

### 3.6 What resources are available to learn more about Bayesian statistics?

#### 3.6 ベイズ統計について理解を深めるための資料

Non-technical introductory references to Bayesian statistics and their application to medicine include Malakoff (1999), Hively (1996), Kadane (1995), Brophy & Joseph (1995), Lilford & Braunholtz (1996), Lewis & Wears (1993), Bland & Altman (1998), and Goodman (1999a, 1999b). Berry (1997) has written for FDA an introduction specifically on Bayesian medical device trials.

ベイズ統計とその医療分野への適用に関する非技術的導入資料としては、Malakoff (1999)、Hively (1996)、Kadane (1995)、Brophy & Joseph (1995)、Lilford & Braunholtz (1996)、Lewis & Wears (1993)、Bland & Altman (1998)、Goodman (1999a, 1999b) 等による文献がある。Berry (1997) はFDA向けに特にベイズ法を用いた医療機器の臨床試験に関する入門書を執筆している。

A comprehensive summary on the use of Bayesian methods to design and analyze clinical trials or perform healthcare evaluations appears in Spiegelhalter, Abrams, & Myles (2004).

ベイズ法を用いた臨床試験のデザインおよび解析または医療評価の実施に関する包括的概要はSpiegelhalter, Abrams, & Myles (2004) の文献で確認できる。

Introductions to Bayesian statistics that do not emphasize medical applications, in order of complexity, are Berry (1996), DeGroot (1986), Stern (1998), Lee (1997), and Gelman, et al. (2004).

Berry (1996)、DeGroot (1986)、Stern (1998)、Lee (1997)、およびGelman, et al. (2004) は、ベイズ統計を医療への適用に焦点を置かず、複雑度の順に紹介している。

References with technical details and statistical terminology are Spiegelhalter, et al. (2000), Spiegelhalter, et al. (1994), Berry & Stangl (1996), Breslow (1990), and Stangl & Berry (1998).

技術的な詳細解説と統計用語を含む参考文献としては、Spiegelhalter, et al. (2000)、Spiegelhalter, et al. (1994)、Berry & Stangl (1996)、Breslow (1990)、およびStangl & Berry (1998) によるものがある。

An overview of Markov Chain Monte Carlo for Bayesian inference may be found in Gamerman (1997). Practical applications appear in Gilks, et al. (1996) and in Congdon (2003).

ベイズ法を用いた推論に関するマルコフ連鎖モンテカルロ法の概要はGamerman (1997) の著書に示されている。実践的応用についてはGilks, et al.

(1996) および Congdon (2003) の著書で述べられている。

Brophy & Joseph (1995) provide a well-known synthesis of three clinical trials using Bayesian methods.

Brophy & Joseph (1995) はベイズ法を用いて、よく知られている3つの臨床試験の結合を行っている。

A list of resources on the Web appears on the International Society for Bayesian Statistics website.<sup>3</sup>

ウェブサイト上の参考文献リストは International Society for Bayesian Statistics のサイト<sup>3</sup>で閲覧できる。

### **3.7 The Bayesian approach should not push aside sound science.**

**3.7 ベイズ法では信頼性のある科学知識を軽視してはならない**  
Scientifically sound clinical trial planning and rigorous trial conduct are important regardless of whether you use the Bayesian or the frequentist approach. We recommend you remain vigilant regarding randomization, concurrent controls, prospective planning, blinding, bias, precision, and all other factors that go into a successful clinical trial. See **Section 5.1: Bayesian trials start with a sound clinical trial design.**

科学的に信頼性のある臨床試験の計画および厳密な試験実施は、ベイズ法と頻度論的手法のいずれを用いるかにかかわらず重要である。無作為化、同時対照、前向き計画、盲検化、バイアス、精度、および臨床試験を成功に導くその他のすべての要素について、常に慎重に配慮することを推奨する。**5.1** ベイズ法を用いた臨床試験は、信頼性のある試験デザインからを参照のこと。

---

<sup>3</sup> <http://www.bayesian.org/>

<sup>3</sup> <http://www.bayesian.org/>

### **3.8 What are the potential benefits of using Bayesian methods?**

#### **3.8 ベイズ法の使用によって得られる可能性のある利点**

##### **Sample size reduction or augmentation**

サンプルサイズの縮小または拡大

The Bayesian methodology may reduce the sample size FDA needs to reach a regulatory decision. You may achieve this reduction by using prior information and interim looks during the course of the trial. When results of a trial are unexpectedly good (or unexpectedly bad) at an interim look, you may be able to stop early and declare success (or failure).

