

Series

海外文献紹介 ― ①

臨床試験の統計解析計画書の内容に関する ガイドライン Guidelines for the Content of Statistical Analysis Plans in Clinical Trials

林 健一*

はじめに

医学研究の信頼性を高めるためには、実施した研究のデザインおよび結果を正確に報告し、研究を再現できるようにすることが必要である。このため、医学研究に関連して作成する文書の内容を定めたガイドライン、すなわちレポート・ガイドラインが発表されており、その多くは EQUATOR (Enhancing the QUALity and Transparency Of health Research) Network のウェブサイト (<http://www.equator-network.org/>) から入手することができる。

このうち臨床試験に関しては、実施計画書（プロトコル）および論文に関するレポート・ガイドラインがすでに公表されている^{1,2)}。しかし、実施した臨床試験の結果から結論を導く際には統計学的な判断が大きな比重を占めており、統計解析の計画および実施に関する詳細な情報を開示することは透明性を高めるうえで重要である。こうした情報は「統計解析計画書 (SAP)」という名称の文書に記載されることが多い。

これまで SAP に関しては、その定義が臨床試験のための統計的原則³⁾に記載されているのみで、記載内容に関するガイドラインは存在しなかった。このため、英国の生物統計家を中心としたグループが 2012 年 11 月から活動を開始し、2017 年末に the Journal of American Medical Association (JAMA) で「臨床試験の SAP の内容に関するガイドライン」を公表するに至った⁴⁾。以上が本ガイドライン作成の背景である。

本ガイドラインの位置づけと特徴

本ガイドラインは、医薬品開発の後期段階に実施するランダム化比較試験の最終解析を想定して作成されたもので、SAP への記載を推奨する必要最小限の項目がまとめられている。各項目の解説と具体的な文例は付録 e Appendix 2 に示されており、JAMA のウェブサイト (https://jamanetwork.com/data/Journals/JAMA/936638/JSC170004suppl1_prod.pdf) から入手することが可能である。

次に、本ガイドラインでは SAP を「プロトコルと合わせて読む文書」と位置づけ、要求された情報がプロトコルに記載されている場合には、SAP でその記載を繰り返す代わりに、プロトコルの記載を参照するようにしている。ただし、記載を参照する場合には、プロトコルがレポート・ガイドライン¹⁾の要求を満たすことが前提となる。

SAP への記載が推奨される項目は表にまとめられており、その和訳を掲載する（表だけを読んでも意味がわかるよう、必要に応じて解説に記載された内容を加筆した）。ここでは SAP を 6 つのセクションに分け、セクションごとに記載すべき項目 (item) を整理している。項目数は 32 である。ただし、項目の中には 4a, 4b, 4c といったように細分化されたものがあり、細目を指す場合、原文では subitem または index という用語を用いている。

構成上の特徴は、緒言および試験方法というセクションを設けていることである。すなわち、プロトコルの記載を長々と繰り返す必要はないものの、SAP だけを読んでも試験を実施する背景や目的、試験デザインの概要

