

原稿作成日： 2017年3月31日

国際誌が求める統計：チェックリストの活用

<教材提供>

AMED 支援「国際誌プロジェクト」 提供

無断転載を禁じます

草案

新谷歩 大阪市立大学医学研究科医療統計学講座教授

加葉田大志朗 大阪市立大学医学研究科医療統計学講座特任助教

査読

大門貴志 兵庫医科大学医療統計学教授

角間辰之 久留米大学バイオ統計センター教授

市川家國 信州大学特任教授

山本紘司 大阪市立大学大学院医学研究科医療統計学講座准教授

石原拓磨 大阪市立大学大学院医学研究科医療統計学講座特任助教

目次

はじめに

Nature : 統計手法の記載におけるチェックリスト

New England Journal of Medicine : 統計手法の記載におけるチェックリスト

Science : 統計手法の記載におけるチェックリスト

The EMBO Journal : 統計手法の記載におけるチェックリスト

JAMA : 統計手法の記載におけるチェックリスト

はじめに

この単元では国際科学雑誌

- ①Nature
- ②New England Journal of Medicine
- ③Science
- ④The EMBO Journal
- ⑤JAMA

が推奨している統計チェックリストから抜粋した内容を噛み砕いて紹介します。基本的なものから発展的な内容まで網羅されており、チェックリストには見慣れない専門用語も多く出てきます。基本的な内容については、この後の単元のテキストで勉強していきますので、ここでは各国際誌がどのようなポイントを押さえるべしとしているのか、その全体像を見ていきましょう。

学習目標

本単元を通じてあなたが修得を目指すものは：

- 国際誌で求められている内容の全体像を把握する

Nature：統計手法の記載におけるチェックリスト

- 比較している群（対称群）を明記すること
- 仮説検定の名称を明記すること
- 用いた全ての統計手法を誤解のないように記述すること
- 適用した手法が適切とする点を記載すること
- 使用されたデータが用いられた検定手法の仮定を満たしていること（例：分布に歪みがないか、症例数が少なくないかなど）
- 多重性によるP値の補正を行なっている場合はその旨を説明すること
- 研究開始時の症例数と、それぞれの解析で使用された症例数を明記すること
- 症例数計算の方法や妥当性について記載すること
- 全ての解析において、解析の対象とした集団について記載すること
- 全ての解析において有意水準を記載すること（例：5%未満）
- 両側・片側検定のどちらを用いたか記載すること
- ランダム化の手順、あるいはバイアスを取り除くために用いた方法について記載すること（特に動物実験において）
- 主解析についての実際のP値を記載すること
- 用いたデータセットの症例数を記載すること
- 平均や中央値など、データの代表値として何が使用されたかを示すこと
- 標準偏差や四分位範囲など、データのばらつきを表すための指標に何が用いられたかを記載すること
- 標準偏差や標準誤差については $a \pm$ 標準偏差 や $a \pm$ 標準誤差 のように記載すること
- 一般的ではない、あるいは複雑な解析手法については *Nature* の幅広いジャンルの読者

が理解できるように詳細を記載すること（もし説明が長くなる場合は補足資料を準備すること）

- データの除去を行なった場合にはその手順の記載と理由を説明すること
- 各解析によって含まれる症例数が異なる場合には、その理由を明記すること
- 割付方法と、その理由と妥当性について明記すること
- 対数（log）変換等データの変換を行なった場合にはその方法と理由を明記すること
- 効果量の提示方法を編集した際（グラフY軸の短縮など）には、その旨と理由を記載すること
- エラーバーは可能な限り各グラフにつけること
- エラーバーが何を示すか明記すること

