

産業界向けガイダンス

医薬品および生物学的製剤の臨床試験におけるベイズ流方法論の利用：ドラフトガイダンス

米国保健福祉省 食品医薬品局 (FDA)

医薬品評価研究センター (CDER)

生物製剤評価研究センター (CBER)

2026 年 1 月 生物統計学

訳：三島 遼、手良向 聡

(京都府立医科大学 大学院医学研究科生物統計学／附属病院臨床研究推進センター)

* 本翻訳は、FDA のウェブサイトに掲載された全文を訳出したものであり、翻訳掲載について FDA の許諾を必要としないものである。

Use of Bayesian Methodology in Clinical Trials of Drugs and Biologics Guidance for Industry¹	医薬品および生物学的製剤の臨床試験におけるベイズ流方法論の利用に関する産業界向けガイダンス¹
This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.	本ドラフトガイダンスは、最終化された際に、本トピックに関する食品医薬品局（FDA または当局）の現在の考え方を表すものである。これはいかなる者に対しても権利を確立するものではなく、FDA または一般市民を拘束するものでもない。適用される法令および規制の要件を満たす場合は、代替的な接近法を利用することができる。代替的な接近法について議論するには、表題ページに記載されている本ガイダンスに責任を負う FDA 職員に連絡されたい。
Footnotes: ¹ This guidance has been prepared by the Office of Biostatistics in the Center for Drug Evaluation and Research and the Division of Biostatistics in the Center for Biologics Evaluation and Research at the Food and Drug Administration.	脚注： ¹ 本ガイダンスは、食品医薬品局における医薬品評価研究センター・生物統計室および生物製剤評価研究センター・生物統計部によって作成された。

I. INTRODUCTION	I. 序論
This document provides guidance to sponsors and applicants submitting	本文書は、臨床試験におけるベイズ流の方法の適切な利用について、新医薬

investigational new drug applications (INDs), new drug applications (NDAs), biologics licensing applications (BLAs), or supplemental applications on the appropriate use of Bayesian methods in clinical trials. Bayesian methods can be used in various ways in clinical trials. For example, Bayesian calculations can be used to govern the timing and adaptation rules for an interim analysis in an adaptive design, to inform design elements (e.g., dose selection) for subsequent clinical trials, or to support primary inference in a trial. The primary focus of this guidance is on the use of Bayesian methods to support primary inference in clinical trials intended to support the effectiveness and safety of drugs.²	品の臨床試験実施申請（IND）、新薬申請（NDA）、生物学的製剤申請（BLA）、または補足申請を提出するスポンサーおよび申請者に対するガイダンスを提供する。ベイズ流の方法は、臨床試験において種々の方法で利用できる。例えば、ベイズ流の計算は、適応的デザインにおける中間解析の時期と適応規則を定めるため、後続の臨床試験のデザイン要素（例えば、用量選択）に情報を与えるため、あるいはある試験における主要な推論を支持するために利用できる。本ガイダンスの主な焦点は、医薬品の有効性および安全性を裏付けることを意図した臨床試験において、主要な推論を支持するためのベイズ流の方法の利用にある。²
In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word <i>should</i> in Agency guidances means that something is suggested or recommended, but not required.	一般に、FDA のガイダンス文書は法的に強制力のある責任を確立するものではない。その代わり、ガイダンスはあるトピックに関する当局の現在の考え方を示すものであり、特定の規制上または法令上の要件が引用されていない限り、単なる推奨として見做されるべきである。当局のガイダンスにおける「should（すべき）」という語の使用は、何かが提案または推奨されているという意味であり、必須であるという意味ではない。
Footnotes: ² For the purposes of this guidance, all references to drugs include both human drugs and biological products unless otherwise specified.	脚注： ² 本ガイダンスの目的のため、医薬品という呼称には、特に断りのない限り、ヒト用医薬品および生物学的製剤の両方が含まれる。

II. BACKGROUND	II. 背景
A. Definition	A. 定義
Bayesian statistics is an approach to estimation or hypothesis testing to draw inference based on the use of Bayes' theorem. In a Bayesian analysis, data collected in a study are combined with a prior distribution that captures the pre-study information about a parameter of interest to form a posterior distribution that expresses the updated, post-study information about the parameter of interest (e.g., the primary estimand³). The prior distribution often represents a summary of information and uncertainty available before the study begins. The posterior distribution can be used for inference and to draw conclusions about efficacy or safety.	ベイズ流統計学は、ベイズの定理の利用に基づいて推論を導くための推定または仮説検定に対する 1 つの接近法である。ベイズ流解析においては、関心のあるパラメータ（例えば、主要な estimand³）に関する更新された試験後情報を表す事後分布を形成するために、試験で収集されたデータが関心のあるパラメータに関する試験前情報を反映した事前分布と結合される。事前分布は、しばしば試験開始前に利用可能な情報および不確実性の要約を表す。事後分布は、推論のため、および効果または安全性に関する結論を導くために利用できる。
Footnote: ³ For discussion of estimands, see the ICH guidance for industry E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials (May 2021). For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-	脚注： ³ Estimand の議論については、ICH 産業界向けガイダンス E9(R1) 臨床試験のための統計的原則: 補遺: 臨床試験における estimand と感度分析（2021 年 5 月）を参照のこと。ガイダンスの最新版については、FDA ガイダンスウェブページ（https://www.fda.gov/regulatory-information/search-fda-guidancedocuments）を確認されたい。

[guidancedocuments.](#)

B. Important Concepts	B. 重要な概念
The following are definitions of important concepts used in this guidance:	以下は、本ガイダンスで用いられる重要な概念の定義である：
• The <i>prior distribution</i> or <i>prior</i> is the pre-study probability distribution for model parameters.	• 事前分布は、モデルパラメータに対する試験前の確率分布である。
• The <i>likelihood function</i> or <i>likelihood</i> describes the quantitative relationship between the parameters of interest and the study data. The mathematical form of the likelihood is determined by the model being used (for example, linear regression, logistic regression, ordinal regression).	• 尤度関数または尤度は、関心のあるパラメータと試験データとの間の定量的関係を記述する。尤度の数学的形式は、用いられるモデル（例えば、線形回帰、ロジスティック回帰、順序回帰）によって決定される。
• The <i>posterior distribution</i> or <i>posterior</i> is the post-study probability distribution for the parameter of interest. It is obtained by combining the prior distribution and the likelihood using Bayes theorem. It quantitatively summarizes what is known about the parameter of interest following collection of study data and can be used to draw inferences on the study hypotheses. Inference is often based on summary measures of this distribution. For example, evaluation of a treatment effect may be informed by the posterior mean to estimate the effect, a credible interval to quantify	• 事後分布は、関心のあるパラメータに対する試験後の確率分布である。これはベイズの定理を用いて事前分布と尤度を結合することによって得られる。これは、試験データの収集後に関心のあるパラメータについて分かっていることを定量的に要約し、試験の仮説に関する推論を導くために利用できる。推論は、しばしば当分布の要約指標に基づく。例えば、治療効果の評価は、効果を推定するための事後平均、推定された効果の周囲における不確実性を定量化するための信用区間、および関連する事後確率（例えば、効果がゼロより大きい事後確率）に基づいて行うことができる。信用区間は、事後

uncertainty around the estimated effect, and relevant posterior probabilities (e.g., the posterior probability that the effect is greater than zero). Credible intervals are intervals of possible values for the unobserved parameter that will contain the parameter value with a specified probability under the posterior distribution (e.g., 95% posterior probability).	分布の下で、特定の確率でパラメータ値を含む（例えば、95%事後確率）、未観測パラメータに対してとり得る値の区間である。
• <i>Bayes theorem</i> is the mathematical rule for combining the prior distribution and likelihood together to form the posterior distribution.	• ベイズの定理 は、事前分布と尤度を結合して事後分布を形成するための数学的規則である。

III. SITUATIONS WHERE BAYESIAN METHODS HAVE BEEN USED	III. ベイズ流の方法が利用されてきた状況
This section discusses settings and specific examples from development programs where Bayesian methods have been used in submissions to the Agency. Most of these case examples focus on the use of borrowing or leveraging of previously available trials or information across populations within a trial. Bayesian methods can also be considered in other settings.	本節では、当局への申請においてベイズ流の方法が利用されてきた設定および開発プログラムの具体例を論じる。これらの事例の多くは、既存の利用可能な試験あるいは単一試験内の集団にわたる情報の借用法または活用法の利用に焦点を当てている。ベイズ流の方法はその他の設定においても考慮することができる。
A. Borrowing from Previous Clinical Trials	A. 先行臨床試験からの借用
Under certain circumstances, an informative prior for a clinical trial analysis can be formed based on results from previous clinical trial(s) of the same drug. For example, borrowing from a previous trial was used in a phase 3	特定の状況下では、同じ医薬品の先行臨床試験から得られた結果に基づいて、臨床試験解析のための情報のある事前分布を形成することができる。例えば、先行試験からの借用が、再発性 <i>Clostridioides difficile</i> 感染症 (CDI)

study to evaluate REBYOTA, a fecal microbiota transplant product, for the prevention of recurrence of <i>Clostridioides difficile</i> infection (CDI) in individuals with recurrent CDI. <i>C. difficile</i> is a common cause of antibiotic-associated diarrhea and colitis and is a major public health burden. The primary analysis of the randomized, double-blind, placebo-controlled phase 3 study to evaluate the effectiveness of REBYOTA used a Bayesian model to formally incorporate data from a previous phase 2 placebo-controlled study of REBYOTA. This analysis supported the effectiveness of REBYOTA, which was approved in 2022.⁴	を有する個人における CDI 再発の予防に対して、糞便微生物叢移植製品である REBYOTA を評価する第 3 相試験において用いられた。<i>C. difficile</i> は抗菌薬関連下痢症および大腸炎の一般的な原因であり、公衆衛生上の重大な負担である。REBYOTA の有効性を評価するランダム化、二重盲検、プラセボ対照第 3 相試験の主要解析は、REBYOTA の先行第 2 相プラセボ対照試験から得られたデータを正式に組み込むためにベイズ流モデルを用いた。当解析は REBYOTA の有効性を支持し、それは 2022 年に承認された。⁴
Footnote: ⁴ See prescribing information for Rebyota (fecal microbiota, live – jslm) suspension (https://www.fda.gov/media/163587/download?attachment).	脚注： ⁴ Rebyota（糞便微生物叢、live – jslm）懸濁液の添付文書を参照のこと。 (https://www.fda.gov/media/163587/download?attachment)。

B. Augmenting a Randomized Concurrent Control Using an External Control or Nonconcurrent Control Data	B.外部対照または非同時対照データを用いたランダム化同時対照の補強
In some cases, it can be challenging to conduct an adequately powered randomized trial due to limited population and/or ethical considerations. Borrowing data from an external or nonconcurrent control to augment the randomized concurrent control may be appealing in these situations, and	場合によっては、限られた対象集団および／または倫理的考慮点のため、適切な検出力をもつランダム化試験を実施することが困難となるかもしれない。ランダム化同時対照を補強するために、外部対照または非同時対照から得られたデータを借用することは、これらの状況において魅力的である可能

Bayesian methods have been proposed to implement such approaches. For example, Bayesian methods have been proposed to augment the randomized concurrent control and leverage nonconcurrent control data in the oncology platform trials GBM AGILE⁵ and Precision Promise⁶, which evaluate marker-targeted treatments for patients with glioblastoma and pancreatic cancer, respectively. The analyses use a Bayesian model to try to account for temporal shifts in efficacy outcomes such as tumor response (Saville et al. 2022).⁷ As another example, a non-inferiority study of pediatric patients with multiple sclerosis⁸ was proposed through the Complex Innovative Trial Design (CID) program⁹ with a prespecified Bayesian analysis to leverage information from historical studies of the active comparator.	性があり、このような接近法を実装するためにベイズ流の方法が提案されてきた。例えば、膠芽腫および膵臓癌を有するそれぞれの患者に対するマーカー標的治療を評価する、がんプラットフォーム試験 GBM AGILE⁵ および Precision Promise⁶ において、ランダム化同時対照を補強し、非同時対照データを活用するためにベイズ流の方法が提案されている。これらの解析は、腫瘍反応といった効果アウトカムにおける時間的ずれを考慮するためにベイズ流モデルを用いている (Saville et al. 2022)。⁷ 別の例として、多発性硬化症⁸を有する小児患者の非劣性試験が、実薬比較の既存試験から得られた情報を活用するために事前規定されたベイズ流解析を伴う、Complex Innovative Trial Design (CID) プログラム⁹を通じて提案された。
Footnotes: ⁵ For additional details on GBM AGILE, see https://www.clinicaltrials.gov/study/NCT03970447. ⁶ For additional details on Precision Promise, see https://www.clinicaltrials.gov/study/NCT04229004. ⁷ See the FDA draft guidance for industry Master Protocols for Drug and Biological Product Development (December 2023) for additional discussion on the use of nonconcurrent control data in platform trials and potential for	脚注： ⁵ GBM AGILE の追加の詳細については、https://www.clinicaltrials.gov/study/NCT03970447 を参照のこと。 ⁶ Precision Promise の追加の詳細については、https://www.clinicaltrials.gov/study/NCT04229004 を参照のこと。 ⁷ プラットフォーム試験における非同時対照データの利用および時間的ずれがバイアスをもたらす可能性に関する追加の議論については、FDA 産業界向けドラフトガイダンス 医薬品および生物学的製剤開発のためのマスター

temporal shifts to lead to bias. When final, this guidance will represent the FDA's current thinking on this topic. ⁸ See CID Case Study: A Study in Pediatric Patients with Multiple Sclerosis. (https://www.fda.gov/media/172313/download). ⁹ Complex Innovative Trial Design Meeting Program (https://www.fda.gov/drugs/development-resources/complex-innovative-trial-design-meeting-program).	プロトコル（2023 年 12 月）を参照のこと。最終化された際、本ガイダンスは本トピックに関する FDA の現在の考え方を表すものとなる。 ⁸ CID 事例研究：多発性硬化症を有する小児患者における試験を参照のこと。（https://www.fda.gov/media/172313/download）。 ⁹ Complex Innovative Trial Design 会議プログラム（https://www.fda.gov/drugs/development-resources/complex-innovative-trial-design-meeting-program）。
---	--