*アラメディック株式会社代表取締役

表 臨床試験の SAP への記載が推奨される項目^{a)}

セクション/項目	細目	内容
セクション 1：管理情報		
臨床試験のタイトルおよび登録	1a	プロトコールと一致した記述的タイトルを示し、タイトルの前またはサブタイトルに「統計解析計画書」と表記する 該当する場合には、臨床試験の頭字語も示す
	1b	臨床試験の登録番号
SAP のバージョン	2	SAP のバージョンと日付
プロトコールのバージョン	3	参照に用いるプロトコールのバージョン
SAP の改訂	4a	SAP の改訂履歴
	4b	SAP を改訂した根拠（改訂ごとに示す）
	4c	中間解析などとの関連を含む SAP の改訂時期
役割および責任	5	SAP の作成に関与した者の氏名、所属、役割
署名欄	6a	SAP の作成担当者
	6b	責任者となる上席統計家
	6c	主任研究者または研究代表者
セクション 2：緒言		
背景および理論的根拠	7	臨床試験を実施する背景および理論的根拠の要約 簡潔でよいので、試験で解明したい疑問と試験実施を妥当と判断した根拠も含める
目的	8	試験固有の目的または検証したい仮説
セクション 3：試験方法		
試験デザイン	9	試験のタイプ（例：並行群間比較、多群比較、クロスオーバー、要因計画）と割り付け比を含む試験デザインの簡潔な記載 可能であれば、試験治療の内容も簡潔に記載する
ランダム化	10	最小化法や層別割り付けなどのランダム化の詳細 （層別因子も含める。SAP に詳細を記載していない場合には、該当する情報が記載されている文書を示す）
目標とする被験者数	11	詳細な被験者数の設定根拠（プロトコールに詳細を記載した場合には SAP で繰り返さず、プロトコールの記載を参照する）
フレームワーク	12	仮説検定のフレームワーク（優越性・同等性・非劣性）を実施する比較ごとに示す
中間解析および早期中止の基準	13a	中間解析に関する情報（解析の内容と実施時期の一覧を含む）
	13b	中間解析のための有意水準の調整方法
	13c	早期中止の基準の詳細
最終解析の実施時期	14	最終解析の実施時期（例：全項目を同時に解析、フォローアップ期間に応じて複数回に分けて解析）
アウトカムの評価時期	15	来院時期の許容範囲を含むアウトカムの測定時期
セクション 4：統計上の原則		
信頼区間と P 値	16	仮説検定の有意水準
	17	多重性の調整方法とその選択根拠、調整する場合は、第 1 種の過誤のコントロール方法の詳細
	18	算出する信頼区間
遵守状況およびプロトコールからの逸脱	19a	試験治療の遵守・不遵守の定義 曝露の程度を含む治療遵守状況の評価方法
	19b	試験治療の遵守状況の集計方法
	19c	プロトコールからの逸脱の定義
	19d	プロトコールからの逸脱の要約方法
解析対象集団	20	解析対象集団の定義（例：intention to treat、プロトコールに適合した集団、試験を完了した集団、安全性の解析対象集団）

表 (つづき)

セクション/項目	細目	内容
セクション 5: 被験者集団		
スクリーニング記録	21	被験者集団の代表可能性を記述するためのスクリーニング記録の報告方法 (記録がある場合)
適格性	22	適格性の基準の要約方法
被験者の組み入れ	23	CONSORT のフローチャートに含める情報
中止・追跡不能	24a	中止のレベル (例: 試験治療を中止, 追跡を中止, またはその両方)
	24b	中止・追跡不能の発生時期
	24c	中止・追跡不能の理由・詳細の要約方法
被験者背景	25a	要約する背景因子の一覧
	25b	被験者背景の要約方法の詳細
セクション 6: 解析		
アウトカムの定義		主要評価項目・副次評価項目ごとに以下の詳細を含めて記載
		アウトカムと評価時期の特定
	26a	該当する場合は, 主要評価項目または重要な副次評価項目の重要度の順位を含める (例: 検定を実施する順序)
	26b	測定方法と単位 (例: 血糖コントロール, HbA1c [mmol/mol または%])
	26c	アウトカムを得るための計算方法または変換方法 (例: ベースラインからの変化量, QOL のスコア, イベント発現までの時間, 対数変換)
解析方法		主要評価項目・副次評価項目ごとに以下の詳細を含めて記載
	27a	用いる解析方法, 治療効果の提示方法
	27b	共変量の調整
	27c	解析方法の前提を確認する方法
	27d	前提とする分布を満たさなかった場合に用いる代替手法の詳細 (例: 正規性, 比例ハザード性)
	27e	計画する感度分析 (実施する場合)
	27f	サブグループの定義を含むサブグループ解析の計画
欠測値	28	欠測値の取り扱い方法とその前提 (例: 多重代入法)
その他の解析	29	実施する場合には, その他の必要な解析の詳細 (例: complier-average causal effect 解析)
有害事象	30	安全性データの要約方法の詳細 (例: 重症度・未知の定義・因果関係に関する情報, 有害事象のコード化または分類方法の詳細, グレード 3/4 のみ集計・incidence case analysis・intervention emergent analysis といった有害事象の解析方法)
統計解析のソフトウェア	31	解析を実施するために用いる統計解析パッケージの詳細
参考文献	32a	標準的ではない解析方法を用いたときの参考文献
	32b	データマネジメント計画に関する文献
	32c	Trial Master File や Statistical Master File に関する文献
	32d	その他の標準業務手順書または遵守すべき文書に関する文献

略語: CONSORT=Consolidated Standards of Reporting Trials, HbA1c=hemoglobin A1c, QOL=quality of life, SAP=statistical analysis plan (統計解析計画書)

*著者らの許可を得て転載

本稿は Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P, Berlin J, Senn S, Day S, Barbachano Y, Loder E. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA 2017; 318: 2337-43. の内容を林健一が解説したものである。表の掲載に際しては American Medical Association から許可を得た。

Copyright © 2017 American Medical Association.