New England Journal of Medicine : 統計手法の記載におけるチェックリスト

- 3群以上の複数群間でそれぞれの2群ごとにアウトカムを比較する場合は、多重検定に伴う第1種の誤りを防止するべく、補正を考慮に入れた検定手法を用いるべきである。正規分布にデータが従う場合、t検定は適切であり、また十分にデータの数が大きくなればt検定の結果はゆるぎないものとなる。データ数が小さく、先行文献においてデータの歪みが示唆されている場合は、ノンパラメトリック検定を適宜用いるべきである。
- 結果は科学性が担保された指標を用いて表すこと。例えば、関連性を示唆する指標としてはオッズ比などが挙げられる。オッズ比は通常2桁の有効数字を用いて表すこと。
- 信頼区間などのデータの不確実性を示す指標は首尾一貫して用いられるべきである。これは結果を包括的に図示する際も同様。
- 非劣性試験のような片側検定が必要とされるような研究デザインを除いて、すべてのP値は両側のものを使用する。0.01より大きいP値については小数点以下2桁まで、P値が0.001から0.01までの間の値について小数点以下3桁まで、P値が0.001より小さい場合は $P < 0.001$ と記載すること。臨床試験の早期中止ルールにP値が用いられる場合や、ゲノムスクリーニング研究にP値が用いられる場合はこの限りではない。
- 無作為化試験の群間を比較する研究対象者の背景を示す表には、表のフットノートに統計的有意差を示唆する（例： $P < 0.05$ ）が、表の中にはP値を載せないものとする。
- 作為化臨床試験の結果の報告に際しては、コンソート（CONSORT）に指示されている様式で被験者のフローチャート図を掲載すること。またコンソートのチェックリストで求められている項目についても報告に含めること。必要であればAppendixなど、論本文体とは別の文書として用意すること。コンソート声明、チェックリスト、流れ図についてはCONSORTのウェブサイトを参照のこと。

Science : 統計手法の記載におけるチェックリスト

- 十分に知識のある読者であれば、論文作成に使用されたデータを利用して結果を検証することが可能となるように、統計解析の詳細を記載しなければならない。
- データの事前編集（データの変換、再コード可、再スケール化、標準化、一定値以上のデータの切り捨て（Truncation）、測定限界値以下の観測値や外れ値の取り扱い、またいかなる観測値の削除や編集も含む）については、正しい知識と正当性のある理由付けが必要である。
- 研究結果の理解に不可欠な変数については、記述統計量を記載すること。記述統計量には症例数や平均や中央値など、どの統計量が使用されたかについても記載すること。

- 連続変数については散布図や、箱ひげ図、ヒストグラムなどのグラフを用いて、または平均や中央値および標準偏差や四分位範囲などデータの中心やばらつきを表す指標を用いて記述すること。
- 正規分布に従う連続変数には、平均や標準偏差を用いること。また分布が正規分布のような左右対称性を持たない連続変数については中央値や最小値・最大値の範囲または四分位範囲を用いること。すべての場合について、平均が用いられたのか、中央値が用いられたのか、どのばらつきの指標が用いられたのかなどを明記すること。
- 症例数が 20 未満のようなデータにおいては、個人情報保護の観点から倫理的に問題がない限りすべてのデータ値を表にして記載することが望ましい。すべての測定値は単位を記載すること。
- t 検定や、ウィルコクソン符号付順位和検定、回帰係数についての Wald 検定など、連続変数の解析に用いられる検定や、正規分布を仮定した 95% の信頼区間、平均と 2×標準偏差、尤度比を用いた信頼区間など、検定に何を用了のかなどを Materials と Methods の章に明記すること。図の凡例中に解析に用いた検定を実験ごとに記載すること。
- それぞれの解析で統計的有意差の判定に用いられた基準について（片側か両側検定か等）についても記載すること。通常は両側検定を用いるべきであるが、片側検定を用いる場合はその正当性の理由づけを行うこと。
- 多重比較により偽陽性の確率を調整するために有意水準の補正法（Bonferroni の調整法やその他）を利用した場合は、その手法についても記載すること。
- ベイズ解析が使用された場合、事前確率の仮定について十分な記載をすること。
- 使用した検定において、その正当性を読者が評価できるよう、それぞれの検定が前提とする仮定・条件が検証されたかについても記載すること。（例：データが正規分布に従っている；生存解析に用いられたデータは比例ハザード性を有するなど）
- 欠損値についての対処法を記載すること。欠損値補完が行われた場合は、その統計的な手法について記載すること。
- 新たな手法や高度な手法が統計解析または計算アルゴリズムで用いられた場合は、読者がそれらを再現できるように参考文献や使用例を添えて十分に説明すること。出版の際には計算に用いられたコードやデータを Supplementary Information として本誌が要求することもある。
- 論文の結論が適切な統計解析によって得られたものであることが分かるように、解析結果については詳細まで記載しておくこと。またその際の制限が生じた事項などについても正直に記載するべきである。
- 平均、相関係数、回帰係数などの点推計値や、平均差、オッズ比、ハザード比など比較の計量化に用いられた指標については標準誤差や、信頼区間など不確実性を表す指標を合わせて記載すること。
- それぞれの統計検定の結果については、十分に統計量と P 値をあわせて記載し、有意差の有無のみに言及を止めないこと。遺伝解析のように多重性に対する補正を行わなければならないというような状況でない限り、P 値については 3 桁以上の有効数字を表記する必要はない。
- いかなる結果でも盲目手法を用いて報告された旨、統計モデルが独立的に確認された旨については、下記の要領を踏まえた詳細な説明の添付が必要である。
 - 1) 解析モデルやデータ処理のステップ、アウトプットの計算アルゴリズム、分類のために利用したカットポイントなどについての詳細。
 - 2) それぞれの解析モデルや予測因子が完全に固定された日付。
 - 3) 盲検化されたデータを保持し、評価を監督した人（たち）の指名。（例えば公平な仲介者など）
 - 4) 解析モデルが固定された後に妥当性確認用データについて修正・追加・除外などが行われていない旨。また妥当性確認用データおよびその一部が、試験で利用されている解析モデルの評価・精錬には利用されていない旨。