ベイズ法ではFDAが規制措置の決定のために必要とするサンプルサイズを縮小することができる。これは、事前知識と臨床試験中の中間調査を用いることによって実現できる。臨床試験の結果が中間調査の時点で予想外に良好（または予想外に不良）であった場合、試験を早期に中止し、試験の成功（または失敗）を表明することができる。

The Bayesian methodology can allow for augmentation of the sample in cases where more information helps FDA make a decision. This can happen if the observed variability of the sample is higher than that used to plan the trial.

ベイズ法では、さらなる知識がFDAの規制措置決定に役立つ場合には、サンプルサイズを拡大することができる。これは、観察されたサンプルのばらつきが試験計画に用いられたものより大きい場合に実現する。

##### **Midcourse changes to the trial design**

試験デザインの途中変更

With appropriate planning, the Bayesian approach can also offer the flexibility of midcourse changes to a trial. Some possibilities include dropping an unfavorable treatment arm or modifications to the randomization scheme. Modifications to the randomization scheme are particularly relevant for an ethically sensitive device or when enrollment becomes problematic for a treatment arm. Bayesian methods can be especially flexible in allowing for changes in the treatment to control randomization ratio during the course of the trial. See Kadane (1996) for a discussion.

適切な計画を立ててベイズ法を用いることにより、臨床試験で柔軟に途中変更を行うことができる。変更の可能性のあるものとしては、好ましくない治療群の脱落、無作為化の方法の修正などがある。無作為化の方法修正は特に倫理的に注意を要する機器の場合や治療群への登録に問題がある場合に行われる。ベイズ法を用いることで、特に試験中の無作為化の比率をコントロールするための治療変更を柔軟に行うことができる。Kadane (1996) の考察

を参照。

### **Exact analysis**

#### 正確な解析

The Bayesian approach can sometimes be used to obtain an exact analysis when the corresponding frequentist analysis is only approximate or is too difficult to implement.

ベイズ法は、対応する頻度論的解析が概算にすぎない場合や実施困難な場合に、正確な解析結果を得るために用いられることがある。

### **3.9 What are the potential difficulties in the Bayesian approach?**

#### **3.9 ベイズ法の使用によって生じる可能性のある問題点**

#### **Extensive preplanning**

##### 広範な事前計画

Planning the design, conduct, and analysis of any trial are always important from a regulatory perspective, but they are crucial for the Bayesian approach. This is because decisions are based on:

- prior information
- information obtained from the trial, and
- the mathematical model used to combine the two.

臨床試験のデザイン、実施および解析の計画は規制措置の観点から常に重要であるが、ベイズ法にとっても不可欠である。これはベイズ法による解析結果が以下に基づいているためである。

- 事前知識
- 試験から得られた情報
- 上記の2つを結合するために使用される数学的モデル

Different choices of prior information or different choices of model can produce different decisions. As a result, in the regulatory setting, the design of a Bayesian clinical trial involves pre-specification of (and agreement on) both the prior information and the model. This includes clinical agreement on the appropriateness of the prior information and statistical agreement on the mathematical model to be used. Since reaching this agreement is often an iterative process, we recommend you meet with FDA early on to discuss and agree upon the basic aspects of the trial design.

事前知識やモデルの選択が異なると、結論も異なってくる。このため、規制措置の決定においては、ベイズ法を用いた臨床試験のデザインには事前知識

とモデルの両方についての事前規定（および合意）が必要である。これには、事前知識の適切性に関する臨床的合意と使用される数学的モデルに関する統計学的合意が含まれる。こうした合意に達するには反復的プロセスを経る場合が多いため、早い段階でFDAと検討会を行い、試験デザインの基本的側面について検討し合意することを推奨する。

A change in the prior information or the model at a later stage of the trial may imperil the scientific validity of the trial results. For this reason, formal agreement meetings may be appropriate when using a Bayesian approach.