C. Pediatric Extrapolation	C. 小児外挿
Pediatric extrapolation is defined in the ICH E11(R1)¹⁰ guideline as “an approach to providing evidence in support of effective and safe use of drugs in the pediatric population when it can be assumed that the course of the disease and the expected response to a medicinal product would be sufficiently similar in the pediatric [target] and reference (adult or other pediatric) population.”	小児外挿は、ICH E11(R1)¹⁰ ガイドラインにおいて、「疾患の経過および医薬品に対して期待される反応が小児（標的）集団と参照（成人または他の小児）集団において十分に類似していると仮定できる場合に、小児集団における医薬品の有効かつ安全な使用を支持する証拠を提供する接近法」として定義されている。
Pediatric extrapolation can extend what is known about the characteristics of interest in the reference population (e.g., efficacy, safety, and/or dosing) to those of the target population based on an assessment of the relevant	小児外挿は、これら 2 つの集団の疾患、薬物薬理学、および療法への反応の関連する類似性の評価に基づいて、参照集団における関心のある特性（例えば、効果、安全性、および／または投与量）について知られていることを標

similarities of disease, drug pharmacology, and response to therapy of the two populations. When some degree of pediatric extrapolation is justified, Bayesian methods can be considered to borrow data from adults in the analysis of a pediatric trial by using an informative prior distribution constructed based on results from previous adult trials. A discussion of safety considerations incorporated into pediatric extrapolation approaches is discussed in other guidance.¹¹ An example of a Bayesian approach to facilitate borrowing can be seen in supportive analyses in recent supplements for empagliflozin¹² and linagliptin¹³ for the treatment of pediatric patients with type 2 diabetes mellitus (T2D). It is critical to consider the relevance¹⁴ of the information from adults when considering borrowing. In these particular cases, the review team concluded that although there are differences in disease progression between pediatric and adult T2D populations, the pathophysiology of pediatric T2D is similar to that in adults and so the information was relevant, and borrowing was justified.	的集団のそれらに拡張することができる。ある程度の小児外挿が正当化されるとき、ベイズ流の方法は、先行成人試験から得られた結果に基づいて構築された情報のある事前分布を用いることによって、小児試験の解析において成人から得られたデータの借用を考慮することができる。小児外挿接近法に組み込まれた安全性の考慮点についての考察は、他のガイダンスにおいて論じられている。¹¹ 借用を促進するベイズ流接近法の例は、2 型糖尿病 (T2D) を有する小児患者の治療のためのエンパグリフロジン¹² およびリナグリプチン¹³ に対する最近の補助資料における補足的解析に見ることができる。借用を考える際には、成人からの情報の関連性¹⁴ を考慮することが極めて重要である。これら特有の事例において、審査チームは、小児と成人の T2D 集団間で疾患進行に違いがあるものの、小児 T2D の病態生理は成人におけるそれと類似しており、よってそれらの情報は関連があり、借用は正当化されると結論付けた。
Footnotes: ¹⁰ See ICH guidance for industry E11(R1) Addendum: Clinical Investigation of Medicinal Products in the Pediatric Population (April 2018). ¹¹ For further discussion on pediatric extrapolation, see ICH guidance for	脚注： ¹⁰ ICH 産業界向けガイダンス E11(R1) 補遺：小児集団における医薬品の臨床試験（2018 年 4 月）を参照のこと。 ¹¹ 小児外挿に関するさらなる議論については、ICH 産業界向けガイダンス

industry E11A Pediatric Extrapolation (December 2024). ¹² See FDA Clinical Review (https://www.fda.gov/media/172973/download); FDA Statistical Review: (https://www.fda.gov/media/172972/download). ¹³ See FDA Clinical Review (https://www.fda.gov/media/172628/download); FDA Statistical Review: (https://www.fda.gov/media/172630/download). ¹⁴ See section V.D.2 for a discussion of factors that impact the relevance of the data.	E11A 小児用医薬品開発における外挿（2024 年 12 月）を参照のこと。 ¹² FDA 臨床審査報告書 (https://www.fda.gov/media/172973/download) ; FDA 統計審査報告書 : (https://www.fda.gov/media/172972/download) を参照のこと。 ¹³ FDA 臨床審査報告書 (https://www.fda.gov/media/172628/download) ; FDA 統計審査報告書 : (https://www.fda.gov/media/172630/download) を参照のこと。 ¹⁴ データの関連性に影響を与える因子の議論については V.D.2 節を参照のこと。
---	---

D. Borrowing Information Across Similar Diseases or Disease Subtypes	D. 類似疾患または疾患亜型間での情報借用
In some cases, distinct diseases or disease subtypes may have similar underlying causes and a history of similar responses to drugs. For example, there are groups of different types of cancer that share a specific molecular alteration and may be expected to respond to drugs targeting that alteration. In such cases, Bayesian methods might be considered to borrow information across the similar diseases or disease subtypes in evaluating an individual drug. For example, Bayesian analyses have been proposed for leveraging	場合によっては、異なる疾患または疾患亜型は、類似した根底にある原因および医薬品に対する類似した反応の既往を有する可能性がある。例えば、特定の分子変化を共有し、その変化を標的とする医薬品に反応することが期待され得る異なる型のがんのグループが存在する。このような場合、個別の医薬品を評価する際に、類似した疾患または疾患亜型間で情報を借用するためにベイズ流の方法が考慮される可能性がある。例えば、共通のマスタープロトコルの下で、複数の疾患または疾患亜型に対して医薬品を評価するバスケ

information about drug effects across related populations in basket trials that evaluate a drug for multiple diseases or disease subtypes under a common master protocol. As another example, a randomized, double-blind, placebo-controlled study in patients with epilepsy with myoclonic-atonic seizures¹⁵ (EMAS) was proposed through the CID program that used a Bayesian primary analysis to borrow information from previous trials evaluating the effect of the drug in related conditions. The proposed approach leveraged data from previously conducted trials for different types of epilepsy using a Bayesian hierarchical model (BHM).	ット試験において、関連する集団間で医薬品の効果に関する情報を活用するためにベイズ流解析が提案されてきた。別の例として、ミオクロニー脱力発作を伴うてんかん¹⁵ (EMAS) 患者におけるランダム化、二重盲検、プラセボ対照試験が、関連する病態における医薬品の効果を評価する先行試験から得られた情報を借用するために、ベイズ流主要解析を用いる CID プログラムを通じて提案された。提案された接近法は、ベイズ流階層モデル (BHM) を用いて、異なる型のでんかんに対して以前に実施された試験から得られたデータを活用した。
Footnote: ¹⁵ See CID Case Study: A Study in Patients with Epilepsy with Myoclonic-Atonic Seizures (https://www.fda.gov/media/172312/download).	脚注： ¹⁵ CID 事例研究：ミオクロニー脱力発作を伴うてんかん患者における試験を参照のこと (https://www.fda.gov/media/172312/download)。

E. Borrowing Information Between Subgroups of a Patient Population (i.e., Subgroup Analysis)	E. 患者集団の部分集団間での情報借用（すなわち、部分集団解析）
It is important to try to understand drug effects in different subgroups of patients. There are statistical approaches that make use of results from every subgroup when estimating the treatment effect in each subgroup. One common approach is shrinkage estimation through a BHM. For a one-way	患者の異なる部分集団における医薬品の効果を理解しようとすることは重要である。各部分集団での治療効果を推定する際、すべての部分集団から得られた結果を利用する統計的接近法が存在する。1つの一般的な接近法は BHM を通じた縮小推定である。一方向 BHM の場合、ある部分集団での推

BHM, the estimated treatment effect in one subgroup is a weighted average of its raw estimated treatment effect (using only the data in that subgroup) and the overall estimated treatment effect. Shrinkage estimation can increase the precision of the subgroup estimates. As an example of this approach, a BHM was used to estimate treatment effects across regions in the Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results trial, which compared liraglutide to placebo in patients with T2D at high risk for cardiovascular disease. The analyses helped clarify potential differences in the drug effects across regions (i.e., Asia, Europe, North America, and the rest of the world).¹⁶ BHMs have also been used for subgroup analyses that appear in some of FDA's Drug Trials Snapshots (Wang et al. 2024). One example is the Rinvoq Drug Trial Snapshot.¹⁷	定治療効果は、その未調整の推定治療効果（その部分集団のデータのみを使用）と全体の推定治療効果との重み付き平均となる。縮小推定は部分集団推定値の精度を向上させることができる。当接近法の例として、心血管系疾患の高いリスクを有する T2D 患者においてリラグルチドとプラセボを比較した、Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results 試験において、地域にわたる治療効果を推定するために BHM が用いられた。その解析は、地域（すなわち、アジア、欧州、北米、および世界の他の地域）にわたる医薬品の効果の潜在的な違いを明らかにするのに役立った。¹⁶ BHM は、FDA の Drug Trials Snapshots のいくつかに現れる部分集団解析にも用いられてきた (Wang et al. 2024)。1 つの例は Rinvoq Drug Trial Snapshot である。¹⁷
Footnotes: ¹⁶ Additional discussion on the use of shrinkage estimation and BHMs for subgroup analysis, and additional details on the BHM model and results in the liraglutide trial, can be found in an FDA impact story: (https://www.fda.gov/drugs/regulatory-science-action/impact-story-using-innovative-statistical-approaches-provide-most-reliable-treatment-outcomes). ¹⁷ Drug Trials Snapshots: RINVOQ Accessed August 29, 2024	脚注： ¹⁶ 部分集団解析のための縮小推定および BHM の利用に関する追加の議論、ならびにリラグルチド試験における BHM および結果に関する追加の詳細は、FDA インパクトストーリーに見出すことができる： (https://www.fda.gov/drugs/regulatory-science-action/impact-story-using-innovative-statistical-approaches-provide-most-reliable-treatment-outcomes)。 ¹⁷ Drug Trials Snapshots: RINVOQ 2024 年 8 月 29 日にアクセスした

(https://www.fda.gov/drugs/drug-approvals-and-databases/drug-trials-snapshots-rinvoq).	(https://www.fda.gov/drugs/drug-approvals-and-databases/drug-trials-snapshots-rinvoq)。
--	--

F. Dose-Finding Trials in Oncology	F. がんにおける用量設定試験
Dose-finding trials for oncology drugs¹⁸ have historically utilized non-randomized dose-escalation trials that seek to identify the maximum tolerated dose (MTD). Dose-escalation designs using Bayesian methods have been proposed with goals such as improving efficiency (e.g., reaching the MTD earlier), optimizing dose selection (i.e., minimizing toxicity and/or improving efficacy), and adding flexibility in terms of cohort sizes and timing of assessments. Designs which aim to identify the MTD include model-based designs (e.g., continual reassessment model [CRM], Bayesian logistic regression model [BLRM]) and model-assisted designs (e.g., Bayesian Optimal Interval Design [BOIN], modified toxicity probability interval [mTPI], mTPI2) (Ji et al. 2010; Quigley and Conway, 2010; Neuenschwander et al., 2008; Yuan et al., 2016;¹⁹ Tighiouart and Rogatko, 2010). Although identifying the MTD has been the traditional paradigm for oncology drug development, for modern targeted therapies, such as kinase inhibitors and antibodies, identifying optimized dosage(s) based on alternative approaches may be more appropriate than selecting the MTD for	抗がん薬の用量設定試験¹⁸は、歴史的に、最大耐容用量 (MTD) を同定しようとする非ランダム化用量漸増試験を利用してきた。ベイズ流の方法を用いる用量漸増デザインは、効率を改善する (例えば、より早く MTD に到達する)、用量選択を最適化する (すなわち、毒性を最小にする、および/または効果を改善する)、およびコホートの大きさと評価の時期に関して柔軟性を加えるといった目標を伴って提案されてきた。MTD の同定を目的とするデザインには、モデル基盤型デザイン (例えば、連続再評価法 [CRM]、ベイズ流ロジスティック回帰モデル [BLRM]) およびモデル支援型デザイン (例えば、ベイズ流最適区間デザイン [BOIN]、改変型毒性確率区間 [mTPI]、mTPI2) が含まれる (Ji et al. 2010; Quigley and Conway, 2010; Neuenschwander et al., 2008; Yuan et al., 2016;¹⁹ Tighiouart and Rogatko, 2010)。MTD の同定は抗がん薬開発の伝統的なパラダイムであったが、キナーゼ阻害剤および抗体といった最新の標的療法に対しては、さらなる開発のために MTD を選択することよりも、代替的な接近法に基づいて最適化された用量を同定することの方がより適切であるかもしれない。²⁰ このような用量の同定を目的とするベイズ流デザインが、がんにおける早期相試験に対

further development. ²⁰ Bayesian designs with the aim of identifying such dosages have been proposed for early-phase trials in oncology (Thall and Cook, 2004; Lin et al. 2020).	して提案されてきた (Thall and Cook, 2004; Lin et al. 2020)。
Footnotes: ¹⁸ See FDA guidance for industry Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases (August 2024). ¹⁹ The BOIN design has received the fit-for-purpose designation for phase 1 dose-finding cancer trials: (https://www.fda.gov/media/155363/download). ²⁰ See footnote 17.	脚注： ¹⁸ FDA 産業界向けガイダンス 腫瘍性疾患の治療のためのヒト用医薬品および生物学的製剤の用量の最適化（2024 年 8 月）を参照のこと。 ¹⁹ BOIN デザインは、第 1 相用量設定がん試験に対する目的適合性の指定を受けた： (https://www.fda.gov/media/155363/download) 。 ²⁰ 脚注 17 を参照のこと。

IV. SUCCESS CRITERIA AND OPERATING CHARACTERISTICS	IV. 成功規準および動作特性
A. Success Criteria: Definition and Role in Regulatory Decision-making	A. 成功規準：定義および規制上の意思決定における役割
Clinical trial design includes pre-specification of criteria for determining whether the primary objectives of the trial have been met. In clinical trials intended to support the effectiveness and safety of drugs, success criteria of this type are useful as a goal to shape other design characteristics, such as	臨床試験デザインには、試験の主要目的が達成されたかどうかを判断するための規準の事前規定が含まれる。医薬品の有効性および安全性を裏付けることを意図した臨床試験において、この種の成功規準は、標本サイズおよび検出力といった他のデザイン特性を形作るための目標として役立つ。このよう

sample size and power. Such criteria serve as a point of discussion between FDA and a sponsor on whether trial results could contribute to substantial evidence of effectiveness. Carefully chosen success criteria are important to trial interpretability and efficiency.	な規準は、試験結果が有効性の実質的な証拠に寄与し得るかどうかについて、FDA とスポンサーとの間の論点として機能する。慎重に選定された成功規準は、試験の解釈可能性および効率性にとって重要である。
In clinical trials intended to support effectiveness and safety that are conducted with an overall frequentist statistical analysis plan, the efficacy success criteria are almost always chosen in such a way that the familywise Type I error rate (FWER) across all primary and secondary estimands is no greater than 0.025, one-sided. In the case of a trial with a single primary estimand, this means performing a statistical hypothesis test at a one-sided significance level of 0.025. For trials with multiple primary and/or secondary endpoints, the criteria can become more complex.²¹	全体的に頻度流の統計解析計画で実施される、有効性と安全性を裏付けることを意図した臨床試験において、効果の成功規準は、ほとんど常に、すべての主要および副次 estimand にわたるファミリーワイズ第 I 種の過誤確率 (FWER) が片側 0.025 を超えないように選定される。単一の主要 estimand を伴う試験の場合、これは片側有意水準 0.025 で統計的仮説検定を行うことを意味する。複数の主要および／または副次評価項目を伴う試験では、当規準はより複雑になり得る。²¹
In clinical trials with Bayesian inference for the primary estimand, this default success criterion may not be applicable or appropriate, such as in the case where there is borrowing of information (see Sections V.D and V.E), so careful specification of alternative success criteria is often critical when using Bayesian analyses. For these Bayesian approaches, specification of a success criterion is most often based on the posterior probability that the true treatment effect size exceeds some threshold. In mathematical notation, such a criterion might take the form $Pr(d > a) > c$, where d is a population-level summary of the size of the treatment effect, a is a minimum threshold for the	主要 estimand に対してベイズ流推論を用いる臨床試験において、情報の借用がある場合 (V.D 節および V.E 節を参照) などには、この既定の成功規準が適用できない、あるいは適切でない可能性があり、よってベイズ流解析を用いるときには、しばしば代替の成功規準の慎重な規定が極めて重要である。これらのベイズ流接近法に対して、成功規準の規定は、ほとんどの場合、真の治療効果の大きさがある閾値を超える事後確率に基づく。数学的記法では、このような規準は $Pr(d > a) > c$ という形式をとる可能性があり、ここに、d は治療効果の大きさの集団水準の要約であり、a は治療効果が有益であると見做される最小の閾値であり、c は有効性の結論を支持する最小

treatment effect to be considered beneficial, and c is a minimum probability level that would support a conclusion of effectiveness. (In some cases, the criterion may be $Pr(d < a) > c$ instead, if lower values of d reflect greater benefit.) Choice of a success criterion of this kind thus means choice of specific values for a and for c. There are a variety of approaches to specifying these thresholds for Bayesian analyses. The choice of which approach to use depends on the trial objectives and specific Bayesian methods used.	の確率水準である（場合によっては、より低いdの値がより大きな有益性を反映する場合、その規準は、代わりに$Pr(d < a) > c$となるかもしれない）。それ故に、この種の成功規準の選定は、aおよびcに対する特定の値の選定を意味する。ベイズ流解析に対してこれらの閾値を規定する様々な接近法が存在する。どの接近法を用いるかに関する選定は、試験目的および利用される特定のベイズ流の方法に依存する。
Footnote: ²¹ See FDA guidance for industry Multiple Endpoints in Clinical Trials (October 2022).	脚注： ²¹ FDA 産業界向けガイダンス 臨床試験における多重エンドポイント（2022年10月）を参照のこと。