The EMBO Journal : 統計手法の記載におけるチェックリスト

- 施行したすべての仮説検定について、その検定の名称と、利用した各時点の症例数、結果として得られたP値などを記載する必要がある。このときP値については統計的に有意か否かではなく、P値そのものを記載する。統計手法に関する記載は基本的に materials と methods に記載するが、図表においても解析手法・症例数・P値などの基本的な情報は記載しておく。
- 記述統計ではデータの中心を示す値（平均や中央値）と散布度を示す値（標準偏差や範囲）を示す必要がある。データの数が少ないような場合には、標準偏差よりも範囲など指標を用いることが望ましい。標準誤差や信頼区間については群間の比較を行う際に記載するのが適切である。
- グラフにはエラーバーとそのエラーバーの名称を記載しておく。さらに±標準偏差や±標準誤差の表記が必要。
- 利用した解析については、その解析についての仮定が満たされているかどうかを明確に記載しておく必要がある。
- 特に正規性を仮定する統計手法を用いるときは、著者はどのように正規性の検証を行ったかを説明しなければならない。データがそうした検定の仮定に従わない場合、ノンパラメトリックな手法を用いるべきである。
- ひとつのデータセットを利用して複数の統計検定を行う場合、第I種の誤りを犯すリスクが高まるが、これを回避する上で、どのような対策をとったのかを記載すべきである。
- それぞれの実験で行われた技術的・生物学的反復の数について明記するべきである。生物的反復 (Biological replicates) は、複数のサンプルからも同じような結果が得られるという外部妥当性の指標となり、技術的反復 (Technical replicates) は、同じサンプルからは同じ結果が得られるという内部妥当性の指標となる。一般的にこの技術的反復については、そこで得られた観測値の平均が計算され解析に利用される。
- 症例数が小さい場合、統計的に正しい手法を選択することがカギとなるが、使用した手法とその妥当性もあわせて明記する必要がある。
- 複雑な実験手技が必要であることが原因で、それぞれ独立した研究対象から多くのデータを蒐集することが困難であることも想定される。そうした場合、統計解析という面では症例数が非常に少なく、統計的に有意と言える基準を超えることができないこともある。そのような少症例数（症例数が5例未満のような場合）については、実際の各観測値についても個別に図示しておくことを推奨する。また症例数が少ない場合には、利用した仮説検定の妥当性についても説明する必要がある。一方、少数の研究対象から反復してデータが得られている場合、それらを統計解析に利用することが可能である。その場合、反復測定の内容などについては詳細に記述しておくことが必要となる。