試験の後期段階で事前知識やモデルの変更を行うと、試験結果の科学的妥当性が損なわれる可能性がある。このため、ベイズ法を用いる場合には正式な合意のための検討会を行うことが適切と考えられる。

Specifically, the identification of the prior information may be an appropriate topic of an agreement meeting.<sup>4</sup>

特に事前知識の確認は、合意検討会の妥当な課題と考えられる。<sup>4</sup>

---

<sup>4</sup> The FDA Modernization Act (FDAMA) provided for two types of early collaboration meetings: agreement meetings and determination meetings. For details, see the FDA Guidance on early collaboration meetings at <http://www.fda.gov/cdrh/ode/guidance/310.html>.

<sup>4</sup> FDAによる近代化法（FDAMA）では、合意検討会と措置決定会の2つの早期合意検討会を提案している。詳細は早期合意検討会に関するFDAガイダンスを参照（<http://www.fda.gov/cdrh/ode/guidance/310.html>）。

### **Extensive model-building**

#### 広範なモデル構築

The Bayesian approach often involves extensive mathematical modeling of a clinical trial, including:

- the probability distributions chosen to reflect the prior information
- the influence of covariates on patient outcomes or missing data
- the relationships between various sources of prior information used in the model.

ベイズ法では多くの場合、以下を含む広範な臨床試験の数学的モデリングを行う。

- 事前知識の反映のために選択された確率分布
- 患者の転帰や欠測値に対する共変量の影響
- モデルで使用される多様な事前知識源の関連性

We recommend you determine modeling choices through close collaboration and agreement with FDA's and your statistical and clinical experts.

FDAおよび自社の統計専門家、臨床専門家との緊密な協業および合意に従ってモデル選択を行うことを推奨する。

### **Specific statistical and computational expertise**

#### 特殊な統計・計算技術

The Bayesian approach often involves specific statistical expertise.

Computer-intensive calculations are often used to:

- analyze trial data
- check model assumptions
- assess prior probabilities at the design stage
- perform simulations to assess probabilities of various outcomes
- estimate sample size.

ベイズ統計ではしばしば特殊な統計技術が必要とされる。

以下については、コンピュータを利用した計算が行われることが多い。

- 試験データの解析
- モデル仮定の確認
- デザイン段階での事前確率評価
- 多様な転帰の確率評価のためのシミュレーション実施
- サンプルサイズの推定

The technical and statistical costs for the above are often offset by the savings of a shorter trial or a more flexible analysis.

上記にかかる技術的・統計的コストは多くの場合、より短期間の試験またはより柔軟性の高い解析による節約で相殺される。

### **Choices regarding prior information**

#### 事前知識に関する選択

An FDA advisory panel may question prior information you and FDA agreed upon beforehand. We recommend you be prepared to clinically and statistically justify choices of prior information. In some cases, we recommend you perform sensitivity analyses to check robustness of models and priors.

FDAの諮問委員会では、企業とFDAがあらかじめ合意した事前知識について質問する場合があります。このため、事前知識の選択について臨床的・統計的な根拠を示すことができるようにしておくことを推奨する。場合によっては、モデルおよび事前知識の信頼性を確認する感度分析の実施を推奨するものとする。

### **Device labeling**

#### 機器のラベリング

Results from a Bayesian trial may be expressed differently from the way trial results are usually described in device labels. We recommend you ensure trial results reported on the device label are easy to understand.

ベイズ法を用いた臨床試験の結果は、通常の試験結果を機器のラベルに示す方法とは異なる方法で示される場合があります。機器のラベルに示す試験結果は、分かりやすいものにすることを推奨する。

### **Checking calculation**

#### 計算の確認

The flexibility of Bayesian models and the complexity of the computational techniques for Bayesian analyses create greater possibility for errors and misunderstandings. FDA, therefore, will carry out a detailed statistical review of a Bayesian submission.

