1.32、パラメータセット2は0.9～2.17、1.5である。補正前のリスク比は1.76であったので、パラメータセット1の方がより大きく補正されている。このことと、 $p(C+|E+) < p(C+|E-)$  はどちらも補正後リスク比が大きい側で生じていること、パラメータセット1では未知の要因のリスク比に非常に大きな値が生じている事、 $p(C+|E+)$ ,  $p(C+|E-)$  に取り得ない値（0未満, 1超）が生じていることに加えて分布の両端の立ち上がりが急であることを合わせて考えると、「パラメータセット1では過剰な補正が生じ、パラメータセット2が適切と考えられる」と説明することはできるであろう。

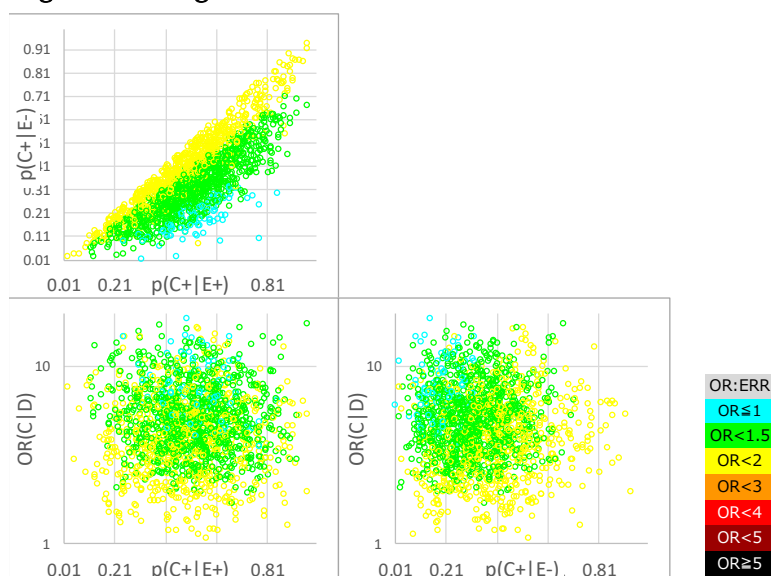
#### ✓ 次に $p(C+|E+) < p(C+|E-)$ の扱いについて検討する。

「5.6.4 パラメータ特性の影響評価（シナリオの反映の確認）」のMCSAのシナリオはcase群の曝露あり（E+）の未知の要因あり（ $p(C+|E+)$ ）がcontrol群の曝露なし（E-）の未知の要因あり（ $p(C+|E-)$ ）より大きいというものであったが、 $p(C+|E+)$ と $p(C+|E-)$ を独立にサンプリングする場合、 $p(C+|E+) < p(C+|E-)$ が生じる [Figure 4-3 参照]。この例では相関を設定したが、それでも  $p(C+|E+) < p(C+|E-)$  はパラメータセット2で3323件（シミュレーションの約16%）生じている。Figure 5-28はこの組合せを除いた結果である。 $RR_{DE} \geq 2$ （橙色）が消えていることが判る。パラメータ分布における  $p(C+|E+) < p(C+|E-)$  の組の分布を確認すると（Figure 5-29の上左から3つのヒストグラムのメッシュの部分）、分布に対して特異な傾

向があるわけではないため、除いても分布に歪みは生じないことが確認できる。この場合の補正後リスク比の分布は Figure 5-30 右端のヒストグラムのようになる。Figure 5-29 の右端のヒストグラムと見比べると、分布の右側がなくなっており、中央値は 1.47 から 1.42、95%頻度区間は 0.91~2.13 から 0.89~1.71 であり  $p(C+|E+) < p(C+|E-)$  を除かない結果より中央値は小さく、95%頻度区間は狭くなっているが、Crude data のリスク比 1.73 が過大である可能性を示唆していることに変わりはない。つまり、この場合は一方を選ばなければならない状況ではなく、何れであっても本研究の結果は過大である可能性があることを説明すべきであろう。

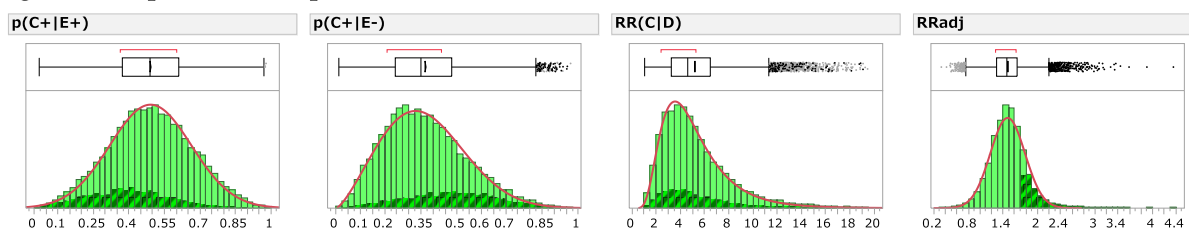
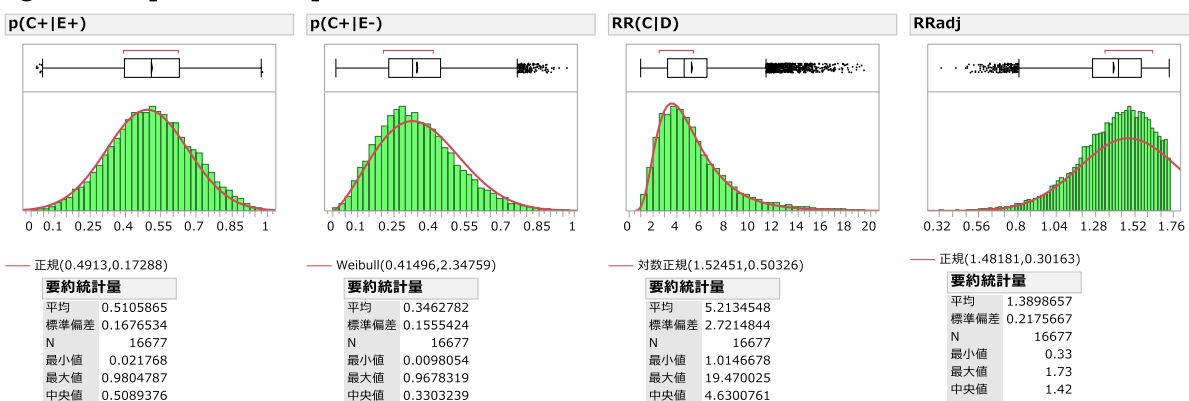
この例は、MCSA において負のセル（頻度）といった不合理ではなく、 $p(C+|E+) < p(C+|E-)$  が生じ得るとするか否かというシナリオの問題<sup>注58</sup>である。しかし、曝露あるいはアウトカムの誤分類、未測定交絡要因において負のセルが生じる場合がある。この場合オッズ比が負の値になるため、ヒストグラムや Box Plot を用いる、或いは計算 Tool、ソフトが出力する情報により確認できる。一般に evidence に基づいたパラメータによる MCSA でこのような異常値が発生する繰り返し（サンプル）数は多くは無いため、単に中央値と 95%頻度区間を示すだけでは気付けない場合がある。特に研究者と解析者が異なる場合、不合理な結果や定義と異なる分布<sup>注59</sup>の使用を研究者が見落とさないよう、解析者はこのようなグラフによる結果提示を心がけることを推奨する。

Figure 5-28 Figure 5-27 から  $P_{C1} < P_{C0}$  とエラーを除いた散布図行列



注58：MEMO:例示の  $P_{C1} < P_{C0}$  は 13%であり、その 1 割程度は逆転することがあり得ると考える事は不自然ではないかもしれない。反対に必ず  $P_{C1} > P_{C0}$  であるとする方が不自然であるかもしれない。

注59：MEMO:シミュレーションから除かれることで結果的に与えた分布と異なる分布が使用される。

Figure 5-29  $p(C+|E+) < p(C+|E-)$  をメッシュで表したFigure 5-30  $p(C+|E+) < p(C+|E-)$  を除いた分布（近似線は除く前と同じ）

### ✓ 負のセルが生じる場合

教科書<sup>2)</sup>には、負のセルが生じるパラメータの範囲が合理的（パラメータとして）と考えるなら、負となったセルを「0」として扱い、PSAの結果の頻度区間の計算に使用する方法<sup>注60</sup>と、そのようなパラメータは現実的ではないとして与える分布を負のセルが生じないように修正する（または結果から除く、自動的に除かれる）の二通りが紹介されている。後者の場合、修正したパラメータを示すあるいは、異常なためシミュレーションから除かれた反復がある場合は最低限その頻度は報告する<sup>78)</sup>、<sup>注61</sup>。また、MCSAのツール（SAS マクロ等）が内部で異常値を処理（シミュレーションから破棄する）している場合もあり得るのでツールの説明、仕様の確認も必要である。研究者はこれらの結果から感度分析パラメータの修正が必要と考える場合があるかもしれない。このような探索的過程は異常値（エラー）への対処では推奨されるべきであるが、それ以外の場合は恣意的な感度分析につながる場合があり慎重に行われる必要がある。何れにしても、PSAにおいて探索的に感度分析パラメータを設定した場合は設定した場合は、その旨と、結果への影響を説明すべきである。

注60：「0」とした場合オッズ比、リスク比の不定となるため、実際には四分表の全てのセルに0.5を加える。これは連続修正ではない。詳細はFleiss JL (2003)<sup>80)</sup>を参照されたい。

注61：Figure 4-33 Multiple bias analysisの結果の例のアウトカムの誤分類の特異度は元の教科書<sup>2)</sup>では最小値を0.8としているが、0.9未満はほとんどがエラーとなるため、0.9に修正している。これに伴い0.9に近い分布を厚くするため、 $\alpha$ を0.7から0.77に修正している。



## ✓ 感度分析の対象とするバイアスの原因と方法について確認する

ここまで、未測定（未知）交絡要因を例としてきたが、実際には曝露やアウトカムの誤分類、選択バイアスなど、複数のバイアスの原因が存在し、感度分析の対象となるが、実際には結果への影響が小さいまたは元々大きい値がより大きくなるなど、意思決定には重要ではない場合があり、結果を提示する必要がない、または実施するとしても簡略化または、感度分析の対象から除くことが可能な場合がある。また、バイアスの原因の数が少なくなると、感度分析の結果も簡潔で理解が容易となる。加えて、未測定交絡要因で曝露への効果修飾がない場合とある場合や選択バイアスで層毎の抽出率を扱うか、オッズ比を扱うか、誤分類は感度・特異度か、陽性的中度・陰性的中度かなどバイアスを計算する方法自体にも複数の選択肢があり得る。これらは得られるバイアスのパラメータ情報源によって異なるが、実際に PSA を行って検討する場合もある。

事例としては、4.1.8 の Figure 4-33 では、未測定交絡要因である喫煙率が群間で偏りが無い（ランダムに異なる）場合と偏りがある場合として共に効果修飾がない場合が検討され、偏りがある場合が Multiple bias modelling に選ばれ、さらにアウトカムの誤分類は除かれている。Multiple bias modelling を行うのではなく、個々の PSA の結果を提示するのであれば、感度分析の目的に照らして必要なもの、例えば結果が容易に覆り得ることを示すのであれば、バイアス（として用いた情報、パラメータ）に根拠があり、設定に理解を得やすく、影響の大きい少数のバイアスの影響を示せば十分であろう。しかし、個々の PSA の結果を提示する計画で感度分析が行われたが、結果の補正の方向が混在していたり、個々の影響が大きくも小さくもないなど全体での評価が必要と思われる場合は Multiple bias modelling を追加することを検討する必要があるかもしれない。

本章ではここまで感度分析の実施について説明を行ってきた。データの処理としてはここで完了であり、あとは結果を整理してまとめ、考察、解釈を行うだけである。

また、筆者が作成した Excel の Tool (ESATJ) を用いて感度分析を行ってきた。このような Tool を用いれば、全くプログラムを書くことなく、固定値を扱うシンプルなものから分布を扱う PSA までの感度分析の結果を得ること自体は容易である。実施が容易な分、途中の計算（パラメータの分布、計算エラー等）、結果を確認し感度分析の目的、シナリオに照らして確認することを推奨する。その際、グラフを作成して視覚的に確認することは有効で GUI を用いて簡便にヒストグラムや Box plot を作成できるソフトの利用は生産性が高まる。本章では ESATJ が出力しないグラフは SAS 社の JMP®13 を用いた。

## 5.7. 感度分析の結果を提示し解釈する

実施した感度分析について論文や報告書にまとめるステップである。読者や意思決定者が理解できることが重要であるが、感度分析になじみがない場合もある。そのため、感度分析とは何か、何を目的として行ったのかといった説明をはじめにしておく理解を得やすいであろう。以下では、その後に行う実施した感度分析の結果提示、解釈などについて説明する。

### 5.7.1. 感度分析の方法を提示する

感度分析は経済系、意思決定論分野で発展してきた経緯があり、読者である医学系、疫学系の研究者、製薬企業、規制当局の担当者及び意思決定者には馴染みがない場合があるので感度分析の方法は可能な限り完全かつ詳細で明快な説明を行う努力が必要である。しかし、実際の研究で感度分析について十分説明されていないことが報告されている<sup>17)</sup>。感度分析は研究データにない仮定を用いるため、対象と方法についてデータの性質、研究デザイン、主解析の手法、限界と関連付けて分かり易く説明することは理解を促し、透明性を得るのに役立つ。このため、主解析の結果の提示には感度分析の対象とする限界（バイアス）について記載する。そして感度分析の項（Methods, Analyses に記載することも Appendix である場合もある）に感度分析の対象としたバイアス（変数）と研究結果の関係（これが目的に当たる）、用いるパラメータの特性、情報源を整理する。つまり、報告書の読者が研究結果の項で結果に伴う限界を理解し、感度分析の項でその限界の影響度の評価方法を理解できるように記述する。

図（DAG や Figure 5-3 のような標本の抽出から割り当てまでのフローチャートなど<sup>11),14),81)</sup>）や表を用いることも有用である。パラメータはその特性（点推定値、最頻値、範囲など）とその根拠（出典）と共に他に妥当な選択肢があればそれを用いなかった理由を示す。

事例として比較的詳細に書かれている 4 件と簡略されている 3 件を簡単に紹介する。実際に行う感度分析の参考になりそうなもの、興味を持ったものは原著を参照いただきたい。

Multiple bias modelling で紹介した Bollaerts K (2016) のインフルエンザワクチンとナルコレプシーの関係を評価した論文<sup>23)</sup>は選択バイアス、未測定交絡、曝露とアウトカムの誤分類など複数のバイアスを扱い、各々 Step1～8 として対象としたバイアス、モデルと数式、パラメータの根拠を Step 毎に説明している。例えば Step1 は選択バイアスを扱い、研究の対象とすべき中絶関連死亡者が実際にその研究に取り込まれている確率について、内部検証研究が理想であるが、実施されておらず、文献報告もされていないため、同様の地域/母集団で実施された妊産婦死亡率の検証研究からのデータと妊産婦死亡率および中絶関連死亡率の研究における母親の死亡の過小評価を調整するために文献で一般的に使用される調整因子から考えられる範囲を踏まえて確率分布として台形分布により設定したことが説明されている。

Baer VL (2007) の新生児への血小板輸血と死亡の関連について、不明な交絡要因の影響の評価に Multidimensional analysis を用いた 7 ページの論文<sup>82)</sup>では、Methods の Sensitivity analysis の項に、不明な交絡要因を含む死亡に対するロジスティックモデルの式を示し、Figure 5-31 を用いてパラメータと死亡リスクの関係を説明し、測定されていない変数には、病気のレベルや遺伝的素因などの要因が含まれている可能性があること、計算の利便性から、平均 0、標準偏差 1 の正規分布<sup>注62)</sup>を仮定したことを説明している。

注62： この感度分析は輸血回数という連続変数に対して行われており、補正パラメータも連続変数である必要がある。

Barron TI (2013) の nested case-control study を用いて乳がんのホルモン療法の遵守/非遵守と早期再発の関係を検討した 9 ページの論文<sup>78)</sup>では、アウトカムである乳がん再発の誤分類を PSA により評価しており、Statistical analysis と並列で Sensitivity analyses の項を立て、PSA の方法、公開されている SAS マクロを使用したこと、パラメータの根拠は同様の方法による外部検証研究に加え、がん登録担当者の意見、医療記録における乳がん再発の同定に関する 3 つの文献によることを説明し、METHODS の項で Statistical analysis の見出しのもと 31 行で感度分析の手法、パラメータ等について出典とともに要約したうえ、差異および非差異誤分類の結果と試行回数、用いたパラメータを Table 5-14 の結果の表の下部にまとめている。

Lash T.L. (2004)<sup>68)</sup>の乳がんの 1 次療法の 5 年間の死亡率の予測を評価した論文はレコードレベルの手法を用いた感度分析を主体とした内容であるが、予測モデルとバイアスの原因（がんのステージに関係する腋窩郭清の扱いの誤り）と誤分類が生じる過程について文章と DAG を用いて説明し、感度分析の方法（三角形の確率密度関数を用いて腋窩郭清の記録のある患者とない患者に分けて誤分類を補正したデータセットを作成し、Cox 比例ハザードモデルによる多変量調整ハザード比の計算を 2,500 回繰り返された）、パラメータの情報源（感度および特異度は 4 つの文献の加重平均）を説明し、結果を 1 つの確率密度グラフと文章により説明している。誤分類が生じる過程（特になぜ特異度が必要か＝偽陽性が生じるのか）の説明、高度な方法であるレコードレベル手法の説明の仕方など参考になるであろう。

Figure 5-31 Baer VL の不明な交絡要因の説明<sup>82)</sup>

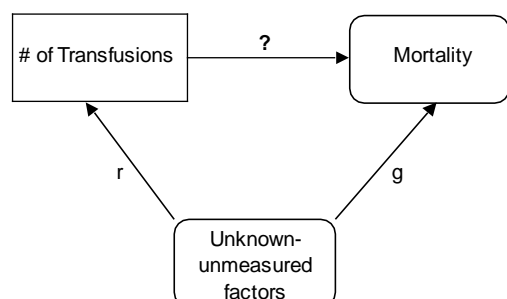


Figure 2 Sensitivity analysis. Relationships are shown schematically between three factors; (1) the number of platelet transfusions received, (2) the mortality rate and (3) unknown-unmeasured factors (such as sickness level of the infant) that could potentially influence both 1 and 2. The solid arrow labeled 'r' is the correlation between the unmeasured factors and the number of platelet transfusions given. The solid arrow labeled 'g' is the correlation between the unmeasured factors and the mortality rate. The dashed arrow with the question mark asks whether there is any direct relationship between the number of platelet transfusions received and the mortality rate.

Table 5-14 再発乳がんの誤分類に関する PSA の結果<sup>78)</sup>

Table 3. Probabilistic sensitivity analyses correcting for non-differential and differential misclassification of breast cancer recurrence					
Sensitivity analysis <sup>a</sup>	Odds ratio 2.5 percentile	Odds ratio median estimate	Odds ratio 97.5 percentile	Change in simulation interval width (b/a) compared with conventional (%)	Proportion of simulations with odds ratio > 1
Persistence—low (X180 day gap)					
Non-differential sensitivity analysis <sup>b</sup>					
Sensitivity analysis only	2.45	3.98	7.78	- 51.5	0.010
Sensitivity analysis with random error	1.39	4.00	12.30	35.4	0.005
Differential sensitivity analysis <sup>c</sup>					
Sensitivity analysis only	2.68	4.26	7.86	- 53.7	0.009
Sensitivity analysis with random error	1.51	4.29	12.79	34.1	0.003
Compliance—low (0%–90%)					
Non-differential sensitivity analysis <sup>b</sup>					
Sensitivity analysis only	1.00	1.36	1.90	- 37.2	0.222
Sensitivity analysis with random error	0.73	1.35	2.60	17.5	0.171
Differential sensitivity analysis <sup>c</sup>					
Sensitivity analysis only	1.08	1.46	2.12	- 38.5	0.234
Sensitivity analysis with random error	0.77	1.46	2.81	19.8	0.124
Cumulative exposure—low (0%–90%)					
Non-differential sensitivity analysis <sup>b</sup>					
Sensitivity analysis only	1.32	1.80	2.61	- 35.7	0.048
Sensitivity analysis with random error	0.92	1.81	3.56	25.3	0.042
Differential sensitivity analysis <sup>c</sup>					
Sensitivity analysis only	1.43	1.94	2.72	- 40.8	0.057
Sensitivity analysis with random error	1.01	1.96	3.67	13.1	0.024
<sup>a</sup> 5000 simulations. <sup>b</sup> Non-differential misclassification of breast cancer recurrence trapezoidal distributions: sensitivity (min ¼ 0.70, mode—1 ¼ 0.75, mode—2 ¼ 0.85, max ¼ 0.90); specificity (min ¼ 0.94, mode—1 ¼ 0.96, mode—2 ¼ 0.98, max ¼ 1.00). <sup>c</sup> Differential misclassification of breast cancer recurrence trapezoidal distributions: sensitivity: low compliance/persistence/exposure (min ¼ 0.70, mode—1 ¼ 0.75, mode—2 ¼ 0.85, max ¼ 0.90); sensitivity: high compliance/persistence/exposure (min ¼ 0.80, mode—1 ¼ 0.85, mode—2 ¼ 0.95, max ¼ 1.00); specificity: low compliance/persistence/exposure (min ¼ 0.94, mode—1 ¼ 0.96, mode—2 ¼ 0.98, max ¼ 1.00); specificity: high compliance/persistence/exposure (min ¼ 0.95, mode—1 ¼ 0.97, mode—2 ¼ 0.99, max ¼ 1.00); correlation between sensitivities and specificities 0.8.					

以上は詳細に感度分析を説明している事例であるが、感度分析の目的、結果によってはより簡便な説明も用い得る。

Youngquist ST (2008)<sup>83)</sup>は cohort study によるβ遮断薬の使用と院外心停止に関する 5 ページの論文であるが、曝露とアウトカムの非差異誤分類について PSA の結果を Statistical analysis の項に誤分類パラメータの感度・特異度のコンパクトな 2 つの表と 25 行程度の文章を用いて説明している。Mitro SD (2015)<sup>84)</sup>の cohort study による血中ビタミン D と子宮平滑筋腫の関係の 6 ページの論文では未確定診断の誤分類の感度分析により結果が否定される可能性を報告しているが、感度分析について図表に整理はされず、Methods の項に Sensitivity analysis の項として PSA について 30 行の文章で説明し、結果を 6 行の文章で提示しているのみである。また、Ekström N (2013)<sup>39)</sup>は 2 型糖尿病患者における低用量アスピリンの脳心血管管及び出血イベントリスクを Cox proportional hazards model を用いて評価した 9 ページの論文であるが、ハザード比が 1.3 の未知の交絡要因の存在を仮定した感度分析について Statistical Method の項に 4 行で記載し、パラメータと結果を online supplementary の table にまとめている。



計算例を示すことは感度分析モデルとパラメータの影響の理解を促すことに役立つ。例えば、曝露の誤分類の場合、パラメータ（感度と特異度）の変化が曝露と結果の分割表の各セルに与える影響を例示する。その際、数式は Appendix2.6) Table 2-2 のように最終的に用いられる要約表（曝露、結果共に 2 値変数の場合は 2×2 の四分表）に整理又はリンク（参考 4 のように表内に関数の左辺を記載し、表外に関数全体を記載する）して示すことを推奨する。また、バイアスの原因毎に感度分析を行うなど、複数の感度分析を行った場合はその各々についてその順序と共に説明する[4.1.8 各種バイアスにおける PSA の設定と結果の例 参照]。

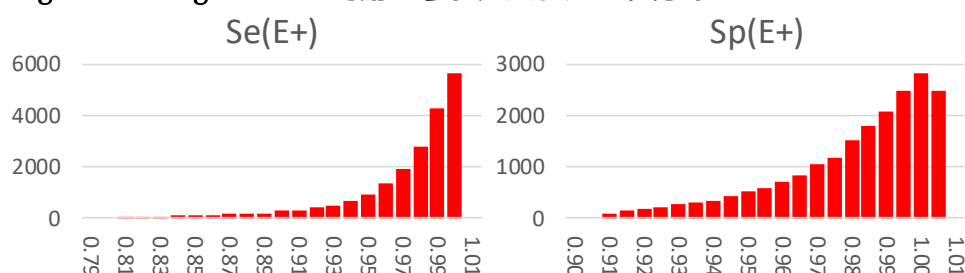
PSA を用いる場合に、分布を表す関数、パラメータを見ても分布形状を容易に想像できない分布が用いられる場合がある。また、Beta 分布、正規分布など一般的な分布であっても実際に用いられている値の範囲（95%区間、最小値、最大値など）の理解が難しい場合がある。このような場合は、ヒストグラムなどにより用いた分布を提示することが最も効果的であると考えられる。例えば、Figure 4-33 Multiple bias analysis の結果の例 の OUTCOME MISCLASSIFICATION の感度 (Se)、特異度 (Sp) の分布は

$$u = \text{Uniform}(0 \sim 1), g = \text{logit}(\alpha) + \beta * \text{logit}(u),$$

$$\text{Se and Sp} = \text{Min} + (\text{Max} - \text{Min}) * \text{expit}(g),$$

という式が用いられ<sup>注63</sup>、Se では $\alpha=0.9$ 、 $\beta=0.8$ 、 $\text{Min}=0.8$ 、 $\text{Max}=1$ 、Sp は $\alpha=0.8$ 、 $\beta=0.8$ 、 $\text{Min}=0.9$ 、 $\text{Max}=1$  が与えられている。一見、感度 (Se) は 0.8~1 の対数分布のようであるが、実際の分布は Figure 5-32 の形状のとおり、感度の 0.85 以下はほとんど発生していない。このような極端に偏った分布形状は四分位やパーセント点でイメージを伝えることは難しく、視覚化が有効であろう。

Figure 5-32 Figure 4-33 で用いられたパラメータ分布



感度分析に用いたシナリオ、モデル（バイアスの計算式、PSA の分布関数、パラメータなど）は唯一の正解ではなく、読者や意思決定者は異なる設定の確認が必要と考えるかもしれない。感度分析の方法の提示が完全であるなら、（手法が raw-data を用いるものでなければ）読者は行われた感度分析を再現することも、異なるパラメータを用いて感度分析を評価することも可能となる。言い換えれば、読者が再現可能な提示を行うことが推奨される。

注63： この関数は Appendix2.4) 参照



### 5.7.2. 感度分析の結果を提示する

単一のパラメータによる単純な感度分析の結果の提示は通常の調整解析結果の提示と同様で容易に行える。しかし、バイアスの原因が1つであってもPSA並びに複数のパラメータを組合せて行う場合の結果提示は容易ではない。しかし、投稿規定により文字数や図表に制約があるため、感度分析の詳細を記述することは一般に困難である上、感度分析の結果は興味深いものもあればそうでない場合もあるが、何れの場合でも何に対して感度分析を行ったかという最低限の事項は示すべきである。

感度分析の個々の手法の結果の見せ方についてはすでに多く紹介してきているので、そちらを参照いただきたいが、報告書には、そのような結果のフルセットではなく、感度分析から得られた影響の意味、価値に合わせてコンパクトな文章、表により整理した提示で足りる場合もある。

そのようなまとめ方の例として、「5.7.1 感度分析の方法を提示する」で紹介した Youngquist ST (2008)<sup>83)</sup> の $\beta$ 遮断薬の使用と院外心停止に関する論文では結果に大きな影響がないため Results の項に8行の文章で感度分析の結果を記載しているに過ぎないが、Mitro SD (2015)<sup>84)</sup> の血中ビタミンDと子宮平滑筋腫の関係を評価した論文は主結果の半分程度の表に感度分析の結果をまとめ、Results の項の文章の半分を費やしている。また、Ekström N (2013)<sup>39)</sup> の2型糖尿病患者におけるアスピリンの心血管疾患予防有効性を評価し、心疾患による死亡リスクは改善されず、非致死性の冠動脈疾患などのリスクの増加が認められた論文では未測定交絡要因の影響を感度分析により評価されたが、Results の最後に13行で感度分析の結果を要約し、詳細はonline supplementary のtableの参照を求めている。このように感度分析の意義に合わせて提示に強弱を付けることは紙面に限りがある以上、合理的である。その場合でもPSAの最も簡潔な結果の提示として最も重要な主解析結果に対応する補正された中央値と95%頻度区間を示すことは可能であろう。しかし幸運にも近年電子ジャーナルが多くなり、WEB上にページ数に縛られずAppendix (Supplement) を公開可能な場合が少なくなく、調整解析の詳細と共に感度分析についても詳細を公開している例がみられる。<sup>51),85)</sup> この場合でも本文に要約を記載すると共に、表に整理する、主解析の結果と形式を揃えるなど解りやすくコンパクトな提示に務めるべきである。