がわかるようにしている。プロトコールの統計解析のセクションに該当する情報だけを記載した SAP も存在していた従来と比べると, これは大きな進歩である。

ガイドラインの内容

1) 管理情報

ここには, 臨床試験のタイトルや登録番号といった文書管理に関する情報を記載する。本セクションの特徴

は、項目2および3でSAPとプロトコールのバージョンを両方記載するようにしたことで、このようにすれば、どのバージョンのプロトコールに対応するSAPなのかが一目でわかる。臨床試験の実施中にはプロトコールが数回改訂されることもあり、SAPとプロトコールそれぞれのバージョンを対比できるようにすることは重要である。

2) 緒言

本セクションでは、最初に臨床試験を実施する背景と理論的根拠を要約する(項目7)。臨床試験を実施する背景と理論的根拠はプロトコールに詳述するため、SAPではその記載を参照すればよい。ただし、「この試験ではどのような疑問を解明しようとするのか」「なぜこの試験の実施を妥当と考えたのか」の2点はSAPにも記載することが要求されている。

続いて、試験の目的や検証する仮説を記載する(項目8)。ここはプロトコールの記載を参照することでもよいが、プロトコールでは医師が理解できる言葉で目的を記載することが多い。したがって、検証的な試験の場合には、帰無仮説・対立仮説といった統計学的な用語を用いて仮説を記載することが望ましい。

3) 試験方法

ここは試験計画の概要を記載するセクションで、並行群間比較といった試験のタイプ・割り付け比・試験治療の概略を記載し(項目9)、ランダム化の詳細を記載する(項目10)。注意したいのは項目10で、ブロックサイズや動的割り付けの詳細なアルゴリズムを記載すると、割り付けの予見可能性を高める恐れがある^{1,3)}。したがって、試験実施中にSAPを公表する場合には、詳細を別文書に示すことが必要である。項目11では、目標とする被験者数の設定根拠を要求しているが、プロトコールに詳細が記載されていれば、それを参照することで差し支えない。

項目12では、実施を予定する比較ごとに、仮説検定のフレームワーク(優越性・非劣性・同等性)を記載する。検証的な試験の場合、主要な仮説は項目8に記載するが、それだけでは不十分なこともある。たとえば、主要評価項目で血糖降下作用が対照薬に劣らないことを検証した後、副次評価項目で体重低下作用が対照薬に優れることを検証する試験では、副次的な比較のフレームワークを示すことも重要である。このため、項目12では、評価項目ごと、比較ごとにフレームワークを示すようになっている。

項目13では中間解析を実施する計画の有無を明記し、実施を計画する場合には、中間解析の内容・実施者・実施時期、有意水準の調整方法、早期中止の基準を示す。

項目14は最終解析の実施時期を明示するように要求したもので、全評価項目を同時に解析するのか、それとも評価項目ごとに最終解析の時期が異なるのかを記載する。バイオマーカーと全生存期間とでは解析時期が異なるといったように、評価項目ごとに解析時期が異なる臨床試験の場合には項目14が重要になる。項目15は各評価項目の観察・測定・判定時期を明示するように要求したもので、プロトコールのスケジュール表を参照することで差し支えない。

4) 統計上の原則

項目16では有意水準および両側検定・片側検定の区別を示し、項目17では多重性の調整方法を記載する。多重性を調整しない場合には、項目17にその旨を明記することが必要となる。さらに、多重性の調整方法が異なれば、結論も異なる可能性があるため、少なくとも主要評価項目に関しては多重性の調整方法の選択根拠を示すことが要求されている。項目18では、評価項目ごとに算出する信頼区間を明らかにする。すべて両側95%信頼区間を算出するのであれば、その旨を記載すればよい。

このセクションの特徴は項目19で、19aでは試験治療の遵守状況を評価する方法を記載し(例：処方した錠剤数を分母として、服薬した錠剤数の割合を算出する)、19bでは遵守状況の集計方法を記載する(例：75%以上の錠剤を服用した被験者数を群別に示す)。続いて、19cで重大な逸脱と軽微な逸脱を定義し、19dでプロトコールからの逸脱の要約方法を記載する(例：逸脱の種類ごとに群別の被験者数を示す)。重大な逸脱と軽微な逸脱を定義することは容易でないが、盲検解除後に決定するとバイアスが生じる恐れがあることから、本ガイドラインでは盲検解除前にこれらを定義すべきとしている。