1. Calibration to Type I Error Rate	1. 第 I 種の過誤確率に対する校正
For some Bayesian approaches, a and c can be chosen such that the overall FWER is controlled at a given level, typically 0.025 one-sided. This is referred to as calibrating the success criteria to Type I error rate. Such an approach may be appropriate for designs where Bayesian approaches are used not to synthesize multiple information sources, but instead to facilitate complex adaptive designs. Calibration to Type I error rate also may be useful in designs with noninformative prior distributions that express a lack of prior	いくつかのベイズ流接近法に対しては、全体的な FWER が所与の水準、一般的に片側 0.025 で制御されるように、aおよびcを選定することが可能である。これは、第 I 種の過誤確率に対して成功規準を校正することと呼ばれる。このような接近法は、ベイズ流の接近法が、複数の情報源を統合するためではなく、複雑な適応的デザインを促進するために用いられるデザインに対しては適切かもしれない。第 I 種の過誤確率に対する校正は、解析に関連する事前情報の欠如を表現する無情報事前分布を用いるデザインにおいても

information relevant to the analysis.	有用かもしれない。
When calibrating Bayesian success criteria to Type I error rate, a is chosen to be 0 for superiority designs or is based on the non-inferiority margin, m , for non-inferiority designs. Determining the appropriate choice of c to control FWER then becomes a computational problem only. For the simplest Bayesian approaches, including some designs with noninformative and/or conjugate prior distributions, c can be derived algebraically. For more complicated designs, including complex adaptive designs, c is instead approximated using clinical trial simulations. ²²	ベイズ流の成功規準を第 I 種の過誤確率に対して較正する際、優越性デザインでは a は 0 と選定され、非劣性デザインでは a は非劣性マージン m に基づく。このとき、FWER を制御するために適切な c の値を決定することは、単なる計算上の問題となる。無情報および／または共役事前分布を用いるデザインを含む、最も単純なベイズ流接近法については、 c を代数的に導出することができる。複雑な適応的デザインを含む、より複雑なデザインについては、その代わりに、 c は臨床試験シミュレーションを用いて概算される。 ²²
Footnote: ²² See FDA guidance for industry Adaptive Designs for Clinical Trials of Drugs and Biologics (December 2019).	脚注: ²² FDA 産業界向けガイダンス 医薬品および生物学的製剤の臨床試験のための適応的デザイン（2019 年 12 月）を参照のこと。

2. Direct Interpretation of Posterior Probability	2. 事後確率の直接的解釈
In Bayesian approaches where the prior distribution has been chosen to provide an accurate summary of the state of belief based on existing information before the trial begins, decision-making can be based on direct interpretation of the posterior probability distribution itself. With a prior	事前分布が、試験開始前の既存情報に基づく信念の状態の正確な要約を提供するように選ばれているベイズ流接近法において、意思決定は、事後確率分布自体の直接的解釈に基づくことができる。このように選ばれた事前分布を用いた場合、事後確率 $Pr(d > a) = c$ であれば、治療効果が a より小さい確率

chosen in this way, if the posterior probability $Pr(d > a) = c$ then the probability that the treatment effect is less than a is less than $1 - c$. For example, if the posterior probability of effectiveness is 0.98, the posterior probability that the treatment is ineffective is 0.02. The choice of success criteria can then be based on a determination of whether a $1 - c$ chance of ineffectiveness is sufficiently small in the specific context.	は $1 - c$ より小さい。例えば、有効性の事後確率が 0.98 であれば、治療が無効である事後確率は 0.02 となる。成功規準の選定は、このとき、 $1 - c$ という無効性の可能性が、特定の文脈において十分に小さいかどうかという決定に基づることができる。
This approach can be appropriate in Bayesian analyses that explicitly leverage external data sources to support decision-making, though it is not limited to that setting. In such scenarios, it is critical that the prior be specified in a way that accurately and comprehensively reflects the external data to ensure credible conclusions.	当接近法は、意思決定を支持するために明示的に外部データ源を活用するベイズ流解析において適切であり得るが、その設定に限定されるものではない。このようなシナリオにおいて、信用できる結論を保証するためには、事前分布が外部データを正確に、かつ包括的に反映する方法で規定されていることが極めて重要である。

3. Success Criteria Based on Benefit-Risk Assessment or Decision-Theoretic Approaches	3. ベネフィット・リスク評価または決定理論的接近法に基づく成功規準
Another general approach specifies success thresholds in a broader context incorporating product risk or additional considerations. The simplest form of such an approach would adopt a choice for the threshold a that ensures sufficient benefit to outweigh known or potential product risks.	別の一般的な接近法は、製品のリスクまたは追加の考慮点を組み込んだ上で、より広範な文脈において成功閾値を規定する。このような接近法の最も単純な形式は、製品の既知または潜在的なリスクを上回るのに十分なベネフィットを保証する閾値 a の選定を採用することである。
More complex approaches might include consideration of external factors	より複雑な接近法は、疾患の重篤性または他の承認済み療法の利用可能性と

such as seriousness of a disease or availability of other approved therapies. A decision-theoretic approach might include assessment of the potential negative consequences of approving an ineffective drug or of not approving an effective drug. The statistical quantification of such negative consequences is sometimes referred to as “loss”. One approach to incorporating such information is to form success criteria that minimize the expected loss. The loss function can include safety as well as effectiveness considerations to incorporate benefit-risk assessment into the formal decision-making process.	いった外的要因の考慮を含む可能性がある。決定理論的接近法は、無効な医薬品を承認すること、あるいは有効な医薬品を承認しないことの潜在的な負の帰結の評価を含むことが可能である。このような負の帰結の統計的定量化は、「損失」と呼ばれることがある。このような情報を組み込むための1つの接近法は、期待損失を最小にする成功規準を形成することである。損失関数は、ベネフィット・リスク評価を正式な意思決定過程に組み込むために、有効性の考慮点だけでなく安全性も含むことができる。
---	---

4. Additional Considerations	4. 追加の考慮点
In trials that include interim decision-making, such as group sequential designs, success criteria should be specified for each decision point. In cases where success criteria are calibrated to Type I error rate, interim success criteria can be specified to ensure overall control of FWER across all decision-points for effectiveness.	グループ逐次デザインといった中間での意思決定を含む試験において、成功規準は各決定時点に対して規定されるべきである。成功規準が第I種の過誤確率に対して較正されている場合、中間での成功規準は、有効性に関するすべての決定時点にわたる FWER の全体的な制御を保証するように規定することができる。
Success criteria should be specified for all primary and key secondary endpoints that use Bayesian approaches. In those trials with success criteria calibrated to Type I error rate, clinical trial simulations may be required to ensure control of FWER across all endpoints. Sample size can be calculated to	成功規準は、ベイズ流接近法を用いるすべての主要評価項目および重要な副次評価項目に対して規定されるべきである。第I種の過誤確率に対して較正された成功規準を用いた試験においては、すべての評価項目にわたる FWER の制御を保証するために、臨床試験シミュレーションが必要とされ

achieve the desired statistical power under the controlled FWER.	るかもしれない。標本サイズを、制御された FWER の下で望ましい統計的検出力を達成するように計算することができる。
--	--

B. Operating Characteristics	B. 動作特性
In the design of a clinical trial, it is important to understand how the trial is likely to perform in terms of supporting correct conclusions and reliable estimation of treatment effects. In trials with frequentist inference, the most important operating characteristics related to hypothesis tests (long-run expectations of trial conclusions under assumptions about true parameter values) are the FWER and the power for the primary and other key endpoints. These quantities are fixed by design before the trial (conventionally, at 0.025 one-sided FWER and 80% or 90% power). This is possible because frequentist inference is based on the conditional probability of observing certain data given fixed assumptions about parameter values, and these probabilities can be calculated pre-test. The most important operating characteristics related to estimation in the frequentist paradigm are bias and mean squared error (MSE) of point estimates, and coverage probability and width of confidence intervals.	臨床試験のデザインにおいて、正しい結論を支持し、信頼性の高い治療効果の推定を行うという観点から、試験がどのような性能を示す可能性が高いかを理解することが重要である。頻度流推論を用いる試験においては、仮説検定に関連する最も重要な動作特性（真のパラメータ値に関する仮定の下での試験の結論の長期的な期待値）は、主要評価項目および他の重要な評価項目に対する FWER と検出力である。これらの量は、試験前にデザインによって固定される（慣例的には、片側 0.025 の FWER および 80% または 90% の検出力で）。これが可能なのは、頻度流推論は、パラメータ値に関する固定された仮定を与えた下で特定のデータを観測する条件付き確率に基づいており、これらの確率は検定を行う前に計算できるからである。頻度流パラダイムにおける推定に関連する最も重要な動作特性は、点推定値のバイアスおよび平均二乗誤差 (MSE)、ならびに信頼区間の被覆確率および幅である。
The situation is somewhat different in trials with Bayesian inference because	ベイズ流推論は事後分布に基づいていることから、ベイズ流推論を用いる試

Bayesian inference is based on the posterior distribution. Operating characteristics of the design and analysis therefore depend on both the prior distribution and the observed data. It is also important to note that the concept of a false positive conclusion in a Bayesian framework is conditional on a positive conclusion and not, as in the frequentist framework, on a true null hypothesis.	験においては、その状況は幾分異なる。したがって、デザインおよび解析の動作特性は、事前分布と観測データの両方に依存する。また、ベイズ流の枠組みにおける偽陽性の結論という概念は、肯定的な結論で条件付けるものであり、頻度流の枠組みでのように真の帰無仮説で条件付けるものではないことに留意することが重要である。
Some trials, including certain complex adaptive design trials, employ Bayesian analysis in an overall frequentist inferential framework (i.e., with calibration to Type I error rate control). In these cases, frequentist operating characteristics are of interest. Other trials, including most trials in which external information is incorporated into an informative prior distribution, use a Bayesian inferential framework that calls for a different approach to quantifying design characteristics. These two cases are discussed in the next two subsections.	ある種の複雑な適応的デザイン試験を含む、いくつかの試験は、全体的な頻度流推論の枠組みにおいてベイズ流解析を採用する（すなわち、第Ⅰ種の過誤確率制御に対する較正を伴う）。これらの場合には、頻度流の動作特性が関心の対象である。外部情報が情報のある事前分布に組み込まれている試験の多くを含む、その他の試験は、デザイン特性の定量化に対する異なる接近法を必要とするベイズ流推論の枠組みを用いる。これらの2つの場合については、次の2つの小節において議論する。

1. Trials Calibrated to Type I Error Rate	1. 第Ⅰ種の過誤確率に対して較正された試験
For trial designs that calibrate Bayesian results to Type I error rate, the primary operating characteristics are the same as those described above for trials with frequentist inference (that is, Type I error rate and power for	第Ⅰ種の過誤確率に対してベイズ流の結果を較正する試験デザインにとって、主要な動作特性は、頻度流推論を用いる試験に対して上述したのと同じである（すなわち、検定については第Ⅰ種の過誤確率と検出力；推定につ

testing; bias and MSE of point estimates, and coverage probability and width of intervals, for estimation). Clinical trial simulations are generally used to estimate or demonstrate control of operating characteristics. Briefly, a large number of simulated trials, conditional on a chosen prior distribution and sample size, are generated under the assumption that the null hypothesis is true or that an alternative hypothesis is true. The proportions of simulated trials in which the null hypothesis is rejected is then used to estimate Type I error rate and power, respectively. An iterative simulation process is often used in which various sample sizes, prior distributions, success criteria, and other design elements (e.g., interim analysis boundaries) are adjusted to obtain desired operating characteristics. Simulations should comprehensively cover the plausible range of assumptions. This includes assumptions about statistical parameters such as the variance or background rate of the endpoint or operational parameters such as the accrual rate.²³	いては点推定値のバイアスと MSE、および区間の被覆確率および幅)。一般に臨床試験シミュレーションは、動作特性を推定するため、あるいはその制御を示すために用いられる。簡潔に言うと、選定された事前分布と標本サイズを条件付けて、帰無仮説が真である、あるいは対立仮説が真であるという仮定の下で、多数のシミュレートされた試験が生成される。このとき、帰無仮説が棄却されるシミュレートされた試験の割合は、それぞれ第 I 種の過誤確率および検出力を推定するために用いられる。望ましい動作特性を得るために、種々の標本サイズ、事前分布、成功規準、および他のデザイン要素（例えば、中間解析での境界値）が調整される反復的なシミュレーション過程が用いられることがよくある。シミュレーションは、尤もらしい範囲の仮定を包括的に含めるべきである。これには、評価項目の分散または背景率といった統計的パラメータ、あるいは集積率といった運営上のパラメータに関する仮定が含まれる。²³
Footnote: ²³ For additional considerations on these and other aspects of simulations, see Section VI.A of the FDA guidance for industry Adaptive Designs for Clinical Trials of Drugs and Biologics (December 2019).	脚注： ²³ これらおよび他の側面のシミュレーションに関する追加の考慮点については、FDA 産業界向けガイダンス 医薬品および生物学的製剤の臨床試験のための適応的デザイン（2019 年 12 月）の VI.A 節を参照のこと。

2. Trials Not Calibrated to Type I Error Rate	2. 第 I 種の過誤確率に対して較正されていない試験
In cases where a sponsor and FDA agree that a study design does not need to be calibrated to the Type I error rate, the accuracy of conclusions depends strongly on the choice of prior distribution, as well as many of the same features required in frequentist inference: an adequately fitting data model, an appropriate experimental design, and accurate measurements. Accordingly, design characteristics are calculated in reference to the prior distribution. For example, Bayesian power is the probability of meeting the success criterion, averaged over a prior distribution (Spiegelhalter et al., 2004). The sample size of a trial is chosen to achieve a desired Bayesian power conditional on a chosen prior distribution and other study design features. Another relevant operating characteristic is the probability of a correct decision (akin to calculating the positive predictive value for a diagnostic test) corresponding to a chosen prior. For example, simulations can be used to calculate the proportion of trials where a positive treatment effect was present, from among those trials that concluded effectiveness. For estimation of treatment effects, relevant operating characteristics include the expected bias and expected MSE of point estimates averaged across a prior. Similarly, the expected coverage probability or width of the corresponding credible interval can be assessed (Adcock, 1988; Joseph and Bélisle, 1997).	スポンサーと FDA が、第 I 種の過誤確率に対して試験デザインを較正する必要がないことに合意する場合、結論の確度は事前分布の選定に強く依存し、それは頻度流推論に要求される多くの同じ特徴（すなわち、適切なデータモデルのあてはめ、適切な試験デザイン、および正確な測定）にも依存する。したがって、デザイン特性は事前分布を参照して計算される。例えば、ベイズ流検出力は、事前分布にわたって平均化された、成功規準を満たす確率である (Spiegelhalter et al., 2004)。試験の標本サイズは、選定された事前分布および他の試験デザインの特徴を条件付けて、望ましいベイズ流検出力を達成するために選ばれる。別の関連する動作特性は、選定された事前分布に対応する正しい決定の確率である (診断検査に対して陽性予測値を計算するのと同様に)。例えば、有効性を結論付けた試験の中から、肯定的な治療効果が認められた試験の割合を計算するために、シミュレーションを利用することができる。治療効果の推定に対して、関連する動作特性には、事前分布にわたって平均化された点推定値の期待バイアスおよび期待 MSE が含まれる。同様に、対応する信用区間の期待被覆確率または期待幅を評価することができる (Adcock, 1988; Joseph and Bélisle, 1997)。
It is always critical for Bayesian analyses to have a prespecified prior, which is	一般的に解析事前分布と呼ばれる、事前規定された事前分布を有しているこ

typically called the analysis prior. In a hypothetical situation where the choice of a prior distribution was unambiguous and clear, Bayesian power and other quantities could be calculated in reference to the same prior distribution that will be used in the final study analysis. In practice, however, there will usually be a range of plausible design priors (sometimes referred to as sampling priors), separate from the analysis prior, that are used as the basis for calculating study design characteristics. In simulation studies, the design prior serves as the prior from which parameter values are drawn, whereas the analysis prior is the prior that is used in the subsequent analysis of the data generated. An example of a design prior is a prior distribution on the treatment effect centered around the minimum clinically important difference (MCID) to evaluate Bayesian power. When the design prior is intended to explore scenarios corresponding only to an efficacious treatment, the design prior might be limited to an interval or a point mass indicating that the treatment is effective. Design priors that explore pessimistic assumptions about treatment effect should also be considered.	とは、ベイズ流解析にとって常に極めて重要である。事前分布の選定が曖昧でなく明確であるという仮想的な状況において、ベイズ流検出力およびその他の量は、試験の最終解析で用いられるのと同じ事前分布を参照して計算することができる。ただし、実践では、解析事前分布とは別個の、試験デザイン特性を計算するための基礎として用いられる、様々な尤もらしいデザイン事前分布（標本抽出事前分布と呼ばれることもある）が通常存在する。シミュレーション研究において、デザイン事前分布は、パラメータ値が抽出される事前分布として機能するのに対し、解析事前分布は、後続の生成されたデータの解析に用いられる事前分布である。デザイン事前分布の一例は、ベイズ流検出力を評価するために、臨床的に重要な最小の差 (MCID) を中心とした治療効果に関する事前分布である。デザイン事前分布が効果のある治療に対応するシナリオのみの探索を意図している場合、デザイン事前分布は、治療が有効であることを指し示す区間または点質量に制限される可能性がある。治療効果に関する悲観的な仮定を探索するデザイン事前分布も考慮されるべきである。
Operating characteristics can be calculated under various plausible design priors. Differences in design characteristics corresponding to different design priors can be used to quantify the sensitivity of the design to the choice of prior or to demonstrate that the probability of making an incorrect decision is very low even when the design and analysis priors do not match. In general,	動作特性は、種々の尤もらしいデザイン事前分布の下で計算することができる。異なるデザイン事前分布に対応するデザイン特性の違いは、事前分布の選定に対するデザインの感度を定量化するため、あるいは、デザイン事前分布と解析事前分布が一致しないときでも誤った決定を行う確率が非常に低いことを示すために用いることができる。一般に、標本サイズが小さい場合、

trial characteristics will be more sensitive to the analysis prior when the sample size is small or when there is an early interim analysis that makes the sample size effectively small.	あるいは、標本サイズを実質的に小さくする早期の中間解析が存在する場合、試験特性は解析事前分布に対してより感度が高くなるであろう。
---	--

