

SSE-DP-2022-1

オッズ比の平方根変換

岩崎 学

統計数理研究所/順天堂大学大学院

2022年11月

SSE-DP(ディスカッションペーパー・シリーズ)は以下のサイトから無料で入手可能です。
<https://stat-expert.ism.ac.jp/training/discussionpaper/>
このディスカッション・ペーパーは、関係者の討論に資するための未定稿の段階にある草稿である。著者の承諾なしに引用・複写することは差し控えられたい。

SSE-DP-2022-1

Square root transformation of the odds ratio

Manabu Iwasaki

The Institute of Statistical Mathematics

November 2022

(Summary)

In medical statistics, risk ratio (RR) and odds ratio (OR) are often used to evaluate the relationship between exposure and outcome. Risk ratios are easier to interpret in practice, but depending on the study design, odds ratios may be only available. The odds ratio is a good approximation of the risk ratio when the baseline risk is very small, but it is not the case when the risk is common. Hence, VanderWeele (2007, 2010) proposed the square root transformation of the odds ratio as an approximation of the risk ratio.

In this article, we first give the relationships between the risk ratio and the odds ratio and the square root transformation of the odds ratio, and evaluate whether the odds ratio and its square root transformation can be an approximation of the risk ratio in terms of the coverage probability of the confidence interval. The simulation results show that the odds ratio approximates the risk ratio well when the baseline risk is less than 0.2, and the square root transformation of the odds ratio approximates the risk ratio well when the baseline risk is between 0.3 and 0.7.

オッズ比の平方根変換

統計数理研究所 特任教授
順天堂大学大学院 客員教授
岩崎 学

(注) 本稿は、2022年11月4日に滋賀大学で開催された「統計エキスパート人材育成—中間報告会」での特別講演「リスク比とオッズ比」(岩崎 学)の内容に加筆したものである。

0. 要約

医療統計では、リスク比とオッズ比が重要な役割を果たす。リスク比はその解釈が容易であるがオッズ比はそうではないのであるが、研究結果がオッズ比で報告されることが多いのが現状である。最近、オッズ比の平方根がリスク比に近似するという論文が出た(VanderWeele, 2017, 2020)。ここでは、その平方根変換の近似の程度を信頼区間の被覆確率の観点から評価する。

本稿の構成は以下のようである。第1節で確率、オッズ、リスク比、オッズ比の定義を確認し、第2節でオッズ比とリスク比の関係を示す。第3節では、オッズ比の平方根変換とその簡単な性質、および信頼区間の被覆確率について述べ、第4節ではシミュレーションの手順とその結果を示す。最後の第5節で簡単なまとめを行う。参考文献では、本文中では陽に言及してはいないが関連する論文をいくつかまとめている。

参考のため、付録に VanderWeele (2017) と関連論文である Bland and Altman (2000) の邦訳を示す。

1. オッズ比とリスク比

1.1. 確率とオッズ、リスク比とオッズ比

ある物質への曝露の有無が何らかの事象の発生の有無に与える影響を考察する(例えば、喫煙が曝露、肺がん罹患が事象)。曝露の有無を表す変数を Z とし、事象の発生の有無を表す変数を Y として、曝露の有無ごとの事象の発生確率 (p) を次のように定義する(表 1.1 参照)。

曝露ありでの事象ありの条件付き確率: $p_1 = P(Y=1 | Z=1)$ (曝露リスク)

曝露なしでの事象ありの条件付き確率: $p_0 = P(Y=1 | Z=0)$ (ベースラインリスク)

表 1.1. 条件付き確率の定義

確率	事象あり ($Y=1$)	事象なし ($Y=0$)	計
曝露あり ($Z=1$)	p_1	$1 - p_1$	1
曝露なし ($Z=0$)	p_0	$1 - p_0$	1

これらの確率により、オッズ (*odds*) が次のように定義される。

$$\text{曝露ありでの事象ありのオッズ} : odds_1 = p_1 / (1 - p_1)$$

$$\text{曝露なしでの事象ありのオッズ} : odds_0 = p_0 / (1 - p_0)$$

この定義により、 $p > 0$ であれば $odds > p$ であることが分かる。また、対数オッズを $\log(odds)$ とする (\log は自然対数)。これらの取り得る値は、

$$0 \leq p \leq 1, \quad 0 \leq odds < \infty, \quad -\infty < \log(odds) < \infty$$

であること、および $p = 0.5$, $odds = 1$, $\log(odds) = 0$ が対応することに注意する。

曝露の有無ごとの事象の生起確率を比較するためのパラメータは以下のようである。

リスク比 (*risk ratio*) : $RR = p_1 / p_0$ 相対リスク (*relative risk*) ともいう

$$\text{オッズ比 (odds ratio)} : OR = \frac{odds_1}{odds_0} = \frac{p_1 / (1 - p_1)}{p_0 / (1 - p_0)} = \frac{p_1(1 - p_0)}{(1 - p_1)p_0}$$

対数オッズ比 (*log odds ratio*) : $\log(OR) = \log(odds_1 / odds_0) = \log(odds_1) - \log(odds_0)$