バイアスの原因が複数ある場合やシナリオが複数（パラメータ間に相関を設定しない/する場合など）ある場合の結果の提示はより複雑となる。このため表に要約されることが多いが情報の多くを犠牲にせざるを得ず、特にパラメータの不確実性の提示が不十分となりやすいことに加えて複雑な表は理解と解釈を困難にする。グラフを用いることは多くの情報を直感的に理解できる場合がある。複数の要因を対象とする場合の感度分析は各々の要因毎の場合と、まとめて（組合せて）行う場合がある。前者の場合は理解が容易で影響が小さく無視可能な要因を判断可能である。その反面、組み合わせによる思いがけない大きな影響や反対に大きな影響が稀であることは情報が得られない。まとめて行うMultiple bias (Probabilistic) modellingでは個々の要因の影響はマスクされ評価と理解が難しくなるため、Figure 4-33 Multiple bias

analysis の結果の例のように個々の要因についての結果など追補的なデータの提示を含めて丁寧な解説を推奨する。この場合、meta-analysis や benefit-risk 評価で用いられる forest plot (Figure 4-33 Multiple bias analysis の結果の例 参照) , tornado diagram (plot) が用いられることがある<sup>14),49),86)</sup>。

事例としては、4.1.8 各種バイアスにおける PSA の設定と結果の例 で紹介した Bollaerts K の論文<sup>23)</sup>は曝露の誤分類、選択バイアス、未測定交絡要因を扱っているが、対象としなかった疾患の誤分類についてもその結果と理由が説明されている。

結果の提示で重要なのは読者がシナリオの影響を誤解無く理解できることである。5.7.1 の簡潔な記載例で紹介し NICU の新生児への血小板輸血と死亡の関心の不明な交絡要因の影響に Miltidimensional analysis を用いて評価した Baer VL の論文<sup>82)</sup>は輸血回数との相関 (R) と死亡のオッズ比 (G) の組み合わせの表を作成し、死亡リスクが大きくなる場合と小さくなる場合を色分けしてシナリオ (パラメータの値) が死亡リスクの指標であるオッズ比にどのように影響するかが数値の色分けによって一目瞭然でありよい例である。

#### ✓ 相関のあるパラメータを用いた場合

次に、結果の提示についてあまり良い事例が見つけれなかったパラメータに相関を用いた場合を Table 4-1 について行われた Figure 4-33 Multiple bias analysis の結果の例の曝露の誤分類についてパラメータを少し変えて PSA を行った場合を例示する。感度が肺癌なしの場合に少し低く、感度がより低い場合がある Table 5-15 のシナリオの感度分析を行った<sup>注64)</sup>結果は

Table 5-16 である。曝露の誤分類によってオッズ比が過少となっている可能性があることが示されているが、同時にオッズ比が 1 以下となる可能性も示されている。

Table 5-15 曝露の誤分類の PSA のパラメータ

Variable Names			Min	Max	$\alpha$	$\beta$	$\rho^*$	Distribution
Outcome	肺癌	Se(肺癌+)	0.8	1	0.9	0.8	—	expit(logit) <sup>#</sup>
Exposure	レジジン樹脂	Se(肺癌-)	0.8	1	0.8	0.75	0.8	Correlate with Se(肺癌+)
Number of simulation =20000		Sp(肺癌+)	0.7	1	0.8	0.8	—	expit(logit) <sup>#</sup>
		Sp(肺癌-)	0.7	1	0.8	0.8	0.8	Correlate with Sp(肺癌+)

\* :  $\rho$  : correlation coefficient

# :  $\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * (\exp(g) / (\exp(g) + 1))$  Triangular

$g = \ln(\alpha / (1 - \alpha)) + \beta * \ln(u / (1 - u))$ ,  $u = \text{Uniform}(0 \sim 1)$  Appendix2. 4) 参照

注64 : PSA にエラーを生じる場合として 4.1 Probabilistic sensitivity analysis の未測定交絡要因の例ではなく曝露の誤分類を用いている。また、例示のためのパラメータであり、根拠となる情報源はない。

Table 5-16 PSA により補正したオッズ比

Analysis	OR Simulation Results	
	Median	2.5th percentile~97.5th percentile
Conventional	1.76	1.2 ~ 2.58
Systematic	1.99	0.71 ~ 14.93
Total Error	2.07	0.7 ~ 15.15

ここでは、例示のため敢えて特異度の最小値を小さく設定している。「感度が肺癌なしの場合に少し低く、感度がより低い場合がある」とシナリオを説明したが、このパラメータからそのことが読み取れるであろうか？ このような記載から用いられたパラメータおよびパラメータと結果の関係を読み取ることは困難である。また、結果に記載されていないが、約 36% でエラーが生じており、与えたパラメータと用いられたパラメータは異なる<sup>注65</sup>。Figure 5-33 は PSA の結果の概略である。この出力では PSA の結果は 12,732 回のサンプリングの結果であり、7,268 回が除かれている（設定は 20,000 回）こと、ヒストグラムから与えたパラメータの分布、下の散布図行列から相関関係、右上の統計量から実際に用いられたパラメータは  $Se(D+): 0.8 \sim 1$ ,  $Se(D-): 0.7 \sim 1$ ,  $Sp(D+): 0.853 \sim 1$ ,  $Sp(D-): 0.91 \sim 1$  であること<sup>注66</sup>、右上の散布図から特異度の影響が大きく、よく見ると補正後のオッズ比が 1 以下になるのは  $Sp(D+)$  が約 0.88 以下、 $Sp(D-)$  が約 0.92 以上であること、また、 $Sp(D-)$  は 0.93 以下で補正後オッズ比が 5 以上になることなどが読み取れる。Figure 5-34 は Figure 5-33 から

- ・ パラメータのヒストグラムから用いられなかったサンプリングを除く
- ・ パラメータの散布図行列は補正後オッズ比の区分ごとに色分け（エラーは灰色）する
- ・ 補正後のオッズ比が 1 以下となる頻度の割合を表示する

を行った。こちらでは、特に  $Sp(D-)$  が 0.9 未満で全てエラーが生じ、他のパラメータにかかわらず、 $Sp(D-)$  が低い場合に補正後オッズ比が大きく（高い場合に小さく）なること、 $Sp(D+)$  と  $Sp(D-)$  が等しい（差異でない）場合でも特異度が低下すると補正後オッズ比はより大きくなることから特異度については慎重な検討が必要であり、対して、感度の影響は小さく、感度について多くの議論をすることは重要ではないこと、補正後オッズ比の頻度区間は 1 を含んでいるが、1 以下となるのは 5% 程度であるため重要ではなさそうであることがわかる。また、このシナリオには  $Se(D+)$  と  $Se(D-)$  間に  $Se(D+) > Se(D-)$  となる相関を設定しているが、Figure 5-34 から相関を用いた影響を読み取るには相関を用いない結果が必要である。全てに相関を用いず、Table 5-17 の設定<sup>注67</sup>による PSA の結果を Figure 5-35 に示す。用いられているパラメータの範囲は同じであるが、補正後のオッズ比の中央値は 1.99 から 1.92 に小さく、頻度区間は 0.71~14.93 から 0.23~24.2 と広く、補正後オッズ比が 1 以下の割合は 18% と増えてい

注65： エラーの扱いについては 94 ページの「負のセルが生じる場合」を参照。

注66： D+はアウトカムである肺癌あり、D-は肺癌なしを表す。

注67： 相関を設定した場合と分布の形状が近くなる「Triangular」を用いている。

る。これらの結果の提示には、Table 5-18 のように統計量を要約し、Figure 5-36 及び Figure 5-37 によりパラメータと結果を視覚化することは結果の理解を促すと考える。この例では、相関の影響は主にばらつきを狭くする働きであるので、相関の設定については、Table 5-18 に相関を用いない場合の補正後オッズ比の統計量のみを提示している。この場合はこの程度の情報で十分説明可能であり、あるいは相関を用いず、シナリオと結果の説明をシンプルにすることも考えられる。

Figure 5-33 PSA の概略

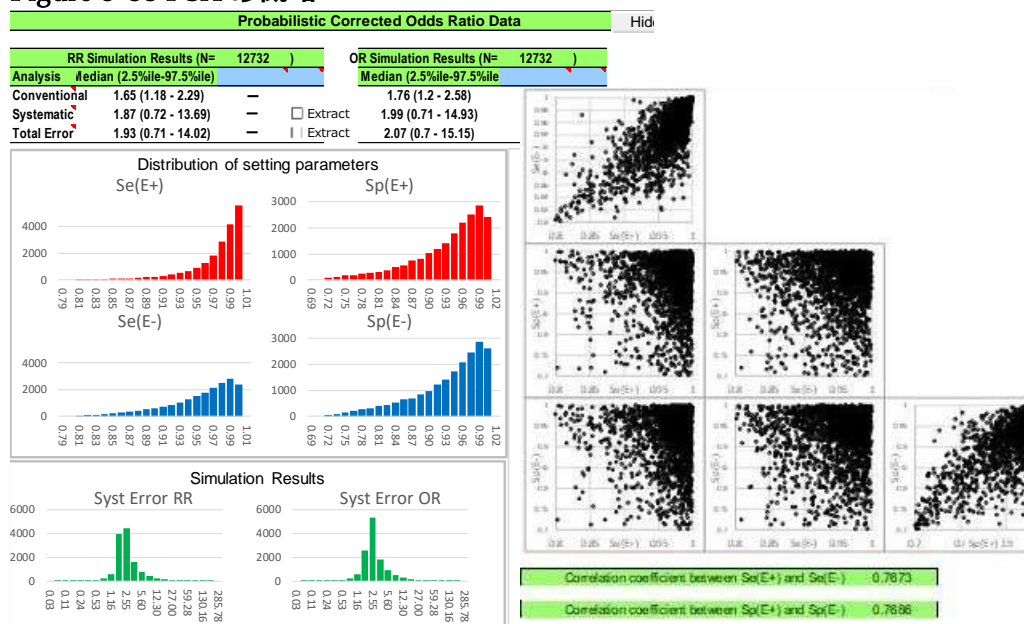


Figure 5-34 工夫した PSA の結果の例

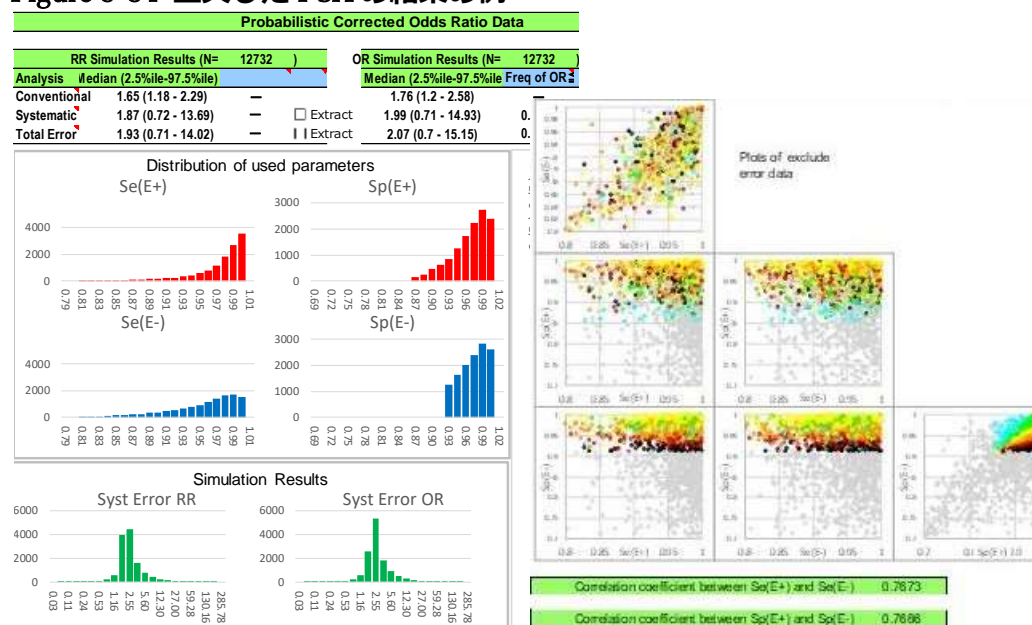




Figure 5-35 相関を用いない PSA の結果

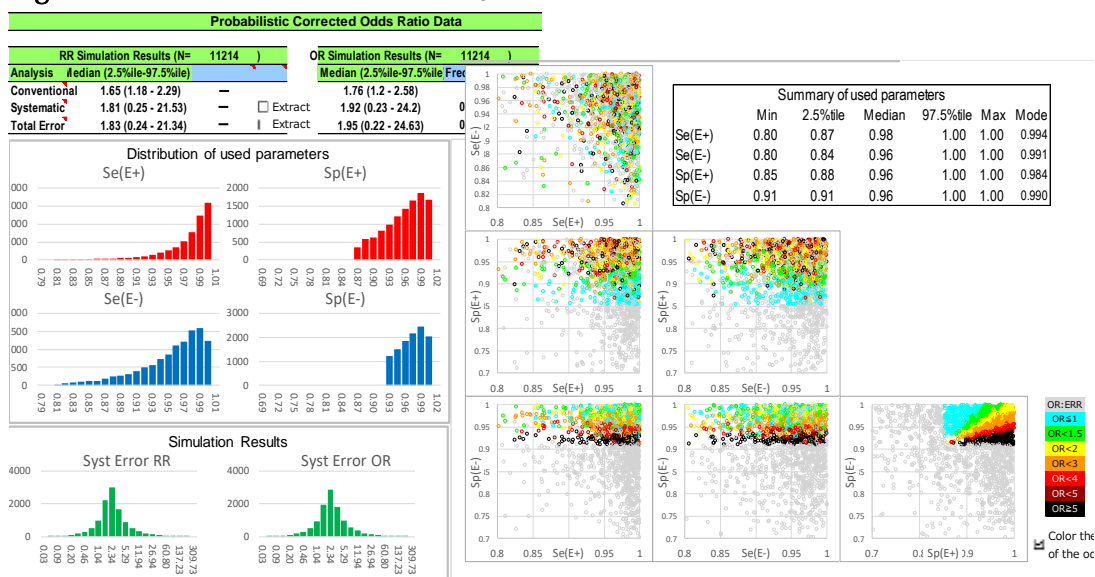


Table 5-17 相関を用いない PSA のパラメータ

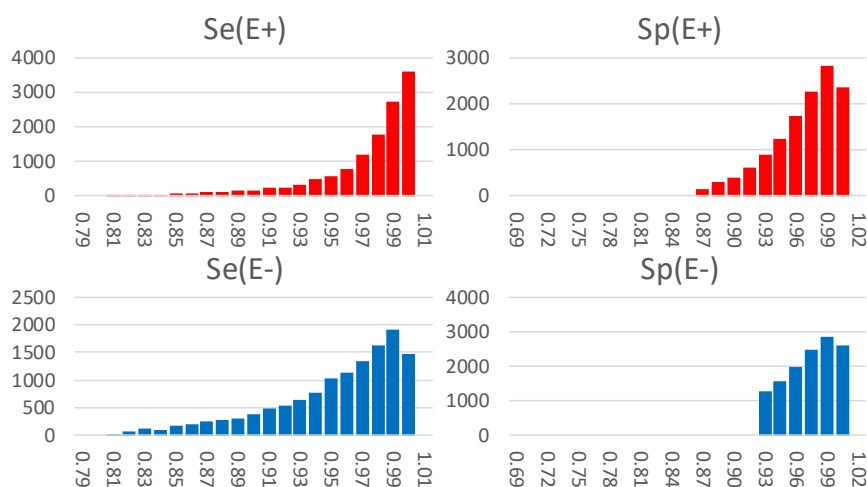
Variable Names			Min	Max	$\alpha$	$\beta$	Distribution
Outcome	肺癌	Se(レジン+)	0.8	1	0.9	0.8	expit(logit)
Exposure	レジン樹脂	Se(レジン-)	0.8	1	0.8	0.75	expit(logit)
Number of simulation =20000		Sp(レジン+)	0.7	1	0.8	0.8	expit(logit)
		Sp(レジン-)	0.7	1	0.8	0.8	expit(logit)

Table 5-18 修正した感度分析の概要

Outcome : 肺癌, Exposure : レジン樹脂, Method : MCSA, Number of simulation=20000									
Setting		Min	Max	$\alpha$	$\beta$	$\rho^*$	Distribution	Used distribution	
								Min~Max	Mode
	Se(レジン+)	0.8	1	0.9	0.8	—	expit(logit)#	0.8~1	0.994
	Se(レジン-)	0.8	1	0.8	0.75	0.8	Correlate with Se(レジン+)	0.80~1	0.991
	Sp(レジン+)	0.7	1	0.8	0.8	—	expit(logit)#	0.85~1	0.984
	Sp(レジン-)	0.7	1	0.8	0.8	0.8	Correlate with Sp(レジン+)	0.91~1	0.990
* : $\rho$ : correlation coefficient    # : $\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * (\exp(g) / (\exp(g) + 1))$ $g = \ln(\alpha / (1 - \alpha)) + \beta * \ln(u / (1 - u))$ , $u = \text{Uniform}(0 \sim 1)$									
Results	Analysis		OR Simulation Results						
			Median	2.5%tile~97.5%tile					
	Conventional		1.76	1.2 ~ 2.58					
	Correlated*	Systematic	1.99	0.71 ~ 14.93					
		Total Error	2.07	0.7 ~ 15.15					
	Independent		1.92	0.23~24.2					
			1.95	0.22~24.63					



Figure 5-36 パラメータの分布（エラーとなったものを除く）



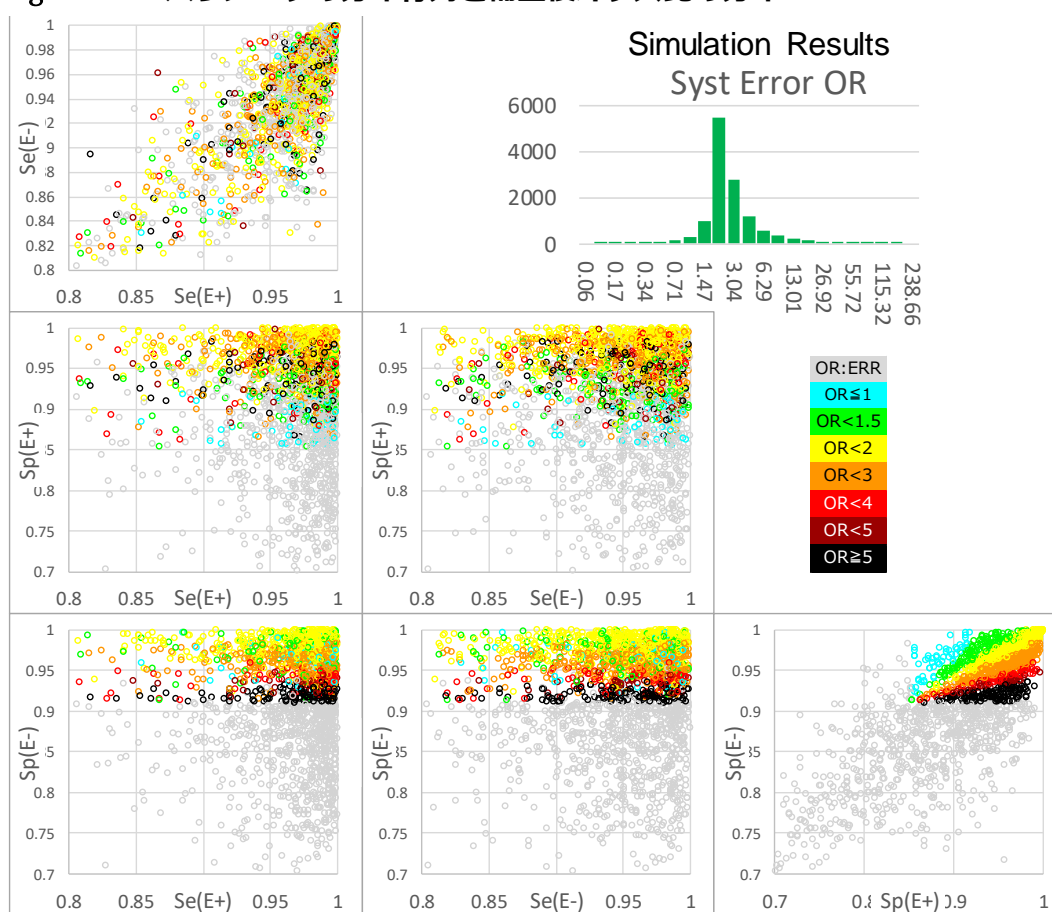
#### ✓ 与えたパラメータ分布と用いられたパラメータ分布が異なる場合

Table 5-18 は Table 5-15,

Table 5-16 を修正した感度分析の概要の提示である。与えたパラメータとエラーとならず用いられたパラメータは表として示すことはできるが、パラメータと補正後の結果の関係をグラフを用いず、文章、統計量（数値表）によりわかりやすく示すことは困難であろう。また、Figure 5-33 のように複数のグラフを並べて視覚的な比較を示す場合はグラフ間で軸や色分けを揃えることは重要である<sup>注68</sup>。このように複数のパラメータを扱う場合は特徴的なパラメータに限定する、カテゴリー化を行うなど 1 つの図表で表す情報量を減らす等の工夫が必要となる。通常、1 つのグラフに表現できる情報数はグラフの軸数と + 1（色）である。5.6.5 モデル、パラメータと補正結果の吟味で説明したように、影響の小さいあるいはトレンドが単調なバイアスの原因を simple sensitivity analysis や multidimensional analysis とする、あるいは感度分析の対象から除くことで結果提示の複雑さを改善できる。なお、感度分析は研究結果を評価する探索的な行為であるので、主要な解析の結果と混同されないように明確に区別する必要がある。<sup>87)</sup>

注68：見やすさのため、散布図はランダムに 2000 回までの試行を表示している。

Figure 5-37 パラメータの分布行列と補正後オッズ比の分布



### 5.7.3. 感度分析結果の解釈

感度分析結果の解釈は discussion, limitation あるいは独立した sensitivity analysis に感度分析についてであることを明示して記載する。残念ながら、解釈の仕方について教科書的な紹介は難しく、弊害が生じることも考えられるため、留意点について説明する。

感度分析の目的の1つは系統誤差と不確実性を考慮させて研究結果の過信、解釈の恣意的な偏りを防ぐことである。このため例えば感度分析による補正オッズ比がより小さくなる場合、研究の結果が有益性を示すものであれば“言い過ぎ”の可能性が低くなる感度分析の結果が重視され、有害性を示す場合は“見落とし”の可能性から研究結果が重視されるべきで、補正オッズ比がより大きくなる場合に有益性を誇張するような扱いは避けるべきであると考え。つまり“conservative”な扱いである。反面、感度分析の目的には不確実性による研究結果の過度に控えめな評価、扱いを改善し意思決定を促す役割もある<sup>注69</sup>。このように場面によって感度分析の結果の位置付けは変わり得る。同様に感度分析そのものも恣意的に偏ったものとなり得る。内部検証研究あるいは外部検証研究が行われているなど、対象とするバイアス、用いるパラメー

注69：例えば、影響があると強く疑われる交絡要因であっても、実際にはそれほど大きく結果は変わらないなど。

タなどが妥当なものであることを説明することでこのような偏りは回避されると思われるが、良い情報源が利用可能でない場合などでも、検討したバイアスの候補、行った感度分析の手順、結果を図表などにより丁寧に説明することで偏らないようにできるであろう<sup>注70</sup>。

感度分析の解釈においては対象要因（バイアス）、条件（シナリオ、情報源）と方法（計算方法、パラメータの分布など）の確認、理解も重要であるので methods の項などで表に整理している場合はこの参照を求め、示されていない場合は整理して記述（表の利用を含む）する。

事例としては、2 型糖尿病患者におけるアスピリンの有効性を評価した Ekström N（2013）<sup>39)</sup>は Discussion の項に 8 行で未測定交絡要因の感度分析の結果の解釈を示し、結果が覆るには相当に大きな交絡要因が必要としているが、Results の項に用いた仮想のハザード比は 1.3 であることが記載されているため読者は感度分析が十分であるか否かを考える事が可能である。しかし、喫煙率はアスピリン使用者と非使用者でほとんど同じであること、調整解析の共変量に用いたことが説明されているが、喫煙によるオッズ比は記載されていない。喫煙のようなよく知られたリスク要因の影響度が示されていれば解釈に役立つであろう<sup>注71</sup>。また、他の研究における同じバイアスの原因の事例を示すことも理解に役立つ<sup>84)</sup>。

感度分析の結果が意思決定に影響を与え得る、つまり研究の主解析の結果と大きく異なる可能性が示された場合は、影響のあるパラメータの説明を加えることが解釈に役立つであろう。例えば、5.7.3 の例のアウトカムの誤分類の結果は補正後のオッズ比が 1 以下となる可能性が示されているが、Figure 5-34 から 1 以下となるのは曝露なしの特異度が 0.999 以下かつ曝露ありの特異度が 0.936 以下の場合に限られており、曝露によらず特異度が同程度であれば 1 以下にならない、つまり条件として限定的であることから実際に 1 以下となることは考えにくいと解釈できるであろう。

また、少数の限られたバイアスのみを扱った場合は、感度分析の対象としなかったものについても理由を示すことを推奨する。例えば、職業によるレジン樹脂曝露の例で曝露の特定法が配属記録とその作業区画から識別されている場合は誤分類の恐れが小さくレジン樹脂曝露を感度分析の対象としないことは合理的と思われるが、自己申告によるものであれば感度分析の結果には疑問が生じる。

更なる検証性のある、あるいはより大規模な研究を求めるコメントは研究結果の報告の結語の常套句であるが、感度分析によって研究の問題点の定量的な提示と意思決定に解決が必要な課題が明確である場合は形式的ではなく具体的な必要性を訴えることが可能である。

注70：教科書などで推奨されているとして本書でここまで説明してきた。恣意的に偏った感度分析は気づかれる。また、感度分析及びその結果の解釈が恣意的であれば、このような説明は困難であろう。

注71：この研究における喫煙のオッズ比が仮に 1.3 であれば、喫煙と同程度の要因が想定され、過少ではないとみなせるが、喫煙のオッズ比が仮に 4 である場合、1.3 は過少で不十分と考えることができるであろう。

## 5.8. 教科書的ではないが現実的で簡便な手順による感度分析

ここまで、感度分析を紹介した Lash T.L. (2014) <sup>4)</sup> にならって教科書的なアプローチを紹介した。この扱いは検証研究の実施も選択肢に含むため、最初から（「一度で」という意味で）最適と考えられる感度分析を行うことを意図している。対象の研究を行った研究者自身が論文、報告書を作成する場合は感度分析が恣意的にならないようにする側面からもこのような“waterfall”型の手順には意味があるであろうが、用いるパラメータの情報収集を除けば、PSAであっても、バイアスの原因を個別に扱う範囲では比較的容易に実施可能であるので、「とりあえず、適当な値で試してみる」という発想は特に感度分析の必要性、根拠となる情報を丁寧に収集する必要がある影響が大きいバイアス（パラメータ）を判断する場合や、企業の安全性評価者など日常的に多くの論文を評価する場合には有用であると考えられる。前者（試してみる）では、容易に入手可能な範囲では根拠となる情報の乏しい、あるいは取得にコスト、リソースを要するパラメータが結果への影響が小さければ、常識の範囲内で大きな（広い）値を用いる、またはバイアスがないとして扱うなど、感度分析で労力を要するパラメータの根拠となる情報の収集を簡略化可能か判断できる。また、後者では法令に従って評価、報告が必要であるため、論文の結果の意味・意義、信頼性・頑健性および外挿性などを評価記録としてまとめ、対応措置の意見を提出することとなるが、概して、バイアスの存在、限界などを文学的、定性的に陳述し保守的に措置保留の結論を述べる場合が多いであろう。しかし、具体的にバイアスがあると考えられるが、その程度など詳細に明確な根拠がなくても具体的なバイアスの仮定の下で感度分析を行うことで、想定するバイアスを定量的に見える化することが可能である。たとえば、Nevriana A (2017) <sup>45)</sup> は不眠症治療薬である Zolpidem および Zopiclone の使用と自動車運転事故リスクとの関係をスウェーデンの住民登録、交通事故及び調剤データベースをリンクさせた二次的データに基づく case-control study と case-crossover study により調査した論文である。結果は、新たに Zolpidem を開始した場合に最もオッズ比が大きく（調整 OR 2.27; 95%信頼区間[CI] 1.21-4.24）、続いて Zolpidem および Zopiclone を併用した場合に大きく、調整された OR は 2.20、95%信頼区間 (CI) は 1.21-4.00 であった。case-crossover study では、Zolpidem または Zopiclone を新たに開始後 2 週間で最も高いリスクを示した (OR 2.66; 95%CI 1.04-6.81)。この報告を安全監視の研究報告として評価する場合、バイアスの存在のみを踏まえると

「国内の使用上の注意には自動車の運転、高所での作業、危険を伴う機械の操作について注意することを記載しているが、不眠症自体が日中の眠気、注意力の低下などを生じて自動車運転事故のリスクになると考えられ、適応による交絡が生じると考えられる。しかし、本研究では例えば index date の前日の睡眠時間や不眠症の有無などのデータは得られておらず、バイアス（適応による交絡）の影響が考えられる。」

