

ご提示いただいた内容は、**Sander Greenland** 氏による論文からの抜粋です。一つは「**Noncollapsibility, confounding, and sparse-data bias. Part 1: The oddities of odds**」、もう一つは「**Noncollapsibility, confounding, and sparse-data bias. Part 2: What should researchers make of persistent controversies about the odds ratio?**」というタイトルの論文です。

これらの論文は、統計的な誤解を防ぐために、統計的検定ではなく推定に焦点を当てるという長年の助言に基づき、アウトカムや効果測定に関する議論を扱っています。特に、オッズやその対数に基づいた尺度が持つ望ましくない性質である **noncollapsibility** に焦点を当てています。

noncollapsibility とは、ある集団で測定された尺度が、その集団の各々人やサブグループで測定された尺度の単純平均と等しくない性質と定義されています。この性質は、リスクの要約やコミュニケーションにとって望ましくありません。

Table 1. Numeric example of odds noncollapsibility

	Male	Female	Total
Died	45	30	75
Survived	5	70	75
Total	50	100	150
Proportion (risk)	0.90	0.30	0.50
Odds	$9/1 = 9.00$	$3/7 = 0.43$	$1/1 = 1$

Part 1 では、この問題がオッズそのものでどのように生じるか、基本的な数値例を用いて説明されています。

例として、死亡または生存という二値のアウトカムを持つ150人の患者の集団が挙げられています。この集団全体の死亡のオッズは $75/75 = 1$ です。もし、この集団を男性50人、女性100人に分け、男性では45人死亡、5人生存、女性では30人死亡、70人生存だったとします。

男性サブグループの死亡オッズは $45/5 = 9/1 = 9$ となります。女性サブグループの死亡オッズは $30/70 = 3/7 \approx 0.43$ となります。

ここで、もし各個人に所属するサブグループのオッズ（男性なら9、女性なら $3/7$ ）を割り当て、これらの個人オッズを単純平均すると、 $\{(50 * 9) + (100 * 3/7)\} / 150 \approx 3.3$ となります。これは、サブグループ化せずに計算した集団全体のオッズである1とは異なります。この不一致が、オッズが個人の単純平均に対して **noncollapsible** であるという性質を示しています。

この数値例からいくつかの重要な点が示されています。

1. **noncollapsibility** は因果的な現象ではなく、基本的な算術的な性質です。上記の例では因果効果は測定されていませんが、集団全体のオッズが個人の平均オッズと等しくならないことが分かります。これは、たとえオッズの変動がサブグループ化変数によって引き起こされたものでなくても起こり得ます。
2. **noncollapsibility** は、任意の大きさのデータセットでも生じ得ます。Table 1の数値を10倍しても、比率、オッズ、および **noncollapsibility** の程度は同じままです。
3. **noncollapsibility** はオッズの大きさに直接比例します。オッズが非常に低い場合、それらはほぼ比率に等しくなり、**noncollapsibility** の程度は限定されます。しかし、いずれかのサブグループでオッズが高い場合、オッズは比率から乖離し、オッズの **noncollapsibility** が生じる可能性があります。Table 1の死亡数を変えずに患者総数を10倍にした例では、死亡比率は0.090と0.030になり、オッズは0.099と0.031となり、オッズは比率に近くなり、**noncollapsibility** の程度は小さくなります。
4. アウトカムをさらに予測する共変量でデータを細分化し続けると、オッズがさらに高くなるサブグループが現れ、**noncollapsibility** はますます大きくなります。最終的には、メンバーが1人だけのサブグループになり、オッズは0または未定義（あるいは無限大）になります。

Table 1. Numeric example of odds-ratio noncollapsibility without confounding

	Males		Females		Collapsed	
	Treated	Untreated	Treated	Untreated	Treated	Untreated
Died	45	30	30	10	75	40
Survived	5	20	70	90	75	110
Totals	50	50	100	100	150	150
Proportions (risks)	0.90	0.60	0.30	0.10	0.500	0.267
Odds	9/1	3/2	3/7	1/9	1/1	4/11
Odds ratios	6.0		3.9		2.75	
Risk ratios	1.5		3.0		1.875	
Risk differences	0.30		0.20		0.233	

Part 2では、このnoncollapsibilityの問題がオッズ比でどのように増幅されるかが説明されています。

noncollapsibilityは因果的現象ではないのに対し、交絡は効果の比較に関わるものであり、適切に定義・説明するには因果的フレームワークが必要です。交絡は、ベースラインの共変量効果が治療効果と混同されることで生じます。交絡が生じるためには、ベースライン共変量がアウトカム自体に影響を持ち、治療と関連している必要があります。

noncollapsibilityと交絡の違いを示すために、Part 2では、治療群と非治療群で共変量（性別）の分布が一致している（交絡がない）状況での数値例が示されています。Table 1（Part 2ではTable 2）に示されているように、治療群と非治療群それぞれで男性50人、女性100人です。男性の死亡率は治療群0.90、非治療群0.60、女性の死亡率は治療群0.30、非治療群0.10です。男性サブグループのオッズ比は $(9/1) / (3/2) = 6.0$ です。女性サブグループのオッズ比は $(3/7) / (1/9) = 2.75$ です。集団全体（男女合計、collapsed）のオッズ比は $(1/1) / (4/11) = 2.75$ です。