1.2. 観測データと研究の種類

観測度数のデータは表 1.2 のように与えられる。

表 1.2. 観測データ

データ	事象あり ($Y=1$)	事象なし ($Y=0$)	計
曝露あり ($Z=1$)	a	b	m
曝露なし ($Z=0$)	c	d	n
計	s	t	N

ここで、前向き研究 (*prospective study*) であるコホート研究 (*cohort study*) と後ろ向き研究 (*retrospective study*) である症例対照研究 (*case-control study*) の区別が重要である。コホート研究では、曝露の有無ごとのサンプルサイズ m および n は固定で、 a および c が観測され、 b と d は $b = m - a$, $d = n - c$ により求められる。標本リスク比 (RR_P^* と表記) および標本オッズ比 (OR_P^* と表記) は

$$(\text{前向き}) \text{ リスク比} : RR_P^* = \frac{a/m}{c/n}, \quad \text{オッズ比} : OR_P^* = \frac{a/b}{c/d} = \frac{ad}{bc}$$

となる。

一方、症例対照研究では、事象の有無ごとのサンプルサイズ s および t が固定で、 a および b が観測され、 c と d は $c = s - a$, $d = t - b$ により求められる。標本リスク比 (RR_R^* と表記) および標本オッズ比 (OR_R^* と表記) は

$$(\text{後ろ向き}) \text{ リスク比} : RR_R^* = \frac{a/s}{b/t}, \quad \text{オッズ比} : OR_R^* = \frac{a/c}{b/d} = \frac{ad}{bc}$$

となる。

前向きと後ろ向きのリスク比 RR_{p^*} と RR_R^* は異なるが、オッズ比 OR_{p^*} と OR_R^* は同じ値となる。そのため以降では、オッズ比では添え字を省略して OR^* と書く。

後ろ向き研究の症例対照研究であっても、前向きのリスク比 $RR = P(Y=1 | Z=1)/P(Y=1 | Z=0)$ が興味あるパラメータであることが多い。しかし、実際に観測されるのは後ろ向きの確率 $P(Z=1 | Y=1)$ および $P(Z=1 | Y=0)$ であるので、オッズ比を基にした議論にならざるを得ない。

2. オッズ比とリスク比の関係

医学研究では、発生確率の比であるリスク比が解釈上分かりやすいが、研究結果がオッズ比で示されることが多い（特にロジスティック回帰の出力として）。しかし、ベースラインリスク p_0 が小さくない場合に、オッズ比をリスク比と解釈するのは問題であるとの指摘がこれまでも多くなされている（付録の A.2 あるいは参考文献に挙げた諸論文を参照）。リスク比 (RR) とオッズ比 (OR) の関係は

$$RR = \frac{OR}{1 + p_0(OR - 1)} \quad (2.1)$$

もしくは

$$OR = \frac{(1 - p_0)RR}{1 - p_0RR} \quad (2.2)$$

である。ここで、(2.1) および (2.2) における RR の存在範囲には注意が必要となる。例えば、 $p_0 = 0.2$ であれば、 $0 \leq p_1 \leq 1$ より $0 \leq RR \leq 5$ となる。

関係式 (2.1) より、 $OR > 1$ であれば $OR \geq RR$ であることが分かる。また、オッズ比からリスク比を求めるためにはベースラインリスク p_0 の値が必要なことも見て取れる。(2.1) および (2.2) から、 p_0 が小さいときは $RR \approx OR$ となることは容易に分かるが、 p_0 がある程度大きいときはその限りではない。

図 2.1 は、式 (2.1) に基づき、 $OR=3$ のときにベースラインリスク p_0 と対応する RR の値を示している。また、図 2.2 は、式 (2.2) に基づいて、 $RR=2$ のときに p_0 と対応する OR の値のグラフである。両図において $p_0 = 0.25$ のときに $RR=2$ 、 $OR=3$ となっている。これらの図からも、 p_0 が大きくなるにつれ、 RR と OR の差が大きくなることが分かる。

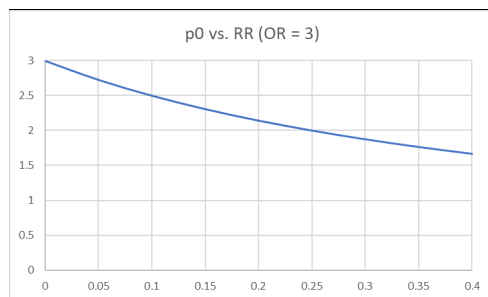


図 2.1. $OR=3$ のときの p_0 (横軸) と RR (縦軸) の関係

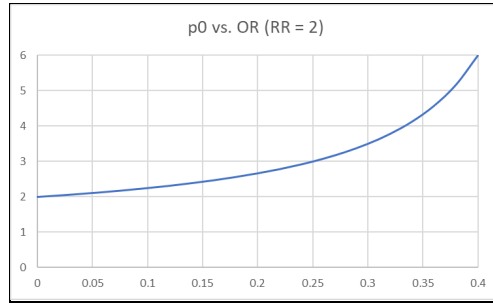


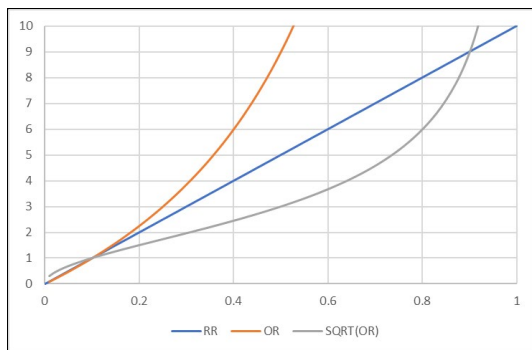
図 2.2. $RR=2$ のときの p_0 (横軸) と OR (縦軸) の関係