ベイズ法のモデルの柔軟性およびベイズ解析のための計算技術の複雑性により、過誤や解釈の相違が生じる可能性が非常に高くなる。このため、FDAではベイズ法を用いた申請について詳細な統計学的検討を行う。

Since the software used in Bayesian analysis is relatively new, FDA will often verify results using alternate software. FDA recommends you submit your data and any programs used for Bayesian statistical analyses electronically.

ベイズ解析に用いられるソフトウェアは比較的新しいものであるため、FDA

は代替ソフトを用いて結果を検証する機会が多い。FDAでは、ベイズ法を用いた統計解析に使用されたデータおよびプログラムの電子的な提出を推奨している。

### **Bayesian and traditional analyses approaches may differ**

#### ベイズ解析と従来の解析法との違い

Two investigators, each with the same data and a different preplanned analysis (one frequentist and one Bayesian), could conceivably reach different conclusions that are both scientifically valid. While the Bayesian approach is often favorable to the investigator with good prior information, the approach can be more conservative than a frequentist approach (e.g., see **Section 5: Planning a Bayesian Clinical Trial**).

同じデータと異なる事前計画解析（一方は頻度論的手法、もう一方はベイズ法）を用いた2人の研究者が、いずれも科学的に信頼性のある異なる結論に達する可能性がある。ベイズ法は良好な事前知識を持つ研究者にとっては都合の良い場合が多いが、この手法は頻度論的手法よりも控えめな結果をもたらす可能性がある（例：5 ベイズ法を用いた臨床試験の計画参照）。

If the results from your pre-planned analysis are not as positive as expected, we recommend you do not switch from a frequentist to a Bayesian analysis (or vice versa). Such post hoc analyses are not scientifically sound and would tend to weaken the validity of the submission.

事前に計画した解析の結果が予想を下回る場合、頻度論的手法からベイズ法へ（もしくはその逆）の切り替えを行わないことを推奨する。そのような事後解析は科学的信頼性に欠け、申請の妥当性を損なう傾向があるためである。

## 4. Bayesian Statistics

### 4. ベイズ統計

#### 4.1 Introduction

##### 4.1 はじめに

The fundamental idea in Bayesian statistics is that one's uncertainty about an unknown quantity of interest is represented by probabilities for possible values of that quantity. For instance, unknown quantities of interest in device trials might be:

- clinical safety and effectiveness endpoints
- a patient's outcome to be observed in the future
- a missing observation on a patient.

ベイズ統計の基本的概念は、不特定量の関心事項に関する不確実性を、その不特定量の関心事項について考えられる値の確率によって示すことである。例えば、医療機器の臨床試験における不特定量の関心事項としては、以下を挙げることができる。

- 臨床的安全性および有効性のエンドポイント
- 患者に将来みられる転帰
- 患者に関する欠測事項

#### Prior distribution and non-informative prior distribution

##### 事前分布および無情報事前分布

Before a trial begins and data are obtained, the investigator assigns prior probabilities to the possible values of the unknown quantity, known as the prior distribution. In principle, the prior can be based on the investigator's personal knowledge of the quantities of interest or on another expert's opinion, etc. If absolutely nothing is known about that quantity, something called a non-informative prior distribution may be specified. In trials undergoing regulatory review, however, the prior distribution is usually based on data from relevant previous trials. 臨床試験の開始およびデータの入手前に、治験責任医師は事前分布として知られる不特定量の関心事項について考えられる値に事前確率を割り当てる。原則として、事前確率は特定数の関心事項に関する治験責任医師の個人的な知識または別の専門家の意見等に基づいて決定される。この関心事項の量についてまったく情報がない場合には、無情報事前分布と呼ばれるものが定められることがある。しかし、規制上の検討が行われる臨床試験では、事前分布は通常関連する過去の臨床試験から得られたデータに基づいている。

## **Bayes' theorem and posterior probabilities**

### ベイズの定理および事後確率

After data are gathered and information becomes available, the prior probabilities are mathematically updated according to a statistical result called Bayes' theorem. The updated probabilities, known as posterior probabilities, are probabilities for values of the unknown quantity after data are observed. This approach is a scientifically valid way of combining previous information (the prior probabilities) with current data. The approach adjusts to changing levels of evidence: today's posterior probabilities become tomorrow's prior probabilities.