といった記述となるであろう。しかし、バイアスを定量的に示すことが可能な感度分析を行うと Figure 5-38 の結果となり、

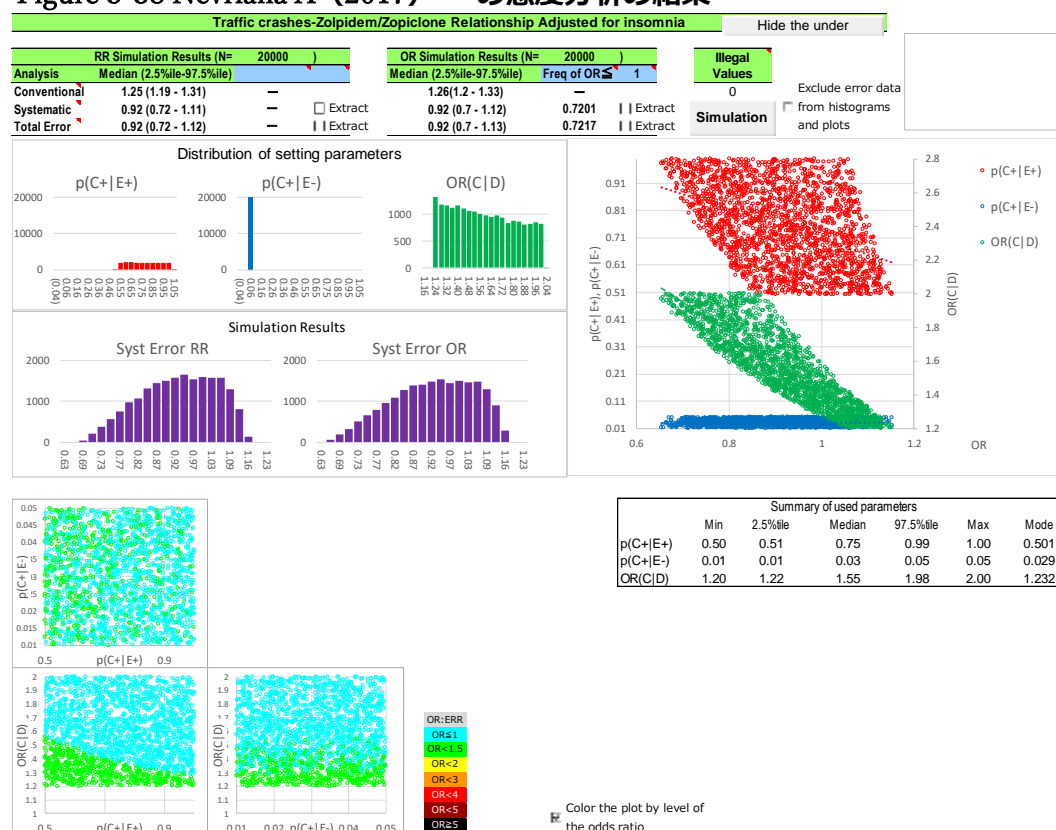
「Lash T.L. (2009) <sup>8)</sup> の定量的感度分析の手法を用いて case-control データに対して不眠症



を未測定交絡要因として補正を試みた。不眠症による交通事故のリスクなど根拠となる統計データは入手していないが、不眠症の事故オッズ比を 1.2~2 (対数一様分布)、不眠症者の割合を曝露群は治療により「不眠」は半数が改善していることを想定して 50%~100% (一様分布)、Zolpidem または Zopiclone の曝露率が case 群 7.9%, control 群 6.1% であることから非曝露群にも未治療の不眠症者は 1%~5% (一様分布) (不眠症の治療率を 50%~80% と仮定) 含まれていると考えられるとした場合、補正後オッズ比は 72% が 1 以下、不眠症の事故オッズ比 1.6 以上では全ての組み合わせで補正後オッズは 1 以下となり、曝露により事故は少なくなり得る結果となった。この結果から、本報告は不眠症者の薬物治療が自動車運転事故を増加させることは示しておらず、むしろ治療しないことによりリスクが増加する可能性が示唆されると解釈することも考えられ、不眠症による交通事故リスク増加を踏まえた判断が必要と考える。」

といった、ベネフィット・リスクを含めた記述が可能である。この例は説明のためにインパクトのある例示としており、当然、このような仮定については議論が必要であるが、このようなパラメータの設定方法であれば、多くの調査は不要で、この分野に習熟している当該医薬品の安全性評価担当者であれば容易であろう。ただし、安全性評価担当者は研究の結果を否定する目的で行わないよう襟を正し、慎重にリスクの報告に向き合う必要があるのは当然である。

Figure 5-38 Nevriana A (2017) <sup>55)</sup> の感度分析の結果





## 6. 感度分析の実施例

### 6.1. 例 1：主に Simple sensitivity analysis を用いたもの

Chang ら (2016)<sup>88)</sup>の論文を用いてアウトカムの誤分類に対する Simple sensitivity analysis を行う。本事例の論文では本書の対象である感度分析は行われていない。著者が論文の査読者からアウトカムの誤分類の可能性について指摘を受け、事後的にバリデーションとして感度分析の項を論文に含めることを想定して例示する。

#### 6.1.1. Chang ら (2016) の論文

はじめに、Chang ら (2016) の論文について紹介する。

米国内科学会 (ACP) の成人 2 型糖尿病患者に対する経口血糖降下薬に関するガイドラインでは、メトホルミンのみが第一選択薬とされている。国内においては、患者の状態をみて医師により治療薬が決定されるが、メトホルミンも第一選択薬として使用される (糖尿病治療ガイド 2016-2017)。一方で、メトホルミンの国内添付文書においては、「重篤な乳酸アシドーシスを起こすことがあり、死亡に至った例も報告されている。乳酸アシドーシスを起こしやすい患者には投与しないこと。」と注意喚起されている。

Chang ら (2016) は、MDV (Medical Data Vision) 社が提供する複数の病院の DPC 及びレセプト情報によるデータベースを用いて、日本人 2 型糖尿病患者におけるメトホルミン使用患者とメトホルミン以外の糖尿病薬使用患者の乳酸アシドーシスの発生率を推定した。解析集団のメトホルミン使用患者は 80,258 人、メトホルミン以外の糖尿病薬使用患者は 252,019 人であり、乳酸アシドーシスの条件を「乳酸測定後に乳酸アシドーシスの診断がなされ、重炭酸ナトリウム投与又は血液透析を受けた」と定義した場合、乳酸アシドーシスが同定された患者はそれぞれ 7 例、23 例であった (Table 6-1)。また、Cox 比例ハザードモデルによるメトホルミン以外の糖尿病薬使用患者に対するメトホルミン使用患者の調整後のハザード比および 95% 信頼区間は 0.92 (0.33-2.55)、傾向スコアマッチングを用いたハザード比および 95% 信頼区間は 0.90 (0.26-3.11) であった。これらの結果により、著者らはメトホルミンによる乳酸アシドーシスのリスク増加はないと結論付けている。

Table 6-1 乳酸アシドーシスが同定された患者の発生率

	LA 患者	Crude		Age-sex adjusted*	
		発生率 <sup>#</sup> 95%CI	リスク比 95%CI	発生率 <sup>#</sup> 95%CI	リスク比 95%CI
メトホルミン 使用患者：133,614 人年 推定 77,545 人	7	5.24 2.50-10.99	0.84 0.36-1.97	5.80 2.68-12.57	1.00 0.41-2.47
メトホルミン以外の 糖尿病薬使用患者：370,555 人年 推定 238,260 人	23	6.21 4.12-9.34	Ref.	5.78 3.73-8.96	Ref.

論文の Table を引用改変

LA 患者：乳酸アシドーシスが同定された患者、#：Incidence: 100,000 人年あたり

\*：ポアソン回帰分析 (年齢区分(18-64, 65-74, 75-), 性別で調整.)

### 6.1.2. 感度分析の必要性

乳酸アシドーシスの症状は嘔吐、頭痛、下痢、気分不良、意識障害（意識レベル低下）など低血糖の症状と重なる部分もあるが、メトホルミンでは稀ではあるがよく知られた副作用であるため（診断の事前確率が高い）、血液ガス分析でアシドーシスや血液検査でアニオンギャップが認められれば乳酸アシドーシス（Lactic acidosis : LA）と診断され、血中乳酸値が測定されることは少なく、測定されている症例では血液透析の経過観察のために事後的に行われている場合も散見される。また、かかりつけ医などを受診、処置後に病院に転送されることも少なくなく<sup>注72</sup>、診断時の検査や処置が病院の診療録にはない場合も少なくない。したがって、論文中の定義「乳酸測定後に LA の診断がなされ、重炭酸ナトリウム投与又は血液透析を受けた」がメトホルミン投与群においては診断の実態を反映していない可能性がある。

このような場合は、乳酸アシドーシスの条件を「乳酸測定後に乳酸アシドーシスの診断がなされ、重炭酸ナトリウム投与又は血液透析を受けた」と定義すると、差異誤分類（メトホルミン投与/非投与の群間で差異がある）が生じていると考えられる。ただし、PH を測定する血液ガス分析は種々の目的で行われるため、検査の実施を乳酸アシドーシスの識別条件とすることも適切ではないと考えられる。

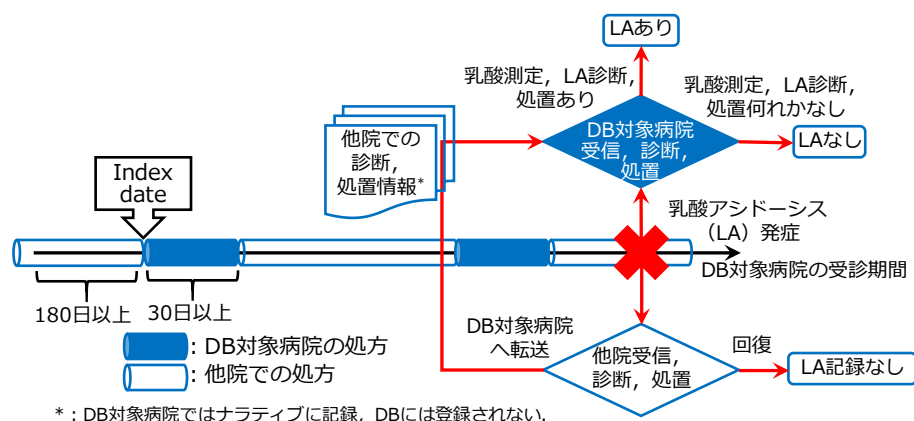
このため、感度分析を本報告の結果が覆る可能性を示す目的で実施する意味があると考える。

以下に、パラメータの設定、誤分類の感度分析について、順に解説する。

### 6.1.3. どのようなバイアスを感度分析の対象とするか？

「6.1.2. 感度分析の必要性」で触れたとおり、本事例ではアウトカムである乳酸アシドーシスの特定に曝露の有無で異なる誤りである選択的な誤分類が生じていると考えられるが、それ以外についても順に確認する。そのため、バイアスが生じやすい例として病診連携により治療されている患者の曝露とアウトカム発生をイメージした Figure 6-1 を作成した。

Figure 6-1 病診連携の行われている患者の曝露とアウトカムのイメージ



注72 : メトホルミン製剤の製造販売会社へのインタビューより

#### 6.1.3.1. 曝露の誤分類について

本件ではメトホルミン製剤を曝露とし、曝露なしを他の糖尿病治療薬の使用と定義されている。曝露記録は病院のDPC、レセプト、処方箋データであり、他院での処方、院外処方の調剤時変更はデータが欠損していると思われる。しかし、本件では病院データにおいて経口糖尿病薬の処方を受けている患者としてコホートが定義され、コホート内での新規イベントが扱われている。病診連携が行われている患者では曝露の誤分類が生じやすいと考えられ、このことへの対処としては曝露定義期間÷処方量（日数）が一定率を超える患者を対象とすることが考えられるが、そのような対処はされていない。また、解析ではCox比例ハザードモデルの時間依存共変量として曝露が扱われているが、その曝露の定義について詳細は説明されていない。そのため、Figure 6-1の例はメトホルミン曝露ありとして扱われたのか否かが明確でない。しかし、メトホルミン製剤は糖尿病治療として広く使用され、また何で治療しているかは重要な情報であることから、病診連携によって曝露が誤ることは少ないと考えられる。

#### 6.1.3.2. アウトカムの誤分類について

本事例ではアウトカムである乳酸アシドーシスの新規発症を「乳酸測定後に乳酸アシドーシスの診断がなされ、重炭酸ナトリウム投与又は血液透析を受けた」と定義され乳酸アシドーシスの診断は、ICD-10コード（のアシドーシス）により絞り込んだのち、テキストサーチにより乳酸アシドーシスの診断名が特定されている。処置が条件にされているため、比較的重篤な乳酸アシドーシスが対象である。このような3種類のすべての条件を満たす必要があるため、6.1.2で説明した通り、乳酸アシドーシス発症が漏れることは少なくないと考えられる。また、反対に一旦は乳酸アシドーシスとされたが、乳酸測定結果から否定された場合は誤って乳酸アシドーシスありとされている場合、ショック状態により酸素不足から血中乳酸値が高くなり乳酸アシドーシスと傷病名が付与されている場合（ショックに伴う二次性の場合）も考えられ、これらの確率はメトホルミン曝露群では乳酸アシドーシス診断の事前確率の高さからより多く生じると考えられる。また、本事例ではアウトカムは必ず有無の何れかに分類されているが、例えば、乳酸アシドーシスと診断され、血液透析を受けているが、乳酸測定がなされていない患者は乳酸アシドーシスなしと断定できず、不明として扱うべきかもしれない（その場合はアウトカムが測定されていないとして扱い、脱落同様に選択バイアスとして扱う）。

よって、アウトカムは差異誤分類が生じていると考えられ、また、アウトカムの誤分類は結果に直接影響し、本事例のアウトカムの件数は少ないことから感度分析の対象とする必要があると考える。

#### 6.1.3.3. 未測定交絡要因について

乳酸アシドーシスのリスクは腎障害以外に、ショック、脱水、疲労、食欲不振、多量の飲酒などであり、腎障害以外は診療情報DBに明示的な記録がなされにくく、未測定交絡要因となり得るが、曝露の有無間での併存（発生）率に大きな差異があるとする根拠はない。メトホルミン投与においては、効果修飾要因である可能性があるが明確ではない。よって、バイアスで

あると考えられるが、影響はアウトカムの誤分類ほどの大きさではなく、本事例の感度分析の目的から、感度分析の対象とする必要性は低いと考える。

#### 6.1.3.4. 選択バイアスについて

本事例では、2010年1月から2014年8月の間に糖尿病治療薬が処方された18歳以上の2型糖尿病患者について、この間の初回糖尿治療薬処方日をindex dateとして、この前180日間に受診記録があり、乳酸アシドーシスの病歴がなく、index date以後30日以上追跡がされている患者を対象とされた。層別解析にNew userの定義があることから、New user designは用いられていないが、メトホルミン製剤投与中の乳酸アシドーシスの発生に曝露期間との特徴のある関係は認められていないことから、重要なバイアスではないと考えられる。本事例では、対象施設においてメトホルミン製剤の処方が行われていること（採用されている）、血中乳酸測定が行われていること、血液透析の設備があることを条件とした説明はなされていない。2型糖尿病治療におけるメトホルミンの位置づけを考えると、すべての施設で処方がされていると見なせるが、同じデータベースで臨床検査値が得られる施設の2008年4月～2012年1月のデータであるが、保険請求に血中乳酸値の測定が1件以上あった施設は約8割であったことから、乳酸アシドーシスであっても、血中乳酸値の測定がなされることがない施設が含まれている可能性がある。血液透析については、本事例のDBに含まれる医療機関はDPC対象施設であることから、血液透析の設備が全くない施設は少ないと考えられる。以上より、選択バイアスの可能性は考えられるが、本事例は絶対的な乳酸アシドーシスの発現率ではなく相対的なリスクを扱っており、これらバイアスが曝露の有無によって差異が生じるとする根拠はないため、重要なバイアスではないと考えられる。

#### 6.1.4. 乳酸アシドーシスの誤分類：乳酸アシドーシスを尤もらしく特定可能するには？

感度分析をおこなう際、重要なのは感度分析に用いる尤もらしいパラメータを設定することである。本事例の場合では、乳酸アシドーシス識別の感度、特異度の設定が必要となるため、情報源を自発報告、レセプトデータベースのバリデーションデータを用いた場合について紹介する。

##### 6.1.4.1. 自発報告に基づく情報

メトホルミン製剤の製造販売会社にインタビューした結果、副作用自発報告による2010年以降の乳酸アシドーシスの報告（アシドーシス含む）のうち、血中乳酸値の測定があったのは約25%、処置経過情報のある患者のうち、炭酸水素ナトリウムの静注又は血液透析の記載があるのは約40%であったとしよう。

血中乳酸値の測定があった患者のうち、炭酸水素ナトリウムの静注又は血液透析の記載が全患者にある場合と40%である場合の乳酸アシドーシスと診断される感度は、それぞれ25%、10%（ $=25\% \times 40\%$ ）となる。

#### 6.1.4.2. バリデーション研究に基づく情報

本研究で用いられたデータベースは一部の施設で臨床検査値データの利用が可能である。このようなデータベースを用いてバリデーション研究を行ったとして、臨床検査値を最も信頼できる基準としてレセプト（DPC）データによる診断と検証し、その結果に基づいてパラメータ設定する方法を以下に記載する。

データベースにおいて、アシドーシスの診断名、メトホルミン投与の記録に加え、医療機関ごとに PH 値および乳酸値の臨床検査値データが利用可能であること、ならびに実際にデータベースに取り込まれていることを確認する。乳酸アシドーシスは血中乳酸値の検査結果のみで疾患を定義できるため、バリデーション研究として、乳酸値 $>45\text{mg/dL}$ である患者を乳酸アシドーシスと定義した。バリデーション研究を、論文中の定義「乳酸測定後に LA の診断がなされ、重炭酸ナトリウム投与又は血液透析を受けた」と比較するためにまとめたのが、Table 6-2 に示す分割表である。さらに、メトホルミン投与患者では乳酸アシドーシスの診断に必ずしも血中乳酸値の測定が用いられないため、もうひとつのバリデーション研究として、 $\text{PH} \leq 7.35$  または乳酸値 $>45\text{mg/dL}$ である患者を乳酸アシドーシスと定義し、Table 6-3 のとおり論文中の定義と比較した。乳酸値が欠測で PH のみが欠測される場合もありうるため、 $\text{PH} \leq 7.35$  で定義されるアシドーシスも広義の乳酸アシドーシスとした<sup>注73</sup>。

Table 6-2 乳酸値>45mg/dL を乳酸アシドーシスとしたときの分割表

		メトホルミン投与あり			メトホルミン投与なし		
アルゴリズム		LA	non LA	計	LA	non LA	計
1 ハ ム ン	LA	1	1	2	2	0	2
	non LA	2	6410	6412	1	7509	7510
	不明 <sup>#</sup>		105			159	

LA: 乳酸アシドーシス, non LA: 乳酸アシドーシスでない

投与又は血液透析が行われている。

Table 6-3 乳酸値>45mg/dL または PH≤7.35 を乳酸アシドーシスとしたときの分割表<sup>注73)</sup>

		メトホルミン投与あり			メトホルミン投与なし		
アルゴリズム		LA	non LA	計	LA	non LA	計
1 ハ 公	LA	1	12	13	2	7	9
	non LA	2	6401	6403	1	7460	7461
	不明 <sup>#</sup>		104			202	

LA: 乳酸アシドーシス, non LA: 乳酸アシドーシスでない

# : 血中乳酸、血液 PH の検査は未実施であるが、乳酸アシドーシス/アシドーシスの診断あるいは重炭酸ナトリウム投与又は血液透析が行われている。

注73： 乳酸値は提供されていないが、PH 値は提供されている医療機関を含むため Table 6-3 の合計は Table 6-2 と合わない。



Table 6-2 と Table 6-3 を基に以下のとおり感度と特異度を算出した。これらが感度分析を実施する際のパラメータとなる。<sup>注74</sup>

Table 6-4 乳酸値>45mg/dL を乳酸アシドーシス【バリデーション】としたときの感度と特異度

	メトホルミン投与あり	メトホルミン投与なし
感度	0.50 (=1/2)	1.00 (=2/2)
特異度	0.9997 (=6410/6412)	0.9999 (=7509/7510)

Table 6-5 乳酸値>45mg/dL または PH≤7.35 を乳酸アシドーシス【バリデーション】としたときの感度と特異度

	メトホルミン投与あり	メトホルミン投与なし
感度	0.0769 (=1/13)	0.22 (=2/9)
特異度	0.9997 (=6401/6403)	0.9999 (=7460/7461)

#### 6.1.5. 誤分類の感度分析

前項で推定した感度分析に用いるパラメータを以下の Table 6-6 にまとめ、各パラメータについて、結果の誤分類に関する Simple sensitivity analysis をおこなう。

Table 6-6 誤分類の感度分析に用いるパラメータ

	メトホルミン投与あり		メトホルミン投与なし	
	感度	特異度	感度	特異度
自発報告①	0.25	1.00	1.00	1.00
自発報告②	0.10	1.00	1.00	1.00
データベース①	0.50	0.9997	1.00	0.9999
データベース②	0.0769	0.9997	0.22	0.9999

手法の詳細は第 1 部 3.4.2 に記載しているが、Lash T.L.ら（2009）の Applying Quantitative Bias Analysis to Epidemiologic Data に基づき、著者らが提供している EXCEL ファイルを利用すると簡単である（<https://sites.google.com/site/biasanalysis/Home> からで入手可能）。

該当するファイル（Ch6\_Misclassification\_2014\_06\_02.xls）に、Table 6-6 の自発報告①のパラメータ（乳酸アシドーシスと診断された患者が 25%のみ同定されたと仮定）を入力するとオ

注74：論文の Figure 1 と Table 2 でメトホルミン投与ありと投与なしの年数が異なることから、メトホルミン投与ありの推定患者数は  $80258 \times 133614 / 138289 = 77545$  人、メトホルミン投与なしの推定患者数は  $252029 \times 370555 / 391970 = 238260$  人と推定し、感度分析をおこなった。

ツズ比は 3.74 となる。同様に、自発報告②、データベース①、データベース②のパラメータを入力すると、補正後オッズ比はそれぞれ 9.36, 121, 173.2 となる<sup>注75</sup>。

### 例：自発報告①のパラメータを用いた解析

Input Bias Parameters				Variable Names		Instructions	
Se (メトホル)	0.25	<	>	Outcome	乳酸アシドーシス	Enter the bias parameters in the blue cells to the left and the observed data in the blue cells below. Cells in green give the results after correcting for outcome misclassification. Note: green cells are expected values and therefore do not have to be integers.	
Se (メトホル)	1.00	<	>	Exposure	メトホルミン		
Sp (メトホル)	1.00	<	>				
Sp (メトホル)	1.00	<	>	Error Check:	No errors		

Data (Enter メトホルミン-乳酸アシドーシス Data in Blue Cells)			
Observed Data			
	メトホル	メトホルミ	Total
乳酸アシドーシス +	7	23	30
乳酸アシドーシス -	77538	238237	315775
Total	77545	238260	

Observed		Measure (95% CI)
RR (メトホルミン-乳酸アシドーシス)	0.94	(0.4 - 2.18)
OR (メトホルミン-乳酸アシドーシス)	0.94	(0.4 - 2.18)

Corrected Data			
	メトホル	メトホルミ	Total
乳酸アシドーシス +	28.0	23.0	51.0
乳酸アシドーシス -	77517.0	238237.0	315754.0
Total	77545	238260	

Corrected		Measure
RR (メトホルミン-乳酸アシドーシス)	3.74	
OR (メトホルミン-乳酸アシドーシス)	3.74	

また、この解析は R のパッケージ「episensr」を用いても解析できる。

プログラム：

```
install.packages("episensr")
library("episensr")
sens1 <- misclassification(matrix(c(7, 23, 77538, 238237),
dimnames = list(c("LA", "non LA"),
c("Metformin +", "Metformin -")),
nrow = 2, byrow = TRUE),
type = "outcome",
bias_parms = c(.25, 1, 1, 1))
print(sens1)
```

アウトプット：

--Observed data--

Outcome: LA

Comparing: Metformin + vs. Metformin -

	Metformin +	Metformin -
LA	7	23
non LA	77538	238237

注75：次のページで説明しているが、データベース①、データベース②の設定では負のセルが生じているため個の値に意味はない。

		2.5%	97.5%
Observed Relative Risk:	0.9351204	0.4012821	2.1791410
Observed Odds Ratio:	0.9351146	0.4012484	2.1792964

Misclassification Bias Corrected Relative Risk: 3.740482

Misclassification Bias Corrected Odds Ratio: 3.741472

ただし、ここで注意が必要である。データベース①、データベース②のパラメータを入力した場合、修正後の四分表に負のセルがあることに気付く。これは、例えば補正值の左上のセル（メトホルミン投与あり、乳酸アシドーシスあり）を算出する際、「観察された左上のセルの人数－観察されたメトホルミン投与ありの人数×（1－メトホルミン投与ありの特異度）」を式に含むことに基づく。本事例のような稀な疾患の場合、特異度が1に近い値であったとしても、稀であることが補正に大きな影響を与え、適切な補正とならない。また、このことは自発報告①、②の設定も不安定な状況下であることも示唆している。したがって、得られた結果の解釈に注意が必要である。なお、データベース②の特異度が1つ下の桁で異なる設定（0.99997と0.99999）とした場合はエラーとならず、補正後オッズ比は1.99、特異度に差がなく何れも0.99999とした場合の補正後オッズ比は2.66である。

### 例：データベース②のパラメータを用いた解析

Input Bias Parameters				Variable Names		Instructions	
Se (メトホルミン+)	0.0769	< >		Outcome	乳酸アシドーシス	Enter the bias parameters in the blue cells to the left and the observed data in the blue cells below. Cells in green give the results after correcting for outcome misclassification. Note: green cells are expected values and therefore do not have to be integers.	
Se (メトホルミン-)	0.22	< >		Exposure	メトホルミン		
Sp (メトホルミン+)	0.9997	< >					
Sp (メトホルミン-)	0.9999	< >					
Error Check:				Negative Cell			
Data (Enter メトホルミン-乳酸アシドーシス Data in Blue Cells)							
Observed Data							
	メトホルミ	メトホルミ	Total				
乳酸アシドーシス +	7 <sup>a</sup>	23 <sup>b</sup>	30				
乳酸アシドーシス -	77538 <sup>c</sup>	238237 <sup>d</sup>	315775				
Total	77545 <sup>m</sup>	238260 <sup>n</sup>					
Corrected Data							
Corrected Data							
	メトホルミ	メトホルミ	Total				
乳酸アシドーシス +	-212.3 <sup>A</sup>	-3.8 <sup>B</sup>	-216.1				
乳酸アシドーシス -	77757.3 <sup>C</sup>	238263.8 <sup>D</sup>	316021.1				
Total	77545 <sup>M</sup>	238260 <sup>N</sup>					
				Corrected		Measure (95% CI)	
				RR (メトホルミン-乳酸アシドーシス)		0.94 (0.4 - 2.18)	
				OR (メトホルミン-乳酸アシドーシス)		0.94 (0.4 - 2.18)	
				Corrected		Measure	
				RR (メトホルミン-乳酸アシドーシス)		173.67	
				OR (メトホルミン-乳酸アシドーシス)		173.2	

## 2. 結果のまとめと提示

論文に感度分析の項を追記する際には、感度分析のパラメータの推定方法および解析方法を記載のうえ、感度分析の結果をまとめるとよい。ただし、感度分析のパラメータ設定の際に仮定を用いていることも多いため、Limitation についての記載を含めることが肝要である。

感度分析の結果を論文中に含め、Limitation を考慮したうえでの結果を解釈することは、研究の価値をより高めるかもしれない。

Chang ら (2016) はメトホルミンによる乳酸アシドーシスのリスク増加はないと結論付けているが、結果の誤分類に対する Simple sensitivity analysis の結果、自発報告に基づく場合のオッズ比は 3.74, 9.36, バリデーションデータに基づく場合のオッズ比は算出できないが、特異度を高くとしたより控えめな設定の補正後オッズ比は 2.66 という結果であった (以下の表を参照)。

バリデーション研究はイベント数が不足しているためパラメータの精度に問題がある可能性があり、また稀な疾患であるが故に補正した値も不安定である可能性があるが、ショック状態で救急搬送され、乳酸アシドーシスを疑い、乳酸測定され、値は正常であるが急性腎不全などの処置の必要性から血液透析が行われる、つまり偽陰性が生じる可能性および乳酸アシドーシスの報告 (アシドーシス含む) のうち、限られた患者にのみ血中乳酸値の測定がされている可能性が考えられることから本事例ではバイアスが生じており、メトホルミン以外の糖尿病薬使用患者に対するメトホルミン使用患者のリスクが高い可能性は否定できない。

結果のまとめ (例)