JAMA : 統計手法の記載におけるチェックリスト

- 論文の Method において、十分に統計学的な知識のある読者であれば著者が解析に用いたデータセットを用いて再現できる程度の詳細を記載しておくこと。特に、あまり一般的に知られていない手法を用いるときには、その手法を最初に紹介した参考文献を誤りなく添えて記載すること。より高度な手法や新たな手法が用いられる場合は、その手法と使い方についての簡単な説明を論文本文中に記載し、詳細をオンライン・サプリメントに掲載することを推奨する。
- Result (結果) の章では、可能な限り結果を数量的に表し、それらを信頼区間などの不確かさの指標とともに記載すること。

- 仮説検定を行ったときには、P 値のみでは実際の差などの重要な情報が得られないため、P 値のみの記載は避けること。
- 観察研究では研究対象となった症例の数を記載しておく。無作為化試験においては無作為化された症例数を記載すること。その際には途中で抜け落ちた症例など追跡不可能な症例の数も記載すること。
- 多変量解析を利用した場合、回帰モデルで考慮した変数についてはすべて記載し、その回帰モデルの妥当性の検討や、モデルのデータへの当てはまり具合について、可能な限り詳細に記載すること。
- 統計用語や略語、記号が使われている場合はその定義を記載すること。相関 (Correlation)、正規 (Normal)、予測因子 (Predictor)、無作為 (Random)、標本 (Sample)、有意 (Significance)、傾向 (Trend) などの統計専門用語については、統計的な枠組み以外で使用することは避けること。結果が統計的に有意でない場合に (P 値 >0.05 の場合など)、「もう少しで有意 (Marginal significance)」や「有意な傾向 (Trend toward significance)」、と言ったあいまいな言い回しを避けること。観察研究においては手法 (Method) と結果 (Result) は関連 (association) や相関 (correlation) という言葉を使用し、できるだけ因果関係 (Cause-and-effect) を示唆する言葉を用いないこと。無作為化試験においては効果 (Effects) や因果関係 (Causal Relationship) のような因果関係を示唆する用語を用いることは可能である。
- 無作為化比較試験では検出力や症例数の計算についての記載が必須である (EQUATOR Network の CONSORT のガイドラインを参照)。観察研究においては、対象となる症例数が固定されている場合には検出力計算は必須ではない。しかし症例数を研究者が設定したのであれば、その正当性について記載すべきである。通常これらの検出力・症例数計算の手順については統計手法の章の先頭におく。
- 利用した解析手法については、一般的に利用される手法については詳細を記載する必要はないが、Method の章で簡単に説明しておくべきである。
- 多変量回帰分析を一概に Multivariate regression として記載するのではなく、2 個以上の説明変数が用いられた回帰分析については Multivariable regression、2 つ以上の従属変数が用いられた場合を Multivariate regression として使い分けること。回帰分析で利用した変数についてはすべて記載し、対数変換のようなデータ変換が行われた場合はそれもあわせて記載すること。またそれらの変数が選ばれた理由について、臨床的、統計学的、あるいはその他の理由によるものかを説明すること。
- 1 人の被験者からデータが繰り返し取得されるなどのクラスターまたは縦断的なデータについては、回帰分析はデータのクラスター内のデータの相関を正しく考慮できるものを選択する必要がある。クラスター内でのデータの相関を考慮しなかった場合、効果量や信頼区間の推計は正しく行うことができない。クラスター内のデータの相関をどのようなモデルで考慮したか等を記載する。例えば、Generalized Estimating Equation (一般化推定方程式、GEE) を用いた解析では、どのような相関構造が仮定されたか (Robust なのかサンドイッチ分散推定量が用いられたか等) を記載する。混合効果モデルを用いた解析では、ランダム効果がどのように用いられたか (切片か傾きか等) を記載すること。固定効果は Conditional Likelihood を用いたと記載すること。共変量 (調整因子) を記述するときに固定効果という言葉は用いないこと。
- 臨床試験からの抜け落ち例や追跡不可能例、または観察研究において入手不可能であったデータ数など、データの欠損について報告すること。何らかの形で、登録された症例がデータの欠損などの理由で最終解析で用いられなかった場合、解析から抜け落ちた症例と考慮された症例の背景比較を行い、Supplementary な結果として報告すること。多重補完法など、統計的に欠損データを補完する方法なども考慮し、欠損のメカニズムが Missing At Random (MAR) であったかどうかを検証すること。Last Observation Carried Forward (LOCF) など最終的に観察されたデータをもってデータを補完するような方法は用いられるべきではない。