データ収集および情報入手後、事前確率はベイズの定理と呼ばれる統計結果に従って数学的に更新される。更新された確率は事後確率として知られ、データ観察後の不特定量の関心事項の値に関する確率である。これは事前知識（事前確率）と最新データを結合する科学的に有効な手法である。この手法により、エビデンスのレベル変化を調整できる。すなわち、現在の事後確率が将来の事前確率となる。

## **The Bayesian paradigm**

### ベイズ情報学

The Bayesian paradigm states that probability is the only measure of one's uncertainty about an unknown quantity. In a Bayesian clinical trial, uncertainty about an endpoint (also called parameter) is quantified according to probabilities, which are updated as information is gathered from the trial.

ベイズ情報学では、確率は不特定の量に関する不確実性の大きさを示すものにすぎないとされている。ベイズ法を用いた臨床試験では、エンドポイント（パラメータとも呼ばれる）に関する不確実性を確率に従って数量化し、それを試験で収集される情報により更新する。

## **Decision rules**

### 決定則

The pre-market evaluation of medical devices aims to demonstrate the safety and effectiveness of a new device. This demonstration is most commonly achieved through statistical hypothesis testing. For Bayesian trials, hypotheses are tested with decision rules. One common type of decision rule considers that a hypothesis has been demonstrated (beyond a reasonable doubt) if its posterior probability is large enough (e.g., 95 or 99 percent).

医療機器の上市前評価は、新しい機器の安全性および有効性の実証を目的としている。この実証は統計学的仮説検定によって行われるのが最も一般的である。ベイズ法を用いた臨床試験では、仮説は決定則に従って検証される。ある一般的な決定則では、ある仮説はその事後確率が十分に高い場合（例：95%または99%）に実証されると考えられる。

The Bayesian approach encompasses a number of key concepts, some of which are not part of the traditional statistical approach. Below, we briefly discuss these concepts and contrast the Bayesian and frequentist approaches.

ベイズ法には多数の重要な概念が含まれており、その中には従来の統計学的手法には含まれていないものもある。このような概念については以下で簡潔に述べ、ベイズ法と頻度論的手法を比較するものとする。

## 4.2 What is a prior distribution?

### 4.2 事前分布とは

Suppose that  $x$  is an endpoint (parameter) of interest in a clinical trial. The initial uncertainty about  $x$  should be described by a probability distribution for  $x$ , called the prior distribution and denoted by  $P(x)$ .

ある臨床試験における関心のあるエンドポイント(パラメータ)を $x$ とする。 $x$ に関する当初の不確実性は、事前分布と呼ばれる $x$ の確率分布 ( $P(x)$ ) によって示される。

As an example, suppose  $x$  is the rate of a serious adverse event. Its possible values will lie between 0 and 1. One prior distribution is the uniform distribution indicating no preference for any value of  $x$ . So the probability that  $x$  lies between 0.1 and 0.2 is the same as the probability that  $x$  lies between 0.4 and 0.5, or between 0.65 and 0.75, or in any interval of length 0.1.

例えば、 $x$ を重篤な有害事象の発生率とする。その可能性のある値は0と1の間となる。1つの事前分布は $x$ がどのような値でも優先されないことを示す一様分布である。したがって、 $x$ の確率が0.1と0.2の間となる確率は、0.4と0.5の間となる確率、0.65と0.75の間となる確率、あるいは他の0.1間隔となる確率と同じである。

Alternatively, the prior distribution might give preference to lower values of  $x$ . For example, the probability that  $x$  lies between 0.2 and 0.3 can be larger than the probability that  $x$  lies between 0.7 and 0.8.

一方、事前分布では $x$ の低値が優先されることもある。例えば、 $x$ が0.2と0.3の間となる確率は $x$ が0.7と0.8の間となる確率よりも高くなる可能性がある。

## 4.3 What is the likelihood of the observed data?

### 4.3 測定値の尤度とは

Now suppose data have been obtained from a clinical trial. The likelihood of these data being observed can be formally expressed in terms of a likelihood function,  $P(\text{data} | x)$ , which is the conditional probability of observing the data, given a specific value of  $x$ , for each possible value of  $x$  (the parameter). The likelihood is

the mathematical and statistical model that reflects the relationships between the observed outcomes in the trial, the covariates, and the endpoint  $x$  of interest.