この例の重要な特徴は、collapsedオッズ比 (2.75) が、男性サブグループのオッズ比 (6.0) と女性サブグループのオッズ比 (2.75) のどちらとも等しくないことです。実際、collapsedオッズ比は女性のオッズ比と偶然一致していますが、男性のオッズ比とは大きく異なります。これは、collapsedオッズ比がサブグループのオッズ比のどのような平均でもないことを示しています。この、性別を条件付けた場合（サブグループ内）に見られるものと、周辺的（無条件）に見られるものの質の差は、治療群と非治療群の間の性別比の不均衡に起因するものではありません。性別はこれらの患者にとって強い死亡リスク要因であり、治療によって影響を受けませんが、性別比のバランスが取れているため、性別による交絡とはなり得ません。

これにより、noncollapsibilityが交絡なしで生じるという見かけ上のパラドックスが生まれます。治療は男性および女性のサブグループの両方で死亡のオッズを3倍以上にしているにもかかわらず、集団全体を考慮すると死亡のオッズを3倍にするには至りません。この現象は、治療群と非治療群で、男性（オッズ比が高いが高死亡率であるため生存者の割合が低い）と女性（オッズ比が低いが生生存者の割合が高い）の重みが異なることで生じます。具体的には、集団全体の治療群のオッズ1/1は、男性と女性のオッズを生生存者数で重み付けした平均です: $\{5(9/1) + 70(3/7)\} / (5 + 70) = (45 + 30) / 75 = 1$ 。ここで男性には $5/75 = 6.7\%$ の重みしか与えられていません。対照的に、集団全体の非治療群のオッズ4/11は、男性と女性のオッズを生生存者数で重み付けした平均です: $\{20(3/2) + 90(1/9)\} / (20 + 90) = (30 + 10) / 110 = 4/11 \approx 0.36$ 。ここでは男性により多くの重み、 $20/110 = 18.2\%$ が与えられています。オッズ比が高い男性が、非治療群のオッズよりも治療群のオッズでより低い割合の生存者を持っているため、collapsedオッズ比は男性および女性のオッズ比の両方よりも小さくなる結果となります。これは、治療が重みに影響を与えることで生じるため、前治療期（ベースライン）における治療群間の違いによって引き起こされる交絡とは異なります。

ハザード（発生）率差および率比も、オッズ比ほどではないものの、交絡なしでnoncollapsibilityを生じることがあります。このようなオッズベースおよびハザードベースの尺度のodd behaviorは、治療や共変量が生存に影響を与える場合には常に予想されるべきものです。対照的に、リスク差およびリスク比は、交絡がない場合にはcollapsibleです。Table 1の例で示されているように、collapsedリスク差はサブグループサイズで重み付けされた共変量別のリスク差の平均であり、collapsedリスク比は各非治療サブグループの死亡数で重み付けされた共変量別のリスク比の平均です。

noncollapsibilityは、アウトカム予測因子で調整する際に、その予測因子が交絡因子や効果修飾因子でない場合でも変化し、輸送性、研究間の比較不可能性、およびsparse-data biasの問題を増大させます。

現代の多重回帰モデルによって課される細かい共変量サブグループ化は、このnoncollapsibilityを悪化させる可能性があります。例えば、10個の二値共変量を持つ多重ロジスティックモデルの実効的なサブグループ数は 2^{10} （100万以上）にもなります。患者数が100万人未満の場合、多くのサブグループには数人しか患者がおらず、オッズ比が0またはゼロセルのために未定義となる可能性があります。モデル制約による平滑化効果はありますが、モデル共変量のアウトカムを予測する能力が高いほど、noncollapsibilityは増加します。その典型的な結果は、さらなる交絡やその他のバイアスが除去されていない場合でも、モデルに予測因子が追加されるにつれて、推定されたオッズ比が1から離れる方向に動くことです。

この「データポイントが少なすぎるのに共変量が多すぎる」ことによる推定値の膨張は、sparse-data biasというトピックで議論されることがよくあります。これは現実のデータ例ではかなり大きくなる可能性があります。しかし、それはしばしば見過ごされるか、交絡除去と誤解され、バイアス除去として扱われますが、実際にはoveradjustmentのアーチファクトです。

したがって、オッズ比を交絡調整のための共変量選択に使用することは誤解を招く可能性があり、また（リスク比の近似である場合を除き）、統計分析の最終的な目標として、あるいは平均的な効果や典型的な効果として扱われるべきではありません。

アウトカムが非常に一般的で、オッズ比がもはやリスク比を近似しない場合、一般的な選択肢は、オッズ比をリスク比に直接変換するとされる式を使用することですが、これはバイアスのあるリスク比を生成します。バイアスが少くない選択肢は、係数がログリスク比として解釈可能なモデルを使用することですが、これらのモデルはアウトカムが一般的な場合に非現実的になる傾向があります。著者（Sander Greenland氏）の推奨する解決策は、ロジスティックモデルを使用し、そこから得られるフィットしたリスクを用いて、共変量別および重み付け平均（標準化）リスク差やリスク比のようなcollapsible measuresを構築することです。