最後に、項目20ではintention to treat、プロトコールに適合した集団、試験を完了した集団、安全性の解析対象集団といったように、解析の対象となる集団を定義する。これは従来のSAPにも記載していた内容である。

5) 被験者集団

臨床試験の結果を示す際には、まずフローチャート(適格性評価から試験完了に至るまでの被験者数の推移)と被験者背景を要約する。これらに該当するのが本セクションで、適格性を評価したスクリーニング記録があれば、その要約方法を示し(項目21)、適格・不適格の判定に関する情報の要約方法を記載する(項目22)。さらに、フローチャートに盛り込む情報を特定する(項目23)。項目24では、試験治療開始後の中止・追跡不能の分類方法(24a)、発生時期の要約方法(24b)、中止・追跡不能の理由の分類方法(24c)を定義する。

項目25は、被験者背景をどのように要約するのかを定

めるもので、最終解析で要約する背景因子の一覧 (25a) と被験者背景の要約方法の詳細 (25b) を記載する。要約方法を記載する際に重要なのは連続量の取り扱いで、背景因子が連続量の場合には、平均値や標準偏差といった記述統計量を示すのか、複数のカテゴリーに分割して各カテゴリーの被験者数と割合を示すのかを明らかにする。

6) 解析

主要評価項目および副次評価項目を解析する際には、各評価項目の定義を明確にしたうえで (項目 26)、評価項目ごとに解析方法を決定することが必要である (項目 27)。あわせて、欠測値の取り扱いを事前に定め (項目 28)、探索的な解析等を実施するのであれば、その内容を明らかにする (項目 29)。項目 26~29 はこうした流れに沿って設定されている。

項目 26 では、主要評価項目・副次評価項目をそれぞれ定義する (26a)。評価項目が反復測定データの場合には測定時点、生存時間データの場合には測定開始時点を明らかにすることが必要である。あわせて、各評価項目の測定単位を明記する (26b)。血清脂質値のように、地域によって測定単位が異なる場合、用いる単位を規定することは特に重要である。さらに、実測値から変化量や変化割合を算出するといったように、測定されたデータを変換する場合には変換方法を記載し、複数の自覚症状の問診結果に基づいて単一のスコアを算出するといった場合には、欠測値の取り扱いも含めたスコアの計算方法を記載する (26c)。

項目 27 では、評価項目ごとに解析対象集団を明確にしたうえで、データの記述方法、群間比較に用いる解析方法、治療効果の推定方法、 P 値や信頼区間を算出するか否かを記載する (27a)。続いて、評価項目ごとに共変量で調整するか否かを示し、調整する場合には、調整に用いる共変量を特定するとともに、調整した解析と調整しない解析のどちらを主要な解析とするのかを記載する (27b)。さらに、用いる解析方法の前提が成立しているか否かを確認する方法 (27c)、前提が成立しなかった場合に用いる代替手法 (27d) を記載する。感度分析を計画する場合には、評価項目ごとに解析対象集団も含めた感度分析の方法を記載し (27e)、サブグループ解析を実施する場合には、解析に用いる背景因子、連続量の分割点、フォレスト・プロットなどの結果の表示方法を記載する (27f)。

項目 28 では、欠測値の取り扱い方法 (推定値で補完する、しない) を明らかにし、推定値で補完する場合には、用いる解析方法と解析の前提条件を記載する。さらに、複数の方法で欠測値を補完する場合には、主要な解析と

副次的な解析を区別する。項目 29 では、探索的な解析などを実施する場合に、その解析方法の詳細と実施する理由を記載する。

以上は主として有効性の評価項目を想定したものであり、安全性の評価項目に関しては項目 30 にデータの要約方法を記載する。具体的には、重症度や因果関係などの取り扱い、有害事象の分類に用いる辞書、解析方法 (例: 基本語レベルで各有害事象が発現した被験者数と発現割合を群別に集計) といった内容を記載する。