3. Additional Considerations	3. 追加の考慮点
The operating characteristics discussed above are typically provided for key objectives of a clinical trial. For any trial and within any development program, there are often multiple other objectives for which data collection is essential. For example, it is essential to generate evidence regarding the safety and tolerability of a product. These other objectives should be considered in the overall trial design and drug development program.	先に議論した動作特性は、一般的に臨床試験の重要な目的に対して提供される。いかなる試験およびいかなる開発プログラム内においても、データ収集が不可欠である複数の他の目的がしばしば存在する。例えば、製品の安全性と忍容性に関する証拠を創出することは不可欠である。これら他の目的は、全体的な試験デザインおよび医薬品開発プログラムにおいて考慮されるべきである。

V. PRIOR DISTRIBUTIONS	V. 事前分布
A. Overview and General Principles	A. 概要および一般原則
Use of a prior distribution is the main feature distinguishing Bayesian from frequentist approaches. Priors allow the analysis to reflect the available information in the particular situation, whether positive, negative, or neutral.	事前分布の利用は、ベイズ流接近法を頻度流接近法と区別する主要な特徴である。事前分布により、肯定的、否定的、または中立的のかにかかわらず、特有の状況で利用可能な情報を解析に反映させることができる。いかなるベイ

With any Bayesian analysis, the prior construction process should be designed, implemented, and documented in a systematic and transparent manner. Sponsors should pre-specify and justify the full details of the proposed prior distribution in the protocol. This justification should address the appropriateness of the prior distribution's influence and the operating characteristics of the design, given the proposed prior. An important distinction is between informative prior distributions that borrow external information into the analysis of the current trial, and noninformative or minimally informative prior distributions that express a stance of general uncertainty. The complexity of the prior construction process will depend on whether an informative prior is used, and on the sources of external information being considered. Informative priors will generally need a greater amount of justification (see Section V.D for details). The following sections provide guidance for these different types of prior distributions and for other aspects of prior distribution evaluation and selection.	ズ流解析においても、事前分布の構築過程は、系統的かつ透明性のある方法で設計され、実装され、文書化されるべきである。スポンサーは、試験実施計画書において、提案する事前分布の完全な詳細を事前規定して正当化すべきである。当正当化は、提案する事前分布を与えた下で、事前分布の影響の適切性およびデザインの動作特性をとり扱うべきである。外部情報を現行試験の解析に借用する情報のある事前分布、および一般的な不確実性の立場を表現する無情報または最小限の情報を有する事前分布の間には重要な区別がある。事前分布の構築過程の複雑さは、情報のある事前分布を用いるかどうか、および考慮されている外部情報源に依存する。情報のある事前分布は、一般に、より多くの正当化を必要とする（詳細は V.D 節を参照）。以下の節は、これら異なる型の事前分布に対する、および事前分布の評価と選択の他の側面に対するガイダンスを提供する。
--	--

B. Noninformative and Minimally Informative Priors	B. 無情報および最小限の情報を有する事前分布
Noninformative and minimally informative priors are typically specified to reflect a stance of general uncertainty regarding the parameters to be estimated. A noninformative prior is designed to use no external information.	無情報および最小限の情報を有する事前分布は、一般的に、推定されるパラメータに関する一般的な不確実性の立場を反映するために規定される。無情報事前分布は、外部情報を利用しないように設計される。このような情報の

There are a variety of approaches that aim to reflect such lack of information, and no single approach is universally preferred to others. A minimally informative prior distribution could be based on general information about a trial and the possible outcomes, such as the plausible magnitudes of change in the endpoints of interest and available information about the variability in previous studies. Noninformative and minimally informative priors are often used in cases where there is no relevant prior information. In many situations, priors that fall into these classes will tend to be overwhelmed by the observed data and therefore have minimal effect on the results such that the final conclusions are dominated by the observed data. In other situations, the available data may not provide much information for a particular parameter and so using a noninformative or minimally informative prior for the parameter may place weight on parameter values we know to be unlikely and may have a large effect on the estimates (for example, see Gelman, 2006). The properties of a particular prior can typically be determined using appropriate simulations during study design.	欠如の反映を目的とする様々な接近法があり、他のものよりも普遍的に好ましい単一の接近法は存在しない。最小限の情報を有する事前分布は、試験に関する一般的な情報、ならびに関心のある評価項目における尤もらしい変化の大きさおよび先行試験におけるばらつきに関する利用可能な情報といった、起こり得るアウトカムに関する一般的な情報に基づくことが可能である。無情報および最小限の情報を有する事前分布は、しばしば関連する事前情報が存在しない場合に用いられる。多くの状況において、これらの組に属する事前分布は、観測データによって圧倒される傾向があり、したがって、結果への影響が最小限であるため、最終結論は観測データによって支配される。他の状況において、利用可能なデータが特有のパラメータに対して多くの情報を提供しないかもしれず、よってパラメータに無情報または最小限の情報を有する事前分布を用いることは、ありそうにないと分かっているパラメータ値に重みを置き、推定値に大きな影響を与える可能性がある(例えば、Gelman, 2006 を参照)。特有の事前分布の特性は、通常、試験デザインの際に適切なシミュレーションを用いて決定することができる。
As with any prior distribution, a noninformative or minimally informative prior may not be invariant to parameter transformation, such that the induced distribution on such a transformation reflects an unintended understanding about the transformed parameter of interest. Care should be taken to understand such induced prior distributions during study design.	いかなる事前分布を用いたとしても、無情報または最小限の情報を有する事前分布はパラメータ変換に対して不変でない可能性があり、このような変換で誘導された分布が、変換された関心のあるパラメータに関する意図とは異なる理解を反映することがある。試験デザインの際には、このような誘導された事前分布を理解するために注意を払うべきである。

Noninformative and minimally informative priors can also play important roles in assessing the influence of informative prior distributions on the analysis. Such priors are often used in the calculation of effective sample size (see Section V.E) as a reference scale against which to measure the influence of another prior. In addition, when an informative prior is specified in the primary analysis, an analysis using a noninformative or minimally informative prior is often helpful in understanding the sensitivity of the results to the choice of prior.	無情報および最小限の情報を有する事前分布は、情報のある事前分布の解析への影響を評価する際に重要な役割を果たすこともできる。このような事前分布は、別の事前分布の影響を測るための参照尺度として、しばしば有効標本サイズの計算 (V.E 節を参照) に用いられる。そのうえ、主要解析において情報のある事前分布が規定されているとき、無情報または最小限の情報を有する事前分布を用いる解析は、事前分布の選定に対する結果の感度を理解する上でしばしば有用である。
--	---

C. Skeptical Priors	C. 懐疑的事前分布
Skeptical priors express skepticism about the presence of very large treatment effects. One scenario where they may be appropriate is when there is prevailing information indicating the need to be more cautious than usual on drawing conclusions in favor of benefit. For example, if there have been a number of failed trials or drug development programs for closely related drugs in a therapeutic area, it would be natural to be skeptical about the potential for benefit with another similar trial or drug. Another possible scenario is in the evaluation of a new drug that is not likely to offer more than incremental improvement over an existing therapy. In this setting, it would be natural to be skeptical of a dramatic improvement over the existing therapy,	懐疑的事前分布は、非常に大きな治療効果の存在に関する懐疑的態度を表現する。それらが適切であるかもしれない1つのシナリオは、便益を支持する結論を導くことに関して、通常よりも慎重になる必要性を指し示す情報が広く存在する場合である。例えば、密接に関連する医薬品に対して、ある治療領域で多数の失敗した試験または医薬品開発プログラムが存在する場合、別の類似した試験または医薬品で便益が得られる可能性について懐疑になることは自然であろう。別のとり得るシナリオは、既存療法を上回るような増分の改善を提供する可能性が低い新薬の評価においてである。当設定において、既存療法を上回る劇的な改善に懐疑になることは自然であり、懐疑的事前分布を解析に用いることが可能である。このような接近法は、医薬品開

and a skeptical prior could be used in the analysis. Such approaches have not been standard practice in drug development but could be considered in relevant circumstances such as those described above.	発において標準的な実践ではなかったが、上述したような関連する状況において考慮することができる。
Skeptical priors may also be considered in trials with an adaptive design. In designs where there is a desire to calibrate to Type I error rate, skeptical priors can be used as an alternative to modification of the decision rule to maintain the desired error rate as the skepticism will counterbalance early random highs or lows and hence early stopping requires even greater early evidence of benefit. Enthusiastic priors (priors containing some degree of positive belief) can be used in a similar way to control early stopping for futility.	懐疑的事前分布は、適応的デザインを伴う試験においても考慮され得る。第I種の過誤確率に対して較正することが望ましいデザインにおいて、懐疑的事前分布は、望ましい過誤確率を維持するための決定規則の改変に代わるものとして用いることができ、というのも、懐疑的態度は早期の偶然による高値または低値を相殺し、それ故早期中止にはより大きな早期の便益の証拠を必要とするためである。熱狂的事前分布（ある程度の肯定的な信念を含む事前分布）は、無益性の早期中止を制御するために同様の方法で利用できる。

D. Informative Priors to Borrow External Information	D. 外部情報を借用するための情報のある事前分布
1. General Recommendations	1. 一般的な推奨事項
When proposing to use an informative prior that borrows external information for inference on the primary estimand, sponsors should provide strong justification that considers feasibility (e.g., of alternative approaches that do not involve borrowing) and the relevance of the available information. Areas where informative priors have been most often proposed include	主要 estimand に関する推論に対して、外部情報を借用する情報のある事前分布を用いることを提案するとき、スポンサーは、（例えば、借用を伴わない代替的接近法の）実行可能性および利用可能な情報の関連性を考慮した強い正当化を提供すべきである。情報のある事前分布が最も頻繁に提案されてきた領域には、小児疾患および希少疾患が含まれる。追加の領域は個別事例

pediatrics and rare diseases. Additional areas can be considered on a case-by-case basis, and FDA advises early discussion of such proposals with the Agency.²⁴ The specific sources of information to be used in the informative prior and the degree of reliance on the information should be justified. The justification should include discussion of the influence of the prior and how it relates to the relevance of the borrowed data for the current trial, the sufficiency of the amount of prospective trial data that will be collected (e.g., for evaluating safety and benefit-risk), and the appropriateness of trial operating characteristics. In some cases, it can be advantageous to determine the details of the borrowing method while still blinded to the results of the trials that will be borrowed. The time needed for FDA and the sponsor to align on an appropriate prior should be considered in the development of the intended trial.	ごとに考慮することができ、FDA はこのような提案について当局と早期に議論するよう助言する。²⁴ 情報のある事前分布に用いられる特定の情報源および情報への依拠の程度は正当化されるべきである。その正当化には、事前分布の影響および事前分布が現行試験に対する借用データの関連性とどのように関係するか、収集される予定の前向き試験データの量の十分性（例えば、安全性とベネフィット・リスクを評価するために）、ならびに試験の動作特性の適切性についての議論を含むべきである。場合によっては、借用される予定の試験結果を盲検化したままで、借用方法の詳細を決定することが利点となる場合がある。適切な事前分布について FDA とスポンサーが合意するために必要な時間は、意図された試験の開発段階で考慮されるべきである。
It is important to consider the possibility of prior-data conflict, defined as the scenario in which the data observed are notably inconsistent with the prior distribution. More formally, prior-data conflict occurs when the prior distribution places its mass “primarily on distributions in the sampling model for which the observed data is surprising (Evans and Moshonov 2006, p 894)”. The potential impact of prior-data conflict on the interpretability of trial results should be explored in simulations at the study design stage by varying the size of the assumed treatment effect over a sufficiently broad	事前分布とデータの対立の可能性を考慮することは重要であり、これは観測データが事前分布と著しく矛盾するシナリオとして定義される。より正式には、事前分布とデータの対立は、事前分布が主たる質量を、「観測データが想定外であるような標本抽出モデルの分布上」におくときに生じる (Evans and Moshonov 2006, p 894)。事前分布とデータの対立が試験結果の解釈可能性に与える潜在的な影響は、標的集団において効果がないことを含む、尤もらしい対立の程度を含んだ十分に広範なシナリオにわたって仮定した治療効果の大きさを変化させることによって、試験設計段階でのシミュレーショ

range of scenarios that cover plausible degrees of conflict, including no effect in the target population. The outcome of these simulations should be considered carefully and discussed in the justification of the prior, and appropriate sensitivity analyses (see Section V.F) should be planned.	ンにおいて探索されるべきである。これらのシミュレーションの結果は慎重に考慮され、事前分布の正当化の際に議論されるべきであり、適切な感度分析 (V.F 節を参照) が計画されるべきである。
Prior construction, including making decisions on how much to borrow, is a multidisciplinary process that requires quantifying the degree of uncertainty in the relevance of the external information to the question of interest. Understanding relevance requires domain knowledge, while quantifying uncertainty and selecting appropriate statistical methodology requires thorough statistical evaluation. Close collaboration between disciplines throughout the process is essential.	どの程度借用するかについての意思決定を含む、事前分布の構築は、関心のある問いに対する外部情報の関連性における不確実性の程度を定量化することを必要とする、多分野にわたる過程である。関連性の理解は領域知識を必要とする一方、不確実性の定量化および適切な統計的方法論の選択は徹底的な統計的評価を必要とする。過程全体を通じた分野間の緊密な協同が不可欠である。
Footnote: ²⁴ See FDA Guidance for Industry Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products (December 2020).	脚注: ²⁴ FDA 産業界向けガイダンス 医薬品および生物学的製剤のための複雑で革新的な試験デザインに関する FDA との意見交換（2020 年 12 月）を参照のこと。