さて、ここで臨床試験からデータが得られたとする。これら測定データの尤度は形式上、尤度関数 $P(\text{data} | x)$ で示されるが、これは $x$ が特定の値であるという前提で $x$ の可能性のある各値（パラメータ）についてデータが測定される条件付き確率である。尤度は臨床試験で観察された転帰、共変量、および関心のあるエンドポイント $x$ の関連性を反映する数学的・統計学的モデルである。

#### 4.4 What is the posterior distribution?

##### 4.4 事後分布とは

The final objective is to obtain the probability of each possible value of the endpoint  $x$  conditional on the observed data, denoted  $P(x | \text{data})$ . Using exclusively the laws of probability, Bayes' theorem combines the prior distribution for  $x$ ,  $P(x)$ , with the likelihood,  $P(\text{data} | x)$ , in order to obtain the posterior distribution for  $x$ ,  $P(x | \text{data})$ . In the Bayesian approach, all available information about  $x$  is summarized by the posterior distribution,  $P(x | \text{data})$ , and all inferences about the endpoint are based on it. 最終的な目的は、測定データを条件とするエンドポイント $x$ の可能性のある各値の確率 ( $P(x | \text{data})$ ) を求めることである。確率の法則のみを用いて、ベイズの定理では $x$ の事前分布 $P(x)$ と尤度 $P(\text{data} | x)$ を結合し、 $x$ の事後分布 $P(x | \text{data})$ を導出する。ベイズ法では、 $x$ について入手可能なすべての情報は事後分布 $P(x | \text{data})$ によってまとめられ、エンドポイントに関するすべての推論はこれに基づいて行われる。

As more data are obtained, more updating can be performed. Consequently, the posterior distribution that has been obtained today may serve as a prior distribution later, when more data are gathered. The more information that is accrued, the less uncertainty there will be about the posterior distribution for  $x$ , and as more and more information is collected from the trial, the influence of the prior will become less and less. If enough data are collected, the relative importance of the prior distribution will be negligible compared to the likelihood.

多くのデータが得られるほど更新も頻繁に行うことができる。この結果、今日得られた事後分布は、後日さらなるデータが収集された際には事前分布となる。得られる情報が多いほど $x$ の事後分布に関する不確実性が少なくなり、臨床試験から収集される情報が多ければ多いほど、事前分布の影響が小さくなる。十分なデータが収集された場合、事前分布の相対的重要性は尤度に比べてごくわずかとなる。

For more introductory material on conditional probability, Bayes' theorem, and Bayesian statistics, see DeGroot (1986), Lee (1997), Berry (1996), Lindley (1985).

For an introduction specific to medical devices, see Berry (1997).  
条件付き確率、ベイズの定理およびベイズ統計に関するさらなる導入資料は、DeGroot (1986)、Lee (1997)、Berry (1996)、Lindley (1985) を参照。  
特に医療機器に関する導入資料はBerry (1997) を参照。

#### 4.5 What is a predictive distribution?

##### 4.5 予測分布とは

The Bayesian approach allows for the derivation of a special type of posterior probability; namely, the probability of future events given outcomes that have already been observed. This probability is called the predictive probability. Collectively, the probabilities for all possible values of future outcomes are called the predictive distribution. Predictive distributions have many uses, including:

- determining when to stop a trial
- helping a physician and patient make decisions by predicting the clinical outcome of the current patient, given the observed outcomes of past patients in a clinical trial
- predicting a clinical outcome from a validated surrogate
- adjusting trial results for missing data (imputation)
- model checking.

ベイズ法では特定の事後確率、つまり、既に観察された転帰を前提とした未来の事象の確率の導出が可能である。この確率を的中率という。未来の転帰の可能性のあるすべての値に関する確率を総じて予測分布と呼ぶ。予測分布には以下を含め、多数の用途がある。

- 臨床試験の中止時期を決定する
- 臨床試験において過去の患者で観察された転帰をもとに現在の患者の臨床転帰を予測することによって医師および患者の意思決定を支援する
- 妥当性が実証された代用例から臨床転帰を推定する
- 欠測値について試験結果を調整する (データ補完)
- モデル確認を行う

These uses are discussed in more detail in **Section 6. Analyzing a Bayesian Clinical Trial**.