3. オッズ比の平方根変換とその評価

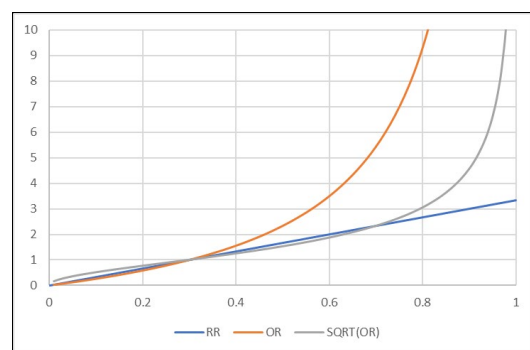
ベースラインリスク p_0 があまり小さくない場合にオッズ比をリスク比に近づける簡単な変換として, VanderWeele (2017) はオッズ比の平方根変換を提案した. また, VanderWeele (2020) は平方根変換を含むオッズ比の変換に加え, ハザード比の変換についても論じ, オッズ比の変換に関しては, ここでの平方根変換がある意味で最適であることを示している. ここではその平方根変換について論じる. なお, VanderWeele (2017) の邦訳を付録 A.1 に示す.

3.1. \sqrt{OR} と RR の数値的關係

オッズ比 OR の対数変換 \sqrt{OR} とリスク比 RR の関係については VanderWeele (2017) がいくつかの結果を示しているが, ここではそれを数式として表現する. まず, RR , OR , \sqrt{OR} のグラフを示す. 図 3.1 はベースラインリスクを $p_0=0.1$ もしくは 0.3 と固定して p_1 を動かした場合の RR , OR , \sqrt{OR} のグラフである. VanderWeele (2017) が言うように, $p_0 = 0.3$ とベースラインリスクがあまり小さくない場合には OR は RR に比べてかなり大きな値となるのに対し \sqrt{OR} は RR にかかなり近くなっている. また, これも VanderWeele (2017) が指摘するように, p_0 と p_1 の値の組み合わせによっては $\sqrt{OR} < RR$ となり得るが, 両者の乖離幅はかなり小さくなっている.



(a) $p_0=0.1$



(b) $p_0=0.3$

図 3.1. p_0 の値を定めた時の RR , OR , \sqrt{OR} のグラフ (横軸: p_1)

\sqrt{OR} と RR の比を取ることに

$$\frac{\sqrt{OR}}{RR} = \frac{\sqrt{\{p_1/(1-p_1)\}/\{p_0/(1-p_0)\}}}{p_1/p_0} = \sqrt{\frac{p_0(1-p_0)}{p_1(1-p_1)}} \quad (3.1)$$

と簡潔な関係式を得る。また、オッズ比 OR とリスク比 RR の比は

$$\frac{OR}{RR} = \frac{\{p_1/(1-p_1)\}/\{p_0/(1-p_0)\}}{p_1/p_0} = \frac{1-p_0}{1-p_1} \quad (3.2)$$

である。図 3.2 は、 p_0 が与えられたとして、 p_1 を動かしたときの比 (3.1) および (3.2) のグラフである。

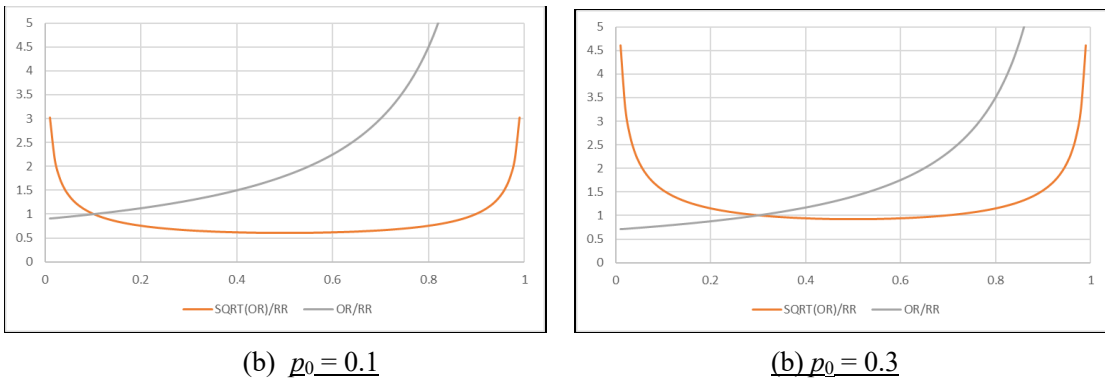


図 3.2. p_0 の値を定めたときの比 \sqrt{OR}/RR と OR/RR のグラフ (横軸: p_1)

関係式 (3.1) からいくつかのことが分かる。まず $\sqrt{OR} \leq RR$ として式を変形すると

$$p_0 \leq p_1 \leq 1 - p_0 \quad (3.3)$$

を得る。すなわち、 p_0 が与えられたとき、 p_1 が (3.3) の範囲では $\sqrt{OR} \leq RR$ となる。ちなみに、等号条件のうち $p_1 = p_0$ のときは $OR = RR = 1$ であり、 $p_1 = 1 - p_0$ では $p_0 + p_1 = 1$ 、すなわち、 w を 0.5 以下の正の数として、 $p_0 = 0.5 - w$ 、 $p_1 = 0.5 + w$ の場合で、これが VanderWeele (2017) が OR の平方根変換を推奨した根拠となっている。

図 3.1 および 3.2 で見て取れるように、(3.3) の範囲では $\sqrt{OR} \leq RR$ となるが、比 \sqrt{OR}/RR は 1 よりどの程度小さくなるのかを考察する。式 (3.1) より、 \sqrt{OR}/RR が最も小さくなるのは分母 $p_1(1-p_1)$ が最も大きくなるときであり、それは $p_1 = 0.5$ によって与えられる。そしてそのときの比は、(3.1) で $p_1 = 0.5$ を代入して $2\sqrt{p_0(1-p_0)}$ となる。 $p_0 =$

0.1 のときは $2\sqrt{p_0(1-p_0)} = 0.6$ 、 $p_0 = 0.3$ のときは $2\sqrt{p_0(1-p_0)} \approx 0.92$ であり、1 からの乖離幅は大きくはないことが数値的に立証される。

3.2. 信頼区間の被覆確率の比較

3.1 項では, RR に対する \sqrt{OR} の数値的な近さが示されているが, その程度はともかくとして \sqrt{OR} が RR を過小評価してしまうというのは気になる点である. また, 統計的観点として, RR の推定において \sqrt{OR} がどの程度有用であるかを信頼区間の被覆確率によって評価する.