2. Identification and Review of Available External Information	2. 利用可能な外部情報の同定および検討
In general, the process of determining a prior should begin with the	一般に、事前分布を決定する過程は、利用可能な関連するすべての外部情報

identification and review of all the available relevant external information. Possible sources of information may include relevant pharmacokinetic, pharmacodynamic, and clinical data (e.g., from previous trials or systematic reviews of trials evaluating the drug in the same or other related conditions), as well as nonclinical data, real-world data, and professional or expert guidelines or consensus opinions. Several factors should be considered in evaluating whether and how much to leverage external information to ensure that conclusions relying on such information are valid, reliable, and interpretable, such as:	の同定および検討から始めるべきである。とり得る情報源には、関連する薬物動態学的データ、薬力学的データ、および臨床データ（例えば、同じまたは他の関連する状態でその医薬品を評価した先行試験または試験の系統的レビューから）、ならびに非臨床データ、リアルワールドデータ、および専門家または専門家集団によるガイドラインあるいは合意形成された意見が含まれるかもしれない。このような情報に依拠する結論が妥当であり、信頼性が高く、解釈可能であることを保証するために、外部情報を活用するかどうか、およびどの程度活用するかを評価する際に考慮すべきいくつかの要因を以下に示す：
• Data quality and reliability: The quality and reliability of information used to construct the prior should be adequate for the type of regulatory decision informed by the analysis. Data from clinical trials designed to support regulatory decisions will typically meet this bar, but other sources may require additional effort to ensure adequate quality and reliability. For real-world data, FDA has issued guidance ^{25 26 27} that discusses the processes and procedures that help ensure quality of the data. Other sources may be necessary in some circumstances, and in these cases comparable steps should be taken to ensure the quality and reliability of the information.	• データの質と信頼性：事前分布を構築するために用いられる情報の質と信頼性は、その解析から情報を得る規制上の決定の種類に対して適切であるべきである。規制上の決定を支持するように設計された臨床試験から得られたデータは一般的に当基準を満たすが、その他の情報源は、適切な質と信頼性を保証するために追加の努力を必要とするかもしれない。リアルワールドデータに対して、FDA はデータの質を保証するのに役立つ過程と手順を論じるガイダンス ^{25 26 27} を発行している。いくつかの状況では、他の情報源が必要となる可能性があり、これらの場合、情報の質と信頼性を保証するために同等の手順が講じられるべきである。
• Pre-specification: Statistical methods should be prespecified to minimize bias. Similarly, sponsors should pre-specify the construction process for the prior, including for the selection of sources for the same reason. As with any	• 事前規定：統計的方法は、バイアスを最小にするために事前規定されるべきである。同様に、スポンサーは、同じ理由で、情報源の選択を含む事前分布の構築過程を事前規定すべきである。あらゆる証拠統合接近法と同様に、

evidence synthesis approach, the standards for inclusion and the intended scope of use of the various sources for the prior should be predefined before starting a systematic search to identify specific sources.	事前分布に対する種々の情報源についての、選択の基準および意図された利用範囲は、特定の情報源を同定するための系統的検索を開始する前に事前定義されるべきである。
• Relevance: The information being leveraged should be relevant to the applicable regulatory question. When multiple information sources are used, not all information may be equally relevant. It is important to consider and discuss any planned approaches to reflect the different degrees of relevance. Factors that may influence relevance include: — Similarity in estimand attributes such as the population (e.g., inclusion or exclusion criteria), endpoint, treatment conditions, or handling of intercurrent events — Any differences in measurement or assessment (e.g., of the endpoint) — Recency of data — Any potentially important changes (e.g., in aspects of standard of care such as concomitant medications) over time	• 関連性：活用される情報は、適用可能な規制上の問いに関連しているべきである。複数の情報源が用いられるとき、すべての情報が等しく関連があるとは限らない。関連性の異なる程度を反映するために、計画されたすべての接近法を考慮して議論することが重要である。関連性に影響を与える可能性のある要因には以下が含まれる： — 対象集団（例えば、選択規準または除外規準）、評価項目、治療条件、または中間事象のとり扱いといった estimand の属性における類似性 — （例えば、評価項目の）測定または評価における何らかの差異 — データの最新性 — （例えば、併用薬のような標準的ケアの側面における）何らかの潜在的に重要な経時的変化
• The design of studies providing the information: Borrowing of information based on randomized controlled comparisons typically relies on fewer and	• 情報を提供する試験のデザイン：ランダム化対照比較に基づく情報の借用は、非ランダム化比較または単一の治療条件（例えば、既存対照データ）に

more plausible assumptions ²⁸ than borrowing of information based on non-randomized comparisons or on a single treatment condition (e.g., historical control data).	基づく情報の借用よりも、一般的に、より少ない、より尤もらしい仮定 ²⁸ に依拠する。
• The availability of patient-level data: Typically, information should come from patient-level data as this allows for a thorough evaluation of the relevance of the external data and the potential to adjust for relevant covariates in the analysis, which is particularly important when inclusion and exclusion criteria are not fully aligned.	• 患者水準データの利用可能性：一般的に、情報は患者水準データから得るべきであり、というのも、これは外部データの関連性および関連する共変量に対して解析で調整する可能性の徹底的な評価を可能にするからであり、これは選択規準と除外規準が完全には整合していない場合にとりわけ重要である。
Footnotes: ²⁵ See FDA guidance for industry Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products (December 2023). ²⁶ See FDA guidance for industry Use of Electronic Health Record Data in Clinical Investigations (July 2018). ²⁷ See FDA guidance for industry Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products (July 2024). ²⁸ When borrowing information based on randomized controlled	脚注： ²⁵ FDA 産業界向けガイダンス リアルワールドデータ：医薬品および生物学的製剤の規制上の意思決定を支持するためのレジストリの評価（2023 年 12 月）を参照のこと。 ²⁶ FDA 産業界向けガイダンス 臨床研究における電子健康記録データの利用（2018 年 7 月）を参照のこと。 ²⁷ FDA 産業界向けガイダンス リアルワールドデータ：医薬品および生物学的製剤の規制上の意思決定を支持するための電子健康記録および医療請求データの評価（2024 年 7 月）を参照のこと。 ²⁸ ランダム化対照比較に基づいて情報を借用する場合でも、情報源間の効果

comparisons, it is still important to evaluate the potential impact of differences in effect modifiers between information sources.	修飾因子の差異の潜在的な影響を評価することは依然として重要である。
---	-----------------------------------

3. Prior Construction	3. 事前分布の構築
After identification and review of the available external information sources, sponsors should decide how to use the relevant information sources. A thorough evaluation of all relevant sources and evidence for informing the prior should be completed, including evidence that may suggest skepticism of the existence or magnitude of a treatment effect. This evaluation should include a prespecified list of criteria and discussion of steps taken to ensure information was not selectively obtained or used to ensure that preferential selection of favorable studies (or the most favorable studies) did not occur. If multiple information sources will be leveraged, an evidence synthesis process should be pre-specified and employed to obtain the most plausible estimate of the responses. Typical best practices for evidence synthesis in systematic reviews and meta-analyses, such as pre-specification of data collection processes, study inclusion or exclusion criteria, and synthesis methods, are relevant here.²⁹	利用可能な外部情報源の同定および検討のあと、スポンサーは関連する情報源をどのように用いるかを決定すべきである。治療効果の存在または大きさについて懐疑的態度を示唆する可能性のある証拠を含む、事前分布に情報を与えるすべての関連する情報源および証拠の徹底的な評価を完了させるべきである。当評価には、良好な結果の研究（または最も良好な結果の研究）の優先的な選択が生じなかったことを保証するために、情報が選択的に取得または利用されなかったことを保証する事前規定された規準の一覧および講じた手順についての議論を含めるべきである。複数の情報源が活用される場合、最も尤もらしい反応の推定値を得るために、証拠統合の過程を事前規定し、使用すべきである。データ収集過程、研究の選択規準または除外規準、および統合方法の事前規定といった、系統的レビューおよびメタアナリシスにおける証拠統合の一般的なベストプラクティスが、ここでは関連している。²⁹
The modeling approach for constructing the prior will depend on several	事前分布を構築する際のモデル構築接近法は、種々のデータ源間の類似性、

factors, including the similarity between the various data sources, the relative amount of information each source contains, and the relationships between the data sources. As the data sources increase in variety and the relationships becomes more varied, the necessary modeling approach becomes more complex. For example, suppose a sponsor is designing a pediatric study that will use an FDA-approved drug as an active control. For the investigational drug, there is available data from completed adult studies, whereas for the drug used as the active control there are pediatric as well as adult data. In this situation, not all data sources would be equally relevant since the adult data would be less relevant than the pediatric data and so a modeling approach that reflects this belief would be preferable. Modeling approaches can also be used to increase the relevance of the prior data (compared to taking the raw results from the previous studies) by adjusting estimates based on covariates or by selecting a subset of the data that more closely aligns with the question of interest. For example, the multiple sclerosis trial discussed in Section III.B modeled the relationship between age and the annualized relapse rate to improve the relevance of the data from the adult trials.	各データ源が含む相対的な情報量、およびデータ源間の関係を含む、いくつかの要因に依存するであろう。データ源が種類を増し、関係がより多様になるにつれて、必要となるモデル構築接近法はより複雑になる。例えば、スポンサーが、FDA 承認済み医薬品を実薬対照として用いる小児試験を設計しているとしよう。試験薬に対しては、完了した成人試験からの利用可能なデータが存在するのに対し、実薬対照として使用される医薬品に対しては、小児データならびに成人データが存在する。この状況において、成人データは小児データよりも関連性が低いことから、すべてのデータ源が等しく関連があるわけではなく、よってこの信念を反映するモデル構築接近法が好ましいであろう。モデル構築接近法はまた、共変量に基づいて推定値を調整することによって、あるいは関心のある問いとより密接に整合するデータの部分集団を選択することによって、(先行試験から得られたそのまゝの結果をとり込むことと比較して) 先行データの関連性を高めるために利用できる。例えば、III.B 節で議論した多発性硬化症試験は、成人試験から得られたデータの関連性を改善するために、年齢と年換算再発率との間の関係にモデルを与えた。
Typically, multiple sources of information and many assumptions ³⁰ underlie a particular prior, and it is crucial to ensure that these are documented in the protocol with a discussion of the supporting evidence for each assumption to facilitate FDA's review. For example, in some cases it may be reasonable to	一般的に、特有の事前分布の根底には複数の情報源および多くの仮定 ³⁰ があり、FDA の審査を促進するためには、各仮定を支持する証拠についての議論とともに、これらの試験実施計画書での文書化を保証することが極めて重要である。例えば、場合によっては、すべてのデータ源が等しく関連があ

assume that all data sources are equally relevant. A typical example of an assumption of equal relevance would be in a primary analysis in which all patients who are enrolled in a single clinical trial are analyzed simultaneously without any regard for possible subgroups. An example of where this assumption should be considered more closely is when pediatric patients (most commonly adolescents) are included in a single trial with adults. The appropriateness of an assumption of equally relevant data from adults and adolescents would be informed by the specific extrapolation concept and plan³¹, and the most appropriate type of relationship should be justified and discussed with the Agency.	ると仮定することが理にかなっているかもしれない。等しい関連性の仮定についての典型的な例は、単一の臨床試験に登録されるすべての患者が、とり得る部分集団を一切考慮することなく、同時に解析される主要解析においてであろう。当仮定がより綿密に考慮されるべき例は、小児患者（最も一般的には青少年）が成人とともに単一の試験に含まれる場合である。成人および青少年から得られる等しく関連があるデータの仮定の適切性は、特定の外挿概念および計画³¹によって情報が与えられ、最も適切な種類の関係が正当化され、当局と議論されるべきである。
In other cases, models that include exchangeability may be reasonable. The outcomes of a group of participants are considered exchangeable in a setting where if the values of the outcomes are revealed, but their labels are not, then the values are not helpful in predicting their labels. When the outcomes of a group of participants are exchangeable then all possible ordering of the outcomes are equally likely prior to observing the values for the outcomes. Statistical models routinely assume that the outcomes of participants within a treatment group are exchangeable or assume that the residuals are exchangeable. Outcomes modeled as independently drawn from the same distribution are exchangeable. Outcomes modeled as random draws without replacement from a set of possible values are also exchangeable. There are	その他の場合には、交換可能性を含むモデルが理にかなっているかもしれない。参加者のあるグループのアウトカムは、アウトカムの値は開示されているが、それらのラベルは開示されていない場合、それらの値がラベルを予測するのに役立たないならば、交換可能であると見做される。参加者のあるグループのアウトカムが交換可能であるとき、アウトカムの値を観測する前には、アウトカムのすべてのとり得る順序付けは等しい可能性を有する。統計モデルは、治療群内の参加者のアウトカムが交換可能であると通常仮定し、あるいは残差が交換可能であると仮定する。同じ分布から独立に抽出されるとしてモデルを与えられたアウトカムは交換可能である。また、とり得る値の集合からランダムに非復元抽出されるとしてモデルを与えられたアウトカムも交換可能である。交換可能なアウトカムを含意する他の条件も存在する

also other conditions that imply exchangeable outcomes (Yuan and Chai, 2008).	(Yuan and Chai, 2008)。
One setting where exchangeability may be considered is in the estimation of treatment effects within subgroups of a population (i.e., subgroup analysis). An assumption of exchangeability may be reasonable if there are no a priori expectations regarding effect modification by the subgroup factor. However, exchangeability of subgroup treatment effects is not always an appropriate choice. For example, if a drug is expected to be more effective in a subgroup of cancer patients who exhibit the molecular target than the complementary subgroup of patients without that target, the two possible orderings of treatment effects with the subgroups would not be equally likely. In some cases where exchangeability of subgroup treatment effects is not appropriate, all known effect modifiers can be included in the model as effect modifiers (e.g., as interaction effects with treatment when modeling individual observations). It then may be appropriate to regard the residual subgroup treatment effects as exchangeable.	交換可能性が考慮され得る 1 つの設定は、集団の部分集団内での治療効果の推定（すなわち、部分集団解析）においてである。部分集団を規定する因子による効果修飾に関して先験的な期待が存在しない場合、交換可能性の仮定は理にかなっているかもしれない。ただし、部分集団での治療効果の交換可能性は常に適切な選定とは限らない。例えば、ある医薬品が、分子標的を有するがん患者の部分集団において、その標的を有さない患者からなる残りの部分集団よりも有効であることが期待される場合、部分集団間における治療効果の 2 つのとり得る順序付けは等しい可能性を有さないであろう。部分集団別治療効果の交換可能性が適切でない場合には、すべての既知の効果修飾因子を効果修飾因子として（例えば、個別観測値にモデルを与える際に治療との交互作用効果として）モデルに含めることができる。このとき、部分集団別治療効果の残差を交換可能と見做すことが適切であるかもしれない。
There are settings where other potential assumptions may be reasonable. For example, there may be cases where it is reasonable to model the relationship between the source and prospective trial data using an assumed bias parameter in the model. This approach, which would allow borrowing of information on the precision while allowing for differences in the treatment	他の潜在的仮定が理にかなっているかもしれない設定が存在する。例えば、モデルに仮定されたバイアスパラメータを用いて、情報源と前向き試験データとの間の関係にモデルを与えることが理にかなっている場合があり得る。当接近法は、治療効果における差異を許容しながら、精度に関する情報の借用を可能にし、バイアスの大きさについての信頼性の高い証拠またはバイア

effect, might be considered if there is reliable evidence of the magnitude of the bias or suspicion of a bias.	スの疑いが存在する場合に考慮され得る。
In other cases, it may be reasonable to use a model with a functional dependence assumption that predicts the distribution of responses using some function based on systematic predictable processes. For example, suppose two distinct populations were studied in two different trials. If a third trial was conducted that included both populations, it may be reasonable to construct the prior expectation for the effect in this trial using a combination of the results from the two previous trials while adjusting this expectation based on the relative sizes of the populations in the prospective trial. This type of assumption can be seen in some pediatric extrapolation settings where attempts are made to adjust for expected differences between adult and pediatric patients. For example, in the development programs for empagliflozin³² and linagliptin³³, the sponsor constructed pharmacometric models based on the previously conducted studies combined with the relevant observed baseline values of the pediatric patients to predict response.	その他の場合には、系統的な予測可能な過程に基づく何らかの関数を用いて反応の分布を予測する、関数的依存関係の仮定を伴うモデルを用いることが理にかなっているかもしれない。例えば、2つの別個の集団が2つの異なる試験で研究されたと想定しよう。第三の試験が両方の集団を含んで実施された場合、当前向き試験における各集団の相対的な大きさに基づいて事前期待を調整しながら、2つの先行試験から得られた結果の組み合わせを用いて、当試験での効果に対する事前期待を構築することは理にかなっているかもしれない。この種の仮定は、成人患者と小児患者との間の期待される差異を調整するために試みがなされている、いくつかの小児外挿の設定において見ることができる。例えば、エンパグリフロジン³²およびリナグリプチン³³の開発プログラムでは、スポンサーは、反応を予測するために、以前に実施された試験と小児患者の関連するベースライン観測値を組み合わせ、薬物動態モデルを構築した。
Finally, there is often a need to implement a discounting approach to reflect any residual uncertainty regarding the question of interest. For example, when using Bayesian methods to implement pediatric extrapolation, if the adult trial results were used to derive the prior without discounting, the evidence from the adults alone would already meet typically used success	最後に、関心のある問いに関して残存する不確実性を反映するために、しばしば割引接近法を実装する必要がある。例えば、小児外挿を実装するためにベイズ流の方法を用いるとき、事前分布を導出するのに成人試験の結果が割引なしで用いられた場合、成人のみから得られた証拠は一般的に用いられる成功規準をすでに満たしている可能性があるとともに、仮に試験が実施さ

criteria and if a trial were conducted, the adult data would often overwhelm the pediatric data regardless of the results in the pediatric study. This would not be consistent with the uncertainty on the question of benefit in pediatric patients. Therefore, it will often be more reasonable to use a prior centered on a degree of benefit similar to what is observed in adults, but with a greater degree of variability and hence uncertainty. Discounting is discussed in more detail in the next section.	れた場合、小児試験での結果にかかわらず、しばしば成人データが小児データを圧倒するであろう。このことは、小児患者における便益についての問いに関する不確実性と整合しないであろう。したがって、成人で観測される便益と類似した一定の便益に中心をおくが、より大きな一定のばらつき、ひいては不確実性を有する事前分布を用いることが、しばしばより理にかなっている。割引については次節でより詳細に議論する。
Footnotes: ²⁹ See FDA draft guidance for industry Meta-Analyses of Randomized Controlled Clinical Trials to Evaluate the Safety of Human Drugs or Biological Products (November 2018). When final, this guidance will represent the FDA's current thinking on this topic. ³⁰ In this discussion, we use the taxonomy of the types of assumed relationships between data sources from Spiegelhalter, David J., Keith R. Abrams, and Jonathan P. Myles. Bayesian approaches to clinical trials and health-care evaluation. Vol. 13. John Wiley & Sons, 2004. ³¹ See the ICH guidance for industry E11A Pediatric Extrapolation (December 2024). ³² See FDA Clinical Review	脚注: ²⁹ FDA 産業界向けドラフトガイダンス ヒト用医薬品または生物学的製剤の安全性を評価するためのランダム化対照臨床試験のメタアナリシス（2018年11月）を参照のこと。最終化された際、本ガイダンスは本トピックに関するFDAの現在の考え方を表すものとなる。 ³⁰ 本議論において、ここでは、Spiegelhalter, David J., Keith R. Abrams, and Jonathan P. Myles. Bayesian approaches to clinical trials and health-care evaluation. Vol. 13. John Wiley & Sons, 2004 の、データ源間に仮定された関係の種類の分類法を用いる。 ³¹ ICH 産業界向けガイダンス E11A 小児用医薬品開発における外挿（2024年12月）を参照のこと。 ³² FDA 臨床審査報告書 (https://www.fda.gov/media/172973/download) ;