これらの用途については6.ベイズ法を用いた臨床試験の解析でさらに詳細に考察する。

## 4.6 What is exchangeability?

### 4.6 交換可能性とは

Exchangeability is a key idea in statistical inference in general, but it is particularly important in the Bayesian approach. Two observations are exchangeable if they provide equivalent statistical information. So, two patients randomly selected from a particular population of patients can be considered exchangeable. If the patients in a clinical trial are exchangeable with the patients in the population for which the device is intended, then the clinical trial can be used to make inferences about the entire population. Otherwise, the trial tells us very little about the larger population. The concept of a representative sample can thus be expressed in terms of exchangeability.

交換可能性とは、一般に統計的推論において重要な概念であるが、特にベイズ法においては重要である。2つの測定データから同等の統計学的情報が得られる場合、これらのデータは交換可能である。したがって、ある特定の患者集団から無作為に選択された2人の患者は交換可能と考えられる。ある臨床試験の患者が医療機器を使用する集団の患者と交換可能な場合、その臨床試験は集団全体に関する推論に用いることができる。そうでない場合、その臨床試験からは、より大きな集団についてほとんど情報を得ることができない。したがって、代表サンプルの概念は交換可能性の観点から示すことができる。

Exchangeability may depend on the statistical model used. If we know, for example, that the adverse event rate for a particular device depends on the patient's body mass index (BMI), then we say that patients are exchangeable conditional on BMI. That is, two patients will provide equivalent statistical information, but only after we account for differences in BMI. Therefore, any discussion of exchangeability should also include a discussion of the statistical models used.

交換可能性は使用する統計モデルによって異なる。例えば特定の医療機器に関する有害事象の発生率が患者の体格指数 (BMI) に左右されることが分かっている場合、患者はBMIを条件として交換可能ということになる。つまり、BMIの差を明らかにして初めて2人の患者から同等の統計学的情報が得られるのである。したがって、交換可能性に関するすべての考察には使用する統計モデルに関する考察も含めるべきである。

We can also think of exchangeability in terms of clinical trials. Two trials are exchangeable if they provide equivalent statistical information about some super-population of clinical trials. Again, the trials may be exchangeable, but only after we account for (that is, condition on) other factors with the appropriate statistical model.

交換可能性については、臨床試験の観点からも考察することができる。2つ

の臨床試験は、それらが一部の臨床試験の超母集団に関して同等の統計学的情報を提供する場合に交換可能である。この場合も、当該試験は適切な統計モデルによってその他の条件を明らかにする（つまり条件を明確にする）ことで初めて交換可能となる。

The use of Bayesian hierarchical models enables us to combine information from different sources that may be exchangeable on some levels but not on others (see **Section 5: Planning a Bayesian Clinical Trial**). If trials are exchangeable, then Bayesian hierarchical models enable us to make full use of the information from all the trials. For technical definitions of exchangeability, see Bernardo & Smith (1995).

ベイズ型階層モデルを用いて、一部のレベルで交換可能であるがその他のレベルではそうではない複数の情報源からの情報を結合することができる（5 ベイズ法を用いた臨床試験の計画参照）。臨床試験が交換可能な場合、ベイズ型階層モデルを用いることによって、それらの試験すべてから得られた情報を十分に活用することができる。交換可能性の技術的定義については、Bernardo & Smith (1995) を参照のこと。

#### 4.7 What is the likelihood principle?

##### 4.7 尤度原理とは

The likelihood principle is important in all of statistics, but it is especially central to the Bayesian approach. The principle states that all information about the endpoint of interest,  $x$ , obtained from a clinical trial, is contained in the likelihood function. In the Bayesian approach, the prior distribution for  $x$  is updated using the information provided by the trial through the likelihood function, and nothing else. Bayesian analysts base all inferences about  $x$  solely on the posterior distribution produced in this manner.