	解析方法・感度分析のパラメータ	解析結果 (95%CI)
Chang ら (2016)	Crude <sup>#1</sup>	リスク比: 0.84 (0.36-1.97)
	ポアソン回帰分析 (年齢区分, 性別で調整)	リスク比: 1.00 (0.41-2.47)
	Cox 比例ハザードモデル	ハザード比: 0.92 (0.33-2.55)
	傾向スコアマッチング <sup>#2</sup>	ハザード比: 0.90 (0.26-3.11)
自発報告に基づく情報①	誤分類の感度分析 $SE_{E+}: 0.25, SP_{E+}: 1.00, SE_{E-}: 1.00, SP_{E-}: 1.00$	オッズ比: 3.74
自発報告に基づく情報②	誤分類の感度分析 $SE_{E+}: 0.10, SP_{E+}: 1.00, SE_{E-}: 1.00, SP_{E-}: 1.00$	オッズ比: 9.36
バリデーションデータに基づく情報①	誤分類の感度分析 $SE_{E+}: 0.50, SP_{E+}: 0.9997, SE_{E-}: 1.00, SP_{E-}: 0.9999$	推定不能 (修正後の四分表に負のセルを含む)
バリデーションデータに基づく情報②	誤分類の感度分析 $SE_{E+}: 0.0769, SE_{E-}: 0.22, SP_{E+}: 0.9997, SP_{E-}: 0.9999$	推定不能 (修正後の四分表に負のセルを含む)
バリデーションデータに基づく情報②'	誤分類の感度分析 $SE_{E+}: 0.0769, SE_{E-}: 0.22, SP_{E+} = SP_{E-} = 0.9999$	オッズ比: 2.66

$SE_{E+}$ : 感度 (メトホルミン投与あり),  $SE_{E-}$ : 感度 (メトホルミン投与なし),

$SP_{E+}$ : 特異度 (メトホルミン投与あり),  $SP_{E-}$ : 特異度 (メトホルミン投与なし)

#1: 論文中の Table 2 のメトホルミン投与ありおよび投与なしの LA 発現 (7/133614 および 23/370555) を基に算出した。

#2: 傾向スコアの層別マッチングに加えて、傾向スコアの層、肝臓に関する病歴、腎臓に関する病歴が調整に用いられた。

#### 論文に感度分析の項を追記する際の注意

Chang ら (2016) はメトホルミンによる乳酸アシドーシスのリスク増加はないと結論付けているが、結果の誤分類に対する Simple sensitivity analysis の結果、自発報告に基づく場合のオッズ比は 3.74, 9.36, バリデーションデータに基づく場合のオッズ比は算出できずという結果であった。バリデーション研究はイベント数が不足しているためパラメータの精度に問題がある可能性があり、また稀な疾患であるが故に補正した値も不安定である可能性があるが、メトホルミン以外の糖尿病

薬使用患者に対するメトホルミン使用患者のリスクが高い可能性は否定できない。そこで、論文に感度分析の項を追記することを考える。

論文に感度分析の項を追記する際には、感度分析のパラメータの推定方法および解析方法を記載のうえ、感度分析の結果をまとめるとよい。ただし、感度分析のパラメータ設定の際に仮定を用いていることも多いため、Limitation についての記載を含めることが肝要である。

感度分析の結果を論文中に含め、Limitation を考慮したうえでの結果を解釈することは、研究の価値をより高めるかもしれない。

## 6.2. 例 2：主に Probabilistic sensitivity analysis (PSA：確率的感度分析) を用いたもの

本項では曝露およびアウトカムの誤分類、未測定（未知）交絡要因に対する感度分析に Probabilistic analysis が用いられた論文を取り上げる。説明を簡潔にするため、未測定（未知）交絡要因のみについて紹介する。

### 6.2.1. S Youngquist ら（2008）の論文

本項で扱う事例は S Youngquist による、院外心停止（out-of-hospital cardiac arrests；OHCA）患者を対象とした、OHCA の原因である無脈性電気活動（Pulseless electrical activity；PEA）または心室細動（ventricular fibrillation；VF）と過去のβ遮断薬服薬の関連を見た、ロサンゼルス内のある公立の病院データ（対象期間：2001 年 1 月 1 日～2006 年 12 月 31 日）を用いた後ろ向きコホート研究である。<sup>83),注76</sup> VF、PEA の頻度とβ遮断薬服薬が関連しているのではないかというリサーチクエスションから本研究実施に至った。

未補正の結果は Table 6-7 のとおりである。「(粗) オッズ比」とその信頼区間はデータに偏りがなく Random error のみを仮定した結果である。

Table 6-7 β遮断薬有無とアウトカムの関連

		β遮断薬 (X)		
		有	無	Total
アウトカム (Y)	PEA	49	51	100
	VF	16	63	79
	Total	65	114	179
(粗) オッズ比 (95%CI)		3.78 (1.93, 7.43)		

注76：コホート研究ではあるが、院外心停止のあった入院患者を抽出し、外傷、薬物によるものなどを除外してアウトカムである PEA と VF の群間で比較されている。つまり、アウトカムが特定され、遡って曝露が確認されている。よって、通常のコホート研究で統計量に用いるリスク比は適切でなく、オッズ比を用いる。



### 6.2.2. 本事例における感度分析の必要性

本研究は、病院 1 施設の診療録を用いた後ろ向きコホート研究であるため、無作為割り付けは行われておらず、結果の一般化並びに内的妥当性の双方で疑義が生じ得る。したがって、生じ得ると考えられる範囲の種々のバイアスについて感度分析を行うことで、結果の頑健性を評価することは意味があると考ええる。

### 6.2.3. Probabilistic sensitivity analysis の実施

Lash T.L. et al<sup>29)</sup>の Chapter8 によると、PSA は以下の 7 つの Step からなる。ここでは、下記手順 (Step 1 ~7) に沿って解説する。

Step1: Identify the likely sources of important bias

Step2: Identify the bias parameters needed to correct for the bias

Step3: Assign probability distributions to each bias parameter

Step4: Randomly sample from each bias parameter distribution

Step5: Use simple bias analysis to correct for the bias

Step6: Save the correct estimate and repeat steps 4 and 5

Step7: Summarize the corrected estimates with a frequency distribution that yields a central tendency and simulation interval

#### Step1. Identify the Sources of Bias

本項で扱う事例では曝露およびアウトカムの非差異誤分類並びに未測定交絡要因について PSA が実施されている。本項では、そのうち未測定（未知）交絡要因について説明を行う。報告に未測定（未知）交絡要因の具体的記載は無く、Confounding by unknown factors と記載のみであるが、ここでは説明のために要因を特定せず何らかの 2 値データ（有，無）の要因とする。以下に DAG<sup>注77</sup>を記載する。

#### Step2. Identify the Bias Parameters

未測定（未知）の交絡因子に対する感度分析での Bias Parameters は以下である。多くの場合、以下の情報は統計データ等外部情報から入手する。

$p_1$  :  $\beta$ 遮断薬有患者での未知交絡要因「有」の割合

$p_0$  :  $\beta$ 遮断薬無患者での未知交絡要因「有」の割合

$OR_{ZY}$  : 未知交絡要因の有無とアウトカムの関連（オッズ比）

注77： 第 1 部 参考 3-6 参照

本事例では極端に大きな交絡要因があった場合の頑健性を確かめる目的で $p_1$ が 30～50%， $p_0$ が 10～30%， $OR_{ZY}$ を 5～20 と仮定した。

### Step3. Assign Probability Distributions to Each Bias Parameters

Bias Parameters へ分布を仮定する。例示のため、ここでは 3 つの Bias Parameters とも Table 6-8 に示す一様分布を仮定する。

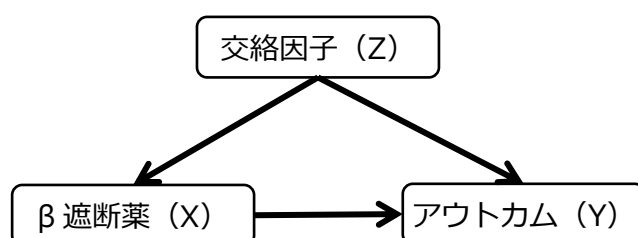


Table 6-8 Bias Parameters

	下限	上限
$p_1$	30%	50%
$p_0$	10%	30%
$OR_{ZY}$	5	20

### Step4. Randomly Sample from Bias Parameter Distributions

$p_1$ ,  $p_0$ ,  $OR_{ZY}$ の 3 つの一様分布から 1 組取り出し得られた乱数値はそれぞれ 36.2%，20.3%，13.5 であった。

### Step5. Use Simple Bias Analysis to Correct for the Bias

本 Step では Step4 で得られた値を用いて Simple Bias Analysis を行う。Step4 で得られた乱数値を用いると、未知交絡要因の「有」と「無」に分けた時、全体の結果は以下の表のようになる。

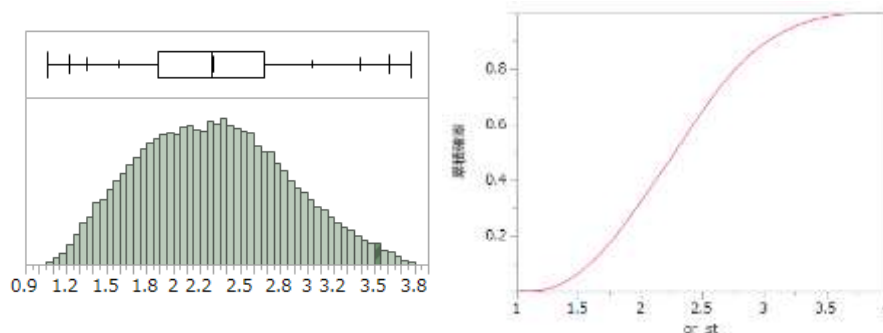
Table 6-9 Simple Bias Analysis の結果

		β遮断薬 (X)					
		Total		未知交絡要因「有」(Z <sub>1</sub> )		未知交絡要因「無」(Z <sub>0</sub> )	
		有	無	有	無	有	無
アウトカム (Y)	PEA	49	51	43.3	39.5	5.7	11.5
	VF	16	63	5.8	12.8	10.2	50.2
	Total	65	114	49.1	52.3	15.9	61.7

その後、β遮断薬有群を標準集団とした標準化法により未測定交絡因子を考慮したオッズ比 ( $OR_{ST}$ ) は 2.42 となる。<sup>注 76</sup>

## Step6. and 7: Resample, Save and Summarize

Step4 及び Step5 を繰り返し（例えば、50,000 回）、集約する。OR<sub>ST</sub>の分布は以下の図のとおりであった。

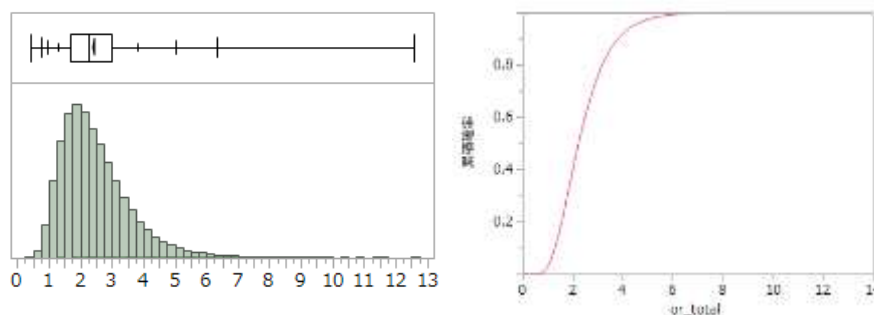
Figure 6-2 OR<sub>ST</sub>の分布

また、以下の表に Table 6-7 の観察された結果であるオッズ比とその 95%信頼区間 (Random error(Conventional result)と表示) 並びに本 Probabilistic analysis により得られた中央値と 2.5%点、及び 97.5%点 (Systematic error と表示) を示す。

Analysis	Median	95% Interval(2.5 <sup>th</sup> , 97.5 <sup>th</sup> percentile)
Random error(Conventional result)	3.78	(1.93, 7.43)
Systematic error	2.28	(1.35, 3.39)

さらに Systematic error 及び Random error を考慮したときの結果は以下のとおりとなる。最終行のとおり、Median は Systematic error と大きく変わらないが、2.5 %点、97.5 %点は更に広がることが分かる。

Analysis	Median	95% Interval(2.5 <sup>th</sup> , 97.5 <sup>th</sup> percentile)
Random error(Conventional result)	3.78	(1.93, 7.43)
Systematic error	2.28	(1.35, 3.39)
Total error analysis	2.24	(0.98, 5.03)



## 6.2.4. 結果のまとめと考察

実施した PSA の設定と結果を Table 6-10 に整理した。未測定交絡要因のオッズ比は 5～20 という大きな値で併存率は必ず PEA 群>VF 群で PEA 群が平均 20%高いとする十分に偏った設定である。OR<sub>total</sub> の 95%頻度区間は 1 を含むが、OR<sub>total</sub> ≤ 1 は 3%に過ぎず (Figure 73), 未測定交絡要因補正後オッズ比は依然としてβ遮断薬曝露にリスクがあることを示し、本研究の結果を支持していると考えられた。

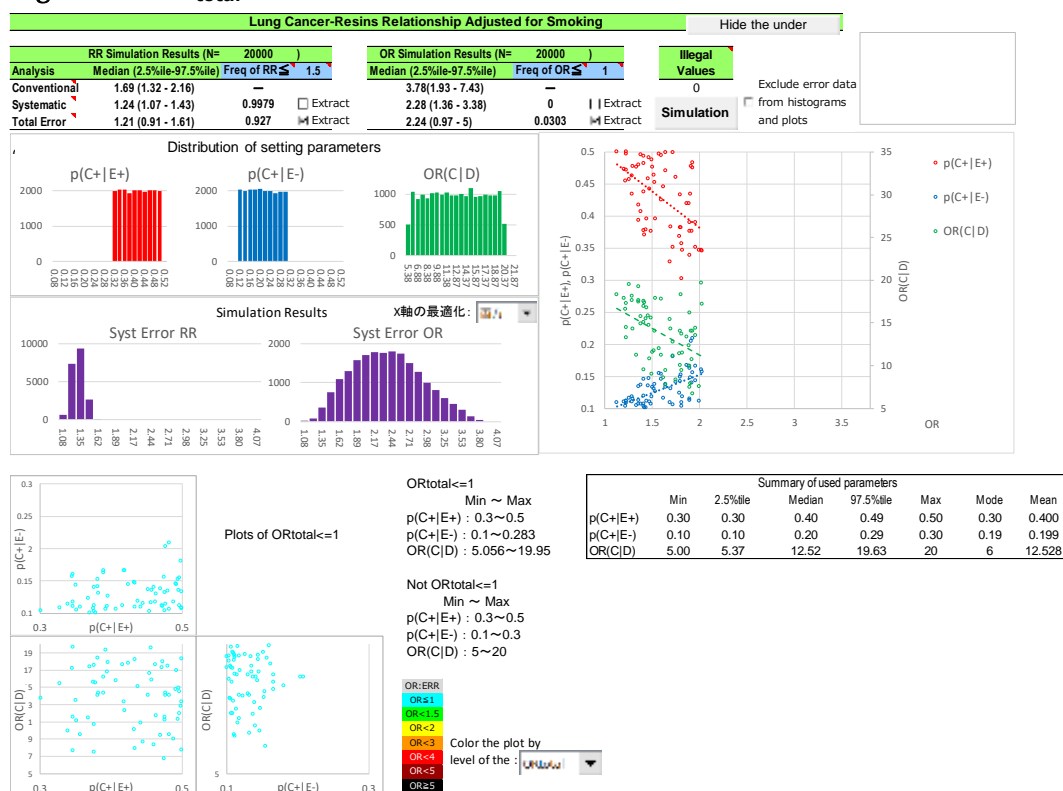
Figure 6-3 ESATJ による結果



Table 6-10 未測定交絡要因の PSA の設定と結果のまとめ

未測定交絡要因	下限	上限	分布形状
未知交絡要因			一様分布
β遮断薬有	30%	50%	
「有」の割合	β遮断薬無	10%	30%
アウトカムとのオッズ比	5	20	
補正後オッズ比 (中央値, 95%頻度区間)	Systematic error	2.28, 1.35～3.39	
	Total error	2.24, 0.98～5.03	

### Figure 6-4 $OR_{total} \leq 1$ となる場合



#### 6.2.5. 読者・意思決定者による追加的考察

本事例では、曝露及びアウトカムの誤分類についても PSA が行われており、十分に慎重な検討が行われているように思われるが、本事例の限界である特定の 1 施設の医療機関の入院患者を親コホートとして、PEA 又は VF の患者のβ遮断薬曝露状況を遡って確認されており、比較する群間でマッチングなど比較妥当性を高めるデザインが用いられていないことを踏まえて考える。

当該医療機関は UCLA の付属病院の Harbor-UCLA Medical Center である。施設の WEB サイト<sup>注78</sup>には、周辺住民は 70 万人であること、支払い能力に関係なく医療サービスが提供されていることが説明されているが、周辺 10km 圏内に 10 件以上の病院がある。Figure 6-5 は地域におけるβ遮断薬の曝露と院外心停止の発生から本研究の対象患者となるまでのフロー図である。PEA は心筋に電気活動はあるが心停止（脈が触れない）の状態であり、除細動などによって回復しないのに対して VF は不整脈の一種で除細動によって回復する。加えて PEA と VF は全く異なる疾患であり、原因、リスク要因も異なるため、発症場所・重症度と医療機関の距離、日常の受診機関、加入保険により搬送される医療機関が選ばれると仮定する。また、β遮断薬は安価な医薬品であり、基礎疾患、医療機関、支払い能力によって曝露確率が異なる可能性がある。これらに加えて本事例の患者数は PEA より一般に発症例が多いと思われる VF の患者数が逆に少ない、また、β遮断薬の使用状況不明が 25%、PEA の 30%、VF の 24%が

注78 : <https://dhs.lacounty.gov/wps/portal/dhs/harbor>



除かれていること、明記されていないが、曝露などは、本人等から聞き取り調査により行われ、死亡した患者が多く除かれていると考えられることから、本事例では Table 6-7 の四分表の各セルの母集団（周辺住民における全ての PEA, VF 発症例）からの抽出率が各々異なる選択バイアスが疑われる。このことから選択バイアスに次のようなシナリオを用いて感度分析を行った結果、ランダム誤差を含まない補正後  $OR_{\text{systematic}}$  の 95% 頻度区間の下限は 1.08 であるが、補正後  $OR_{\text{total}}$  の 95% 頻度区間の下限は 0.8,  $OR_{\text{total}} \leq 1$  は 7.4% であった。

補正後のオッズ比が 1 以下となるのは、4.1.8.1 数式 4-2 より

$$\text{Selection } OR = \frac{S(\text{PEA} + |\beta \text{ blocker+})/S(\text{PEA} + |\beta \text{ blocker-})}{S(\text{VF} + |\beta \text{ blocker+})/S(\text{VF} + |\beta \text{ blocker-})} > OR = 3.78$$

の場合である。

このシナリオ (Table 6-11) は、VF の  $\beta$  遮断薬非使用患者 (Y-|X-) の抽出率 10~20% を基準に、 $\beta$  遮断薬使用患者 (Y-|X+) は -5%、PEA の  $\beta$  遮断薬非使用患者 (Y+|X-) を +5%、PEA の  $\beta$  遮断薬使用患者 (Y+|X+) は非使用患者の +5% の差異を設定したもので、抽出率に根拠はないが、周辺の病院数から基本となる抽出率を仮定し、VF の患者数が PEA の患者数と同じか少し多く (1.2 倍程度) 仮定した。 $\beta$  遮断薬使用患者の抽出率は頑健性の確認を目的としてバイアスが生じる設定としたものでランダムエラーを含めた Selection OR が 3.78 以上となる場合が生じる結果となった (Table 6-12)。この設定がそれほど極端と思われないなら、「結果が覆る可能性が示唆された。」と評価されるであろう。このような抽出率の妥当性は地域の救急搬送の統計などからある程度評価可能であると考ええる。

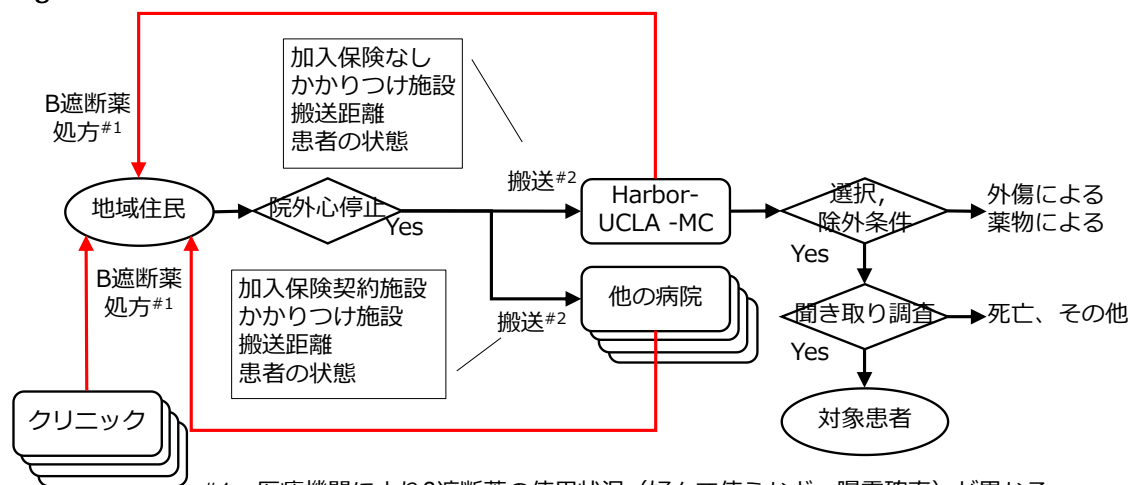
Table 6-11 シナリオを反映したパラメータ

抽出率	S(VF+  $\beta$ blocker-)~Uniform(0.1~0.2),
	S(VF+  $\beta$ blocker+)~Uniform(0.05~0.15),
	S(PEA+  $\beta$ blocker-)~Uniform(0.15~0.25),
	S(PEA+  $\beta$ blocker+)~Uniform(0.2~0.3),
相関	S(PEA+  $\beta$ blocker+)と S(PEA+  $\beta$ blocker-)の相関係数=0.7,
	S(PEA+  $\beta$ blocker+)と S(VF+  $\beta$ blocker+)の相関係数=0.7,
	S(PEA+  $\beta$ blocker-)と S(VF+  $\beta$ blocker+)の相関係数=0.7,

Table 6-12 PSA の結果

Analysis	Median	95% Interval(2.5 <sup>th</sup> , 97.5 <sup>th</sup> percentile)
Random error(Conventional result)	3.78	(1.93, 7.43)
Systematic error	2.01	(1.08, 3.31)
Total error analysis	1.96	(0.8, 4.64)

Figure 6-5 院外心停止発生から研究対象となるフロー



#1：医療機関によりβ遮断薬の使用状況（好んで使うなど、曝露確率）が異なる。  
 #2：患者状態により搬送先を選ぶ優先度は搬送距離、保険の扱いなど異なる。

## 終わりに

本書では、曝露とアウトカムが2値の値の場合の四分表に要約された結果について用いられる感度分析の手順と方法を事例を交えながら紹介しました。基本的で単純な手法ですが、薬剤疫学研究の文献の評価や医療情報データベースを用いた薬剤疫学研究を自ら行う際に利用可能な場面は少なくないでしょう。特に市販医薬品の安全性にかかわる研究はアウトカムが低頻度であるという技術的課題及び倫理的な課題から観察研究に頼らざるを得ませんが、単にバイアスがあるから、測定に信頼性がない（曝露やアウトカムに誤分類がある）からといった理由で結果が控えめに扱われてきたのではないのでしょうか。しかし、感度分析、特に確率的感度分析（PSA）により、研究結果の限界を定量的に可視化することが可能になります。ExcelのToolなどを用いれば、議論の最中に異なる意見、基準を用いた結果をその場で提示することも容易です。当然、その場合は定性的な言葉、文脈であった意見（反論）を数値化する必要が生じ、可視化されるため、これまで“空中戦”となっていた議論の論点や基準が明確になります。感度分析が活用され、より良い、そして迅速な意思決定につながることを期待しています。

## 執筆者・タスクフォースメンバー

日本製薬工業協会 医薬品評価委員会 データサイエンス部会

## 2016 年度タスクフォース 2

- |         |             |
|---------|-------------|
| ○ 兼山 達也 | 大日本住友製薬株式会社 |
| ○ 小森 哲志 | バイエル薬品株式会社  |
| 塩境 一仁   | 第一三共株式会社    |
| 中水流 嘉臣  | ファイザー株式会社   |

## 担当副部会長

小山 暢之 第一三共株式会社

○：タスクフォースリーダー兼推進委員

## レビューアー

## 第1部

## データサイエンス部会

- |        |             |
|--------|-------------|
| 山本 英晴  | 中外製薬株式会社    |
| 酒井 弘憲  | エーザイ株式会社    |
| 菅波 秀規  | 興和株式会社      |
| 土屋 悟   | 大日本住友製薬株式会社 |
| 奥山 ことば | MSD 株式会社    |

## 第2部

## データサイエンス部会

- |        |             |
|--------|-------------|
| 山本 英晴  | 中外製薬株式会社    |
| 酒井 弘憲  | エーザイ株式会社    |
| 松下 泰之  | 第一三共株式会社    |
| 土屋 悟   | 大日本住友製薬株式会社 |
| 奥山 ことば | MSD 株式会社    |

## ESATJ (Excel Tool)

## データサイエンス部会

- |       |           |
|-------|-----------|
| 淡路 直人 | 中外製薬株式会社  |
| 土綿 慎一 | ファイザー株式会社 |
| 松下 泰之 | 第一三共株式会社  |

## REFERENCE

- 1) Rosenbaum, Paul R. Design of Observational Studies. Springer 2009.
- 2) Rothman K.J., Greenland S., Lash T.L. Modern Epidemiology, Third Edition. Lippincott, Williams, & Wilkins. 2008.
- 3) Lin NX, Logan S, Henley WE. Bias and Sensitivity Analysis When Estimating Treatment Effects from the Cox Model with Omitted Covariates. *Biometrics*. 2013 Dec;69(4):850-60.
- 4) Lash T.L., Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. 2014 Dec;43(6):1969-85.
- 5) Uddin MJ, Groenwold RH, Ali MS, de Boer A, Roes KC, Chowdhury MA, Klungel OH. Methods to control for unmeasured confounding in pharmacoepidemiology: an overview. *Int J Clin Pharm*. 2016 Jun;38(3):714-23.
- 6) Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol*. 1996 Dec;25(6):1107-16.
- 7) Greenland S. Multiple-bias modelling for analysis of observational data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2005. 168:267-306.
- 8) Lash T.L., Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer. 2009.
- 9) Greenland S, Salvan A, Wegman DH, Hallock MF, Smith TJ. A case-control study of cancer mortality at a transformer-assembly facility. *Int Arch Occup Environ Health*. 1994;66(1):49-54.
- 10) Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making*. 1985 Summer;5(2):157-77.
- 11) Steenland K, Greenland S. *Am J Epidemiol*. 2004 Aug 15;160(4):384-92. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer.
- 12) Greenland S. Bayesian perspectives for epidemiological research: III. Bias analysis via missing-data methods. *Int J Epidemiol*. 2009 Dec;38(6):1662-73.
- 13) Gustafson P, McCandless LC, Levy AR, Richardson S. Simplified Bayesian sensitivity analysis for mismeasured and unobserved confounders. *Biometrics*. 2010 Dec;66(4):1129-37.
- 14) Corrao G, Nicotra F, Parodi A, Zambon A, Soranna D, Heiman F, Merlino L, Mancina G. External adjustment for unmeasured confounders improved drug-outcome association estimates based on health care utilization data. *J Clin Epidemiol*. 2012 Nov;65(11):1190-9.
- 15) Corrao G, Nicotra F, Parodi A, Zambon A, Heiman F, Merlino L, Fortino I, Cesana G, Mancina G. Cardiovascular protection by initial and subsequent combination of antihypertensive drugs in daily life practice. *Hypertension*. 2011 Oct;58(4):566-72.
- 16) Franchi M, Asciutto R, Nicotra F, Merlino L, La Vecchia C, Corrao G, Bosetti C. Metformin, other antidiabetic drugs, and endometrial cancer risk: a nested case-control study within Italian healthcare utilization databases. *Eur J Cancer Prev*. 2016 May 24.
- 17) Hunnicutt JN, Ulbricht CM, Chrysanthopoulou SA, Lapane KL. Probabilistic bias analysis in pharmacoepidemiology and comparative effectiveness research: a systematic review. *Pharmacoepidemiol Drug Saf*. 2016 Dec;25(12):1343-1353.
- 18) Greenland S. 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. *J Am Stat Assoc*, 2003, vol. 98 : 47-54.
- 19) Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*. 2006 May;15(5):291-303.
- 20) Lash T.L., Schmidt M, Jensen AØ, Engebjerg MC. Methods to apply probabilistic bias analysis to summary estimates of association. *Pharmacoepidemiol Drug Saf*. 2010 Jun;19(6):638-44.
- 21) Fox MP, Lash T.L., Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. *Int J Epidemiol*. 2005 Dec;34(6):1370-6.
- 22) Lash T.L., Fink AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology*. 2003 Jul;14(4):451-8. (SAS code)