- 無作為化研究・観察研究によらず主評価項目・副次的評価項目・サブグループや感度解析などの計画については計画段階で記載しておくことが求められる。計画時点で定められていなかった解析については「Post Hoc（後付け）」の解析であることを明示する。
- 2個以上の主評価項目が用いられた場合は、そのP値について多重性による補正を行うこと。副次的な評価項目においては多重性によるP値の補正を行うか、その結果を探索的なもの（あくまでも仮説を生み出すものであって仮説を検証するものではない）として報告すること。
- 無作為化比較試験においては、すべての統計解析計画内容について Method の章に記載し、オンライン Supplement で報告すること。また無作為化比較試験においては原則 Intention-to-treat の手法を用いて解析すること。完全な Intention-to-treat の法則から外れた場合は“modified intention-to-treat（改変された intention-to-treat）”である旨を記載し、どの点が改変されたかを明記すること。
- Method の章の終わりには、統計的有意差を判定した基準などについて簡潔に記載すること。また解析に用いたソフトウェアについても、そのバージョン・製造者・利用したパッケージ名等を記載しておく。ただし、解析手法としてソフトウェアのコマンド名などを直接記載しないこと（例：SAS proc mixed was used to fit a linear mixed-effects model）。解析に用いられたコードなどを記載する場合はオンライン Supplement を利用すること。
- 解析は EQUATOR Reporting Guidelines で記載された注意点を踏まえて、研究計画で記載された通り遂行し、後付けの解析は Post Hoc（後付け）であることを明示すること。
- 可能な限り計量的な結果（頻度、率）については、信頼区間など不確実性の度合い（測定誤差など）をあらわす指標とともに報告すること。
- Mean（平均）や SD（標準偏差）はデータが正規分布に従うときに使用し、そうでない場合は Median（中央値）や Interquartile ranges（IQRs）を用いること。
- P 値のみでは実際にどの程度の差があったのかなどの情報が得られないため、頻度や差などの比較結果もあわせて記載することを推奨する（例：0.8%（95%CI：0.2% to 1.8%）；P = 0.13）。
- P 値を報告する場合、0.001 未満の P 値は“P<.001”と記載し、0.001 から 0.01 の P 値は小数点第 3 桁まで、0.1 以上の場合は小数点以下第 2 桁まで、0.99 より大きい場合は“P>.99.”と記載すること。遺伝解析などで得られるような極端に小さい P 値については、 $P = 1 \times 10^{-5}$ などと表記することも可能。原則として t 統計量・F 統計量・ χ^2 のような統計量や自由度については記載の必要はない。
- 副次的な解析やサブグループ解析の場合、検定が多重になるため、関連が実際はないのにあるとしてしまう第 1 の誤りをどのようにして回避したか（有意水準の補正等）を記載すること。そのような記載がない場合の解析結果については探索的または後付けの解析として扱うこと。
- 複数群を比較する無作為化比較試験における群間の背景比較において、P 値を計算することは科学的に妥当ではないゆえ、避けること。比較群間での背景情報の差を表す場合は P 値ではなく、臨床的に重要な偏りであるかどうか、またその偏りを多変量解析などで考慮したか否かなどを報告すべきである。

本単元は日本医療研究開発機構：研究公正高度化モデルである「医系国際誌が規範とする研究の信頼性にかかる倫理教育プログラム」（略称：AMED 国際誌プロジェクト）によって作成された教材です。作成および査読等に参加した専門家の方々の氏名は、冒頭に掲載されています。

参考 HP

①Nature

(<http://image.sciencenet.cn/olddata/kexue.com.cn/upload/blog/file/2010/12/2010128212513557501.pdf>; visited on 2018.02.11)

②New England Journal of Medicine (<http://www.nejm.org/page/author-center/manuscript-submission#electronic>; visited on 2018.02.11)

③Science (<http://www.sciencemag.org/authors/science-editorial-policies>; visited 2018.02.11)

④The EMBO Journal (<http://emboj.embopress.org/authorguide#embargopolicy>; visited on 2018.02.11)

⑤JAMA (<http://jamanetwork.com/journals/jama/pages/instructions-for-authors>; visited on 2018.02.11)