観測度数が表 1.2 のように与えられるとき, OR^* を標本オッズ比とすると, $\log(OR^*)$ の標準誤差 (SE) は, 近似的に

$$SE = SE[\log(OR^*)] \approx \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}} \quad (3.4)$$

で与えられる. ここで $\log(OR^*)$ の正規近似を用いると, $\log(OR)$ の近似的な 95%信頼区間は

$$(C, D) = (\log(OR^*) - 1.96SE, \log(OR^*) + 1.96SE)$$

となる. よって, これを指数変換して, OR の 95%信頼区間が

$$(\exp(C), \exp(D)) \quad (3.5)$$

と得られる.

平方根変換した \sqrt{OR} に関しては, $\log(\sqrt{OR^*}) = 0.5 \cdot \log(OR^*)$ であるので, SE を (3.4)

の値として, $SE[\log(\sqrt{OR^*})] = 0.5 \cdot SE$ となる. これより, $\log(\sqrt{OR})$ の 95%信頼区間は

$$(E, F) = (0.5 \log(OR^*) - 1.96 \cdot 0.5 \cdot SE, 0.5 \log(OR^*) + 1.96 \cdot 0.5 \cdot SE)$$

なり, \sqrt{OR} の 95%信頼区間は, 指数変換により

$$(\exp(E), \exp(F)) \quad (3.6)$$

となる.

\sqrt{OR} が RR をうまく近似しているのであれば, \sqrt{OR} の信頼区間 (3.6) が RR を中に含む確率が大きくなるであろう. そしてそれは, OR^* に基づく信頼区間 (3.5) が RR を中に含む確率に比べてどの程度大きいのであるのかが気になる点である.

そこで, 95%信頼区間に関し, 次の3つの確率 (被覆確率 coverage probability) をシミュレーションによって求める.

(a) $(\exp(C), \exp(D))$ が OR を含む確率

(b) $(\exp(C), \exp(D))$ が RR を含む確率

(c) $(\exp(E), \exp(F))$ が RR を含む確率

これらの確率のうち (a) は, OR の近似的な信頼区間 (3.5) の被覆確率が信頼度 (95%) に近いかどうかの確認である. そして (b) は OR が RR の近似になり得るかの評価, (c) は \sqrt{OR} が RR の近似になり得るかの評価である.

4. シミュレーションの手順と結果

4.1. シミュレーション手順

シミュレーションは以下の手順で行う。

- (i) 確率 p_0 , p_1 およびサンプルサイズ m , n を定める。そして、真値を $RR = p_1/p_0$, $OR = \{p_0/(1-p_0)\}/\{p_1/(1-p_1)\}$ とする。
- (ii) 二項分布 $B(m, p_0)$ に従う乱数 a および $B(n, p_1)$ に従う乱数 c を生成し, $b = m - a$, $d = n - c$ を求める (表 1.2 の分割表を 1 つ生成する)。
- (iii) オッズ比 $OR^* = ad/bc$ を計算し, (3.5) および (3.6) を用いて信頼区間 ($\exp(C)$, $\exp(D)$) および ($\exp(E)$, $\exp(F)$) を求める。
- (iv) 3.2 項の (a), (b), (c) に対応し,
 - (a) ($\exp(C)$, $\exp(D)$) が OR を含むかどうかを調べる
 - (b) ($\exp(C)$, $\exp(D)$) が RR を含むかどうかを調べる
 - (c) ($\exp(E)$, $\exp(F)$) が RR を含むかどうかを調べる

以上の (i)~(iv) を M 回繰り返す, (a), (b), (c) の割合 (被覆確率の推定値) を計算する。

4.2. シミュレーション結果

ここでは, $m = n = 100$ と固定し, $p_0 = 0.1$ (0.1) 0.9, $p_1 = 0.1$ (0.1) 0.9 と動かして, $M = 1000$ 回繰り返す。結果を表 4.1 から表 4.4 に示していて, いずれも 1000 回の繰り返し中の頻度である。

表 4.1 は区間 ($\exp(C)$, $\exp(D)$) が真値 OR を含んだ頻度である。表 4.2 は区間 ($\exp(C)$, $\exp(D)$) が RR を含んだ頻度で, 被覆確率 90%以上を青く, 50%未満を赤く塗ってある。表 4.3 は区間 ($\exp(E)$, $\exp(F)$) が RR を含んだ頻度で, ここでも被覆確率 90%以上を青く, 50%未満を赤く塗ってある。そして表 4.4 は表 4.3 の頻度から表 4.2 の頻度を引いたもので, 差がマイナスとなったセルを黄色く塗った。