- 23) Bollaerts K, Shinde V, Dos Santos G, Ferreira G, Bauchau V, Cohet C, Verstraeten T. Application of Probabilistic Multiple-Bias Analyses to a Cohort- and a Case-Control Study on the Association between Pandemrix™ and Narcolepsy. PLoS One. 2016 Feb 22;11(2):e0149289.
- 24) Gerdtts C, Ahern J. Multiple-bias analysis as a technique to address systematic error in measures of abortion-related mortality. Popul Health Metr. 2016 Mar 22;14:9.
- 25) Takeuchi A, Matsuyama Y, Ohashi Y, and Ueshima H. Monte Carlo sensitivity analysis for adjusting multiple-bias in the longitudinal cardiovascular study. Japanese Journal of Biometrics 2010; 31: 63-76.
- 26) Momoli F, Siemiatycki J, McBride ML, Parent MÉ, Richardson L, Bedard D, Platt R, Vrijheid M, Cardis E, Krewski D. Probabilistic Multiple-Bias Modeling Applied to the Canadian Data From the Interphone Study of Mobile Phone Use and Risk of Glioma, Meningioma, Acoustic Neuroma, and Parotid Gland Tumors. Am J Epidemiol. 2017 Oct 1;186(7):885-893.
- 27) Schneeweiss S, Avorn J., A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005 Apr;58(4):323-37.
- 28) Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005 Apr;58(4):323-37.
- 29) Lash T.L., Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer. 2009.
- 30) McCandless LC, Gustafson P, Levy AR., A sensitivity analysis using information about measured confounders yielded improved uncertainty assessments for unmeasured confounding. J Clin Epidemiol. 2008 Mar;61(3):247-55.
- 31) Höfler M, Lieb R, Wittchen HU. Estimating causal effects from observational data with a model for multiple bias. Int J Methods Psychiatr Res. 2007;16(2):77-87.
- 32) Schisterman EF, Cole SR, Platt RW., Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009 Jul;20(4):488-95.
- 33) Sterne JA , Hernán MA , Reeves BC , et al . ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
- 34) Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoepidemiology and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report--Part II. Value Health. 2009 Nov-Dec;12(8):1053-61.
- 35) ENCePP Guide on Methodological Standards in Pharmacoepidemiology  
[http://www.encepp.eu/standards\\_and\\_guidances/methodologicalGuide4\\_2\\_2.shtml](http://www.encepp.eu/standards_and_guidances/methodologicalGuide4_2_2.shtml)
- 36) Delgado-Rodríguez M, Llorca J. , Bias. J Epidemiol Community Health. 2004 Aug;58(8):635-41.
- 37) 景山 茂, 久保田潔, 薬剤疫学の基礎と実践 第2版 医薬ジャーナル社 2016
- 38) Chen Q, Galfalvy H, Duan N. Effects of disease misclassification on exposure-disease association. Am J Public Health. 2013 May;103(5):e67-73.
- 39) Ekström N, Cederholm J, Zethelius B, Eliasson B, Fhärm E, Rolandsson O, Miftaraj M, Svensson AM, Gudbjörnsdóttir S., Aspirin treatment and risk of first incident cardiovascular diseases in patients with type 2 diabetes: an observational study from the Swedish National Diabetes Register. BMJ Open. 2013 Apr 20;3(4). pii: e002688.
- 40) Ahrens K, Lash T.L., Louik C, Mitchell AA, Werler MM. Correcting for exposure misclassification using survival analysis with a time-varying exposure. Ann Epidemiol. 2012 Nov;22(11):799-806.
- 41) 医療情報のデータベース等を用いた医薬品の安全性評価における薬剤疫学研究の実施に関するガイドライン 独立行政法人 医薬品医療機器総合機構 初版 <https://www.pmda.go.jp/files/000147250.pdf>
- 42) 「日本における傷病名を中心とするレセプト情報から得られる指標のバリデーションに関するタスクフォース」報告書 日本薬剤疫学会 <http://www.jspe.jp/committee/pdf/validationtrr120180528.pdf>
- 43) Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data.  
<http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm243537.pdf>

- 44) Pekkanen J, Sunyer J, Chinn S., Nondifferential disease misclassification may bias incidence risk ratios away from the null. *J Clin Epidemiol.* 2006 Mar;59(3):281-9.
- 45) Nevriana A, Möller J, Laflamme L, Monárrez-Espino J. New, Occasional, and Frequent Use of Zolpidem or Zopiclone (Alone and in Combination) and the Risk of Injurious Road Traffic Crashes in Older Adult Drivers: A Population-Based Case-Control and Case-Crossover Study. *CNS Drugs.* 2017 Aug;31(8):711-722.
- 46) Johnson CY, Flanders WD, Strickland MJ, Honein MA, Howards PP. Potential sensitivity of bias analysis results to incorrect assumptions of nondifferential or differential binary exposure misclassification. *Epidemiology.* 2014 Nov;25(6):902-9.
- 47) Stürmer T, Glynn RJ, Rothman KJ, Avorn J, Schneeweiss S. Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. *Med Care* 2007; 45(10S): S158-65.
- 48) Duffy SW, Maximovitch DM, Day NE, External validation, repeat determination, and precision of risk estimation in misclassified exposure data in epidemiology. *J Epidemiol Community Health.* 1992 Dec;46(6):620-4.
- 49) Yue Zhao, Amy H. Herring, Haibo Zhou, Mirza W. Ali, Gary G. Koch, A MULTIPLE IMPUTATION METHOD FOR SENSITIVITY ANALYSES OF TIME-TO-EVENT DATA WITH POSSIBLY INFORMATIVE CENSORING. *J Biopharm Stat.* Author manuscript; available in PMC 2014 May 5.
- 50) Huybrechts KF, Hernández-Díaz S, Avorn J. Antidepressant use in pregnancy and the risk of cardiac defects. *N Engl J Med.* 2014 Sep 18;371(12):1168-9.
- 51) Palmsten K, Huybrechts KF, Kowal MK, Mogun H, Hernández-Díaz S. Validity of maternal and infant outcomes within nationwide Medicaid data. *Pharmacoepidemiol Drug Saf.* 2014 Jun;23(6):646-55.
- 52) Widdifield J, Ivers NM, Young J, Green D, Jaakkimainen L, Butt DA, et al. Development and validation of an administrative data algorithm to estimate the disease burden and epidemiology of multiple sclerosis in Ontario, Canada. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2015;21(8):1045-54.
- 53) Kumamaru H, Judd SE, Curtis JR, Ramachandran R, Hardy NC, Rhodes JD, Safford MM, Kissela BM, Howard G, Jalbert JJ, Brott TG, Setoguchi S. Validity of claims-based stroke algorithms in contemporary Medicare data: REGARDS study linked with medicare claims. *Circ Cardiovasc Qual Outcomes.* 2014 Jul;7(4):611-9.
- 54) Howard VJ, Cushman M, Pulley L, Gomez CR, Go RC, Prineas RJ, Graham A, Moy CS, Howard G. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology.* 2005;25(3):135-43.
- 55) Bhatt H, Safford M, Glasser S. Coronary heart disease risk factors and outcomes in the twenty-first century: findings from the REasons for Geographic and Racial Differences in Stroke (REGARDS) Study. *Curr Hypertens Rep.* 2015 Apr;17(4):541.
- 56) Xie F, Colantonio LD, Curtis JR, Safford MM, Levitan EB, Howard G, Muntner P. Linkage of a Population-Based Cohort With Primary Data Collection to Medicare Claims: The Reasons for Geographic and Racial Differences in Stroke Study. *Am J Epidemiol.* 2016 Oct 1;184(7):532-544
- 57) Strom B.L., Kimmel S.E., Hennessy S. *Pharmacoepidemiology* 5th Edition. Wiley, 2012.
- 58) Holcroft CA, Spiegelman D., Design of validation studies for estimating the odds ratio of exposure-disease relationships when exposure is misclassified. *Biometrics.* 1999 Dec;55(4):1193-201.
- 59) Lash T.L., Ahern TP. Bias analysis to guide new data collection. *Int J Biostat.* 2012 Jan 6;8(2).
- 60) Chan AW, Fung K, Tran JM, Kitchen J, Austin PC, Weinstock MA, Rochon PA. Application of Recursive Partitioning to Derive and Validate a Claims-Based Algorithm for Identifying Keratinocyte Carcinoma (Nonmelanoma Skin Cancer). *JAMA Dermatol.* 2016 Oct 1;152(10):1122-1127.
- 61) Thomas D, Stram D, Dwyer J., Exposure measurement error: influence on exposure-disease. Relationships and methods of correction. *Annu Rev Public Health.* 1993;14:69-93.
- 62) Brenner H, Savitz DA., The effects of sensitivity and specificity of case selection on validity, sample size, precision, and power in hospital-based case-control studies. *Am J Epidemiol.* 1990 Jul;132(1):181-92.
- 63) Walker AM. Anamorphic analysis. Sampling and estimation for covariate effects when both exposure and disease are known. *Biometrics.* 1982;38:1025-1032.
- 64) Flanders WD, Greenland S. Analytic methods for two-stage case-control studies and other stratified designs. *Stat Med.* 1991;10:739-747.

- 65) 山口 拓洋他. 医療情報データベースを用いた静脈血栓塞栓症発症、出血性イベントのバリデーション研究. *Iyakuhin Johogaku* Vol. 17 (2015) No. 2 August p. 87-93, UMIN000012667 医療情報データベースを用いた心原性脳塞栓症・静脈血栓塞栓症発症の実態に関する疫学的研究.
- 66) Shiro T, Hiroshi H, Akiko I, Teruhiko M, Takanori Y, Takayuki H. Validation Study of Claims-based Definitions of Suspected Atypical Femoral Fractures Using Clinical Information. *Jpn J Pharmacoepidemiol* 2016 ; 21 (1) : 13-19.
- 67) Medicare Current Beneficiary Survey <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/index.html>
- 68) Lash T.L., Silliman RA. A sensitivity analysis to separate bias due to confounding from bias due to predicting misclassification by a variable that does both. *Epidemiology*. 2000 Sep;11(5):544-9.
- 69) Miyashita N, Kawai Y, Tanaka T, Akaike H, Teranishi H, Wakabayashi T, Nakano T, Ouchi K, Okimoto N. Detection failure rate of chest radiography for the identification of nursing and healthcare-associated pneumonia. *J Infect Chemother*. 2015 Jul;21(7):492-6.
- 70) Carnahan RM, Moores KG. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned. *Pharmacoepidemiol Drug Saf*. 2012 Jan;21 Suppl 1:82-9.
- 71) Carnahan RM. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative data: summary of findings and suggestions for future research. *Pharmacoepidemiol Drug Saf*. 2012 Jan;21 Suppl 1:90-9.
- 72) Nordstrom BL, Norman HS, Dube TJ, Wilcox MA, Walker AM. Identification of abacavir hypersensitivity reaction in health care claims data. *Pharmacoepidemiol Drug Saf*. 2007 Mar;16(3):289-96.
- 73) Tversky A, Kahneman D, Judgment under Uncertainty: Heuristics and Biases. *Science*. 1974 Sep 27;185(4157):1124-31.
- 74) Dawson NV, Physician judgment in clinical settings: methodological influences and cognitive performance. *Clin Chem*. 1993 Jul;39(7):1468-78.
- 75) Phillips CV, LaPole LM. Quantifying errors without random sampling. *BMC Med Res Methodol* 2003;3:9.
- 76) Ding P, VanderWeele TJ., Sensitivity Analysis Without Assumptions. *Epidemiology*. 2016 May;27(3):368-77.
- 77) Gilbert R, Martin RM, Donovan J, Lane JA, Hamdy F, Neal DE, Metcalfe C., Misclassification of outcome in case-control studies: Methods for sensitivity analysis. *Stat Methods Med Res*. 2016 Oct;25(5):2377-2393.
- 78) Barron TI, Cahir C, Sharp L, Bennett K., A nested case-control study of adjuvant hormonal therapy persistence and compliance, and early breast cancer recurrence in women with stage I-III breast cancer. *Br J Cancer*. 2013 Sep 17;109(6):1513-21.
- 79) M.M. Joffe, Confounding by indication: the case of calcium channel blockers. *Pharmacoepidemiol Drug Saf*, 9 (2000), pp. 37-41.
- 80) Fleiss JL. Statistical Methods for Rates and Proportions. Third Edition Wiley-Interscience. 2003.
- 81) Arfè A, Nicotra F, Ghirardi A, Simonetti M, Lapi F, Sturkenboom M, Corrao G. A probabilistic bias analysis for misclassified categorical exposures, with application to oral anti-hyperglycaemic drugs. *Pharmacoepidemiol Drug Saf*. 2016 Dec;25(12):1443-1450.
- 82) Baer VL, Lambert DK, Henry E, Snow GL, Sola-Visner MC, Christensen RD. Do platelet transfusions in the NICU adversely affect survival? Analysis of 1600 thrombocytopenic neonates in a multihospital healthcare system. *J Perinatol*. 2007 Dec;27(12):790-6.
- 83) Youngquist ST, Kaji AH, Niemann JT, Beta-blocker use and the changing epidemiology of out-of-hospital cardiac arrest rhythms. *Resuscitation*. 2008 Mar;76(3):376-80.
- 84) Mitro SD, Zota AR, Vitamin D and uterine leiomyoma among a sample of US women: Findings from NHANES, 2001-2006. *Reprod Toxicol*. 2015 Nov;57:81-6.
- 85) Jain A, Marshall J, Buikema A, Bancroft T, Kelly JP, Newschaffer CJ. Autism occurrence by MMR vaccine status among US children with older siblings with and without autism. *JAMA*. 2015 Apr 21;313(15):1534-40.

- 86) Stonebraker JS, Farrugia A, Gathmann B; ESID Registry Working Party, Orange JS. Modeling primary immunodeficiency disease epidemiology and its treatment to estimate latent therapeutic demand for immunoglobulin. J Clin Immunol. 2014 Feb;34(2):233-44.
- 87) Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3) [https://www.ema.europa.eu/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3\\_en.pdf](https://www.ema.europa.eu/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf)
- 88) Chia-Hsien Chang. Epidemiology of lactic acidosis in type 2 diabetes patients with metformin in Japan. Pharmacoepidemiol Drug Saf. 2016 Oct;25(10):1196-1203.
- 89) Kim SY. et.al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. J Clin Epidemiol. 2013 Apr;66(4):408-14.
- 90) The ROBINS-I tool (Risk Of Bias In Non-randomized Studies - of Interventions)  
<https://sites.google.com/site/riskofbiastool/>
- 91) Jurek AM, Greenland S, Maldonado G. Brief Report: How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? Int J Epidemiol (2008) 37 (2): 382-385

## Appendix1. 感度分析の必要性

感度分析の必要性の評価はそのまま感度分析の目的を定めることである。ここでは、まず扱いが比較的容易と思われる科学的視点として研究の質による評価について説明し、続けて定型的に扱うことが難しいベネフィット・リスク評価、意思決定（論）の視点として研究の意義や公衆衛生など社会的意思決定における重要性による評価を Lash T.L. (2014) <sup>4)</sup> の分類に習い、感度分析が必要不可欠な場合、推奨される場合、不可欠ではない場合に区分して説明する。

### 1) 研究の質による評価：計画～解析における精度とバイアスから

観察研究は対象者の選択（参加）、追跡/脱落、データの測定（欠測）、測定誤差及び未測定交絡因子等の偏りから種々のバイアスの影響（系統的誤差）を受けるが、アウトカムの有無を類推している場合など、中には結果に重大な影響を及ぼす場合がある。観察研究におけるバイアスリスク評価ツールである RoBANS (Risk of Bias Assessment Tool for Nonrandomized Studies) <sup>89)</sup> は Table 1-1 のように参加者（研究対象者）の選択、交絡変数、曝露の測定、アウトカム評価の盲検化、アウトカムデータの不完全さ、選択的なアウトカム報告の6つのドメインについてバイアスリスクを低い、高い、不明の3段階に分ける基準を示している。例えば、曝露の測定では、自己申告による場合、明らかな想起バイアスがある場合を「高い」としている。また、Cochrane はより詳細な The ROBINS-I tool (Risk Of Bias In Non-randomized Studies - of Interventions) <sup>90)</sup> を公開している。このようなツールを用いることは研究の計画段階で感度分析の必要性を検討できるのみならず、研究計画の改善にも役立つ。また、データ収集後の解析段階であれば、データ欠損・打ち切りの一様性（比較の群間、主要な共変量に対する）、既存の知見との一貫性（結果、アウトカム及び主要な共変量の頻度）から判断することとなる。

この様に研究の質を評価した結果、研究対象者や参加者が偏っている、対象者の脱落がランダムでない、測定誤差・誤り、測定されていない重要な交絡因子の存在など影響がありそうではあるが、結果への影響の質あるいは程度が想定できない、収集されたデータによる調整解析などでは評価できない場合は感度分析を行うことで影響を見積もることが出来る。特に結果がリスク比、オッズ比やハザード比で表される研究においては、アウトカムの特異度<sup>注79)</sup>は僅かな偏りであっても結果に大きな変化を生じさせ得るので注意が必要である。

なお、研究におけるデータの扱い、解析方法（モデル、リンク関数等）に複数の選択肢がある場合、例えば臨床検査値（連続変数）から異常やイベントの有無を判定するカットオフ値、調整解析に用いる変数等に明確で臨床的な根拠がなく、選択の余地がある場合はこ

注<sup>79)</sup>：特異度については第1部 3.2.1, 3.4.4 参照



これらの条件の影響を評価するため異なる条件を用いた解析も感度分析と呼ばれることがあるが、これらは測定された値の扱いを変える、測定されたデータに基づく手法であるので本書では感度分析に含めない。

Table 1-1 RoBANS : The risk-of-bias assessment tool for nonrandomized studies<sup>89)</sup>

Domain	Details	Risk of bias*
Selection of participants	Selection bias caused by the inadequate selection of participants	Low, High, Unclear
Confounding variables	Selection bias caused by the inadequate confirmation and consideration of confounding variable	Low, High, Unclear
Measurement of exposure	Performance bias caused by the inadequate measurement of exposure	Low, High, Unclear
Blinding of outcome assessments	Detection bias caused by the inadequate blinding of outcome assessments	Low, High, Unclear
Incomplete outcome data	Attrition bias caused by the inadequate handling of incomplete outcome data	Low, High, Unclear
Selective outcome reporting	Reporting bias caused by the selective reporting of outcomes	Low, High, Unclear

\*：評価の基準は原論文の Appendix を参照

## 2) 研究の重要性による評価：結果の社会的影響から

この視点で考える際に研究結果を利用した判断（意思決定）が誤っていた場合に生じるリスク（害）として考える事が有効である。これにより、研究の結果が好ましい作用（ベネフィット）を評価する場合は過大評価でないことが重要であり、統計解析の視点では第1種の過誤（誤って差があるとする）を重視することに該当する。反対に研究の結果が好ましくない作用（リスク）を評価する場合は過小評価でないことが重要であり、第2種の過誤（差を見逃す）を重視することになる。

### 1. 感度分析が必要不可欠である場合

#### A. 報告により公衆衛生上慎重さが要求される行動の変容や意思決定が行われる場合

結果を検証する追加の研究を実施する猶予がなく、例えば代替え治療のない医薬品の広範なアクセス制限あるいは販売中止を伴う場合、代替え治療の総合的なリスクが高い場合や意思決定による措置に社会的に多くのコストを伴う場合などであり、第1種の過誤、第2種の過誤の両方の重視が求められる。このような場合、意思決定の主要な根拠となる報告は内的妥当性並びに外的妥当性（公衆衛生における影響）の両方に信頼性・頑健性が求められる。このような直接的な意思決定や政策を伴う因果推論に対する感度分析は比較的シンプルな方法であっても用いる仮定に妥当性が求められ、複数の仮定、分析が必要であり、多くのパラメータを扱う必要が生じる。さらにバイ

アスの原因を個別に評価するにとどまらず、同時に扱う多変量の確率的な感度分析も求められる場合がある。

B. 示されたリスクの小ささにより公衆衛生上重要な対策を保留する場合

汎用される医薬品の使用による死亡リスク増加等、リスクが大きく確かであれば使用制限など重要な対策が講じられるが、示されたリスク情報（増加リスク、オッズ比等）が小さいため措置が保留される場合があるが、バイアスにより報告されたリスクが真の値より過少となっている場合がある。また、曝露やアウトカムの誤分類は比較する群間でわずかな偏りでもリスク比などの結果への多様な影響を生じる<sup>91)</sup>ため実際に感度分析を行い確認することを推奨する。[第1部 3.2.2, 3.4.4, 3.4.6 参照]

C. 具体的な意思決定のために複数の報告の統合解析が行われる場合

重要な政策や行動変容の判断を目的として複数の報告の統合が行われる場合、個々の研究の感度分析以外に統合研究としての感度分析、特にレポーティングバイアスの感度分析が必要となる。

D. 報告の信頼性の不足などにより追加の研究を行う場合

報告の信頼性・精度や情報が不十分なため措置の必要性などが判断できず、新たに追加の研究を行う場合は先行研究の十分な感度分析を行い、その結果を新たな研究の計画に反映させることは有用であり当然である。<sup>59)</sup>

## 2. 感度分析が推奨される場合

A. それ程大きくはない効果（リスク差等）に基づき対策を求める場合

過去に報告がないまたは報告があっても結論付けられていない因果関係を報告し、公衆衛生に関わる対策の意思決定を求める場合である。効果の種類にもよるが、効果がそれ程大きくなくバイアスによる不確実性（効果はバイアスによるもの）が懸念される場合、通常意思決定は慎重に行われる。効果が市販医薬品の安全性リスク等の場合、真の効果が報告に対して小さい場合には過剰な対策を講じることとなるが、反対に真の効果が大きい場合に対策を講じないことは見逃しの過誤を生じさせることとなる。このため、示された効果がリスクの場合には対策を講じる意思決定が行われやすく、更に第1部 3.2.2 で説明しているが、比較する群間で曝露、アウトカムの誤分類の偏りがわずかな場合や偏りが無い場合でもオッズ比等などに影響が生じる場合があるため感度分析の実施が推奨される<sup>91)</sup>。

B. 過去の報告と異なる相当に大きな効果の結果に基づき対策を求める場合

これまでの報告とは異なる大きな効果（リスク等）との因果関係が認められ、公衆衛生に関わる対策を求める場合である。この場合、過去の報告とは異なる結果となった原因を含め感度分析が求められる。

### 3. 感度分析が不可欠ではない場合

下記のような場合は結果について限界、制約などを説明する必要はあるが、感度分析までは求められない。

#### A. 純粋な記述統計のみの場合

収集されたデータを要約するのみで、そこから何らかの推測や結論を導かない場合である。ただし、この結果を見た者がここから何かを推論する場合のため、限界や前提条件については丁寧に述べられるべきである。

#### B. 因果関係評価を行わない場合

記述統計に近いが、何らかの仮説（この場合 1 つとは限らないが少数の）に基づく研究から、追加の研究の必要性あるいは新たな仮説の可能性を提示するのみで、因果関係の結論や何らかの対策の提案を伴わない報告であり、研究の報告で最も多くみられるものである。

#### C. 研究のデザインが不適切または問題が多く結果の妥当性が低い場合

曝露及びアウトカムの誤分類並びに未測定交絡要因の何れも見逃さないバイアスがある場合や、オッズ比等の結果はある程度大きな値以内で精度が不十分な場合である。前者の感度分析の結果は不確実性が大きく有益な情報が得られることは少ない<sup>29)</sup>。後者は例えばオッズ比は 2 であるが、95%信頼区間は 0.37～10.7 のような場合である。このような研究は標本サイズが不足しておりデザインの段階でサンプルサイズの設定に誤り（アウトカム数の見積もりなど）があったか、抽出条件あるいは用いたアウトカムの定義がサンプルサイズの設定に用いた物と異なる場合などである。何れにしてもデザイン～データ収集を改善すべき研究である。

#### D. 過去に同様の仮説の研究が複数行われ、結果に一貫性があり大きな値である場合

同様の仮説の研究が異なる研究者、データソース、研究デザインによって複数行われ、層別解析を含めて結果に一貫性がある場合であり、例えば示されたオッズ比の値などが意思決定において議論の対象とならないほど大きく、既に当該仮説に対して一定の結論が得られている場合である。

#### E. 研究の副次的結果に対して

研究の主要な仮説について以外に副次的な結果について報告書に記述することは少なくない。このような記述は主要な結果の補完あるいは新たな仮説生成の種となるものであり、感度分析までは求められない。

#### F. 探索的な解析を行う場合

事前に明確な仮説を設定せず、探索的な解析を行った研究の報告において感度分析は必要ではない。ただし、当然であるが、探索的な研究の結果であることを明示し、結論的なことは述べられていない場合である。

以上、感度分析が必要不可欠な場合から不可欠ではない場合まで説明した。初めに説明した通り、感度分析は常に必要ではないが、純粹な記述統計のみの場合や因果関係評価を行わない場合など明らかに不要な場合以外の要否は明確でなく、要否の判断には感度分析の入り口であるバイアスの影響の見積もりが必要な場合も少なくない。そのため、感度分析に用いるバイアスの情報の入手や感度分析の計算プログラムの作成などにリソースがかからない範囲であれば、積極的に実施するほうが合理的であると考え。この点については「第1章7）教科書的ではないが現実的で簡便な手順による感度分析」で説明している。

また、感度分析の要否以上に、行う感度分析に求める精度や、説明力（理解しやすさ、透明性）を考えるうえで、ここで説明してきた状況、場面は関係する。概ね、感度分析が必要不可欠な場合は精度と透明性がより求められ、リソースを要しない簡便な感度分析では不十分な場合が多いであろう。そのような場合は条件を変化させて感度分析を行う PSA（MCSA）が有用である。説明力は結果の見せ方も関係するが、感度分析の対象とするバイアスの特定とその理由が重要であるので、5.3 以下で説明している。また、精度や説明力を左右するバイアスの情報の入手は「5.6.1 感度分析パラメータに関する情報源と収集」で扱った。

## Appendix2. 分布に基づく乱数の生成

以下、特に断らない場合、「Table 4-1 職業によるレジン樹脂曝露（E）と肺癌発生の case-control 研究」を元に説明している。見やすさのため、Table 2-1 に Table 4-1 の値を入れて再掲する。

Table 2-1 数値例 再掲)

	Crude data			C=1			C=0		
	X=1	X=0	Total	X=1	X=0	Total	X=1	X=0	Total
cases	A <sub>1+</sub> =45	A <sub>0+</sub> =94	M <sub>1+</sub> =139	A <sub>11</sub>	A <sub>01</sub>	M <sub>11</sub>	A <sub>1+</sub> - A <sub>11</sub>	A <sub>0+</sub> - A <sub>01</sub>	M <sub>1+</sub> - M <sub>11</sub>
controls	B <sub>1+</sub> =257	B <sub>0+</sub> =945	M <sub>0+</sub> =1202	B <sub>11</sub>	B <sub>01</sub>	M <sub>01</sub>	B <sub>1+</sub> - B <sub>11</sub>	B <sub>0+</sub> - B <sub>01</sub>	M <sub>0+</sub> - M <sub>01</sub>
X=1: レジン樹脂曝露あり, X=0: レジン樹脂曝露なし, C=1: 喫煙あり, C=0: 喫煙なし									

MCSA (PSA) など、パラメータを変化させて simulation を行う場合にモデル（式）に与えるパラメータを任意の分布からのサンプリングが行われる。4.1 Probabilistic sensitivity analysis で説明しているように、感度分析では数学的に一般的な一様分布、正規分布、β分布以外に三角分布 (Triangular Distribution)、台形分布 (Trapezoidal Distribution) などが用いられるが、この2つの分布はソフトウェアに関数化されていない場合、プログラムを作成する必要が生じる場合がある。一様分布、正規分布以外について紹介する。なお、ESATJ にはこれら以外も利用できる。 Appendix3.1) 11(オ)

### 1) 台形分布 (Trapezoidal), 三角分布 (Triangular)

台形分布は、任意の最小値～最大値の間の値 (x とする) の密度が台形を形成するもので、Figure 2-1 のように図示され、最小値 (min) ～最大値 (max) の間で最頻値 (mod1) まで単調に増加し、mod1～mod2 は一様で、mod2～max まで単調に減少する。ここで、 $\text{mod1} \leq x \leq \text{mod2}$  の x は最頻値となる。

このような分布に基づく乱数を得るには、多くのソフトが生成できる一様乱数を利用する。

p を 0 から 1 の一様分布に属する乱数とする。  $0 \leq p \leq 1$

例えば、Microsoft Excel ではワークシート関数 RAND() により p を生成できる。

任意の最小値 (min) と最大値 (max) の一様分布 u は、p を用いて

$$u = \text{min} + (\text{max} - \text{min})p \cdots \cdots \textcircled{1}$$

により求めることができる。

台形分布を Figure 2-1 に示す。任意の min, max 及び最頻値の区間 (mod1～mod2) を持つ台形分布の確率密度は黄色と灰色の部分で表せられる。これに対して、min と mod1 の中点を最



小値,  $\max$  と  $\text{mod2}$  の中点を最大値とする一様分布（赤線）を考えるとこの赤線内と台形の面積は同じであり, その一様分布乱数  $u$  は①から,

$$u = \frac{(\min + \text{mod1})}{2} + \frac{(\max + \text{mod2} - \min - \text{mod1})p}{2}$$

により得られる。(高さ=1 としている場合, 面積=x 軸上の点となる)

この台形分布となる乱数を  $s$  とすると  $\text{mod1} \leq u \leq \text{mod2}$  においては,

$$s = u$$

である.  $u < \text{mod1}$  の場合は, 台形左端の  $\min \sim s$  を底辺とする三角形の面積が  $u \times 1$  (高さ=1 としているので) となる点  $s$  を求める.

$$2u - (\min + \text{mod1}) = (s - \min) \frac{s - \min}{\text{mod1} - \min}$$

$$[2u - (\min + \text{mod1})](\text{mod1} - \min) = (s - \min)^2$$

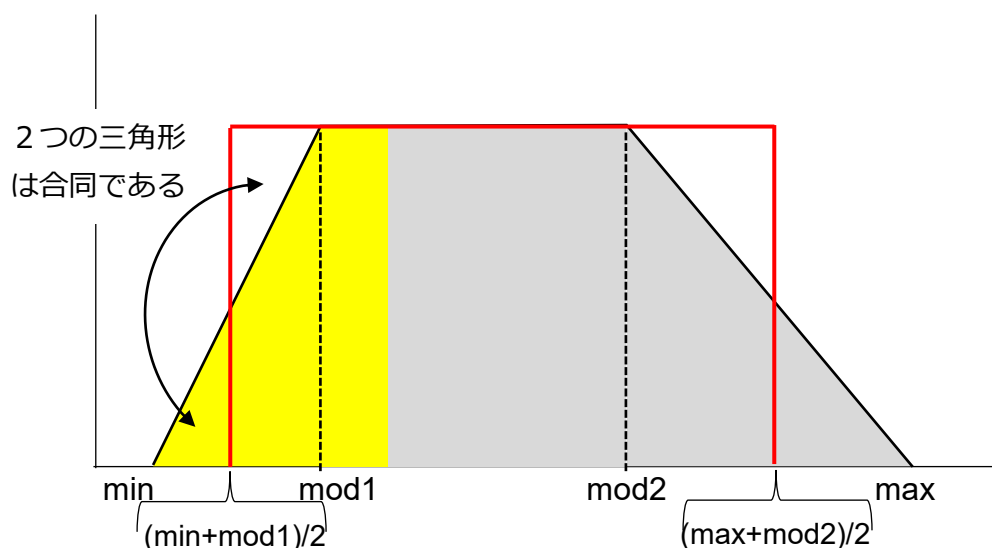
$$s = \min + \sqrt{(2u - \min - \text{mod1})(\text{mod1} - \min)}$$

同様に  $u > \text{mod2}$  の場合は一様分布の  $u \sim (\max + \text{mod2}) / 2$  までの面積と等しい台形の右端の  $s \sim \max$  を底辺とする面積から

$$s = \max - \sqrt{2 \cdot (\max - \text{mod2})(u - \text{mod2})}$$

により求められる.