(https://www.fda.gov/media/172973/download); FDA Statistical Review (https://www.fda.gov/media/172972/download). ³³ See FDA Clinical Review (https://www.fda.gov/media/172628/download); FDA Statistical Review: https://www.fda.gov/media/172630/download).	FDA 統計審査報告書： (https://www.fda.gov/media/172972/download) を参照のこと。 ³³ FDA 臨床審査報告書 (https://www.fda.gov/media/172628/download) ; FDA 統計審査報告書： https://www.fda.gov/media/172630/download) を参照のこと。
---	---

4. Discounting	4. 割引
Many prior discounting approaches fall into one of two categories – either static discounting where the method borrows the same information regardless of the observed trial data or dynamic discounting (sometimes called adaptive discounting) where the amount that is borrowed is determined by some measure of the similarity of the data, such as the difference between the observed and prior means. Static discounting methods tend to be easy to design and implement, as there are fewer decisions to be made, while dynamic discounting methods can provide protection against prior-data conflict because they borrow less when the data are less similar and more when the data are more similar. Dynamic approaches are used in many cases due to the more advantageous operating characteristics, such as with respect to bias and MSE, resulting from the lesser borrowing in cases of prior-data	事前分布の割引接近法の多くは、2つのカテゴリのうちの1つに分類される – 観測された試験データにかかわらず同じ情報を借用する方法である静的割引、あるいは借用される量が、観測平均と事前平均との差異といった、データの類似性に関する何らかの指標によって決定される動的割引（適応的割引と呼ばれることもある）のいずれか。静的割引法は、なすべき決定が少ないため、設計と実装が容易である傾向がある一方、動的割引法は、データの類似性が低い場合にはより少なく借用し、データの類似性が高い場合にはより多く借用することから、事前分布とデータの対立に対する防御策を提供することができる。動的接近法は、事前分布とデータの対立がある場合に借用をより少なくすることで、バイアスおよび MSE といった観点で、動作特性がより有利になるため、多くの場合に用いられる。動的割引法の課題は、例えば、用いられる類似性指標および当類似性指標に基づいて借用を減少させる

conflict. The challenge with dynamic discounting methods is that they introduce more parameters that need to be specified, for example, the similarity measure to be used and the rate at which borrowing declines based on this similarity measure.	速度など、その方法が規定の必要なより多くのパラメータを導入することにある。
There are many specific methods proposed in the scientific literature for implementing discounting. With any approach, the goal should be to identify parameter values that correspond to a reasonable degree of borrowing. What is reasonable will depend on the situation and the relevance of the borrowed data. A prior that is too informative risks overwhelming the data collected in the target population, regardless of what is observed. Leveraging too little information means not taking advantage of the available data and more information will be required in the target population, making the trial less efficient. Hence, finding the appropriate balance in the informativeness of the prior is crucial. Model parameter values that determine the degree of discounting and their interpretation depend on the method being used and may not always be easily compared across models. Hence, it is often useful to use more interpretable metrics (such as those discussed in Section V.D).	割引を実装するための多くの特定の方法が、科学文献において提案されている。いかなる接近法を用いても、その目標は理にかなった借用の程度に対応するパラメータ値の同定であるべきである。何が理にかなっているかは、状況および借用されるデータの関連性に依存する。情報が多すぎる事前分布は、何が観測されるかにかかわらず、標的集団において収集されたデータを圧倒するおそれがある。少なすぎる情報の活用は、利用可能なデータの利点を活かしていないことを意味し、より多くの情報が標的集団に必要とされ、試験の効率を下げるであろう。それ故、事前分布の情報量の大きさに適切な均衡を見出すことは極めて重要である。割引の程度を決定するモデルパラメータ値およびそれらの解釈は、利用される方法に依存し、モデル間で常に容易に比較できるとは限らない。それ故、(V.E 節で議論されるような)より解釈可能な評価指標を用いることがしばしば有用である。
In some cases, it may be helpful to conduct a formal expert elicitation exercise (O' Hagan et al., 2006) with subject matter experts to incorporate the degree of consensus in the level of borrowing. It is important to consult with the Agency on the scope and utility of such an exercise before conducting it, and	場合によっては、借用の水準に合意形成の程度を組み込むために、主題に関する専門家集団とともに、正式な専門家からの誘出作業 (O' Hagan et al., 2006) を行うことが有用かもしれない。このような作業を行う前に、その対象範囲と有用性について当局に相談し、誘出のための接近法を事前規定する

to prespecify the approach for elicitation.	ことが重要である。
The following paragraphs discuss considerations for the use of a few different methods for discounting but are not intended to provide an exhaustive list of methods. Applicability of a discounting method will be determined on a case-by-case basis and should be discussed with the Agency.	以下の段落では、割引のためのいくつかの異なる方法の利用に関する考慮点を論じるが、方法の網羅的な一覧を提供することは意図していない。割引法の適用可能性は、個別事例ごとに決定され、当局と議論されるべきである。
One approach uses power priors that are constructed by raising a prior distribution based on external data to a fixed power less than one, sometimes called a discounting factor. The discounting factor adjusts the relative informativeness of a single external subject's data compared to a single current subject's data. A discounting factor of 1 corresponds to equal weighting of external and prospective trial subjects, a discounting factor of 0 corresponds to no borrowing of external information, and discounting factors in between correspond to borrowing with each external subject contributing less information than each trial subject. Once the discounting factor is chosen, the degree of borrowing is fixed, making this a static borrowing approach. Advantages of the power prior approach include simple implementation and a seemingly intuitive approach to discounting. A disadvantage is that it is static and therefore may have worse operating characteristics than dynamic approaches.	1つの接近法は、外部データに基づく事前分布を1未満の固定値でべき乗することによって構築される、べき乗事前分布を用いることであり、その固定値は割引因数と呼ばれることがある。割引因数は、単一の現行対象者のデータと比較した、単一の外部対象者のデータの相対的な情報量の大きさを調整する。割引因数1は、外部対象者と前向き試験対象者の等しい重み付けに対応し、割引因数0は、外部情報の借用なしに対応し、0と1の間の割引因数は、各外部対象者の寄与する情報が各試験対象者よりも少ない借用に対応する。いったん割引因数が選ばれると、借用の程度は固定され、これは静的借用接近法となる。べき乗事前分布接近法の利点には、単純な実装および一見直感的にみえる割引の接近法であることが含まれる。欠点は、静的であるため、動的接近法よりも悪い動作特性をもつ可能性があることである。
Although the basic power prior approach is a static borrowing approach,	基本的なべき乗事前分布接近法は静的借用接近法ではあるが、観測データと

there are extensions proposed in the literature such as commensurate priors and supervised power priors which make the degree of borrowing dependent on different measures of the similarity between the observed and prior data. As the difference between the observed and prior means grows larger, less information will be borrowed. These approaches can allow control of how fast the discounting occurs as the difference grows, and determining an appropriate rate is one of the main challenges in using this approach.	事前データとの間の類似性の異なる指標に借用の程度を依存させる commensurate 事前分布および supervised べき乗事前分布といった、文献で提案されている拡張が存在する。観測平均と事前平均との差異が大きくなるにつれて、より少ない情報が借用される。これらの接近法は、その差異が大きくなるにつれて割引がどのくらいの速さで生じるかを制御することを可能にするが、適切な率の決定が当接近法を用いる際の主要な課題の1つである。
Using a mixture prior is another simple way to implement a dynamic borrowing approach. Two or more individual prior components are combined with prespecified weights, and the result is a prior distribution that will adaptively reflect the similarity between the individual components and the observed data. The chosen weights must sum to one and can be interpreted as the probability of applicability of the particular data source. For example, a mixture prior might be constructed by combining an informative component based on estimates obtained from previous trial data with a noninformative component. The degree to which the resulting posterior will borrow the previous trial data will depend on the degree of similarity between the observed and previous data. The main difficulty is determining the level of weighting for informative components and hence the strength of borrowing. The advantages of this approach are simple programming implementation and well-developed methods for estimating the degree of borrowing.	混合事前分布の利用は、動的借用接近法を実装するための別の単純な方法である。2つ以上の個別の事前分布の構成要素が事前規定された重みを用いて結合され、その結果として個別の構成要素と観測データとの間の類似性を適応的に反映する事前分布である。選定された重みは合計が1にならなければならない、特有のデータ源の適用可能性の確率として解釈することができる。例えば、混合事前分布は、先行試験データから得られた推定値に基づく情報のある構成要素を、無情報の構成要素と組み合わせることによって構築され得る。結果として生じる事後分布が先行試験データを借用する程度は、観測データと先行データとの間の類似性の程度に依存する。主な困難は、情報のある構成要素に対する重み付けの水準、ひいては借用の強さを決定することである。当接近法の利点は、単純なプログラミング実装および借用の程度を推定する方法が十分に開発されていることである。

Other discounting approaches include Bayesian hierarchical models and elastic priors. Bayesian hierarchical models are the main method used to implement exchangeable models. They induce borrowing by assuming that the parameters for various groups are at some level drawn from a common distribution. For example, in a basket trial evaluating a drug in multiple related diseases, a possible starting point for the modeling process might be that the treatment effects for the diseases are a random sample drawn from a single common distribution. Elastic priors (Jiang et al., 2023) are implemented in two steps. First, a measure of the degree of similarity called a congruence measure is used to quantify the strength of similarity between the external and current data. Then a function called an elastic function is used to map the congruence measure to a degree of borrowing. Like the power prior approach, the degree of borrowing takes values between 0 (no borrowing) and 1 (full borrowing). Elastic priors are flexible methods as they can use any of an extremely broad range of different functions to implement discounting that adapts as fast or as slow as desired based on the level of observed conflict between the prior and current data. However, this flexibility may make it harder to justify a particular choice of function.	その他の割引接近法には、ベイズ流階層モデルおよび elastic 事前分布が含まれる。ベイズ流階層モデルは、交換可能モデルを実装するために用いられる主要な方法である。当モデルは、種々のグループに対するパラメータが、ある水準において共通の分布から抽出されると仮定することによって借用を誘導する。例えば、複数の関連する疾患で医薬品を評価するバスケット試験において、モデル構築過程のとり得る出発点は、疾患に対する治療効果が単一の共通の分布から抽出されたランダム標本であるという可能性がある。elastic 事前分布 (Jiang et al., 2023) は 2 つの手順で実装される。まず、一致性指標と呼ばれる類似性の程度の指標が、外部データと現行データとの間の類似性の強さを定量化するために用いられる。次に、elastic 関数と呼ばれる関数が、一致性指標を借用の程度に写像するために用いられる。べき乗事前分布接近法と同様に、借用の程度は 0 (借用なし) と 1 (完全な借用) との間の値をとる。elastic 事前分布は、事前分布と現行データとの間に観測された対立の水準に基づいて、望むだけ速くまたは遅く適応する割引を実装するために、極めて広範囲の異なる関数のいずれでも使用できるため、柔軟な方法である。ただし、この柔軟性は、関数の特有の選択を正当化することをより困難にするかもしれない。
---	--

E. Quantifying the influence of the prior distribution	E. 事前分布の影響の定量化
--	----------------

The influence of a prior distribution should be discussed and documented. The influence can be measured in multiple ways using different metrics. Metrics that have been used in the regulatory setting include:	事前分布の影響は、議論され文書化されるべきである。その影響は、異なる評価指標を用いて複数の方法で測定することができる。規制上の設定において用いられてきた評価指標には以下が含まれる：
• The estimated treatment effect/difference or parameter(s) of interest based solely on the prior distribution. This is usually the mean of the prior distribution.	• 事前分布のみに基づく、関心のある推定治療効果／差またはパラメータ。これは通常、事前分布の平均である。
• Effective sample size (ESS): a measure of the information in a probability distribution in terms of the equivalent number of patients in the target population. For example, when applied to the prior alone we would call it the prior ESS. Note that this number may be larger than the sample size from the source population if there is a larger amount of variability in the target population than the reference population. Multiple methods that rely on different summary measures have been proposed for the ESS (Malec, 2001; Morita et al., 2008; Neuenschwander et al, 2020). In general, it is important to quantify and summarize the ESS for the entire plausible range of outcomes while also presenting relevant summary statistics such as the maximum and mean ESS.	• 有効標本サイズ（ESS）：標的集団における同等な患者数という観点からの確率分布における情報量の指標。例えば、事前分布のみに適用されるとき、これを事前 ESS と呼ぶ。参照集団よりも標的集団においてより大きな量のばらつきが存在する場合、この数は源泉集団の標本サイズより大きい可能性があることに留意されたい。ESS に対しては、異なる要約指標に依拠する複数の方法が提案されてきた（Malec, 2001; Morita et al., 2008; Neuenschwander et al, 2020）。一般に、最大 ESS および平均 ESS といった関連する要約統計量も提示しながら、アウトカムの尤もらしい範囲全体に対して ESS を定量化して要約することが重要である。
These metrics are helpful to compare candidate priors. When an informative prior is used, values of these metrics should be compared to values obtained when using noninformative or minimally informative priors (or other relevant	これらの評価指標は、候補となる事前分布を比較するのに役立つ。情報のある事前分布を用いるとき、情報のある事前分布の影響を理解するために、当評価指標の値は、無情報または最小限の情報を有する事前分布（または他の