これらの表から次のことが読み取れる。

- (1) 表 4.1 より, OR に基づく近似的な信頼区間は概ね信頼度 (95%) を達成している。ただしここではサンプルサイズが $m = n = 100$ と比較的大きいので $\log(OR)$ の正規近似がうまく働いているが, もう少し小さなサンプルサイズではどうかという問題は残る。
- (2) 表 4.2 より, ベースラインリスク p_0 および曝露リスク p_1 が小さいときは (概ね 0.2 以下), OR は RR の良い近似となるため, OR に基づく信頼区間が RR を含む確率は概ね信頼度程度となる。しかし, p_0 が小さくても (0.1 程度でも) p_1 が大きいときはその限りではない。
- (3) 表 4.3 より, p_0 が小さい場合には, 平方根変換はうまく機能しない。しかし, p_0 および p_1 が中程度 (0.3 ~ 0.7) の場合には, OR の平方根変換に基づく信頼区間が RR を含む確率は 90%を超え, 比較的うまく機能しているといえる。

- (4) 表 4.4 の OR そのものとその平方根変換 \sqrt{OR} の比較では, p_0 および p_1 が小さい場合には OR そのものがよく, p_0 および p_1 がある程度大きい場合には平方根変換がよくなるといえる. ただし, 上述の (2) および (3) で示したように, どちらも悪いという場合もあるので注意が必要である.

5. おわりに

リスク比 RR をオッズ比 OR が近似し得るか, また OR の平方根変換 \sqrt{OR} が RR の近似となり得るかを考察した. 特に, それらの数値的な近さに加え, ここでは, 信頼区間の被覆確率の観点から評価した. ベースラインリスク p_0 および曝露リスク p_1 が小さい場合には, OR は RR の良い近似となることは, 被覆確率の見地からも確かめられたが, p_0 が小さくても p_1 が大きい場合にはその限りでないことが分かった.

また, p_0 がある程度大きいときは OR は RR の近似とは見做せないことが, 数値的検討ならびに被覆確率の観点からも確かめられた. OR の平方根変換は, p_0 が小さいときにはうまく機能しないが, p_0 および p_1 がある程度大きい ($0.3 \sim 0.7$) ときはそれなりに機能することが確かめられた.

本稿は, VanderWeele (2017) の数値的検討を, 信頼区間の被覆確率の観点から評価したもので, OR の平方根変換の妥当性をより詳しく見ている.

表 4.1. (exp(C), exp(D)) が OR を含む頻度

p0 p1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	962	972	948	959	959	960	945	962	954
0.2	957	949	955	948	946	952	955	947	953
0.3	952	953	939	944	942	955	947	958	954
0.4	958	959	943	933	956	945	958	959	960
0.5	955	955	952	947	940	953	954	947	954
0.6	972	946	953	949	940	944	939	953	963
0.7	945	962	945	948	948	947	961	951	953
0.8	956	952	964	932	943	946	953	951	969
0.9	960	956	957	957	961	968	950	964	956

表 4.2. (exp(C), exp(D)) が RR を含む頻度

p0 p1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	962	957	918	826	666	387	157	13	0
0.2	946	949	932	865	683	440	112	13	0
0.3	904	934	939	909	780	526	177	18	0
0.4	852	871	905	933	903	703	382	50	0
0.5	669	701	808	907	940	881	592	154	4
0.6	410	417	527	723	904	944	839	393	28
0.7	155	141	196	324	590	853	961	745	181
0.8	17	10	9	65	145	427	757	951	636
0.9	0	0	0	0	4	19	187	625	956

表 4.3. (exp(E), exp(F)) が RR を含む頻度

p0 p1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	962	710	476	347	355	365	474	719	954
0.2	700	949	881	746	738	739	885	947	680
0.3	455	858	939	927	907	922	947	876	383
0.4	358	763	916	933	946	945	936	756	221
0.5	348	719	918	942	940	952	923	740	188
0.6	355	751	924	949	941	944	931	760	242
0.7	482	861	945	914	907	920	961	860	397
0.8	706	952	883	780	720	781	871	951	746
0.9	960	722	398	240	190	244	410	721	956

表 4.4. 表 4.3 の頻度から表 4.2 の頻度を引いたもの

p0 p1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0	-247	-442	-479	-311	-22	317	706	954
0.2	-246	0	-51	-119	55	299	773	934	680
0.3	-449	-76	0	18	127	396	770	858	383
0.4	-494	-108	11	0	43	242	554	706	221
0.5	-321	18	110	35	0	71	331	586	184
0.6	-55	334	397	226	37	0	92	367	214
0.7	327	720	749	590	317	67	0	115	216
0.8	689	942	874	715	575	354	114	0	110
0.9	960	722	398	240	186	225	223	96	0