Figure 2-1 台形分布



三角形分布は台形分布が  $\text{mod1} = \text{mod2}$  の特殊な場合として定義できるので, 同じ式を用いて分布関数を作成できる.

## 2) ベータ分布

ベータ分布は様々な分布の表現が可能<sup>注80</sup>でかつ滑らかであるが、Beta(0.5,0.5)の例のように2峰性になる場合があり $\alpha$ 及び $\beta$ の設定には注意を要する。

$\alpha$ 及び $\beta$ はBeta分布を用いる場合はヒストグラムなどグラフによって分布を確認することでこのような問題は回避できる。また、 $\beta$ 分布の平均を $\mu$ 、分散を $var$ 、標準偏差を $\sigma$ 、最頻値を $mode$ 、中央値を $mid$ とすると、( $\alpha > 1$ ,  $\beta > 1$ の場合)

$$\mu = \frac{\alpha}{\alpha + \beta} \quad var = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad var = \sigma^2 \quad mode = \frac{\alpha - 1}{\alpha + \beta - 2}$$

$$mid \approx \frac{\alpha - 1/3}{\alpha + \beta - 2/3} \quad \alpha = \mu \left( \frac{\mu(1 - \mu)}{\sigma^2} - 1 \right) \quad \beta = (1 - \mu) \left( \frac{\mu(1 - \mu)}{\sigma^2} - 1 \right)$$

である。パラメータの分布形状を考えると、平均 ( $\mu$ ) と最頻値 ( $mode$ ) を仮定できれば  $mode = \dots$  と  $\mu = \dots$  の式より、

$$\alpha = \frac{2mode - 1}{\frac{mode}{\mu} - 1} = \frac{(2mode - 1)\mu}{mode - \mu} \quad \beta = \frac{\alpha}{\mu} - \alpha$$

平均 ( $\mu$ ) と分散 ( $var$ ) を仮定できれば

$$\alpha = \mu^2(1 - \mu) \frac{1}{var} - \mu$$

により、 $\alpha$ と $\beta$ を求めることができる。

Beta分布は  $0 < x < 1$  となる分布であるが、任意の最小値 ( $min$ )、最大値 ( $max$ ) を取る分布を  $x'$  とする場合は次により変換する。

$$x' = (max - min)x + min$$

この場合の平均値を $\mu'$ 、最頻値を  $mode'$  とすると、元のBeta分布の $\mu$ および  $mode$  は

$$\mu = \frac{\mu' - min}{max - min} \quad mode = \frac{mode' - min}{max - min}$$

である。

例えば、教科書<sup>2)</sup>で Greenland S. がレジジン曝露と肺癌の関連の事例で曝露の誤分類に用いているパラメータは

$$Se = 0.8 + 0.2 \cdot \text{expit}(g_1), \quad Sp = 0.8 + 0.2 \cdot \text{expit}(g_2),$$

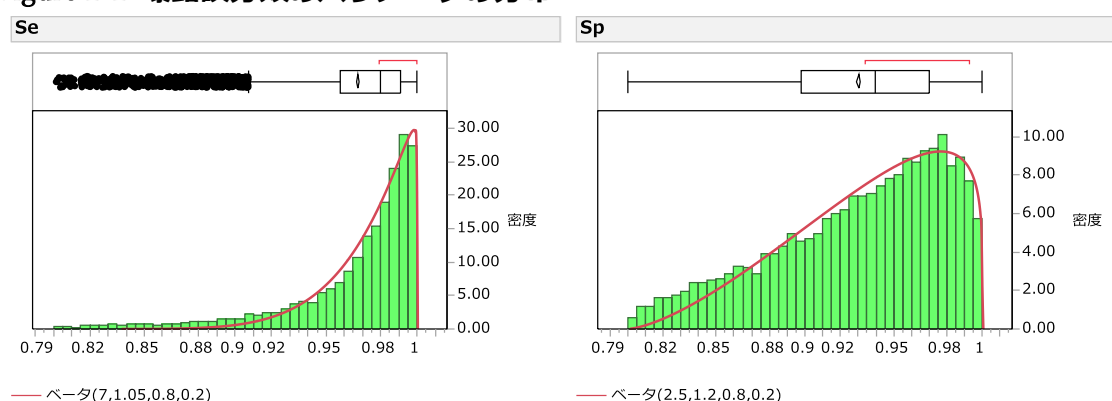
$$g_1 = \text{logit}(0.9) + 0.8 \cdot \text{logit}(u_1), \quad g_2 = \text{logit}(0.7) + 0.8 \cdot \text{logit}(u_2), \quad \text{注81}$$

注80：ベータ分布自体は  $0 \sim 1$  の範囲であるが、加減乗除によりそれ以外に対応できる。例えば  $1 \sim 5$  の範囲であれば、ベータ分布の値  $\times 4 + 1$  とすればよい。

注81：この関数は Appendix 2.4) 参照

である。読者が感度分析に用いる Tool (ソフト) ではこのような関数はメニューにないかもしれないが、Beta 分布で代用可能である。Figure 2-2 はこの式を用いて発生させた乱数のヒストグラムである。赤色の曲線は、Beta 分布の乱数  $b_1 \sim \text{Beta}(8, 1.2)$ ,  $b_2 \sim \text{Beta}(2.5, 1.2)$  を用いて、 $Se = 0.8 + (1 - 0.8) b_1$ ,  $Sp = 0.8 + (1 - 0.8) b_2$  による確率密度曲線である。このように実用上十分な近似を得られる。

Figure 2-2 曝露誤分類のパラメータの分布



### 3) 対数正規分布

選択バイアス、未測定交絡要因、共変量の誤分類では、パラメータにオッズ比あるいはリスク比を用いるが、これらは 0 より大きく、無限大の対数分布に属する。このため、パラメータの分布にも対数正規分布を仮定する場合がある。対数変換したパラメータを平均を  $\mu$  (つまり、パラメータの平均は  $\exp(\mu)$ )、標準偏差を SE、標準正規分布  $N(0, 1^2)$  の乱数を  $n_i$  とするとパラメータの分布は

$$\text{Exp}[\mu + n_i \cdot \text{SE}]$$

となる。分布の 95%信頼区間は

$$\text{下限 (2.5\%点)} = \text{Exp}[\mu - 1.96 \cdot \text{SE}],$$

$$\text{上限 (97.5\%点)} = \text{Exp}[\mu + 1.96 \cdot \text{SE}]$$

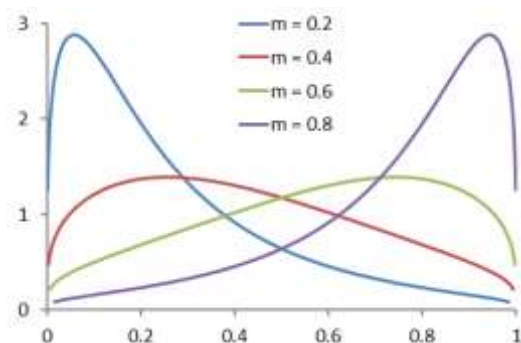
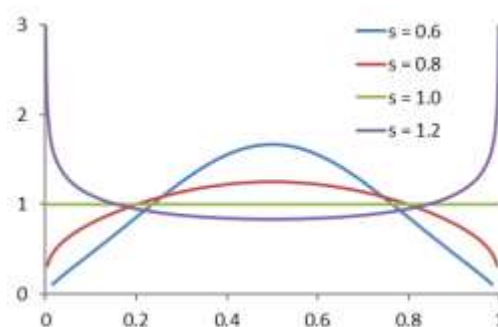
である。

### 4) ロジット・ロジスティック (logit-logistic) 分布

任意のパラメータ  $m$ ,  $s$  を与え、 $0 \sim 1$  の一様分布に従う乱数  $u_i$  を以下のように変換すると、ロジスティック分布に従う乱数  $w_i$  が得られる。

$$w_i = \text{logit}(m) + s \cdot \text{logit}(u_i), \text{ ただし } s > 0, \text{logit}(u_i) = \ln[u/(1 - u_i)]$$

この乱数をさらに  $p = \text{expit}(w_i) = \exp(w_i) / \{1 + \exp(w_i)\}$  と（逆）変換するとロジット・ロジスティック分布に従う乱数が得られる。  $w$  の分布の中央値が  $\text{logit}(m)$  となることから、  $p$  の分布の中央値は  $m$  となる。例えば  $s = 0.8$  と固定して  $m$  の値を 0.2, 0.4, 0.6, 0.8 等と変化させると分布の中心が移動する。

Figure 2-4 パラメータ  $m$  を変化させた例Figure 2-3 パラメータ  $s$  を変化させた例

一方、  $s$  はスケールパラメータである。1 より大きくすると  $p$  の密度が 0 や 1 といった両極に集中し、逆に 1 より小さくすると中央値付近に密度が集中する。例えば  $m = 0.5$  と固定して  $s$  の値を 0.6, 0.8, 1.0, 1.2 等と変化させると分布の密度が中心から端へ移動する。

実際に用いる場合は、得られた  $p$  の分布と想定している内容と一致しているかグラフ等で確認するとよい。

本文中では 5.6.4, 5.7.1 等で用いているが、感度や特異度の分布を設定する際にロジット・ロジスティック分布を用いることがある。例えば、感度  $Se$  について、0.8 から 1 という範囲の値を取り、中央値が 0.98 となる分布を想定したいとする。このような場合、例えば以下のように乱数を発生させることができる。

- ・ 一様乱数  $u$  を発生させ、  $h_i = \text{logit}(0.9) + 0.8 \cdot \text{logit}(u_i)$  と変換する。
- ・ 更に  $Se_i = 0.8 + 0.2 \cdot \text{expit}(h_i) = 0.8 + 0.2 \cdot \exp(h_i) / \{1 + \exp(h_i)\}$  と変換する

最小値は 0.8 であるので固定値として加算し、最大が 1 であるので、  $1 - 0.8 = 0.2$  を  $\text{expit}(h_i)$  に掛けている。この例では、一様乱数を変換する際、  $h$  の中央値が  $m = \text{logit}(0.9)$  と設定されている。このため、  $h_i$  を逆変換した  $\text{expit}(h_i)$  の中央値は 0.9 となる。よって、感度  $Se$  の中央値は  $0.8 + 0.2 \times 0.9 = 0.98$  となる。また、一様乱数を変換する際  $s = 0.8$  としているが、これは変換後の値の分布を見ながら経験的に定めたものである。

## 5) 相関のある分布

相関のある乱数の作り方について教科書などに詳しい説明が少なく、得られる分布に若干の“癖”があるので、詳しく説明する。

### ✓ 相関係数を用いる方法

4.1.3.パラメータの分布 で扱った分布をはじめ、PSA について書かれた教科書、論文の多くで 0～1 の一様分布をもとにして任意の形状の分布を発生させている。相関のある一様分布を用いることによって、相関を持つ任意の形状の分布も得られる。

任意の相関係数  $\text{corr}$  の一様分布を得る式は Fox MP. (2005)<sup>21)</sup>より、相関のある乱数を  $P_{1i}$ ,  $P_{2i}$ , 元となる一様分布の乱数を  $u_{0i} \sim \text{Uniform}(0 \sim 1)$ , 誤差  $u_{1i} \sim \text{Uniform}(0 \sim 1)$ ,  $u_{2i} \sim \text{Uniform}(0 \sim 1)$  とすると,  $P_{1i}$  および  $P_{2i}$  のうち  $\text{corr}$  の割合が同じ  $u_{0i}$  に,  $1 - \text{corr}$  の割合が各々  $u_{1i}$ ,  $u_{2i}$  とすることで得られる。

$$P_i = \sqrt{\text{corr}} \cdot u_{0i} + \sqrt{1 - \text{corr}} \cdot u_{1i}$$

結果を Uniform (0～1) とするため  $u_{0i}$ ,  $u_{1i}$ ,  $u_{2i}$  を一旦 logit 変換  $\text{logit}(p) = \ln[p/(1-p)]$  を行い各々  $g_u$ ,  $g_{e1}$ ,  $g_{e2}$  としたのち元に戻す expit 変換  $\text{expit}(p) = \exp(p)/[1 + \exp(p)]$  すると相関する分布となる  $P_{1i}$ ,  $P_{2i}$  は

$$P_{1i} = \frac{\text{Exp}(\sqrt{\text{corr}} \cdot g_u + \sqrt{1 - \text{corr}} \cdot g_{e1})}{1 + \text{Exp}(\sqrt{\text{corr}} \cdot g_u + \sqrt{1 - \text{corr}} \cdot g_{e1})} \quad P_{2i} = \frac{\text{Exp}(\sqrt{\text{corr}} \cdot g_u + \sqrt{1 - \text{corr}} \cdot g_{e2})}{1 + \text{Exp}(\sqrt{\text{corr}} \cdot g_u + \sqrt{1 - \text{corr}} \cdot g_{e2})}$$

により得られる。この式より,  $\text{corr}$  を 0.1～0.9 (0.1 刻み) で変化させ,  $P_1$ ,  $P_2$  各々 2000 個の乱数を 200 回生成した結果の Pearson の相関係数と  $\text{corr}$  の差の分布を Figure 2-5 に示す。x 軸  $\text{corr}$  が与えた相関係数パラメータであるが,  $\text{corr} > 0.5$  では実際の相関係数は全てが  $\text{corr}$  より小さな値となっている。

確認のため,  $u_1$  と  $u_2$  の相関係数を Figure 2-6 に示す。特に偏りは認められない。Figure 2-5 と同じ  $u_{0i}$ ,  $u_{1i}$  及び  $u_{2i}$  を用いて  $\beta$  分布 ( $\alpha=10$ ,  $\beta=2$ ) 間および Trapezoidal ( $\text{min}=0$ ,  $\text{mode1}=0.5$ ,  $\text{mode2}=0.8$ ,  $\text{max}=1$ ) 間の積率相関係数と  $\text{corr}$  の差の分布を Figure 2-8, Figure 2-7 に示す。Figure 2-5 と同じ傾向の分布となっている。このように, 相関のある一様分布を用いることで, 様々な形状の分布においても相関のある乱数列を得ることができる。Figure 2-8, Figure 2-7 のとおり, 分布形状の異なる乱数間でも相関関係は継承されている。ただし, この式では負の相関 ( $\text{corr} < 0$ ) を設定することはできない。負の相関の場合は,  $\text{corr}$  に絶対値を与え,  $1 - P_i$  とする。



Figure 2-5 相関のある一様分布の Pearson 積率相関係数と与えた相関係数の差の分布

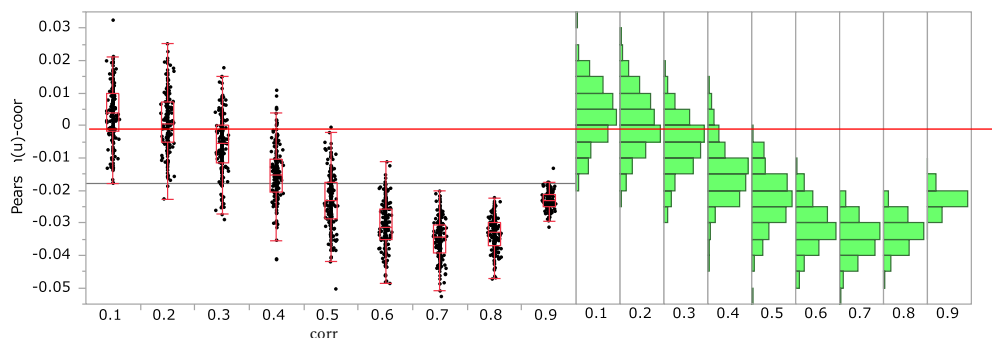
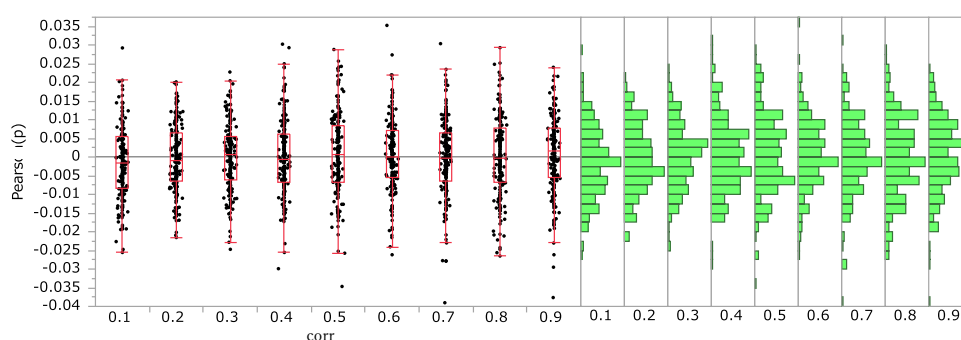
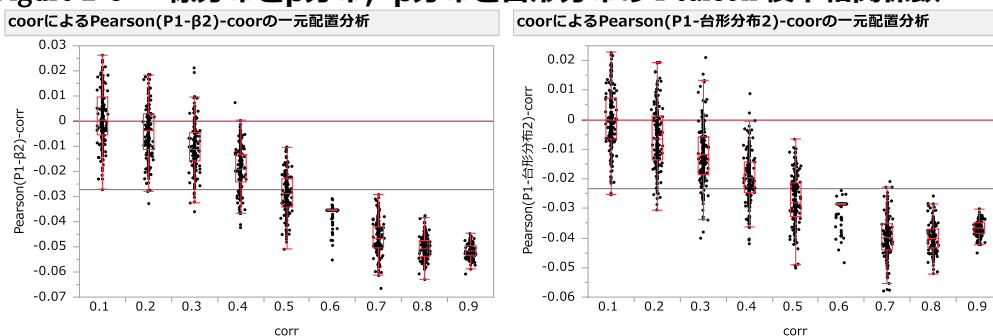
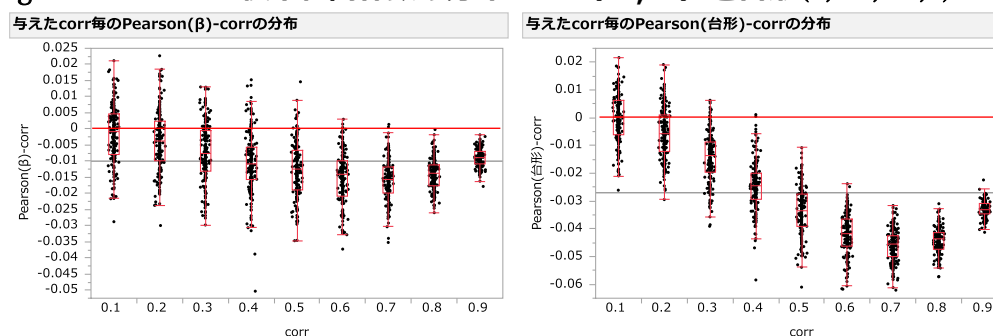
Figure 2-6 相関する一様乱数の式の  $P_1$  と  $P_2$  の Pearson 積率相関係数の差の分布Figure 2-8 一様分布と $\beta$ 分布,  $\beta$ 分布と台形分布の Pearson 積率相関係数

Figure 2-7 Pearson 積率相関係数の分布-Beat (10, 2) と台形(0,0.5,0.8,1)



## ✓ 基準となる分布に値を加える方法

例えば未測定交絡要因の併存率が曝露なしより平均 10%多い、あるいは平均 1.2 倍であるといった仮定を置く場合は相関に基づく方法よりも基準となる分布に値を加える又は掛ける方が直感的である。その場合、基準となる分布は一様分布やベータ分布など、既に説明した方法で作成し、これに正規分布を加える又は掛けることで対応できる。未測定交絡要因について曝露なしが  $p(C+|E-)$  0.2~0.5 の一様分布

( $\text{uniform}(0.2,0.5)$ ), 曝露あり ( $p(C+|E+)$ ) が平均+10%となるよう、曝露なし+平均 0.1, 標準偏差 0.05 の正規分布 ( $p(C+|E-)+\text{normal}(0.1,0.05)$ ) のシナリオを ESATJ では Figure 2-9 の設定で実施できる。作成されたパラメータの分布は Figure 2-10 のとおりで、直線のあてはめからシナリオが正しく反映されていることが確認できる。同様にして曝露あり ( $p(C+|E+)$ ) が  $p(C+|E-)$  の 1.2 倍となり、かつ  $p(C+|E+) \geq p(C+|E-)$  となるよう、曝露なし×平均 1.2, 標準偏差 0.1 の正規分布 ( $p(C+|E-)\times\text{normal}(1.2,0.1)$ ) のシナリオのパラメータの分布は Figure 2-11 のとおりである。何れの方法でも正規分布の標準偏差により  $p(C+|E+) \geq p(C+|E-)$  となるよう調整している。また、相関係数を用いる方法では分布の形状（一様分布, 正規分布, 対数正規分布, 台形分布など）は元となる分布と同じ形状となるが、こちらの方法は元となる分布と異なり、分布の両端は正規分布様となる。

Figure 2-9 ESATJ 平均 10%大きいパラメータの設定例

Select Method **Setting of Probabilistic** Submit UN PRO

Bias Parameters	Min/Mean	Max/ Mod1/ Mod2/ SD	$\alpha$	$\beta$	Distribution
p (Smoking+ Resins+)	0.1	0.05			Pr2+Normal
p (Smoking+ Resins-)	0.20	0.5			Uniform
<b>OR</b> (Smoking-Lung Cancer)	8.60	0.28			Log-normal

↑ RR, OR or RD Selection FALSE #Simulation 20000 re-sampling

Figure 2-10  $p(C+|E-)=\text{uniform}(0.2,0.5)$ ,  $p(C+|E-)+\text{normal}(0.1,0.05)$ の分布

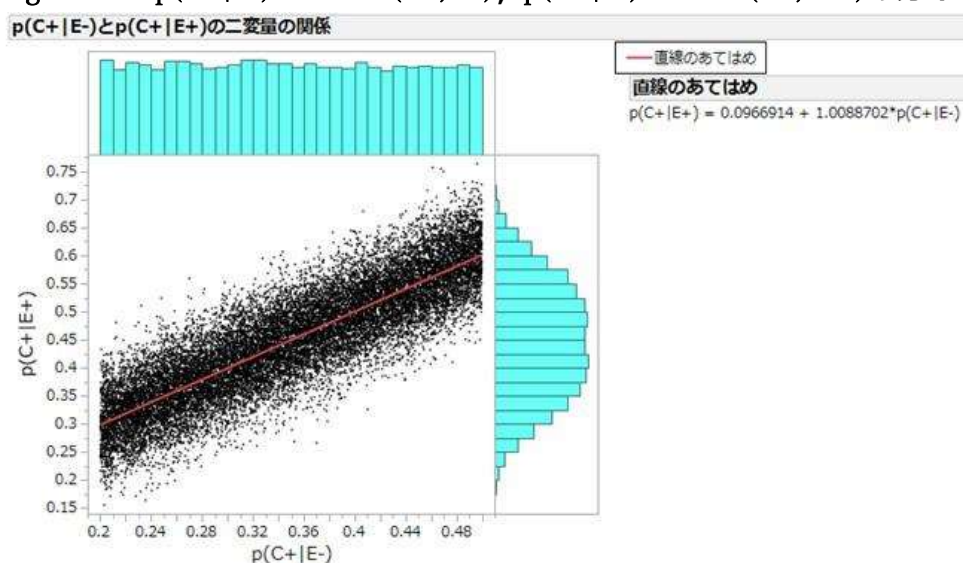
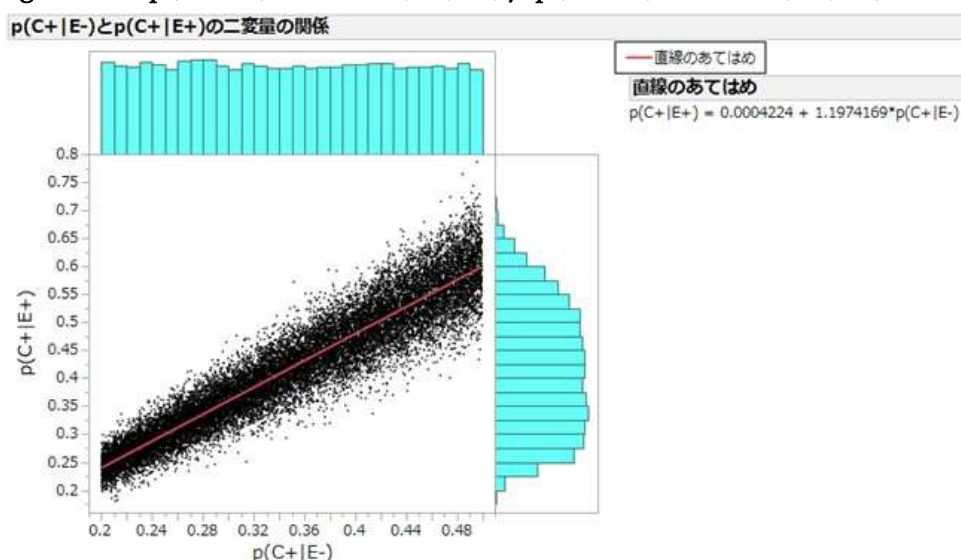


Figure 2-11  $p(C+|E-) = \text{uniform}(0.2, 0.5)$ ,  $p(C+|E-) \times \text{normal}(1.2, 0.1)$  の分布

## 6) 分布に関する注意事項

パラメータの設定には次の事に注意が必要である。

## ① PSA を行う場合の差異誤分類について

層により誤分類の確率が異なると仮定して、例えば層によって感度の最小と最大が  $Se_1 \sim \text{Uniform}(0.7 \sim 0.9)$ ,  $Se_0 \sim \text{Uniform}(0.8 \sim 1)$  の一様分布を設定するような場合、0.8～0.9の間は単にランダムな変動（誤差）として扱われ、「差異誤分類がある」（偏りがある）というシナリオは反映されない。このような場合は、相関係数を用いて、すべての範囲で  $Se_1 < Se_0$  となる確率が高くなるよう設定することでシナリオを反映できる。

