

「治療」2018年7月号『プライマリ・ケア臨床研究』正誤表

いつも小社出版物をご利用いただき誠にありがとうございます。

「治療」2018年7月号『プライマリ・ケア臨床研究』に以下の誤りがございました。深くお詫びするとともに、ここに訂正いたします。誠に申し訳ございませんでした。

①838頁， 本文3,4行目

【誤】 食事療法を行った群は体重が平均100g, 対象群では平均50kg, それぞれ減少し

【正】 食事療法を行った群は体重が平均100g, 対照群では平均50g, それぞれ減少し

②839頁， 追記1行目

【誤】 生物統計家 Doaglas G Altman 先生

【正】 生物統計家 Douglas G Altman 先生

次ページより『生物統計家との連携 —研究の価値を高めムダを減らすために—』(836-839頁)の修正版を提示させていただきます。

2018年7月現在

生物統計家との連携

— 研究の価値を高め ムダを減らすために —

市川周平

三重大学医学部名張地域医療学講座



プライマリ・ケア医と臨床研究のかかわりについてどうあるべきか？



プライマリ・ケア医が臨床研究を行う際は、その研究が社会にもたらす利益を最大にするために、ただ研究をするだけでなく、よりよい研究を行う必要がある¹⁾。そのために、研究の設計段階から統計家と連携することが望ましい。その際、統計解析の方法だけでなく、アウトカムや例数の設計など、研究計画の個々の要素を統計家とともに吟味することで、当該研究で抽出できる効果を増加し、バイアスの影響を小さくすることができる。

プライマリ・ケア医が取り扱う研究テーマは、患者や一般市民にとって身近なものが多い。それゆえ、プライマリ・ケア医こそ価値の高い研究を行うことが求められる。

はじめに

臨床研究を推進することは、日本のプライマリ・ケア領域の喫緊の課題である。とくに、総合診療専門医や家庭医療専門医を取得するためには研修期間中に研究をテーマとしてポートフォリオを書くことが必須とされており、専攻医は上級医と二人三脚で研究に取り組んでいることだろう。しかし、そうして生み出された研究プロダクトがどれほどの価値をもつか、考えたことがある方はどれくらいいるだろうか？

近年、研究方法論の専門家の間では、どのようにして臨床研究の価値を高めるかが重要な関心事となっている。2009年には85%の生物医学研究がムダ (research waste) だとの推計が報告され²⁾、2014年にはその損失額は2000億米ドルに上るとされた³⁾。臨床研究へ投資できる社会的資源には限りがあるため、研究者は1つ1つの研究の価値を可能な限り高めなければならない。これは、プライマリ・ケアの臨床研究でも例外ではない。ムダ研究の抑制を目指したLancet誌の特集⁴⁾では、研究の価値を高めるための1つの方法として、研究の早期から統計家を巻き込むことを推奨している。研究の価値を高めるためのポイントは複数存在するが、本稿では研究の早期に考慮すべき点のうち統計家の支援が必要な点に絞って概説する。

I 効果とバイアスと研究デザイン

研究の価値を高めるひとつの要件は、効果 (effect size) が大きくバイアスの影響が小さいことである。大きな効果を検出し、バイアスの影響を低減するためには、研究の設計段階での工夫が必要である。介入研究の場合、臨床有意性を考慮したアウトカムや例数の設計、ランダム化や盲検化などを研究開始前に設計しておくことが、効果やバイアスの多寡に致命的に影響する。お粗末な研究デザインで僅かな効果しか得られなかったものを事後的な統計解析で取り返すことはできないし、それはむしろ研究の再現可能性を低める錯乱 (spin) となり得る。

研究デザインの過程では、研究の型の選択も含め、統計学的な吟味が必要な事柄が多い。たとえば、アウトカムの設定は研究デザインの選択や統計解析の方法と密接に関連する。欠損値の取り扱いや高度な統計学的判断が必要であり、研究方法や解釈の妥当性や透明性を大きく左右する。例数設計は、とくに介入研究では研究の妥当性を左右し得る。近年では、傾向スコアや回帰分断モデル、中断時系列デザインなどの擬似実験が臨床研究で応用されるようになり、因果推論でのバイアスの低減に貢献している。これらを適切に設計するためには、研究の早期の段階から統計家の関与が必要である。

II 例数設計

例数設計 (sample size estimation) とは、統計学的有意性を得るために必要な例数を研究の開始前に予め見積もることである。

例数設計の目的の1つは、試験参加者の数を必要最低限に抑えることである。臨床研究で頻用される有意確率や信頼区間を用いた頻度主義統計学での仮説検定は、患者や研究参加者を多く研究に組み入れるほど、統計学的有意性を得やすい。その一方で、効果が立証されていない治療法を試行される患者は、可能な限り少数に抑えるべきである。事前に例数設計を行うことで、これらの一見相反する要求を両立することができる。

例数設計のもう1つの目的は、研究の手続きや解釈の妥当性を担保することである。例数設計を行わない場合、「主たる要因に効果がない」、「主たる要因は効果を有するが例数が足りないために検出できない」のいずれであるかを峻別できない。例数設計を行ったとしても、設計した例数をリクルートすることができなければ同じことである。結果を解釈できなければ、研究者、研究参加者、出資者のいずれにとっても、研究にかけた労力と時間と資金がムダになる。