参考文献

- Atare, J. and Maucort-Boulch, D. (2016). Odds ratio, hazard ratio and relative risk. *Metodoloski Zvezki*, **13**, 59-67.
- Bland, J. M. and Altman, D. G. (2000). The odds ratio. *BMJ*, **320**, 1468.
- Brumback, B. and Berg, A. (2008). On effect-measure modification: Relationships among changes in the relative risk, odds ratio, and risk difference. *Statistics in Medicine*, **27**, 3453-3465.
- Chinn, S. (2000). A simple method for converting odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, **19**, 3127-3131.
- Glas, A. S., Lijmer, J. G., Prins, M. H., Bonsel, G. J. and Bossuyt, P. M. M. (2003). The diagnostic odds ratio: a single indicator of test performance. *Journal of Clinical Epidemiology*, **56**, 1129-1135.
- Grant, R. L. (2014). Converting an odds ratio to a range of plausible relative risks for better communication of research findings. *BMJ*, **348**, f7450.
- Knol, M. J., Le Cessie, S., Alga, A., Vandenbroucke, J. P. and Groenwold, R. H. H. (2012). Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. *Canadian Medical Association Journal*, **184**, 895-899.
- Kraemer, H. C. (2004). Reconsidering the odds ratio as a measure of 2×2 association in a population. *Statistics in Medicine*, **23**, 257-270.
- Kraemer, H. C. (2006). Correlation coefficients in medical research: from product moment correlation to the odds ratio. *Statistical Methods in Medical Research*, **15**, 525-545.
- Newcombe, R. G. (2006). A deficiency of the odds ratio as a measure of effect size. *Statistics in Medicine*, **25**, 4235-4240.
- Nurminen, M. (1995). To use or not to use the odds ratio in epidemiologic analyses? *European Journal of Epidemiology*, **11**, 365-371.
- Pepe, M. S., Janes, H., Lomgton, G., Leisenring, W. and Newcomb, P. (2004). Limitations of the odds ratio in gauging the performance of a diagnostic prognostic, or screening marker. *American Journal of Epidemiology*, **159**, 882-890.
- Simon, S. D. (2001). Understanding the odds ratio and relative risk. *Journal of Andrology*, **22**, 533-536.
- VanderWeele, T. (2017). On a square-root transformation of the odds ratio for a common outcome. *Epidemiology*, **28**, e58-e60.
- VanderWeele, T. J. (2020). Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. *Biometrics*, **76**, 746-752.
- Zhang, J. and Yu, K. F. (1998). What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA*, **280**, 1690-1691.

付録 (論文邦訳)

A.1 VanderWeele (2017)

コホートデータの分析において、発生が稀ではない一般的なアウトカムでは、曝露とそのアウトカムに関連するオッズ比 (OR) の大きさが、対応するリスク比 (RR) を大幅に上回ることはよく知られている。アウトカムが稀な場合 (10%がしばしばカットオフ値として用いられる)、 OR は RR に精度よく近似するため、しばしば RR として解釈される。しかし、アウトカムの発生確率がある程度大きな場合、 OR を RR として解釈すると、 RR を大幅に誇張することになり、決して妥当な使い方とは言えない。ロジスティック回帰は多変量解析のツールとして使用されることが多く、アウトカムが一般の場合であっても OR による結果報告は日常的に行われているのである。共変量調整を考慮することによりそれらの一般的なアウトカムに対する RR を推定する方法が数多く開発されているが^{1,2}、これらの方法の使用頻度は依然として低い²。一般的なアウトカムに対する OR の報告は、臨床医学の文献では依然として頻繁に行われている。このような状況では、 OR の大きさを直感的に理解することはより困難である。このレターは、一般的な結果に対する OR の単純な変換法を提案するもので、多くの状況において、はるかに RR に近い値をもたらす。このレターの目的は、一般的なアウトカムに対する RR の推定のための方法を使用すべきではないとするものではなく、むしろ一般的なアウトカムに対する OR が実際に論文で報告されている場合に、その解釈を支援することを意図するものである。

提案する変換は単純で、単に OR の推定値の平方根を取ることである。したがって、 RR の近似値として、 OR の2は1.41になり、 OR の4は2になり、 OR の9は3になり、といった具合になる。この変換の動機を簡単に説明し、次に RR をより近似する値としての性能に関連するいくつかの特性を議論する。

まず、曝露者のアウトカムの発生確率が0.5より w だけ大きく、非曝露者のアウトカムの確率が0.5より同じ値 w だけ小さく、曝露者と非曝露者の確率がそれぞれ $p_1 = 0.5 + w$ と $p_0 = 0.5 - w$ であるという設定を考えよう。このとき、 $RR = p_1/p_0 = (0.5 + w)/(0.5 - w)$ となり、 $OR = \frac{p_1(1-p_0)}{(1-p_1)p_0} = \frac{(0.5+w)^2}{(0.5-w)^2}$ となる。すなわち、 OR は RR のちょうど2乗であり、

OR の平方根をとると RR が復元される。

この同じ変換が、アウトカムが一般の場合、アウトカム確率のほとんどの値で驚くほどよく働くことが分かる。

まず、 $p_1 > p_0$ となるような因果的な曝露から始めよう。まず、 p_0 と p_1 がともに0.2以上0.8以下であるとする。この場合、 OR は最大で400%も増大するが (例えば、 $p_0=0.2, p_1=0.8, RR=4, OR=16$)、 RR より \sqrt{OR} が増大するのは25%である (例えば、 $p_0=0.5, p_1=0.8, RR=1.6, \sqrt{OR}=2$)。アウトカム確率 p_0 と p_1 が0.2~0.8であれば、 OR の平方根の RR からの乖離は最大で25%である。

一方、 p_0 と p_1 の両方が0.1から0.9の間であれば、 OR の増大分は900%であるが (た

たとえば, $p_0 = 0.1$, $p_1 = 0.9$, $RR = 9$, $OR = 81$), オッズ比の平方根は RR に対してせいぜい 67%の増大にとどまる (たとえば, $p_0 = 0.5$, $p_1 = 0.9$, $RR = 1.8$, $\text{sqrt}(OR) = 3$). 平方根変換は, 数値の増大を劇的に減少させ, 上記のように曝露リスクと非曝露リスクが平均して 0.5 である場合, 変換によりバイアスは完全に消滅する. アウトカムの確率が 0.9 を超えると, さらに大きなインフレーションが起こり得るが, 平方根変換は RR の近似値としてまだ改善をもたらす.