priors) to understand the effect of the informative prior. These metrics may be more easily understood than prior parameter values that depend on the specific model being used.	関連する事前分布) を用いたときに得られる値と比較されるべきである。これらの評価指標は、使用されている特定のモデルに依存する事前分布のパラメータ値よりも容易に理解されるかもしれない。
It can also be helpful to consider extensions of these metrics. For example, one might determine the ratio of the prior effective sample size to the intended prospective trial size or to the total amount of information (prior effective sample size plus prospective trial size) to measure the relative amount of information borrowed in an analysis. If a dynamic prior is used, then the effective sample size contribution of the prior will change as information is accrued and so it is important to reassess this measure following completion of the trial. After data are collected, additional measures regarding the prior's impact may be useful. Examples include measures of the conflict between the prior and the observed data and updated measures quantifying the strength of borrowing.	これらの評価指標の拡張を考慮することも役立ち得る。例えば、解析において借用された情報の相対的な量を測るために、意図された前向き試験のサイズに対する事前 ESS の比、あるいは総情報量（事前 ESS と前向き試験のサイズの和）に対する事前 ESS の比を決定することができる。動的事前分布が用いられる場合、事前分布の ESS への寄与は情報が集積されるにつれて変化するため、試験の完了後に当指標を再評価することが重要である。データが収集されたあとは、事前分布の影響に関する付加的な指標が有用かもしれない。それらの例には、事前分布と観測データとの間の対立に関する指標、および借用の強さを定量化する更新された指標が含まれる。
In some cases, the type I error rate has been proposed as a way to measure the degree of information in the informative prior. Borrowing information in an analysis will typically lead to an inflation in the Type I error rate compared to the nominal value (most often, a one-sided level of 0.025) used in the absence of borrowing. This degree of inflation has been proposed as one way of measuring the influence of a prior (Pennello and Thompson, 2007). However, there are multiple issues that make this method a poor way of	場合によっては、第 I 種の過誤確率が、情報のある事前分布における情報量の程度を測る 1 つの方法として提案されてきた。解析において情報を借用することは、借用のない場合に用いられる名目値（多くの場合、片側水準 0.025）と比較して、一般的に第 I 種の過誤確率の増大をもたらすであろう。この増大の程度が事前分布の影響を測る 1 つの方法として提案されてきた (Pennello and Thompson, 2007)。ただし、当方法が借用の程度を測る方法として不十分である複数の課題が存在することから、この目的のために

measuring the degree of borrowing, and therefore it is not recommended for this purpose. First, by borrowing information, one is assuming that the borrowed information is relevant. Evaluating the degree of borrowing based on the expected outcome when there is no effect is philosophically inconsistent given a prior which assumes a non-zero effect. Second, dynamic methods lessen the impact of borrowing when effects that are smaller and close to the null value are observed. Consequently, borrowing will be low when the observed data are incompatible with the prior, mitigating any potential increase in Type I error rate. Hence, the Type I error rate alone will not provide a complete assessment of the influence of the prior.	は推奨されない。まず、情報を借用することによって、借用された情報が関連していると仮定していることになる。ゼロでない効果を仮定する事前分布を与えた下で、効果がないときに期待されるアウトカムに基づいて借用の程度を評価することは、哲学的に矛盾している。次に、動的方法は、帰無仮説の値に近いより小さな効果が観測されるときに借用の影響を小さくする。結果として、観測データが事前分布と整合しないときに借用は少なくなり、第 I 種の過誤確率における潜在的な増加を緩和する。それ故、第 I 種の過誤確率のみでは、事前分布の影響の完全な評価を提供しないであろう。
--	--

F. Sensitivity analyses	F. 感度分析
A necessary part of statistical inference in clinical trials is evaluating the sensitivity of the results and conclusions to plausible deviations from important analysis assumptions. In a Bayesian analysis, the choice of prior distribution is a critical assumption that can affect the results and conclusions, particularly when the prior distribution is informative and is associated with a large ESS. Therefore, sensitivity analyses should be planned that utilize a range of alternative reasonable choices for the prior distributions. For example, the amount or strength of borrowing could be varied (see Section IV.B on	臨床試験における統計的推論の必要な部分は、重要な解析の仮定からの尤もらしい逸脱に対する結果と結論の感度を評価することである。ベイズ流解析において、事前分布の選定は結果と結論に影響を与え得る極めて重要な仮定であり、とりわけ事前分布が情報を有していて、大きな ESS と関連する場合にはそうである。したがって、感度分析は、事前分布に対する様々な代替の理にかなった選定を利用するように計画されるべきである。例えば、借用の量または強さを変化させることができる（追加の例については、動作特性に関する IV.B 節を参照）。（それぞれの事後分布に基づく）このような解析の結

operating characteristics for additional examples). Comparing the results and conclusions (based on the respective posterior distributions) of such analyses can help one understand the sensitivity of the primary results and conclusions. Some approaches can build uncertainty about specific assumptions into the prior itself.	果と結論を比較することは、主要な結果および結論の感度を理解するのに役立ち得る。いくつかの接近法では、特定の仮定に関する不確実性を事前分布自体に組み込むことが可能である。
---	---

VI. ESTIMANDS AND MISSING DATA IN A BAYESIAN SETTING	VI. ベイズ流設定における estimand および欠測データ
The general considerations related to estimands and missing data described in the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) E9 and E9(R1) guidances are relevant for a trial that uses Bayesian methods as with any other trial. These considerations include the importance of constructing a relevant primary estimand of interest that can be estimated with plausible assumptions, implementing design and conduct approaches to prevent missing data, pre-specifying an appropriate primary analysis approach (i.e., estimator), documenting missing data assumptions, and planning sensitivity analyses to evaluate robustness to violations in those assumptions. Bayesian methods can be used to address missing data in the primary analysis or in sensitivity analyses.	医薬品規制調和国際会議 (ICH) E9 および E9(R1) ガイダンスに記述された、estimand および欠測データに関連する一般的考慮点は、他のあらゆる試験と同様に、ベイズ流の方法を用いる試験においても関連がある。これらの考慮点には、尤もらしい仮定を伴って推定できる関心のある関連する主要な estimand を構築することの重要性、欠測データを防ぐためのデザインおよび実施における接近法を実装すること、適切な主要解析の接近法（すなわち、推定量）を事前規定すること、欠測データの仮定を文書化すること、およびそれらの仮定の違反に対する頑健性を評価するために感度分析を計画することが含まれる。ベイズ流の方法は、主要解析または感度分析において欠測データを取り扱うために用いることができる。

There are additional important considerations in trials that use Bayesian methods to borrow external information. In particular, it is critical to consider whether there are any differences in the estimands and estimators between the external information source (e.g., previous trial) and the prospective trial. Ideally, the same primary estimands and estimators should be used in analyzing both data sources. The properties of the parameter estimates depend on the approaches used, and lack of alignment in the approaches between the data sources can make the external data less relevant and affect considerations about the degree of borrowing. However, FDA recognizes that it may not be feasible or advisable to use the same approach in all situations.	外部情報を借用するためにベイズ流の方法を用いる試験においては、追加の重要な考慮点が存在する。特に、外部情報源（例えば、先行試験）と前向き試験との間で estimand および推定量に何らかの差異があるかどうかを考慮することは極めて重要である。理想的には、両方のデータ源を解析する際に、同じ主要 estimand および推定量が用いられるべきである。パラメータ推定値の特性は用いられる接近法に依存しており、データ源間の接近法における整合性の欠如は、外部データの関連性を低下させ、借用の程度に関する考慮点に影響を与える可能性がある。ただし、FDA は、すべての状況において同じ接近法を用いることが、必ずしも実行可能または推奨されるとは限らないことを認識している。
In general, where there is a lack of alignment between the estimands or estimators, sponsors can consider reanalyzing the external data using the approach that is planned for the prospective trial. This will produce the most compatible estimates but may not be possible if patient-level data are inaccessible or if a particular estimand or estimator that will be used in the prospective trial relies on data that were not collected in the external data source (for example, patient outcome data after treatment discontinuation may be necessary for some approaches and are not always collected). Reanalyzing completed studies risks overfitting to observed random highs, thereby introducing bias. Sponsors should have early discussions with the Agency about the planned estimands, estimators, and approaches for	一般に、estimand または推定量の間に整合性の欠如が存在する場合、スポンサーは、前向き試験のために計画されている接近法を用いて外部データを再解析することを考慮できる。これにより、最も整合性のある推定値が得られるが、患者水準データにアクセス不可能な場合、あるいは前向き試験で用いられる特有の estimand または推定量が、外部データ源で収集されていないデータに依拠する場合（例えば、いくつかの接近法に対して、治療中止後の患者アウトカムデータが必要となる可能性があるが、必ずしも常に収集されるとは限らない）、これは可能ではないかもしれない。完了した試験の再解析は、観測された偶然による高値に過剰適合するおそれがあり、その結果、バイアスを導入するおそれがある。スポンサーは、計画された estimand、推定量、および借用される外部データの解析において欠測データ

handling missing data in the analyses of external data that will be borrowed, and any differences relative to the approaches planned for the prospective trial data.	をとり扱うための接近法、ならびに前向き試験データで計画されたそれらの接近法と比較したあらゆる相違点について、当局と早期に議論の場をもつべきである。
--	---

VII. SOFTWARE AND COMPUTATION	VII. ソフトウェアおよび計算
When performing Bayesian inference, it is often necessary to rely on various approximate sampling algorithms to perform statistical inference. These sampling algorithms are implemented in a variety of general and specialized statistical software packages. These algorithms may also be implemented in specific situations using a variety of general programming languages. As stated in the Statistical Software Clarifying Statement, ³⁴ "FDA does not require use of any specific software for statistical analyses." As noted in the statement and in ICH guidance E9 Statistical Principles for Clinical Trials, ³⁵ "computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available."	ベイズ流推論を行う際には、しばしば統計的推論を行うための種々の近似標本抽出アルゴリズムに依拠することが必要となる。これらの標本抽出アルゴリズムは、様々な汎用および特殊な統計ソフトウェアパッケージにおいて実装されている。当アルゴリズムは、特定の状況では、様々な汎用プログラミング言語を用いて実装されることもある。統計ソフトウェアに関する明確化声明で述べられているように、 ³⁴ 「FDA は統計解析において特定のソフトウェアの使用を要求しない」。同声明および ICH ガイダンス E9 臨床試験のための統計的原則に述べられているように、 ³⁵ 「データマネジメントおよび統計解析に使用されるコンピュータソフトウェアは信頼性が高いものであるべきであり、適切なソフトウェア検証手順の文書化が利用可能であるべきである」。
When performing Bayesian analyses, it is critical to ensure that the sampling algorithm being employed is reliable for the specific model (Gelman et al., 2020). Bayesian methods are dependent on algorithms to approximate the	ベイズ流解析を行う場合、使用されている標本抽出アルゴリズムが、特定のモデルに対して信頼性が高いことを保証することは極めて重要である (Gelman et al., 2020)。ベイズ流の方法は、事後分布を近似するためのアル

posterior distribution and approximations can fail for a variety of reasons. For example, a Markov Chain Monte Carlo (MCMC) algorithm may not adequately sample from a particular part of the posterior distribution, or an approximation used in an algorithm may not be accurate. Hence it can be necessary, especially for more complex models, to evaluate the reliability and accuracy of Bayesian computation using simulations before final selection of the model. In cases where sampling or convergence issues are encountered, more detailed documentation of the reliability of the final selected model should be provided, e.g., in a simulation report, including a discussion of any steps taken to ensure better sampling or convergence. The documentation should include the range of scenarios evaluated and the code used to implement the model to allow verification of the sampling properties.	ゴリズムに依存しており、近似は様々な理由で失敗する可能性がある。例えば、マルコフ連鎖モンテカルロ (MCMC) アルゴリズムは、事後分布の特有の部分から適切に標本抽出を行わない可能性があり、あるいはアルゴリズムで用いられる近似が正確ではない可能性がある。それ故、複雑なモデルに対しては特に、モデルの最終選択前に、シミュレーションを用いてベイズ流計算の信頼性と正確性を評価することが必要な場合がある。標本抽出または収束の問題に遭遇する場合には、最終的に選択されたモデルの信頼性のより詳細な文書化が、例えばシミュレーション報告書において、提供されるべきであり、これには、より良好な標本抽出または収束を保証するために講じたあらゆる手順についての議論が含まれる。当文書化には、評価したシナリオの範囲および標本抽出特性の検証を可能にするモデルの実装に用いたコードを含むべきである。
Footnotes: ³⁴ See FDA Statistical Software Clarifying Statement (https://www.fda.gov/media/161196/download). ³⁵ See ICH guidance for industry E9 Statistical Principles for Clinical Trials (September 1998).	脚注: ³⁴ FDA 統計ソフトウェアに関する明確化声明を参照のこと (https://www.fda.gov/media/161196/download)。 ³⁵ ICH 産業界向けガイダンス E9 臨床試験のための統計的原則（1998 年 9 月）を参照のこと。

VIII. DOCUMENTING AND REPORTING BAYESIAN ANALYSES	VIII. ベイズ流解析の文書化および報告
Clear documentation is necessary for FDA to review Bayesian proposals. The study design, estimands, and analyses should be pre-specified and justified in a protocol and consider applicable guidances.³⁶ A clinical study report should describe the design, analysis, and results. As with non-Bayesian approaches, an assessment of parametric assumptions associated with the likelihood should be provided in the clinical study report using appropriate diagnostics. Details specific to Bayesian approaches are discussed below.	FDA がベイズ流の提案を審査するために、明確な文書化が必要となる。試験デザイン、estimand、および解析は、試験実施計画書において事前規定され、正当化されるべきであり、適用可能なガイダンスを考慮すべきである。 ³⁶ 試験総括報告書は、デザイン、解析、および結果について記述すべきである。ベイズ流ではない接近法と同様に、尤度に関連するパラメトリックな仮定の評価は、適切な診断を用いて、試験総括報告書において提供されるべきである。ベイズ流接近法に特定の詳細は以下で議論される。
Footnote: ³⁶ For example, see the ICH guidances for industry E9 Statistical Principles for Clinical Trials (September 1998) and E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials (May 2021), and the FDA guidance for industry Adaptive Designs for Clinical Trials of Drugs and Biologics (December 2019).	脚注: ³⁶ 例えば、ICH 産業界向けガイダンス E9 臨床試験のための統計的原則（1998 年 9 月）および E9(R1) 臨床試験のための統計的原則: 補遺: 臨床試験における estimand と感度分析（2021 年 5 月）、ならびに FDA 産業界向けガイダンス 医薬品および生物学的製剤の臨床試験のための適応的デザイン（2019 年 12 月）を参照のこと。

A. Documenting Plans for Bayesian Analyses	A. ベイズ流解析に対する計画の文書化
The protocol should describe and justify the design and the planned statistical analysis methods. For Bayesian methods, this includes detailed	試験実施計画書は、デザインおよび計画された統計解析方法を記述し、正当化すべきである。ベイズ流の方法に関して、これには、提案された事前分布

information to support the proposed prior distribution and any external information borrowing, likelihood function, success criteria, and trial operating characteristics. Additional details or supporting information may be provided in additional documents such as a statistical analysis plan or simulation report. All relevant information should be submitted to FDA during the design stage and as early as possible to ensure sufficient time for FDA feedback prior to initiation of the trial.	および外部情報の借用を支持する詳細な情報、尤度関数、成功規準、ならびに試験の動作特性が含まれる。追加の詳細または補足情報は、統計解析計画書またはシミュレーション報告書といった追加文書において提供されるかもしれない。試験開始前に FDA からフィードバックを得るための十分な時間を保証するために、関連するすべての情報は、デザイン段階において、できるだけ早期に FDA に提出されるべきである。
The following information should be provided to support the proposed prior distribution (see Section V for further discussion):	提案された事前分布を支持するために、以下の情報が提供されるべきである（さらなる議論については V 節を参照）：
• A detailed description of the proposed prior distribution, including explicit specification of prior parameterization and underlying assumptions. If a function of parameter(s) is of interest, this should include a description of any induced priors.	• 事前分布のパラメータ化および根底にある仮定の明示的な規定を含む、提案された事前分布の詳細な記述。パラメータの関数が関心の対象である場合、これには誘導される事前分布についての記述を含めるべきである。
• A rationale for the plausibility of the assumptions.	• 仮定の尤もらしさに関する合理的根拠。
• Any data or other information that informed the prior distribution.	• 事前分布に情報を与えるあらゆるデータまたは他の情報。
If the trial will use an informative prior to borrow external information for inference on the primary estimand the following additional information should also be provided:	試験が、主要 estimand に関する推論に対して、外部情報を借用するために情報のある事前分布を用いる場合、以下の追加情報も提供されるべきである：