## ② パラメータの有効桁数について

稀な事象の誤分類のモデルで用いる特異度のように掛け合わせる相手（「陰性」として測定された値）が万単位など非常に大きな値の場合、小数点以下第4位の値によってさえ結果が大きく変わる場合がある。特に割合に関するパラメータが丸めにより「1」となる場合に注意が必要である。数値例として Table 2-2 を用いる。補正前のオッズ比は 1.76 であるが、曝露の有無に関わらず感度=0.95、特異度=0.99995 の場合の補正後オッズ比は 2.53 となる。しかし、パラメータの有効桁数を小数点以下3位とした場合の特異度の 1.000 を用いた場合の補正後オッズ比は 1.76 となり補正前と同じである。

Table 2-2 曝露の特異度の有効桁数に関する数値例

		E=1	E=0
Crude data#	D*=1	$A_1^* = 45$	$A_0^* = 94$
	D*=0	$B_1^* = 257000$	$B_0^* = 945000$
誤分類	D=1	$A_1 = [Sp_1 A_1^* - (1 - Sp_1) B_1^*] / (Se_1 + Sp_1 - 1) = 32$	$A_0 = [Sp_0 A_0^* - (1 - Sp_0) B_0^*] / (Se_0 + Sp_0 - 1) = 47$
	D=0	$B_1 = (A_1^* + B_1^*) - A_1 = 257013$	$B_0 = (A_0^* + B_0^*) - A_0 = 945047$
# : Table 4-1 の D- (control) の抽出率を 1/1,000 と仮定したコホート研究の場合の数値			

### ③ リスク比やオッズ比に分布を当てはめる場合について

リスク比やオッズ比等のように対数正規分布を取る変数には設定する分布が「1」を含むか否か、あるいは例えばオッズ比=5 という点推定が得られており、その  $1/2 \sim 2$  倍である  $2.5 \sim 10$  という範囲を設定する場合、1 を挟んで 0.1 と 10 が、あるいは 5 を挟んで 2.5 と 10 が等距離という対数軸を想定しているため、Figure 2-12 のように常用対数 ( $\log_{10}(\text{OR})$  等) に対して分布を当てはめる必要がある。Table 4-1 のレジン樹脂曝露と肺癌のデータに未測定交絡要因として  $p(\text{C}+|\text{E}+) \sim \text{uniform}(0.4, 0.7)$ 、 $p(\text{C}+|\text{E}-) \sim \text{uniform}(0.2, 0.5)$  として交絡要因のオッズ比を  $\text{uniform}(1, 10)$  及び  $\text{Log-uniform}(1, 10)$  とした場合、補正後オッズ比と  $0.025^{\text{th}} \sim 0.975^{\text{th}}$  は  $1.37$  ( $0.87 \sim 1.83$ ) と  $1.49$  ( $0.92 \sim 1.81$ ) であり、単に一様分布とした場合は左側に偏った結果となっている。実際には Figure 2-13 と Figure 2-14 (下部中央) のように、補正後オッズ比の分布形状が大きく異なっている。

Figure 2-12 比に対する 0.1～10 の一様分布乱数の分布例

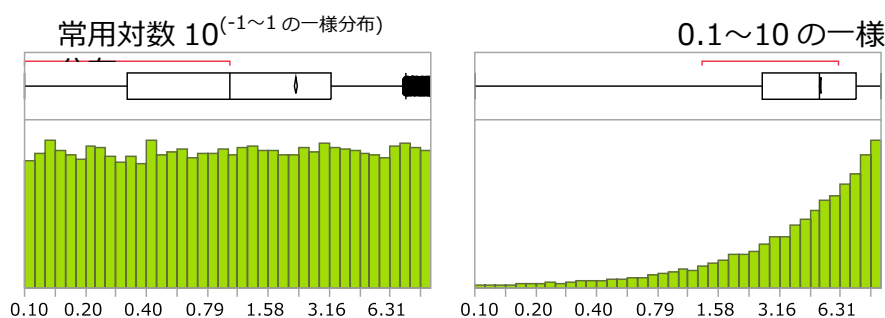


Figure 2-13 未測定お交絡要因のオッズ比の分布に一様分布を用いた場合

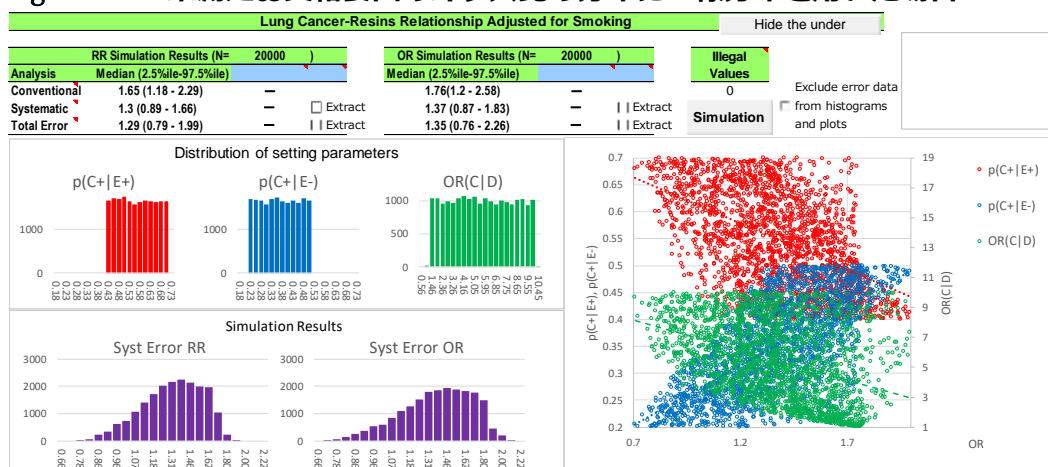
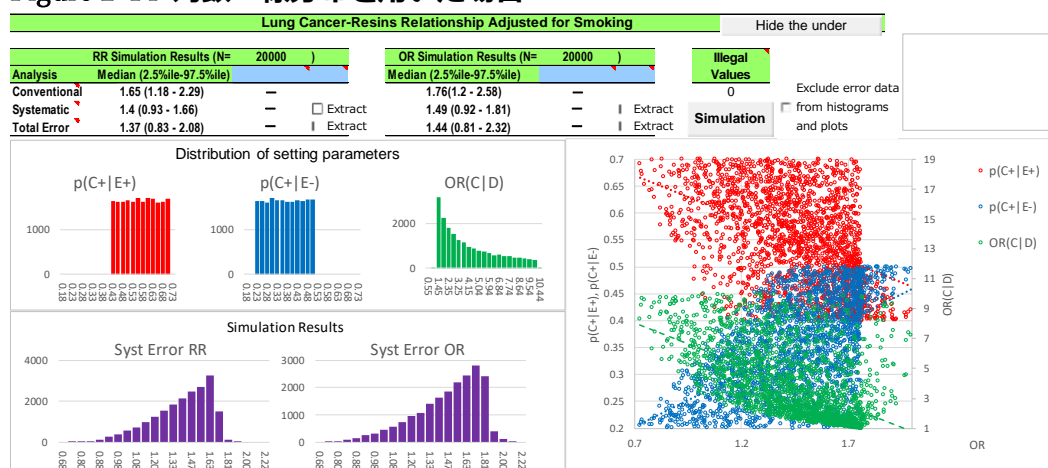


Figure 2-14 対数一様分布を用いた場合

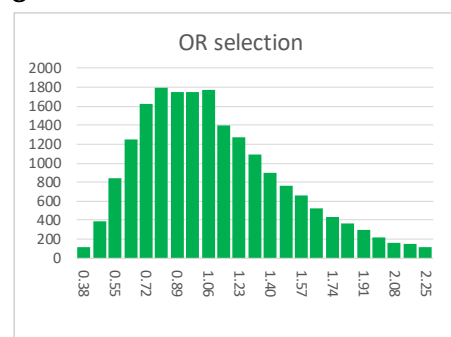


加えて、選択バイアスの Selection OR<sup>注82</sup>をパラメータとする場合の一様分布の使用は注意が必要である。Figure 2-15 は、Table 2-3 の四分表の各セルの抽出率  $S_{11}$ ,  $S_{10}$ ,  $S_{01}$ ,  $S_{00}$  すべてに 0.1~0.2 の一様分布を用いた場合の Selection OR の分布である。数式から予想される通り、Selection OR が一様分布を取ることは現実には生じない。この時の ORselection の計算上の最小値、中央値、最大値は 0.25, 1 と 4 であるが、シミュレーションでは 0.27, 1, 0.38 前後、2.5%点と 97.5%点は 0.46, 2.16 である。Table 2-3 のデータに  $S_{11}$ ,  $S_{10}$ ,  $S_{01}$ ,  $S_{00}$  すべてに 0.1~0.2 の一様分布を用いた場合と Selection OR に 0.29~3.84 の一様分布、対数一様分布およびシミュレーションの分布に近い  $\beta$  分布 (min:0.25, max:4,  $\alpha$ :2.5,  $\beta$ :8)<sup>注83</sup>を与えた場合の補正後オッズ比の分布を Figure 2-16 に示す。Selection OR に一様分布を用いた場合のオッズ比の分布は各セル抽出率に一様分布を用いた場合に比べて大きく偏ったものとなる。対数一様分布を用いた場合、このグラフではわかりにくいですが、補正後オッズ比も対数一様分布となっている。各セル抽出率に一様分布を用いた場合の結果とは異なるが、この場合は目的により適切な場合がある。

Table 2-3 選択バイアスを補正したデータ

	Crude data	
	X=1	X=0
cases	$A_{1+}/S_{11}$	$A_{0+}/S_{10}$
controls	$B_{1+}/S_{01}$	$B_{0+}/S_{00}$

Figure 2-15 選択バイアスパラメータの関



注82: Lash T.L. (2009)<sup>8)</sup>で使用され、WEBに公開されている Excel workbook<sup>Appendix3.1)</sup>の Multidimensional および Probabilistic sensitivity analysis では選択バイアスのバイアスパラメータに  $S_{DE}$  ではなく  $(S_{11} S_{00}) / (S_{10} S_{01})$  の計算結果である Selection OR を用いている。

注83: 得られた分布の最小値、2.5%点、中央値、97.5%点、最大値は 0.26, 0.42, 1.09, 2.21 であった。



Figure 2-16 選択バイアスのパラメータの違いによる補正後オッズ比の分布



## Appendix3. 感度分析を行う Tool, macro

### 1) Excel workbook

#### 1. 本書を作成するにあたって作成した Excel book (ESATJ)

次の「2Applying Quantitative Bias Analysis to Epidemiologic Data で紹介されている。」で紹介している Excel book を参考に、1つのファイルでパラメータに固定した値を用いる Simple sensitivity analysis から、複数の PSA の結果を併合する Multiple bias analysis まで行える Tool を日本製薬工業協会の WEB の本書が公開されているページに公開している。この Tool は次の特徴があり、Apache License, Version 2.0

(<http://www.apache.org/licenses/LICENSE-2.0>) に同意することで利用できる。

- ・ 誤分類について感度+特異度、または陽性的中度+陰性的中度の何れかを指定することができる。
- ・ PSA の分布関数に標準で 11 種類が組み込まれており、さらにユーザーが Worksheet のセルに入力した式を使用することができる（パラメータ間に任意の相関を設定可能）ため、利用できる分布関数に制限がない。
- ・ パラメータ間に複雑に相関を設定できる。
- ・ PSA の結果のグラフ表示が充実している。
- ・ PSA の試行回数を 2 万回としても数分で終了する。（最大 10 万回）

#### ESATJ のメニュー

感度分析の主なバイアスの分類ごとにシートが分かれている。

#### (ア) EDATJ を使い始める前に

ESATJ はマクロ有効 Excel Book ファイルです。また、Excel のアドインである「分析ツール」と「分析ツール-VBA」を使用しています。開いたときにセキュリティの警告等が表示された場合、メッセージバーで「コンテンツの有効化」をクリックください。「コンテンツの有効化」直後、開いたとき又は感度分析の実行時にエラーメッセージで止まる場合、アドインの設定がない場合が考えられます。一旦、Excel を終了してから、(Excel 2016 の場合) Excel のみ起動→空の新規ブックを開く（新規作成）→ファイル→オプション→アドイン→設定 で開いたダイアログで「分析ツール」と「分析ツール-VBA」にチェックを入れて閉じる→ESATJ を開く→空の新規ブックを閉じる を操作ください。

#### (イ) 使い方（基本・共通）

MULTIPLE BIAS, Visualization for Distribution 以外の各シートの使い方は基本的に同じで、感度分析の対象となる四分表のデータを入力して用いる感度分析の手法を Simple

sensitivity analysis (SSA), Multidimensional analysis (MA), Probabilistic sensitivity analysis (PSA) から選択して必要なパラメータを入力する。手法を選択すると、Worksheet の入力欄、ボタンの表示などが変わります。

SSA は入力するとすぐに（ボタン操作なく）計算結果が表示されます。他の 2 つは Submit ボタンを押します。PSA は Submit でまず MCSA のパラメータのリストを作成する。リスト作成後、確認メッセージが表示されるので、そのまま続けて感度分析の計算を行うことも、いったん中止してパラメータの分布を確認後、感度分析の計算を再開することもできる。再開する場合は、Simulation ボタンを押す。

右上の【Set Example data】を押すと、Modern Epidemiology の Chapter 19 で用いられている例（本書の の例）の設定が入力される。Worksheet は、入力欄以外が改変されないよう、保護をしているが、PSA のグラフの軸の修正が必要な場合などのために【UN PROTECT】ボタンを押すことで保護を解除できる。保護を解除しても、感度分析の手法を選んだり、他の何れかのボタンを押すことで、保護が設定される。

#### (ウ) SSA（画面例：EXPOSURE MISCLASSIFICATION）

入力欄は水色で統一している。[Select Method]のリストから「Setting of Single」を選び、変数名、四分表のデータ、[The parameters are]にパラメータの種類（感度、特異度は「Se/Sp」、陽性的中度、陰性的中度は「Ppv/Npv」）を選び、曝露の誤分類のパラメータを入力します。[Corrected Variance?] のチェックは補正した結果に四分表に基づく信頼区間を計算する/しないを指定する。計算結果は下部の[Corrected Measure]欄の黄緑色のセルに表示される。

EXPOSURE MISCLASSIFICATION						
This spreadsheet can be used to conduct a simple sensitivity analysis to correct for exposure misclassification. The example follows Greenland, S., and T. L. Lash. 2008. chapter 19 Bias analysis. In Modern Epidemiology, K. J. Rothman, S. Greenland, and T. L. Lash, 3rd ed					<input type="button" value="Set Example data"/>	
					<input type="button" value="Clear Data"/>	
Instructions						
Enter the bias parameters in the blue cells to the left and the observed data in the blue cells below. Cells in green give the results after correcting for exposure misclassification. Note: green cells are expected values and therefore do not have to be integers.						
Input Bias Parameters				Confirm distribution		
The parameters are		<input type="text" value="Ppv/Npv"/>	Select Method	<input type="text" value="Setting of Single"/>		
				<input type="button" value="UN PROTECT"/>		
Variable Names		Bias Parameters				
Outcome	<input type="text" value="Lung Cancer"/>	Ppv (Lung Cancer+)	<input type="text" value="0.80"/>			
Exposure	<input type="text" value="Resins"/>	Ppv (Lung Cancer-)	<input type="text" value="0.80"/>			
		Npv (Lung Cancer+)	<input type="text" value="0.80"/>			
Error Check	<input type="text" value="No errors found"/>	Npv (Lung Cancer-)	<input type="text" value="0.80"/>			
		<input type="checkbox"/> Corrected Variance?				
Data (Enter Resins-Lung Cancer Data in Blue Cells)			Corrected Data			
Observed Data			Corrected Data			
	Resins +	Resins -	Total	Resins +	Resins -	Total
Lung Cancer +	<input type="text" value="45"/>	<input type="text" value="94"/>	<input type="text" value="139"/>	<input type="text" value="54.8"/>	<input type="text" value="84.2"/>	<input type="text" value="139"/>
Lung Cancer -	<input type="text" value="257"/>	<input type="text" value="945"/>	<input type="text" value="1202"/>	<input type="text" value="394.6"/>	<input type="text" value="807.4"/>	<input type="text" value="1202"/>
Total	<input type="text" value="302"/>	<input type="text" value="1039"/>		<input type="text" value="449.4"/>	<input type="text" value="891.6"/>	
Observed		Measure (95% CI)		Corrected Measure		
RR (Resins-Lung Cancer)		1.65 (1.18 - 2.29)		RR (Resins-Lung Cancer)	1.29 (0.69 - 2.43)	
OR (Resins-Lung Cancer)		1.76 (1.2 - 2.58)		OR (Resins-Lung Cancer)	1.33 (0.71 - 2.51)	

## (工) MA (画面例 UNMEASURED CONFOUNDING with effect modification)

[Select Method]のリストから「Setting of Multidimensional」を選択する。

セル H14 に交絡要因の種類（オッズ比=OR, リスク比=RR, リスク差=RD）を選び, [Setting of Probabilistic→]欄に補正後の値の種類（オッズ比=OR, リスク比=RR）を選び, [Min]の列に各パラメータの最小値, [Increment]の列に Min から増やす 1 区間の値, [#Cycle]の列に区間の数（Min のみの場合は「0」を設定する）を入力する. 例の  $p(\text{Smoking+}|\text{Resins+})$  は 0.4, 0.45, 0.5, 0.55, 0.6 になる. 等区間でない設定をする場合は, 【Input in List】ボタンを押して, 表示される入力フォームに半角カンマ区切りで入力します. 何れも設定を入力後に【Submit】ボタンを押すと計算される.

## UNMEASURED CONFOUNDING with effect modification (OR, RR or RD)

This spreadsheet can be used to conduct a simple sensitivity analysis to correct for unknown or unmeasured confounding. The example follows Greenland, S., and T. L. Lash. 2008. chapter 19 Bias analysis. In Modern Epidemiology, K. J. Rothman, S. Greenland, and T. L. Lash, 3rd ed

Set Example data

Clear Data

## Instructions

Enter bias parameters in blue cells to the right and the crude data in the blue cells below. Cells in green give the results after adjusting for the unmeasured confounder.

## Input Bias Parameters

Confirm distribution

Select Method

Setting of Multidimensional

Submit

UN PROTECT

## Variable Names

Outcome	Lung Cancer
Exposure	Resins
Confounder	Smoking

## Bias Parameters

	Min	#Cycle	ment	Max
p (Smoking+ Resins+)	0.40	4	0.05	
p (Smoking+ Resins-)	0.20	4	0.05	
OR (Smoking-Lung Cancer E+)	2	6	1	
OR (Smoking-Lung Cancer E-)	2	6	1	

Error Check

No errors found

↑ RR, OR or RD Selection

Corrected Measure→

↑ Input in List

## Data (Enter Crude Resins-Lung Cancer Data in Blue Cells)

	Total		Smoking +		Smoking -	
	Resins +	Resins -	Resins +	Resins -	Resins +	Resins -
Lung Cancer +	45	94	25.7	31.3	19.3	62.7
Lung Cancer -	257	945	102.8	189.0	154.2	756.0
Total	302	1039	128.5	220.3	173.5	818.7

2.75

## Crude and Unmeasured Confounder Specific Measures of Resins-Lung Cancer Relationship

	Crude	Measure (95% CI)	Smoking +	Smoking -
OR (Resins-Lung Cancer)	1.76	(1.2 - 2.58)	1.51	1.51
RR (Resins-Lung Cancer)	1.65	(1.18 - 2.29)	1.41	1.45
RD (Resins-Lung Cancer)	0.059	(0.015 to 0.102)	0.06	0.035

## Resins-Lung Cancer Relationship Adjusted for Smoking

## Standardized Morbidity Ratio/ Difference

SMR <sub>OR</sub>	1.51	OR <sub>c</sub>	1.17
SMR <sub>RR</sub>	1.43	RR <sub>c</sub>	1.15
SMD <sub>RD</sub>	0.010	RD <sub>c</sub>	5.839

## Mantel-Haenszel

MH <sub>OR</sub>	1.51	OR <sub>c</sub>	1.17
MH <sub>RR</sub>	1.43	RR <sub>c</sub>	1.15
MH <sub>RD</sub>	0.043	RD <sub>c</sub>	0.015

UnmeasuredConfound\_EMM Setting of Multidimensional

p(E+|D+) 0.2, 0.25, 0.3, 0.4, 0.5, 0.6

p(E+|D-) 0.2, 0.25, 0.3, 0.4, 0.5, 0.6

\* OR (E+) 2, 2.5, 3, 4, 5, 10

RR (E-) 1.5, 2, 2.5, 3, 4, 5, 10

OK Cancel

## Multidimensional OR Resins-Lung Cancer Relationship Adjusted for Smoking

Hide the under

IF RR OR or RR and MH-1 THEN -1/MH

	RR	p	Smoking+ Resins-					Smoking+ Resins+					Smoking+ Resins-					Smoking+ Resins+					Smoking+ Resins-					Smoking+ Resins+				
			0.20	0.25	0.30	0.35	0.40	0.20	0.25	0.30	0.35	0.40	0.20	0.25	0.30	0.35	0.40	0.20	0.25	0.30	0.35	0.40	0.20	0.25	0.30	0.35	0.40	0.20	0.25	0.30	0.35	0.40
0	2.00	0.40	1.51	1.57	1.63	1.70	1.76	1.82	1.89	1.96	2.03	2.10	1.51	1.57	1.63	1.70	1.76	1.82	1.89	1.96	2.03	2.10	1.51	1.57	1.63	1.70	1.76	1.82	1.89	1.96	2.03	2.10
1		0.45	1.56	1.62	1.68	1.74	1.80	1.86	1.92	1.99	2.05	2.12	1.56	1.62	1.68	1.74	1.80	1.86	1.92	1.99	2.05	2.12	1.56	1.62	1.68	1.74	1.80	1.86	1.92	1.99	2.05	2.12
2		0.50	1.61	1.67	1.73	1.79	1.84	1.90	1.96	2.02	2.08	2.14	1.61	1.67	1.73	1.79	1.84	1.90	1.96	2.02	2.08	2.14	1.61	1.67	1.73	1.79	1.84	1.90	1.96	2.02	2.08	2.14
3		0.55	1.66	1.72	1.78	1.84	1.89	1.95	2.01	2.07	2.13	2.19	1.66	1.72	1.78	1.84	1.89	1.95	2.01	2.07	2.13	2.19	1.66	1.72	1.78	1.84	1.89	1.95	2.01	2.07	2.13	2.19
4		0.60	1.71	1.77	1.83	1.89	1.94	2.00	2.06	2.12	2.18	2.24	1.71	1.77	1.83	1.89	1.94	2.00	2.06	2.12	2.18	2.24	1.71	1.77	1.83	1.89	1.94	2.00	2.06	2.12	2.18	2.24
0	3	0.40	1.66	1.73	1.80	1.87	1.93	1.99	2.05	2.11	2.17	2.23	1.66	1.73	1.80	1.87	1.93	1.99	2.05	2.11	2.17	2.23	1.66	1.73	1.80	1.87	1.93	1.99	2.05	2.11	2.17	2.23
1		0.45	1.70	1.77	1.84	1.91	1.97	2.03	2.09	2.15	2.21	2.27	1.70	1.77	1.84	1.91	1.97	2.03	2.09	2.15	2.21	2.27	1.70	1.77	1.84	1.91	1.97	2.03	2.09	2.15	2.21	2.27
2		0.50	1.75	1.82	1.89	1.95	2.02	2.08	2.14	2.20	2.26	2.32	1.75	1.82	1.89	1.95	2.02	2.08	2.14	2.20	2.26	2.32	1.75	1.82	1.89	1.95	2.02	2.08	2.14	2.20	2.26	2.32
3		0.55	1.80	1.87	1.94	2.00	2.06	2.12	2.18	2.24	2.30	2.36	1.80	1.87	1.94	2.00	2.06	2.12	2.18	2.24	2.30	2.36	1.80	1.87	1.94	2.00	2.06	2.12	2.18	2.24	2.30	2.36
4		0.60	1.85	1.92	1.99	2.05	2.11	2.17	2.23	2.29	2.35	2.41	1.85	1.92	1.99	2.05	2.11	2.17	2.23	2.29	2.35	2.41	1.85	1.92	1.99	2.05	2.11	2.17	2.23	2.29	2.35	2.41
0	4	0.40	1.82	1.90	1.98	2.06	2.13	2.20	2.27	2.34	2.41	2.48	1.82	1.90	1.98	2.06	2.13	2.20	2.27	2.34	2.41	2.48	1.82	1.90	1.98	2.06	2.13	2.20	2.27	2.34	2.41	2.48
1		0.45	1.86	1.94	2.02	2.09	2.16	2.23	2.30	2.37	2.44	2.51	1.86	1.94	2.02	2.09	2.16	2.23	2.30	2.37	2.44	2.51	1.86	1.94	2.02	2.09	2.16	2.23	2.30	2.37	2.44	2.51
2		0.50	1.91	1.99	2.07	2.14	2.21	2.28	2.35	2.42	2.49	2.56	1.91	1.99	2.07	2.14	2.21	2.28	2.35	2.42	2.49	2.56	1.91	1.99	2.07	2.14	2.21	2.28	2.35	2.42	2.49	2.56
3		0.55	1.96	2.04	2.11	2.18	2.25	2.32	2.39	2.46	2.53	2.60	1.96	2.04	2.11	2.18	2.25	2.32	2.39	2.46	2.53	2.60	1.96	2.04	2.11	2.18	2.25	2.32	2.39	2.46	2.53	2.60
4		0.60	2.01	2.09	2.16	2.23	2.30	2.37	2.44	2.51	2.58	2.65	2.01	2.09	2.16	2.23	2.30	2.37	2.44	2.51	2.58	2.65	2.01	2.09	2.16	2.23	2.30	2.37	2.44	2.51	2.58	2.65
0	5	0.40	1.90	1.98	2.06	2.14	2.21	2.28	2.35	2.42	2.49	2.56	1.90	1.98	2.06	2.14	2.21	2.28	2.35	2.42	2.49	2.56	1.90	1.98	2.06	2.14	2.21	2.28	2.35	2.42	2.49	2.56

(オ) PSA (画面例 : OUTCOME MISCLASSIFICATION)

A) 手法の選択

[Select Method]のリストから「Setting of Probabilistic」を選択する.

B) データの入力

[Data]欄に感度分析を行う四分表のデータを入力する.

C) パラメータの種類を選択

[The parameters are]にパラメータの種類 (感度, 特異度は「Se/Sp」, 陽性的中度, 陰性的中度は「Ppv/Npv」) を選ぶ.

D) 分布形状の選択

各パラメータの分布形状を[Distribution]列をリストから選ぶとパラメータの入力欄が変わる.

