臨床試験の結果に対する批判的吟味

はじめに

臨床試験の結果には、試験期間中に得られたすべてのデータ、計測、統計解析が含まれます。試験対象母集団、ベースライン データ、参加者に対する治療の効果を測定する方法、および試験参加者に生じた有害事象に関する説明が含まれます。臨床試験の結果と分析はさまざまな経路、特に学会や医学雑誌で公表されます。

読者は、特に、記載されているエビデンスのレベルを評価して、公表文献に含まれる誤りの考えられるあらゆる原因を特定するため、臨床試験結果を批判的な目で精読をする必要があります。読者は、利用可能な最良の情報源から得た関連情報を考慮に入れる必要があります。読者は、利用可能なツール、たとえば PubMed を使用して関連論文を特定するために文献を検索できます。また、読者は、信頼できる機関 (たとえば、EMA、FDA、または国内患者団体や国際患者団体) により公表された文献を考慮してもよいでしょう。

以下の項目では、臨床試験結果のレビューに際し、批判的な読者が直面する可能性がある疑問について説明しています。

試験は信頼できますか?

目的と仮説の特徴が明確であるか考えてください。
試験結果がより大きな母集団に一般化できる可能性はありますか?読者は、試験結果を適用可能な対象について考える必要が
あります。対象となった標本集団の特性について、説明する必要があります。試験で実施されたすべての治療の詳細について明記されており、探索的治療は読者の疑問と関連がありますか? 治療により患者が享受する可能性が高いベネフィットとリスクは何ですか?
あらゆる利益相反について考えてください – 研究の信頼性と客観性に影響を与える可能性があるでしょうか。

試験方法は、記載されている仮説を評価するのに適切ですか?

標準治療は、現在一般に行われている治療に相当する適切な比較対象ですか?プラセボ、選択可能な治療法、最良の支援的ケア、または既存対照群ですか?
試験対象母集団を明確に定義すべきです。全母集団またはサブセットのいずれが評価されたのか、および選択に際し生じうるバイアスがあるかどうかが明確である必要があります。試験から脱落したあらゆる患者の妥当性および脱落理由について考え
てください。
対照群が適切に整合するか、および除外基準が妥当かどうかを評価してください。試験のエンドポイントが適切に定義され、意味があるものですか? 試験が主要エンドポイントに対してどのように検出力を保持したのか、明確ですか? 試験は、必要な数の事象を記録して結果の評価を行えるような十分な期間に実施されましたか?

結果は説得力のあるものですか?

結果は、明確かつ客観的、十分詳細に記載されるべきです – たとえば、疾病のステージ、年齢、性別、および/または考
えられるあらゆる交絡因子により結果が分類されるべきです。
結果がどの程度の説得力を持つか、統計解析は適切か、記載内容以外の結果をもたらした他の原因が存在するかどうかを考えてください。
試験中のフォローアップの欠測率、および無反応者がどのように扱われたか – たとえば、無反応者が治療失敗とみなされたか、または解析に別に含まれているのかどうかを特定してください。
あらゆるバイアスについて確認してください。研究者がリスクを抑制または減少させたかどうかを評価してください。

考察セクションは説得力がありますか?

考察には、当初の仮説を支持するものに限らず、試験のすべての結果が含まれるべきです。
考察には、当初の目的が満たされたか、および研究課題が解決したかどうかが記載されるべきです。
著者が考えられるバイアスを除外しているか、および試験の推定される限界を明記したかどうかを評価してください。
試験結果の一般化が誤った形で適用されていないかどうかを確認してください。
既存の文献に合致するかどうかを確認してください (常に課題が同じ他の出版物を探してください)。

示された効果は臨床的に意味のあるものですか?

主張されている効果は臨床的に意味があるか、批判的に評価してください – 患者の健康上、有意な効果が認められますか?たとえば、統計学的に有意な効果でありながら、さほどではないため、患者にとって臨床的に意味がない可能性がありま
す。試験の規模が大きくなるほど、小さくても意義のある効果が検出できるようになります。効果が統計学的に有意であっても臨床的に意味がない場合は、臨床試験の規模または検出力が大きすぎる結果である可能性があります。
一方、エビデンスの欠落は、効果が全くないことを意味しません。試験群間で統計学的に有意な差が見られない場合、これは比較された治療法が同等であることを意味しません。これは、統計的検定では、仮説を支持するエビデンスを評価するのでは
なく帰無仮説が真でないことを支持するエビデンス (帰無仮説を支持するエビデンス) の評価を試みるためです。言い換
えれば、統計的検定では帰無仮説の検証を試みます。治療の有効性が本当に異なる場合であっても、偶然の動き (第二種の
過誤)、または十分な量の情報を入手できない (試験規模が小さく検出力に欠ける) ため、統計的検定では有意でない可
能性があります。

結論は妥当ですか?

著者により提供された結論は、入手可能なデータによりサポートされる必要があります。結論が、記載された試験の意図およ
び目的と関連していることを確認してください。

□□□□

A2-4.35.2-v1.1