• A strong justification for borrowing that considers feasibility (e.g., of alternative approaches that do not involve borrowing) and the relevance of the available information.	• （例えば、借用を伴わない代替的な接近法の）実行可能性および利用可能な情報の関連性を考慮した、借用に対する強い正当化。
• A thorough evaluation of all the available relevant evidence for informing the prior, including evidence that may suggest skepticism. This should include a discussion of steps taken to ensure information was not selectively obtained or used.	• 懐疑的態度を示唆する可能性のある証拠を含む、事前分布に情報を与えるために利用可能なすべての関連する証拠の徹底的な評価。これには、情報が選択的に取得または利用されなかったことを保証するために講じた手順についての議論を含むべきである。
• A detailed description of each external information source used to inform the prior, as well as any source that was excluded and why, including a discussion of data quality and reliability, relevance, type (e.g., patient-level vs. summary-level) and completeness of data available, and a rationale for the degree of borrowing incorporating such factors.	• 事前分布に情報を与えるために用いた各外部情報源の詳細な記述、ならびに除外されたあらゆる情報源および理由。これには、データの質と信頼性、関連性、種類（例えば、患者水準 対 要約水準）および利用可能なデータの完全性についての議論、ならびにそのような要因を組み込んだ上での借用の程度に対する合理的根拠が含まれる。
• If multiple information sources are leveraged, a description of how the data will be synthesized and the modeling approach for their synthesis.	• 複数の情報源が活用される場合、データがどのように統合されるか、およびそれらの統合のためのモデル構築接近法についての記述。
• If a discounting method is incorporated, a description and justification of the selected approach, including a rationale for any parameters (e.g., that impact the degree of borrowing).	• 割引法が組み込まれている場合、（例えば、借用の程度に影響を与える）あらゆるパラメータに対する合理的根拠を含む、選択された接近法の記述および正当化。
• A discussion of how the proposed analysis will handle prior-data conflict	• 提案された解析が、事前分布とデータの対立をどのようにとり扱うかにつ

(e.g., supported by simulation results).	いての（例えば、シミュレーション結果によって支持される）議論。
• Quantification and discussion of the degree of influence of the prior distribution (e.g., based on the ESS and other metrics).	• 事前分布の影響の程度の（例えば、ESS および他の評価指標に基づく）定量化および議論。
• Planned sensitivity analyses that utilize a range of alternative reasonable choices for the prior distribution (e.g., that vary the amount or strength of borrowing from the prior distribution).	• 事前分布に対する様々な代替の理にかなった選定を利用する（例えば、事前分布からの借用の量または強さを変化させる）計画された感度分析。
The following information should be provided to support the proposed success criteria (see Section IV.A for further discussion):	提案された成功規準を支持するために、以下の情報が提供されるべきである（さらなる議論については IV.A 節を参照）：
• Specification of the success criteria for all primary and key secondary estimands. For success criteria based on the posterior probability, both the treatment effect threshold and the minimum probability of effectiveness should be discussed and justified.	• すべての主要および重要な副次 estimand に対する成功規準の規定。事後確率に基づく成功規準に対しては、治療効果の閾値および有効性の最小確率の両方が議論され、正当化されるべきである。
• For trials that include interim analyses with the potential for early stopping for efficacy, specification of the success criteria for each decision point.	• 効果に対する早期中止の可能性を伴う中間解析を含む試験に対しては、各決定時点に対する成功規準の規定。
• For trials with success criteria calibrated to Type I error rate, justification that the choice of success criteria leads to control of the FWER (e.g., based on a detailed report of simulations used to estimate the Type I error rate).	• 第 I 種の過誤確率に対して校正された成功規準をもつ試験に対しては、成功規準の選定が FWER の制御をもたらすことの（例えば、第 I 種の過誤確率を推定するために用いたシミュレーションの詳細な報告書に基づく）正当

	化。
The following information should be provided to describe and support the appropriateness of the operating characteristics of the trial (see Section IV.B for further discussion):	試験の動作特性の適切性を記述して支持するために、以下の情報が提供されるべきである（さらなる議論については IV.B 節を参照）：
• A discussion of how operating characteristics are estimated, with sufficient detail to facilitate FDA's verification of the values and evaluation of the proposed sample size and other design elements.	• FDA による、値の検証および提案された標本サイズと他のデザイン要素の評価を促進するための十分な詳細を伴う、動作特性がどのように推定されたかについての議論。
• For study designs calibrated to Type I error rate, an evaluation of the Type I error rate and power under a range of plausible effect sizes, and estimation properties including bias and MSE of point estimates and coverage probability and width of intervals.	• 第 I 種の過誤確率に対して較正された試験デザインに対しては、様々な尤もらしい効果の大きさの下での第 I 種の過誤確率と検出力の評価、ならびに点推定値のバイアスと MSE および区間の被覆確率と幅を含む推定特性。
• For study designs not calibrated to Type I error rate, an evaluation of Bayesian power, expected bias, expected interval width, expected MSE, and other quantities to assess the accuracy of conclusions and treatment effect estimates under relevant priors.	• 第 I 種の過誤確率に対して較正されていない試験デザインに対しては、関連する事前分布の下での結論および治療効果推定値の正確性を評価するための、ベイズ流検出力、期待バイアス、期待区間幅、期待 MSE、およびその他の量の評価。
• When simulations are required to estimate operating characteristics, a comprehensive simulation report. This report should include simulation code and a detailed description of the simulation design, implementation, and results. The scenarios and assumptions used in the simulation should be pre-	• 動作特性を推定するためにシミュレーションが必要とされる場合、包括的なシミュレーション報告書。当報告書には、シミュレーションコードならびにシミュレーションの設計、実装、および結果の詳細な記述を含むべきである。シミュレーションで用いられるシナリオと仮定は、事前規定され、包括

specified, comprehensive, and plausible.	的かつ尤もらしいものであるべきである。
For Bayesian analyses performed using approximate sampling algorithms (see Section VII), sponsors should provide a summary of the simulation setup. For example, for MCMC approaches, sponsors should provide a general summary of the length of warmup or burn-in, number of iterations, number of chains, convergence diagnostics, and any other important algorithm-specific settings (e.g., proposal distribution for the Metropolis algorithm, target proposal acceptance probability when relevant for software implementation of Hamiltonian Monte Carlo). If computation requires specialized software, the software should be cited, and details of planned implementation provided. If such computation is required, there should be a description of how implementation issues (e.g., lack of convergence) will be addressed in the analysis.	近似標本抽出アルゴリズムを用いて行われるベイズ流解析に対しては (VII 節を参照)、スポンサーはシミュレーション設定の要約を提供すべきである。例えば、MCMC 接近法に対しては、スポンサーは、ウォームアップまたはバーンインの長さ、反復回数、連鎖数、収束診断、およびその他の重要なアルゴリズム固有の設定（例えば、メトロポリスアルゴリズムの提案分布、ハミルトニアンモンテカルロのソフトウェア実装に関連する場合には目標提案受容確率）の一般的な要約を提供すべきである。計算が特殊なソフトウェアを必要とする場合、そのソフトウェアが引用され、計画された実装の詳細が提供されるべきである。このような計算が必要とされる場合、実装上の問題（例えば、収束の欠如）が解析でどのようにとり扱われるかについての記述があるべきである。
For complex design proposals such as those with borrowing of external information based on informative priors, comparisons should be made between the operating characteristics of the proposed design and analysis plan and a variety of alternatives, including simpler designs. These comparisons should be documented as they help understand the potential advantages and limitations of the proposed design features. FDA also recommends that sponsors considering complex designs request a meeting with FDA prior to initiation of the trial. Including a detailed discussion of the	情報のある事前分布に基づく外部情報の借用を伴うような複雑なデザインの提案に対しては、提案されたデザインおよび解析計画と、より単純なデザインを含む様々な代替案との間で動作特性の比較がなされるべきである。これらの比較は、提案されたデザインの特徴の潜在的利点および限界を理解するのに役立つことから、文書化されるべきである。FDA はまた、複雑なデザインを考慮しているスポンサーが試験開始前に FDA との会合を要請することを推奨する。会合の資料一式に、提案されたデザインおよび代替のデザインとの比較についての詳細な議論を含めることは、FDA 審査官が試験の動

proposed design and comparisons against alternative designs in meeting packages can help FDA reviewers understand and assess the appropriateness of trial operating characteristics.	作特性の適切性を理解し、評価するのに役立ち得る。
--	--------------------------

B. Reporting Results from Bayesian Analyses	B. ベイズ流解析から得られた結果の報告
For completed clinical trials, sponsors should submit a clinical study report describing the design, analysis plan, and results from the trial. In addition to the typical content, the report of a trial with Bayesian analyses should include the following:	完了した臨床試験に対しては、スポンサーは、試験のデザイン、解析計画、および結果を記述する試験総括報告書を提出すべきである。一般的な内容に加えて、ベイズ流解析を伴う試験の報告書には以下を含むべきである：
• The principal aspects of the design and analysis plan, as described in Section VIII.A.	• VIII.A 節で記述されているような、デザインと解析計画の主要な側面。
• The results of the planned primary and secondary statistical analyses, including the treatment effect estimates, uncertainty in the estimates (e.g., with a credible interval), and whether pre-specified success criteria were met. The marginal posterior distributions of the estimands of interest should be described, including measures of location and variation. The results of sensitivity analyses, including quantifying and summarizing the sensitivity of results and conclusions to alternative reasonable choices for the prior distribution (see Section V.F). In the case of borrowing of external	• 治療効果推定値、その推定値における不確実性（例えば、信用区間を伴って）、および事前規定された成功規準が満たされたかどうかを含む、計画された主要および副次統計解析の結果。関心のある estimand の周辺事後分布は、位置とばらつきの指標を含めて記述されるべきである。事前分布に対する代替の理にかなった選定に対する、結果と結論の感度を定量化し、要約することを含む、感度分析の結果（V.F 節を参照）。主要解析において情報のある事前分布を伴う外部情報の借用がある場合には、異なる借用の程度を用

information with an informative prior in the primary analysis, the posterior results should be shown for other prior choices with different degrees of borrowing.	いた他の事前分布の選定に対する事後の結果についても示すべきである。
• The results from model checking, including an assessment of prior-data conflict and comparisons of the model predictions to the observed data.	• 事前分布とデータの対立の評価およびモデル予測値と観測データとの比較を含む、モデル点検から得られた結果。
• A report of sampling convergence diagnostics. For Bayesian analyses performed using non-direct sampling algorithms, this should include a description of the simulation settings as implemented. Any deviations from the planned implementation and a rationale for such deviations should be discussed.	• 標本抽出の収束診断の報告。直接的でない標本抽出アルゴリズムを用いて行われるベイズ流解析に対しては、これには、実装されたシミュレーション設定についての記述を含めるべきである。計画された実装からのあらゆる逸脱、およびこのような逸脱に対する合理的根拠が議論されるべきである。
• The software used to conduct analyses, including version details. Documented code should be provided for all primary and key secondary analyses and for any sensitivity analyses described in the study report. For Bayesian analyses performed using MCMC, sufficient detail should be provided such that the results are reproducible. This requires reporting of the seed number(s) used during chain initiation.	• バージョンの詳細を含む、解析を実施するために使用されたソフトウェア。文書化されたコードは、すべての主要および重要な副次解析に対して、ならびに試験報告書に記述されたあらゆる感度分析に対して提供されるべきである。MCMC を用いて行われるベイズ流解析に対しては、結果が再現可能となるように十分な詳細が提供されるべきである。これには、連鎖の開始時に使用されたシード番号の報告が必要とされる。
• A discussion of the overall conclusions about the evidence related to the key trial objectives and analyses. This should briefly comment on the model chosen and the sensitivity of results and conclusions to alternative plausible	• 重要な試験目的および解析に関連する証拠についての全体的な結論の議論。ここでは、選定されたモデル、ならびに代替の尤もらしい仮定および事前分布の規定に対する結果と結論の感度に関して簡潔に言及すべきである。

assumptions and prior specifications.	
In some cases, it may be reasonable to include highly technical information (e.g., reports of sampling convergence diagnostics) in an Appendix to the study report or in a separate dedicated document.	場合によっては、高度に技術的な情報（例えば、標本抽出の収束診断の報告）を、試験報告書の付録または別個の専用文書に含めることが理にかなっているかもしれない。

IX. GENERAL REFERENCES	IX. 一般的な参考文献
Adcock, C. J. "A Bayesian approach to calculating sample sizes." <i>Journal of the Royal Statistical Society: Series D (The Statistician)</i> 37, no. 4-5 (1988): 433-439. Evans, Michael, and Hadas Moshonov. "Checking for prior-data conflict." (2006): 893-914. Gelman, Andrew. "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)." (2006): 515-534. Gelman, Andrew, Aki Vehtari, Daniel Simpson, Charles C. Margossian, Bob Carpenter, Yuling Yao, Lauren Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. "Bayesian workflow." <i>arXiv preprint arXiv:2011.01808</i> (2020). Ji, Yuan, Ping Liu, Yisheng Li, and B. Nebiyu Bekele. "A modified toxicity probability interval method for dose-finding trials." <i>Clinical trials</i> 7, no. 6 (2010): 653-663. Jiang, Liyun, Lei Nie, and Ying Yuan. "Elastic priors to dynamically borrow information from historical data in clinical trials." <i>Biometrics</i> 79, no. 1 (2023): 49-60.	

- Joseph, Lawrence, and Patrick Bélisle. "Bayesian sample size determination for normal means and differences between normal means." *Journal of the Royal Statistical Society Series D: The Statistician* 46, no. 2 (1997): 209-226.
- Lin, Ruitao, Yanhong Zhou, Fangrong Yan, Daniel Li, and Ying Yuan. "BOIN12: Bayesian optimal interval phase I/II trial design for utility-based dose finding in immunotherapy and targeted therapies." *JCO precision oncology* 4 (2020): 1393-1402.
- Malec, Donald. "A closer look at combining data among a small number of binomial experiments." *Statistics in medicine* 20, no. 12 (2001): 1811-1824.
- Morita, Satoshi, Peter F. Thall, and Peter Müller. "Determining the effective sample size of a parametric prior." *Biometrics* 64, no. 2 (2008): 595-602.
- Neuenschwander, Beat, Michael Branson, and Thomas Gsponer. "Critical aspects of the Bayesian approach to phase I cancer trials." *Statistics in medicine* 27, no. 13 (2008): 2420-2439.
- Neuenschwander, Beat, Sebastian Weber, Heinz Schmidli, and Anthony O'Hagan. "Predictively consistent prior effective sample sizes." *Biometrics* 76, no. 2 (2020): 578-587.
- O'Hagan, Anthony, Caitlin E. Buck, Alireza Daneshkhah, J. Richard Eiser, Paul H. Garthwaite, David J. Jenkinson, Jeremy E. Oakley, and Tim Rakow. "Uncertain judgements: eliciting experts' probabilities." (2006).
- O'Quigley, John, and Mark Conaway. "Continual reassessment and related dose-finding designs." *Statistical science: a review journal of the Institute of Mathematical Statistics* 25, no. 2 (2010): 202.
- Pennello, Gene, and Laura Thompson. "Experience with reviewing Bayesian medical device trials." *Journal of Biopharmaceutical Statistics* 18, no. 1 (2007): 81-115.

Saville, Benjamin R., Donald A. Berry, Nicholas S. Berry, Kert Viele, and Scott M. Berry. "The Bayesian time machine: accounting for temporal drift in multi-arm platform trials." *Clinical Trials* 19, no. 5 (2022): 490-501.

Spiegelhalter, David J., Keith R. Abrams, and Jonathan P. Myles. *Bayesian approaches to clinical trials and health-care evaluation*. Vol. 13. John Wiley & Sons, 2004.

Thall, Peter F., and John D. Cook. "Dose-finding based on efficacy–toxicity trade-offs." *Biometrics* 60, no. 3 (2004): 684-693.

Tighiouart, Mourad, and André Rogatko, "Dose Finding with Escalation with Overdose Control (EWOC) in Cancer Clinical Trials". *Statistical Science*, 2010, 25: 217–226.

Wang, Yun, Wenda Tu, William Koh, James Travis, Robert Abugov, Kiya Hamilton, Mengjie Zheng, Roberto Crackel, Pablo Bonangelino, and Mark Rothmann. "Bayesian Hierarchical Models for Subgroup Analysis." *Pharmaceutical Statistics* 23, no. 6 (2024): 1065-1083.

Yuan, Ao, and Gen Xiang Chai. "Optimal adaptive generalized Pólya urn design for multi-arm clinical trials." *Journal of Multivariate Analysis* 99, no. 1 (2008): 1-24.

Yuan, Ying, Kenneth R. Hess, Susan G. Hilsenbeck, and Mark R. Gilbert. "Bayesian optimal interval design: a simple and well-performing design for phase I oncology trials." *Clinical Cancer Research* 22, no. 17 (2016): 4291-4301.