このように、例数設計は研究倫理および統計学の双方の見地から必須なのである⁵⁾。例数設計では、欠損の取り扱いや臨床的有意性など高度な統計学的判断を求められるため、研究の計画段階で統計家に相談しておくといよい。

Ⅲ 臨床有意性と統計学的有意症 (significantosis)

臨床有意性とは得られた結果の臨床的または医学的な見地からの有意性であり、臨床研究を行う際に本来追及すべきものである。たとえば、肥満の成人を対象に、体重減少を目的とした4週間の食事療法の有効性の比較試験を行ったところ、食事療法を行った群は体重が平均100g、対照群では平均50g、それぞれ減少し、減少量に群間の統計学的有意性が認められたとする。この食事療法に臨床的な意味を見出せるかと問われれば、多くの医師や医療従事者が否と答えるだろう。この例のように、介入効果が臨床有意性を主張できないほど小さいにもかかわらず、統計学的に有意が得られることは決して珍しいことではない。統計学的有意性が認められても、臨床有意性を伴っていなければ医学や医療へのインパクトを主張することはできない。臨床有意性が認められても、統計学的有意性が認められなければその結果から何かを主張することはできない。研究者が臨床有意性を見失い、統計学的有意性の有無に一喜一憂する様は、統計学的有意症 (Significantosis)⁶⁾として古くから批判されてきた。

臨床研究の結果を適切に解釈するためには、統計学的有意性と臨床有意性の双方から吟味しなければならない。そのためには、臨床的な有意性を主張できるようなアウトカムを設定すると良い。先ほどの食事療法の比較試験であれば、体重がどの程度減少したら臨床的な意味があるとみなせるかを検討したうえでしきい値を設定し、そのしきい値を超えて低減した人の割合や、しきい値を超えて低減するまでに要する期間をアウトカムとして設定することなどが代替案として考えられる。また、QOLや患者報告アウトカムを取り扱う場合は、検出可能な最小限の変化 (smallest detectable change) や臨床的に重要な最小限の差 (minimal clinically important difference) といった尺度特性を研究計画の段階で織り込んでおくとうい。

あわせて、有意確率 (p 値) や信頼区間に頼らない仮説検証の方法に慣れておくこともお勧めする。有意確率や帰無仮説検定は頻繁に誤用されるため、古くから批判されてきた。また、p 値や信頼区間を用いた仮説検証は、p-hacking などの spin の誘因となり得る。アメリカ統計協会は2016年に声明を発表し、帰無仮説検定や有意確率を用いた安直な意思決定に警鐘を鳴らしている⁷⁾。また、Basic and Applied Social Psychology 誌は、p 値や信頼区間を用いた論文の掲載を拒否している⁸⁾。こうした動きがプライマリ・ケア領域にまで波及するかは不明確だが、そうした動きがあることは理解しておくとうい。p 値や信頼区間を代替する手法として、大規模データを用いて記述統計を丹念に行うことや、ベイズ主義統計学を用いることなどが提案されている。とくにベイズ主義統計学は臨床有意性を考慮した仮説検証と非常に相性が良いため、ぜひ挑戦していただきたい。

おわりに

本稿では、研究の早期から統計家を巻き込むべきポイントを概説してきた。プライマリ・ケア医が取り扱う研究テーマは、患者や一般市民にとって身近なものが多い。それゆえ、プライマリ・ケア医こそ価値の高い臨床研究を行うことが求められる。プライマリ・ケアに携わる医療従事者が生物統計の専門家である必要はないが、統計家と良好な関係を築きながら彼らのもつ技術と見識を活用してほしい。

追記

本稿執筆中に生物統計家Douglas G Altman先生が逝去された。Altman先生は医学研究での統計解析の誤用を戒め⁹⁾、研究の透明性を高めるための取り組みを数十年にわたり続けてこられた。彼の金言を引いて、筆を置くこととする。

“To maximise the benefit to society, you need to not just do research but do it well.”
Evidence Live 2017のPreview Video¹⁾より

参考文献

- 1) Center for Evidence-Based Medicine, Oxford university : Doug Altman – Scandal of Poor Medical Research.
<https://www.youtube.com/watch?v=ZwDNPldQO1Q>
- 2) Chalmers I, Glasziou P : Avoidable waste in the production and reporting of research evidence. *Lancet*, 374 : 86-89, 2009.
- 3) Macleod MR, Michie S, Roberts I, et al : Biomedical research : increasing value, reducing waste. *Lancet*, 383 : 101-104, 2014.
- 4) Ioannidis JPA, Greenland S, Hlatky MA, et al : Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*, 383 : 166-175, 2014.
- 5) Schulz KF, Altman DG, Moher D, CONSORT Group : CONSORT 2010 statement : updated guidelines for reporting parallel group randomised trials. *BMJ*, 340 : c332, 2010.
- 6) 佐久間 昭 : 5. 臨床薬効評価試験の計画と解析—とくに繰り返しの多重推論について—. *臨床薬理*, 17 : 447-450, 1986.
- 7) Wasserstein RL, Lazar NA : The ASA’s statement on p-Values : Context, Process, and Purpose. *Am Statistician*, 70 : 129-133, 2016.
- 8) Trafimow D, Marks M : Editorial. *Basic Appl Soc Psych*, 37 : 1-2, 2015.
- 9) Altman DG : The scandal of poor medical research. *BMJ*, 308 : 283, 1994.