平方根変換は, 實際上, 常に RR に向かって OR を減少させる. しかし状況によっては, $\text{sqrt}(OR)$ が RR より小さくなるように過修正することもある (たとえば, $p_0 = 0.3$, $p_1 = 0.5$, $RR = 1.67$, $\text{sqrt}(OR) = 1.52$) しかし再び, p_0 と p_1 が 0.2 から 0.8 の間ならば, 減少幅は最大で 1/1.25 倍 (つまり 20%減) であり, p_0 と p_1 が 0.1~0.9 では, 最大で 1/1.67 倍 (つまり 40%減) となっている. $\text{sqrt}(OR)$ が RR よりも小さくなるこれらの状況においてさえ, $\text{sqrt}(OR)$ が RR よりも減少する程度は, 大部分の設定において, OR が RR よりも過大となる程度より小さくなるであろう. 図 1 では, 両方の確率が 0.1 以上であるとき, このような結果が得られる確率の値が黒い領域としてプロットされている. 両方の確率が 0.1 以上のとき, 約 93%の確率で, OR の増大率が $\text{sqrt}(OR)$ の減少率を上回っている. 両方の確率が 0.2 以上の場合, 可能なアウトカム確率の 99%でそうなっている. 両方の確率が 0.25 以上の場合, 常にそれが成り立つ. $p_1 < p_0$ の曝露についても, 上記のすべての主張と同様の記述が成り立つ.

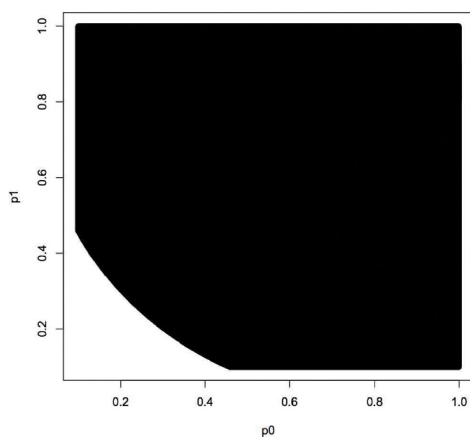


図 1

比率の尺度は, 公衆衛生上の重要性の尺度を得る目的で, 過剰相対リスク尺度に変換されることがある^{3,4}. この目的のためには, RR と OR の比率や $\text{sqrt}(OR)$ ではなく, これらの量間の差異が問題となる. ここでもまた, 大多数の状況で平方根変換が優れている. それは常に OR を RR に向かって減少させる. 時には減少しすぎることもあるが, その場合でも, ほとんどの場合, 絶対差 $|\text{sqrt}(OR) - RR|$ は絶対差 $|OR - RR|$ より小さい. 因果的な曝露の場合, 両方の確率が 0.2 から 0.8 のとき, OR の絶対差は 12 と大きくなることがあるが, $\text{sqrt}(OR)$ では 0.55 大きくなるのみである. 両方のアウトカム確率が 0.1 から 0.9 のとき, OR

の絶対差は 72 となることがあるが \sqrt{OR} では 2.43 大きくなるだけである。アウトカム確率が 0.1 から 0.3 のときは、 \sqrt{OR} は 0.5 ほど大きくなる。原因となる曝露については、両方のアウトカム確率が 0.1 以上であれば 95%の確率で、両方のアウトカム確率が 0.2 以上であれば 99%の確率で平方根変換の方が差の絶対値が小さくなる。保護的曝露 $p_1 < p_0$ では、平方根変換は、両方のアウトカム確率が 0.1 より大きい場合、90%ではより小さい絶対差となり、両方のアウトカム確率が 0.2 より大きい場合、98%では絶対差はより小さくなる。

ここでも、平方根変換はほとんどすべてのシナリオで RR に近く、 RR のある程度妥当な近似値を提供する。経験則として、アウトカムの有病率が 20%以上の場合、平方根近似が望ましいことが示唆されるであろう。したがってこの変換は、ランダム化試験、コホート、横断データ、または累積サンプリングによるケースコントロール・データで使用可能であろう³。しかし、罹患密度サンプリングによるケースコントロール研究では、罹患率比の直接推定値が得られるので、罹患率比と比例ハザードモデルに関するさらなる議論は、eAppendix に記載されている。ここで提案する変換は、メタアナリシスの結果を解釈する際にも有用であると思われる。メタアナリシスでは、標準化効果量と対数オッズ比の間で一般に近似的な変換が行われ^{5,6}。メタアナリシスで使用される変換は、アウトカムが一般である場合にのみ適用でき、アウトカム確率が非常に小さいか非常に大きい場合にはうまく機能しない⁷。繰り返しになるが、このレターの目的は、一般のアウトカムについて RR を推定する方法を置き換えることではなく、むしろ文献ですでに報告されている一般的なアウトカムについての OR 推定値の解釈を助けることである。

参考文献

1. Knol MJ, le Cessie S, Algra A, Vandenbroucke JP, Groenwold RHH. Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. *Canadian Medical Association Journal*. 2012; 184:895–899. [PubMed: 22158397]
2. Yelland LN, Salter AB, Ryan P. Relative risk estimation in randomized controlled trials: a comparison of methods for independent observations. *International Journal of Biostatistics*. 2011; 7(1)
3. Rothman, KJ., Greenland, S., Lash, TL. *Modern Epidemiology*. 3. Lippincott; 2008.
4. VanderWeele, TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press; New York: 2015.
5. Borenstein, M., Hedges, LV., Higgins, JPT., Rothstein, HR. *Introduction to Meta-Analysis*. 1. Wiley; 2009.
6. Hasselband V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psychological Bulletin*. 1995; 117:167–178. [PubMed: 7870860]
7. Anzures-Cabrera J, Sarpatwari A, Higgins JPT. Expressing findings from meta-analyses of continuous outcomes in terms of risks. *Statistics in Medicine*. 2011; 30:2967–2985.