項目 31 では、解析に用いるソフトウェアの名称を記載する。なお、ソフトウェアのバージョンは試験実施中に更新される可能性があるため、解析実施時のバージョンは統計解析報告書に示すことで差し支えない。項目 32 は参考文献を記載するもので、標準的ではない解析手法を用いたときの参考文献 (32a)、データマネジメント計画に関する文献 (32b)、Trial Master File や Statistical Master File に関する文献 (32c)、標準業務手順書などに関する文献 (32d) を記載する。

おわりに

最近では、臨床試験の被験者レベルのデータ (匿名化を図ったもの) を提供しようという動きが高まっており、2019 年 1 月以降に被験者のリクルートを開始する臨床試験に関しては、臨床試験の登録時にデータ提供の可否を明らかにすることとなっている⁵⁾。さらに、データ提供、プロトコル、SAP の公表を論文掲載の条件とする医学雑誌もみられるようになってきている。こうした流れに従って、被験者レベルのデータが利用可能になり、SAP が公表されると、臨床試験成績は完全に再現可能となる。これはデータの捏造・改竄を防ぐうえで有効に機能することは間違いなく、この観点からも SAP を公表する意義は大きい。

なお、本ガイドラインの末尾にも記載されているように、ICH E9 ガイドラインの補遺⁶⁾ は今後の SAP 作成に大きな影響を及ぼすものである。本ガイドラインでは想定する逸脱を項目 19 に記載するようにしており、これは ICH E9 の補遺と一致するものである。その一方で、項目 20 は従来の考え方に従って設定されている。このため、今後 SAP を作成する際には、ICH E9 の補遺を含む他のガイドラインの通知・改訂にも注意を払うことが必要である。

文 献

- 1) Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials.

- BMJ 2013; 346: e7586.
- 2) Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c869.
 - 3) 厚生省医薬安全局審査管理課長通知. 臨床試験のための統計的原則. 平成 10 年 11 月 30 日付医薬審第 1047 号. <http://www.pmda.go.jp/> から入手可能, アクセス日 2018 年 10 月 1 日.
 - 4) Gamble C, Krishan A, Stocken D, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017; 318: 2337-43.
 - 5) International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. Updated December 2017. <http://www.icmje.org/> から入手可能, アクセス日 2018 年 10 月 1 日.
 - 6) International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Harmonised Guideline: Estimands and Sensitivity Analysis in Clinical Trials E9 (R1). Step 2 version dated 16 June 2017. <http://www.ich.org/> から入手可能, アクセス日 2018 年 10 月 1 日.

第 4 回 JSCTR 認定 GCP エキスパート® 資格更新者

[順不同]

須崎 友紀 (武田薬品工業株式会社)	延原あゆみ (富士フイルム富山化学株式会社 開発推進部)
若林 由美 (高知大学医学部附属病院 次世代医療創造センター)	内田 高子 (富士フイルム富山化学株式会社 開発推進部)
玖須さつき (横浜市立大学附属病院 次世代臨床研究センター)	武田 美和 (富士フイルム富山化学株式会社 開発推進部)
中村由美子 (横浜市立大学附属病院 次世代臨床研究センター)	工藤 隆道 (富士フイルム富山化学株式会社 臨床開発部)
金子美智子 (静岡県立静岡がんセンター 薬剤部)	平野 隆司 (国立長寿医療研究センター 薬剤部)
田邊 純生 (東京大学大学院医学系研究科 医学部 研究倫理支援室)	堀田 和男 (愛知医科大学病院 臨床研究支援センター/ 薬剤部)
亀田有希子 (北里大学北里研究所病院 研究部)	桑田 純子 (株式会社アレグロ)
中村 和美 (国立病院機構名古屋医療センター 臨床研究企画管理部)	中村 徹也 (小野薬品工業株式会社 開発業務部QC管理課)
福永 修司 (ネクスレージ株式会社)	櫻澤 有紀 (栃木県立がんセンター 薬剤部)
細野 弘朗 (富士フイルム富山化学株式会社 クリニカルサイエンス部)	表 正宏 (エヌ・エム・ピー ビジネスサポート株式会社 精密機器事業部)
廣山昭太郎 (富士フイルム富山化学株式会社 臨床開発部)	小野寺 崇 (東京大学医学部附属病院 臨床研究支援センター)
田中健太郎 (メドトロニックソファモアダネック株式会社)	榎本 裕介 (日本メドトロニック株式会社 メドトロニックコアテクニカルソリューションズ モニタリング部門)
平尾 知之 (富士フイルム富山化学株式会社 クリニカルサイエンス部)	横田 崇 (東北大学病院 臨床研究監理センター)