・ 分布関数

Trapezoidal and Triangular is

If Triangular then  $\text{Mod2} = \text{Mod1}$

$u = \text{Uniform}(0 \sim 1)$ ,  $v = (\text{Min} + \text{Mod1} + u(\text{Max} + \text{Mod2} - \text{Min} - \text{Mod1})) / 2$

If  $v < \text{Mod1}$  then  $p = \text{Min} + \text{Sqrt}[(\text{Mod1} - \text{Min})(2v - \text{Min} - \text{Mod1})]$

If  $v > \text{Mod2}$  then  $p = \text{Max} + \text{Sqrt}[2(\text{Max} - \text{Mod2})(v - \text{Mod2})]$

Else  $p = v$

Trapezoidal or Triangular =  $p$

Normal is

$\text{Normal} = \text{Mean} + \text{Norm.Inv}(\text{Uniform}(0 \sim 1), 0, \text{SD})$

Weibull is

$\text{tp1} = \beta * \{[-1 * \text{Log}(0.000005)]^{(1/\alpha)}\}$ ,  $\text{tp2} = (\text{Max} - \text{Min}) / \text{tp1}$

$u = \text{Uniform}(0 \sim 1)$

$\text{Weibull} = \text{Min} + \text{tp2} * \{\beta * [\text{Log}(1/(1 - u))^{(1/\alpha)}]\}$

Log-normal is

If  $\text{Mean} = 0$  Then  $\text{Mean} = 1$ ,  $u = \text{Uniform}(0 \sim 1)$

$\text{Log-normal} = \exp\{\ln(\text{Mean}) + \text{Norm.Inv}(u, 0, \text{SD})\}$

Log-uniform is

$u = \text{Uniform}(\ln(\text{Min}) \sim \ln(\text{Max}))$ ,  $\text{Log-uniform} = \exp(u)$

Logit is

$u = \text{Uniform}(0 \sim 1)$ ,  $\text{logit}(u) = \ln[u/(1-u)] = \ln(u) - \ln(1-u)$



expit(logit) is

$$u = \text{Uniform}(0 \sim 1), \quad g = \ln[\alpha / (1 - \alpha)] + \beta * \ln[u / (1 - u)]$$

$$\text{expit}(\text{logit}) = \text{Min} + (\text{Max} - \text{Min}) * [\exp(g) / (\exp(g) + 1)]$$

OR<sub>Rec</sub>=OR<sub>Exposure-Confounder</sub> is

$$n = \text{normal}(\ln(\text{mean}), \text{SD}), \quad p = p(\text{Confounder} + | \text{Exposure} -) = \exp(n),$$

$$q = p(\text{Confounder} + | \text{Exposure} -)$$

$$\text{OR}_{\text{Rec}} = 1 / \{ [(1 - p) / (p * \exp(n))] + 1 \}$$

$$\text{Pr2} + \text{Normal} = \text{Parameter1} + \text{Mean} + \text{Norm.Inv}(\text{Uniform}(0 \sim 1), 0, \text{SD})$$

$$\text{Pr2} \times \text{Normal} = \text{Parameter1} \times (\text{Mean} + \text{Norm.Inv}(\text{Uniform}(0 \sim 1), 0, \text{SD}))$$

RatioDistF= Distribution of ratio by F distribution is

n= Number of validation-ssample, x=Number of occurrences  
t=Total number of layers, u= Uniform(0~1)

If n<100 or n/t<0.1 then

$$d1 = 2(n - x + 1), \quad d2 = 2x$$

$$\text{RatioDistF} = d2 / (d1 * F.\text{inv}(u, d1, d2) + d2)$$

$$\sim \text{RatioDistF} \geq \text{min and RatioDistF} \leq \text{max}$$

Else 'finite correction

$$p = x / n$$

$$pp = p + \text{Norm.inv}(u, 0, 1) * \text{Sqrt}((1 - n / t) * p * (1 - p) / n)$$

Expression is input formula yourse

If Exsample When logit(p)

p~normal(1,0.1) of Unmeasured Confound OR<sub>CD</sub> then

Formura of cell P13 is "=Norm.inv(RAND(),1, 0.1)"

Formura of cell M13 is "=ln(P13/(1-P13))"

#### E) 分布の形状を確認する

検証研究や文献，過去の経験などからパラメータに関する情報を得て分布を決める場合でも，分布関数を用いて表現される分布の形状の確認は必要である．また，信頼できる唯一の情報ではなく，複数の情報を総合して分布を決める（作る）必要がある場合も少なくない．更に，このように得られている情報，例えば感度について平均値と標準偏差が得られているが，ここから分布を作ると 0 未満や 1 以上が生じてしまうなど，不都合な場合もある．

ESATJ では「Distribution」シートで、D)で紹介している分布関数に様々なパラメータを設定して実際の分布を確認することができる。Figure 3-1 はロジット (expit(logit)) の例である。分布を表すヒストグラムに最小、最大、平均、最頻値、パーセント点などのテーブルが表示されている。更に、このような任意の関数をベータ分布で近似できるパラメータをシミュレーションで得られる。平均値 0.9、標準偏差 0.05 の正規分布に最小値 0.7、最大値 1 の条件で  $\beta$  分布をあてはめた場合、Figure 3-2 のようになる。正規分布は約 2.5% が 1 以上となっており、例えば感度のパラメータに用いる場合、1 以下の分布とすることが望ましい。このため、右側の図でベータ分布の範囲を指定し、さらに最頻値を重視した近似を行うオプションを指定している。結果は第 1 パラメータ (a) に 5.535、第 2 パラメータ ( $\beta$ ) に 2.76 を得られ、Figure 3-2 の右の累積確率分布と左のヒストグラムに追加された曲線からあてはまりを確認できる。

Figure 3-1 分布の確認の例

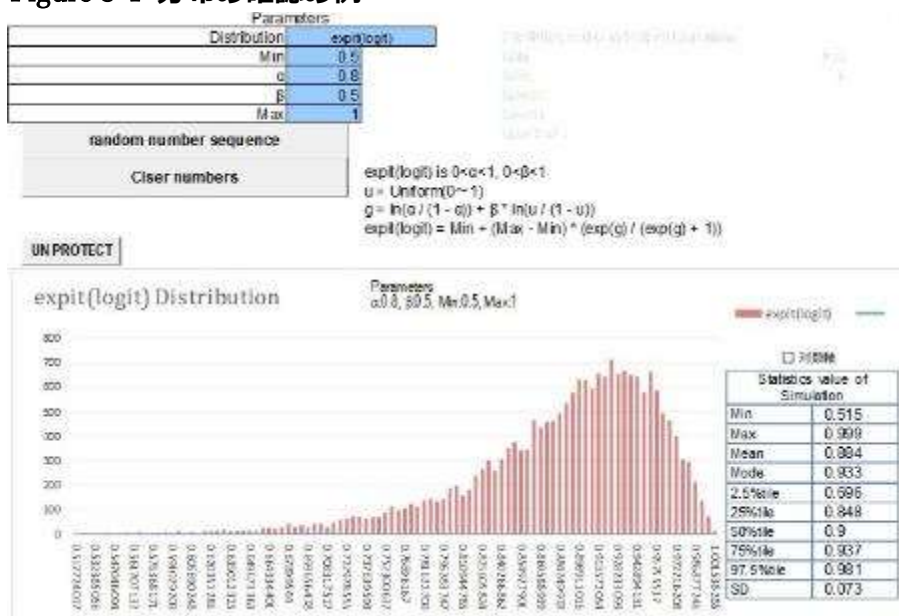
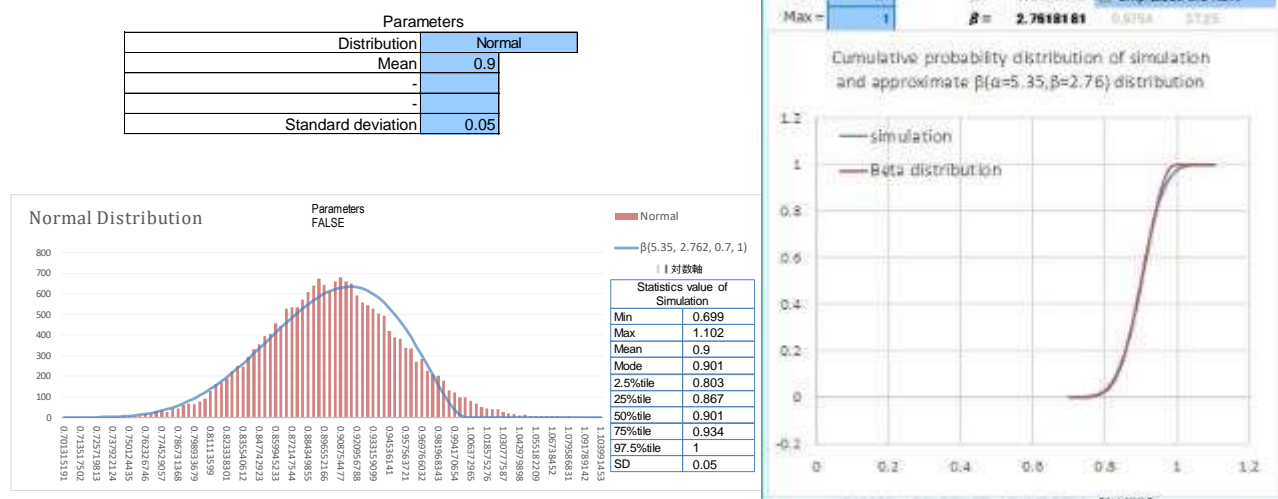


Figure 3-2 正規分布をベータ分布で近似した例



## F) パラメータの入力

分布形状に合わせて、分布の設定を[Min], [Max]などに入力します。分布の確認は [Distribution]列の見出しをクリックすると、「MENU」シートの分布の説明が表示される。また、右上の[Confirm distribution]をクリックすると、「Distribution」シートが表示される。相関するパラメータは、パラメータの入力欄の順番で上側のパラメータは分布形状の設定を行い、下側の相関するパラメータの[Distribution]列に相関するパラメータを「Correlate 1」～「Correlate 3」から選びます<sup>注84</sup>。例は相関係数=1 とすることで、 $Se(Resins-) = Se(Resins+)$ ,  $Sp(Resins-) = Sp(Resins+)$ となる設定である。相関する分布の形状 (Distribution) は同じものを使用するが、その最小, 最大, 形状のパラメータは個別に異なるものを設定できる。

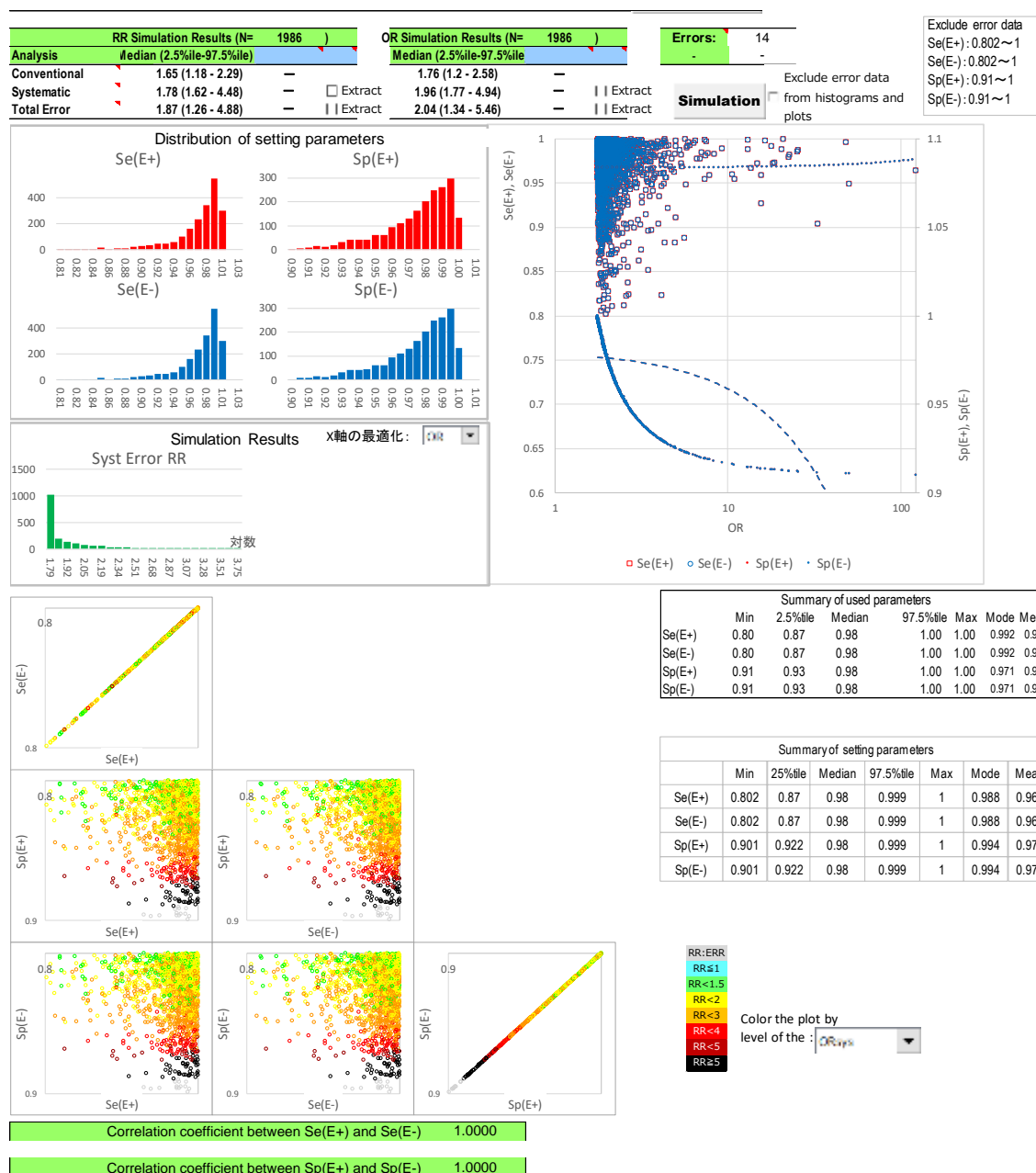
パラメータには、本来取り得ない値となる分布 (感度に負の値や1より大きな値) を設定するべきではないが、validation study に基づいて平均と分散による分布設定を行うなど、致し方ない場合がある。そのような場合に、エラーとするのではなく、0 から1の間の値が得られるまで再抽出を行いたい場合は[re-sampling]をチェックする。ただし、この場合は分布の中央値が変わるため、補正後のオッズ比などの中央値が影響を受ける。感度、特異度などが0未満または1を超えても無視し、負のセルが生じた場合もエラーとせず小さな値として補正後の値の頻度に用いる場合は[Handling of Negative Cells]にリストから適切な値を指定する。なお、教科書には0.5を用いることが説明されている<sup>注60</sup>。

Input Bias Parameters				Confirm distribution			
The parameters are <b>Se/Sp</b>		Select Method	<b>Setting of Probabilistic</b>		Submit		UN PROTECT
Variable Names		Bias Parameters		Min	Max	1/α	2/β
Outcome	Lung Cancer	Se(Resins+)		0.80	1	0.9	0.8
Exposure	Resins	Se(Resins-)		0.80	1	0.9	0.8
Error Check <b>Negative Cell</b>		Sp(Resins+)		0.90	1	0.8	0.8
Handling of Negative Cells <b>ERR</b>		Sp(Resins-)		0.90	1	0.8	0.8
		<input type="checkbox"/> Corrected Variance?					
				2000		#Simulation	
						<input type="checkbox"/> re-sampling	

Data (Enter Resins-Lung Cancer Data in Blue Cells)				Corrected Data			
Observed Data				Corrected Data			
	Resins +	Resins -	Total	Resins +	Resins -	Total	
Lung Cancer +	45	94	139	21.1	-14.1	7.00	
Lung Cancer -	257	945	1202	280.9	1053.1	1334.00	
Total	302	1039		302.0	1039.0		

注<sup>84</sup>： 詳細は 4.1.7 相関したパラメータ分布 参照



## G) パラメータの乱数リスト作成

データ、パラメータを設定後、[←#Simulation]にシミュレーションの回数を設定して【Submit】ボタンを押すと、PSAのパラメータの乱数リストが作成される。続けて感度分析の計算を行うか確認が表示されるが、【キャンセル】を押して一旦中止し、パラメータの分布をグラフで確認する。

H) パラメータの分布が良ければ、【Simulation】ボタンを押す。終了のメッセージが表示されたら、結果を確認する（エラーがあると、その数が表示されます）。

- ・ エラーの確認

マイナスのセルが生じるなど、エラーとなった simulation 回数が[Illegal Values] に表示される。パラメータのグラフは、エラーとなったものも表示されているが、[Exclude error data from histograms and plots]をチェックすることでヒストグラムおよび散布図からエラーとなったパラメータを除いて表示できる。

- ・ パラメータと補正後のオッズ比との関係の確認

下部の[Color the plot by level of the]のリストを選択すると、散布図行列を選択した指標（OR<sub>total</sub> など）で色分けして表示できる。（少し時間がかかります。画面例は、選択のある状態である。）

- ・ オッズ比、リスク比が任意の値以下/以上となる場合の確認

[RR Simulation Results]（又は[OR Simulation Results]）欄の水色のセルで「Freq of RR ≥」又は「Freq of RR ≤」を選び、その右に閾値とする値を入力すると、その下に条件に合う割合が表示される。その右の[Extract]をチェックすると、散布図を条件の合う組み合わせのみで表示される。（ヒストグラムは変わらない）。

- ・ 任意の条件の確認

Worksheet の下部（概ね 120 行辺り）にパラメータと計算したオッズ比などのリストが表示されている。これのフィルタと散布図はリンクしているので、任意の場合に表示を制限できる。また、このリストを別の Excel Book にコピーして特定のパラメータ（範囲）の場合の補正後の値の統計量を求めることができる。

## (カ) MULTIPLE BIAS

他の worksheet で実施した感度分析の結果を併合しますので、先に各バイアスの PSA が完了している必要がある。

### A) 補正後の値の種類などを指定する。

[#Simulation]にシミュレーションの回数、[Combinse value]に OR, RR 何れか併合する値を指定する。

### B) 併合するモデルを指定する。

表の[↓ SELECT COMBINED ORDER]に併合するモデルの順番を入力する。

表には、各 Worksheets で実施されている PSA の設定が表示されている。結果に表示するが、併合には使用しないモデルには「NO」を入力する。ブランクにすると結果に表示さない。また、[SELECTION BIAS]は各セルの抽出率を用いず、抽出のオッズ比を用いる場合は[Use Selection OR]をチェックする。

### C) 【Run Simulation】 ボタンを押す。

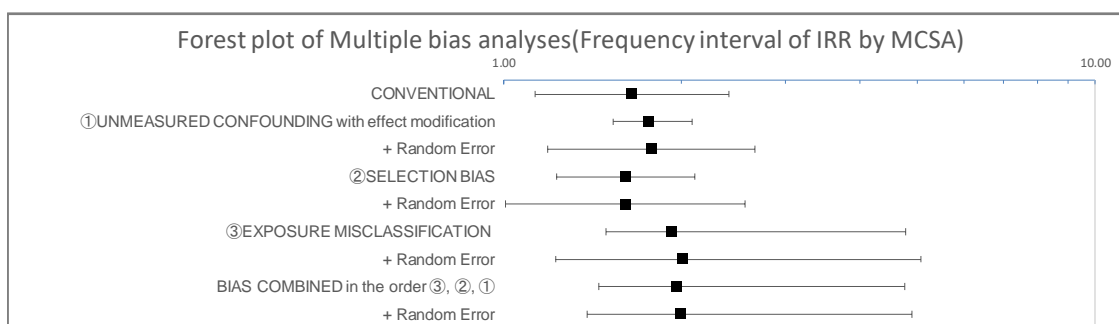
### D) forest plots など結果が表示される。



	Variable Names
Outcome	Asthma
Exposure	powder dust
Unmeasured Confounder	unknown
Confounder	-

Observed Data	powder dust	
	+	-
Asthma+	476	450
Asthma-	9286	15536
	9762	15986
	25748	

↓ SELECT COMBINED ORDER		Bias Parameters	Min	Mode1/ $\alpha$	Mode2/ $\beta$	Max	$\rho$	Distribution
NO	UNMEASURED CONFOUNDING without effect	p(unknown+ powder dust+)	0.01			1.00		Uniform
		p(unknown+ powder dust-)	0.01			1.00		Uniform
		RR(unknown-Asthma)	2.00			10.00		Uniform
3	UNMEASURED CONFOUNDING with effect modification	p(Smoking+ Resins+)	0.00			0.10		p(C+ E+)+Normal
		p(Smoking+ Resins-)	0.40			0.70		Uniform
		OR(Smoking-Lung Cancer E+)	8.66			0.28		Log-normal
		OR(Smoking-Lung Cancer E-)	8.66			0.28	1.00	Correlate with OR(C-D)
2	SELECTION BIAS	S(Lung Cancer+ Resins+)	0.10			0.20		Uniform
		S(Lung Cancer+ Resins-)	0.10			0.20		Uniform
		S(Lung Cancer- Resins+)	2.00			0.50		Uniform
		S(Lung Cancer- Resins-)	1.00			0.50		Uniform
		-						
1	EXPOSURE MISCLASSIFICATION	Se (Lung Cancer+)	0.98	0.90	0.80	1.00		expit(logit)
		Se (Lung Cancer-)	0.98	1.00	0.80	1.00	0.80	Correlate with Se(D+)
		Sp (Lung Cancer+)	0.95	1.00	0.80	1.00		expit(logit)
		Sp (Lung Cancer-)	0.95	1.00	0.80	1.00	0.80	Correlate with Sp(D+)
NO	OUTCOME MISCLASSIFICATION	Se(Resins+)	0.80	0.90	0.80	1.00		expit(logit)
		Se(Resins-)	0.80	1.00		1.00	1.00	Correlate with SeE
		Sp(Resins+)	0.90	1.00	0.80	1.00		expit(logit)
		Sp(Resins-)	0.90	1.00		1.00	1.00	Correlate with SpE
	CONFOUNDER MISCLASSIFICATION							



	Whithout Incorporating Random Error			Whith Incorporating Random Error		
	Median	Percentiles 2.5th-97.5th	Ratio of Limits	Median	Percentiles 2.5th-97.5th	Ratio of Limits
CONVENTIONAL	1.65	- -	-	1.65	1.13 ~ 2.40	2.13
UNMEASURED CONFOUNDING without effect modification ① p(C+ E+)~p(C+ E+)+Normal(0,0.1), p(C+ E+)~Uniform(0.4,0.7), OR(C-D)~Log-normal(8.66,0.28),	1.76	1.53 ~ 2.08	1.36	1.78	1.19 ~ 2.66	2.24
BIAS COMBINED	1.96	1.45 ~ 4.77	3.29			
SELECTION BIAS ② S(D+ E+)~Uniform(0.1,0.2), S(D+ E-)~Uniform(0.1,0.2), S(C- E+)~Uniform(2,0.5), S(C- E-)~Uniform(1,0.5),	1.61	1.23 ~ 2.10	1.71	1.61	1.01 ~ 2.56	2.54
BIAS COMBINED	1.92	1.49 ~ 4.67	3.13			
EXPOSURE MISCLASSIFICATION ③ Se(D+)~expit(logit)(0.98,0.9,0.8,1) Se(D-)~Correlate with Se(D+)(0.98,1,0.8,1,0.8) Sp(D+)~expit(logit)(0.95,1,0.8,1) Sp(D-)~Correlate with Sp(D+)(0.95,1,0.8,1,0.8)	1.92	1.49 ~ 4.78	3.21	2.00	1.22 ~ 5.07	4.15
BIAS COMBINED	1.92	1.49 ~ 4.78	3.21			
BIAS COMBINED in the order ③, ②, ①	1.96	1.45 ~ 4.77	3.30	1.98	1.38 ~ 4.82	3.50
Uniform(min,max), Log-normal(mean,SD), expit(logit)(min, $\alpha$ , $\beta$ ,max), Correlate with Parameter(Pr1,Pr2,Pr3,Pr4,p)						

2. Applying Quantitative Bias Analysis to Epidemiologic Data で紹介されている.

同様の SAS macro (は 2) SAS macro の 4, 5, R Package (は 4) R Package の 2 を参照.

<https://sites.google.com/site/biasanalysis/Home>

- Selection bias :  
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxiaWFzYW5hbHlzaXN8Z3g6NGZmOGFmMWUyYzhmOTY1OA>
- Unmeasured Confounding:  
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxiaWFzYW5hbHlzaXN8Z3g6N2NmMjU3YmJkMzIyOGNINQ>
- Misclassification:  
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxiaWFzYW5hbHlzaXN8Z3g6NTAxZGQ2NDI2ZGI0ZWnkMw>
- Multidimensional:  
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxiaWFzYW5hbHlzaXN8Z3g6NDViYmU0OTI3MmUxNGM4YQ>
- Probabilistic sensitivity analysis (person time):  
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxiaWFzYW5hbHlzaXN8Z3g6NzAwZDdhMDBiNDI4NWE3Yw>
- Probabilistic sensitivity analysis:  
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxiaWFzYW5hbHlzaXN8Z3g6NDY5M2YyNTNiODJiZmE0MA>

3. Division of Pharmacoepidemiology & Pharmacoeconomics Department of Medicine Brigham and Women's Hospital Harvard Medical School

[http://www.drugpi.org/dope-downloads/#Pharmacoepidemiology Toolbox](http://www.drugpi.org/dope-downloads/#Pharmacoepidemiology%20Toolbox)

Sensitivity Analysis of Confounding: [http://www.drugpi.org/wp-](http://www.drugpi.org/wp-content/uploads/2011/04/Sensi.Ana_.Confounding.Apr08.xls)

[content/uploads/2011/04/Sensi.Ana\\_.Confounding.Apr08.xls](http://www.drugpi.org/wp-content/uploads/2011/04/Sensi.Ana_.Confounding.Apr08.xls)

## 2) SAS macro

4. Applying Quantitative Bias Analysis to Epidemiologic Data <sup>29)</sup>

<https://sites.google.com/site/biasanalysis/> の

chapter9multiplebiasanalysis.sas

<https://sites.google.com/site/biasanalysis/Home/chapter9multiplebiasanalysis.sas?attredirects=0&d=1>

5. A method to automate probabilistic sensitivity analyses of misclassified binary variables. <sup>21)</sup>

<https://sites.google.com/site/biasanalysis/sensmac>

6. Semi-automated sensitivity analysis to assess systematic errors in observational data <sup>22)</sup>

<https://sites.google.com/site/biasanalysis/multiple-bias-model-lash-tl-fink-a>

### 3) STATA

Modern Epidemiology 3<sup>rd</sup> ed , Bias analysis の手法を STATA で行うコード.

Orsini, N., Bellocco, R., Bottai, M., Wolk, A. and Greenland, S. A tool for deterministic and probabilistic sensitivity analysis of epidemiologic studies, The Stata Journal, 2008, 8, 29–48.

<http://nicolaorsini.altervista.org/stata/tutorial/e/episens.htm>

### 4) R Package

1. causalsens:

Selection Bias Approach to Sensitivity Analysis for Causal Effects.

The causalsens package provides functions to perform sensitivity analyses and to study how various assumptions about selection bias affects estimates of causal effects.

<https://cran.r-project.org/web/packages/causalsens/index.html>

2. episensr:

Basic Sensitivity Analysis of Epidemiological Results.

Basic sensitivity analysis of the observed relative risks adjusting for unmeasured confounding and misclassification of the exposure/outcome, or both. It follows the bias analysis methods and examples from the book by Lash T.L., Fox M.P, and Fink A.K. "Applying Quantitative Bias Analysis to Epidemiologic Data", ('Springer', 2009).

<https://cran.r-project.org/web/packages/episensr/index.html>

3. evidenceFactors:

Reporting Tools for Sensitivity Analysis of Evidence Factors in Observational Studies.

The tools and algorithms implemented in evidenceFactors package are discussed in detail in Karmakar et. al. (2016). For general discussion of sensitivity analyses in observational studies, see Chapter 4 of Rosenbaum (2002).

<https://cran.r-project.org/web/packages/evidenceFactors/index.html>

4. multisensi:

Multivariate Sensitivity Analysis.

This package generalises sensitivity analysis to simulation models with multivariate output. It makes it easy to run a series of independent sensitivity analyses on a set of output variables and to plot the results. Alternatively, it allows to apply sensitivity analyses to the

variables resulting from the application of a multivariate method (such as PCA or splines or polynomial regression) to the output data (Lamboni et al., 2009).

<https://cran.r-project.org/web/packages/multisensi/index.html>

5. obsSens:

Sensitivity analysis for Observational studies

Observational studies are limited in that there could be an unmeasured variable related to both the response variable and the primary predictor. If this unmeasured variable were included in the analyses it would change the relationship (possibly changing the conclusions). Sensitivity analysis is a way to see how much of a relationship needs to exist with the unmeasured variable before the conclusions change. This package provides tools for doing a sensitivity analysis for regression (linear, logistic, and cox) style models.

<https://cran.r-project.org/web/packages/obsSens/index.html>

6. SBSA:

Simplified Bayesian Sensitivity Analysis

Simplified Bayesian sensitivity analysis of models with partially observed confounders. The SBSA package is an implementation of algorithms for simplified Bayesian sensitivity analysis described in Gustafson et al (2010).

<https://cran.r-project.org/web/packages/SBSA/index.html>

7. sensitivity2x2xk:

Sensitivity Analysis for 2x2xk Tables in Observational Studies.

Performs exact or approximate adaptive or nonadaptive Cochran-Mantel-Haenszel-Birch tests and sensitivity analyses for one or two 2x2xk tables in observational studies.

<https://cran.r-project.org/web/packages/sensitivity2x2xk/index.html>

8. SensitivityCaseControl:

Sensitivity Analysis for Case-Control Studies.

This package performs sensitivity analysis for case-control studies in which some cases may meet a more narrow definition of being a case compared to other cases which only meet a broad definition. The sensitivity analyses are described in Small, Cheng, Halloran and Rosenbaum (2013, "Case Definition and Sensitivity Analysis", Journal of the American Statistical Association, 1457-1468). The functions sens.analysis.mh and sens.analysis.aberrant.rank provide sensitivity analyses based on the Mantel-Haenszel test statistic and aberrant rank test statistic as described in Rosenbaum (1991, "Sensitivity Analysis for Matched Case Control Studies", Biometrics); see also Section 1 of Small et al. The function adaptive.case.test provides adaptive inferences as described in Section 5 of Small et al. The function adaptive.noether.brown provides a sensitivity analysis for a matched cohort study based on an adaptive test. The other functions in the package are

internal functions.

<https://cran.r-project.org/web/packages/SensitivityCaseControl/index.html>

9. sensitivityfull:

Sensitivity Analysis for Full Matching in Observational Studies.

Sensitivity to unmeasured biases in an observational study that is a full match.

<https://cran.r-project.org/web/packages/sensitivityfull/index.html>

10. SimpleTable:

Bayesian Inference and Sensitivity Analysis for Causal Effects from 2 x 2 and 2 x 2 x K Tables in the Presence of Unmeasured Confounding

SimpleTable provides a series of methods to conduct Bayesian inference and sensitivity analysis for causal effects from 2 x 2 and 2 x 2 x K tables when unmeasured confounding is present or suspected.

<https://cran.r-project.org/web/packages/SimpleTable/index.html>

11. treatSens:

Sensitivity Analysis for Causal Inference. Utilities to investigate sensitivity to unmeasured confounding in parametric models with either binary or continuous treatment.

<https://cran.r-project.org/web/packages/treatSens/index.html>