A.2. Bland and Altman (2000)

近年、医学論文ではオッズ比が広く使われるようになり、BMJ 誌にもいくつか研究結果が掲載されているはずである。その理由は3つある。まず、2つの2値変数（「はい」か「いいえ」）の間の関係の推定値（信頼区間付き）を提供することである。次に、ロジスティック回帰を用いて、その関係に他の変数が及ぼす影響を調べることができる。第3に、それらは症例対照研究において特別でかつ非常に便利な解釈が可能となる。

オッズは確率を表現する方法で、特に賭け事によく知られている。たとえば、サイコロを1回投げて6が出るオッズは1対5つまり1/5である。オッズとは、対象となる事象が発生する確率と、発生しない確率の比のことである。これは、関心のある事象が発生する回数と発生しない回数の比によって推定されることが多い。表1は、11歳の子供における花粉症 (hay fever) と湿疹 (eczema) の有病率を示す横断調査のデータである¹。湿疹のある子供が花粉症も持つ確率は比 141/561 (25.1%) で推定される。オッズは 141/420 と推定される。同様に、湿疹のない子供が花粉症になる確率は 928/14453 (6.4%) で、オッズは 928/13525 と推定される。比率の差 $141/561 - 928/14453 = 0.187$ (18.7%)、比率の比 $(141/561)/(928/14453) = 3.91$ (相対リスクともいう)、オッズ比 $(141/420)/(928/13525) = 4.89$ によって両群を比較することができる。

表1 湿疹の有無と花粉症の有無の関連性 (11歳の子供)

観測度数	花粉症あり (Y=1)	花粉症なし (Y=0)	計
湿疹あり (Z=1)	a = 141	b = 420	m = 561
湿疹なし (Z=0)	c = 928	d = 13535	n = 14453
計	s = 1069	t = 13945	N = 15522

では、逆に、花粉症の子供が湿疹を併発する確率はどのくらいかと考えてみよう。比率は 141/1069 (13.2%)、確率は 141/928 となる。花粉症でない子供の場合、湿疹がある比率は 420/13945 (3.0%)、確率は 420/13525 である。これらの比率を比較すると、差は $141/1069 - 420/13945 = 0.102$ (10.2%)、比 (相対リスク) は $(141/1069)/(420/13945) = 4.38$ 、オッズ比は $(141/928)/(420/13525) = 4.89$ である。オッズ比は表のどちらを見ても同じだが、差と比は異なる。その理由は簡単にわかる。

この2つのオッズ比は

$$\frac{141/420}{928/13525}, \frac{141/928}{420/13535}$$

であり、両者の分母分子を並べ替えるといずれも

$$\frac{141 \times 13525}{928 \times 420}$$

となる。オッズ比は、行と列のカテゴリの順番を入れ替えても、同じ値となる。行だけあ

るいは列だけの順番を入れ替えると、オッズ比は逆数 $1/4.89=0.204$ となる。これらの性質から、オッズ比は関係の強さを示す有用な指標となる。

標本オッズ比は負になることがなく、下限は限定されるが上限は限定されないで、歪んだ分布になる。しかし、対数オッズ比は任意の値をとることができ²、ほぼ正規分布となる。また、変数の1つのカテゴリーの順序を逆にすると、単純に対数オッズ比の符号が逆になるという便利な性質がある：すなわち $\log(4.89)=1.59, \log(0.204)=-1.59$ である。

対数オッズ比の標準誤差により信頼区間を計算することができる。対数オッズ比の標準誤差は、4つの度数の逆数の和の平方根で簡単に推定できる。ここでの例の場合

$$SE[\log OR] = \sqrt{\frac{1}{141} + \frac{1}{420} + \frac{1}{928} + \frac{1}{13535}} \approx 0.103$$

となる。

対数オッズ比の95%信頼区間は、推定値の両側の標準誤差を1.96倍して求められる。この例では、対数オッズ比は $\log_e(4.89)=1.588$ であり、信頼区間は $1.588 \pm 1.96 \times 0.103$ の計算から1.386から1.790になる。これらの限界値を指数関数化して、オッズ比自体の95%信頼区間を、 $\exp(1.386)=4.00$ から $\exp(1.790)=5.99$ として与えることができる²。観察されたオッズ比である4.89は、オッズ比の尺度が非対称であるために、信頼区間の中央にはない。このため、グラフでは、オッズ比はしばしば対数尺度でプロットされる。2つの変量間に関係がないときオッズ比は1である。オッズ比が1であるという帰無仮説は、 2×2 の分割表に対する通常のカイ2乗検定で検定することができる。

オッズ比はその有用性にもかかわらず、解釈の難しさを引き起こすことがある³。この議論を見直し、またロジスティック回帰や症例対照研究でのオッズ比についても今後議論していきたい。

データを提供していただいた Barbara Butland に感謝する。

参考文献

1. Strachan DP, Butland BK, Anderson HR. Incidence and prognosis of asthma and wheezing illness from early childhood to age 33 in a national British cohort. *BMJ*. 1996; 312: 1195-9.
2. Bland JM, Altman DG. Transforming data. *BMJ* 1996; 312: 770.
3. Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! *Evidence-Based Medicine*, 1996; 1: 164-6.