尤度原理はすべての統計において重要であるが、特にベイズ法においては不可欠である。尤度原理では、臨床試験から得られる関心のあるエンドポイント  $x$  に関するすべての情報は尤度関数に含まれる、とされる。ベイズ法では  $x$  の事前分布を臨床試験から得られる情報を用いて尤度関数のみによって更新する。ベイズ解析では  $x$  に関するすべての推論をこのようにして得られた事後分布にのみ基づいて行う。

A trial can be altered in many ways without changing the likelihood function. As long as the modification schemes are pre-specified in the trial design, adherence to the likelihood principle allows for flexibility in conducting Bayesian clinical trials, in particular with respect to:

- modifications to sample size

- interim looks for the purpose of possibly stopping the trial early
- adaptive designs in general.

臨床試験は尤度関数を変更することなく、さまざまに修正することができる。修正計画が試験デザインで事前に示されている限り、尤度原理を遵守することによって、ベイズ法を用いた臨床試験の実施において特に以下の点で柔軟性がもたらされる。

- サンプルサイズの変更
- 試験の早期中止の可能性評価を目的とした中間調査
- 適応できるデザイン全般

For more on the topic, see Berger & Wolpert (1988), Berger & Berry (1988), Irony (1993), and Barlow, et al., (1989).

尤度原理の詳細についてはBerger & Wolpert (1988)、Berger & Berry (1988)、Irony (1993)、およびBarlow, et al. (1989) 参照。

## 4.8 How do the Bayesian and frequentist approaches differ?

### 4.8 ベイズ法と頻度論的手法の違い

As outlined above, Bayesian analysts base all inferences on the posterior distribution, which (in adherence to the likelihood principle) is the product only of the prior and the likelihood function. Although the frequentist approach makes extensive use of the likelihood function, it does not always strictly adhere to the likelihood principle. For example, the interpretation of a frequentist p-value is based on outcomes that might have occurred but were not actually observed in the trial; that is, on something external to the likelihood.

前述の通り、ベイズ解析ではすべての推論を、(尤度原理を遵守した) 事前分布および尤度関数のみによって得られる事後分布に基づいて行う。頻度論的手法では尤度関数を大いに活用するが、常に尤度原理を遵守しているわけではない。例えば、頻度論的p値の解釈は、起こる可能性はあったが実際に試験では観察されなかった転帰、つまり尤度の範囲外にあるものに基づいて行われる。

Another way of saying this is that Bayesian inferences are based on the “parameter space” (the posterior distribution), while frequentist inferences are based on the “sample space” (the set of possible outcomes of a trial).

言い換えると、ベイズ法による推論は「パラメータスペース」(事後分布)に基づいているが、頻度論的手法による推論は「サンプルスペース」(臨床試験で考えられる一連の転帰)に基づいているといえる。

## 5. Planning a Bayesian Clinical Trial

### 5. ベイズ法を用いた臨床試験の計画

#### 5.1 Bayesian trials start with a sound clinical trial design

##### 5.1 ベイズ法を用いた臨床試験は信頼性のある試験デザインから

The basic tenets of good trial design are the same for both Bayesian and frequentist trials. Parts of a comprehensive trial protocol include:

- objectives of the trial
- endpoints to be evaluated
- conditions under which the trial will be conducted
- population that will be investigated
- planned statistical analysis.

優れた試験デザインの基本的理念は、ベイズ法を用いた臨床試験、頻度論的手法を用いた臨床試験のいずれにおいても同じである。包括的な試験実施計画書の一部には以下が含まれる。

- 試験の目的
- 評価項目（エンドポイント）
- 試験の実施条件
- 試験対象となる集団
- 統計解析計画

We recommend you follow the principles of good clinical trial design and execution, including minimizing bias. Randomization minimizes bias that can be introduced in the selection of which patients get which treatment. Randomization allows concrete statements about the probability of imbalances in covariates due to chance alone. For reasonably large sample sizes, randomization ensures some degree of balance for all covariates, including those not measured in the study.

バイアスを最小限に抑えることを含め、優れた臨床試験のデザインおよび実施の原則に従うことを推奨する。無作為化により、患者の治療割り付けにおいて導入され得るバイアスを最小限に抑えることができる。無作為化をすることによって、偶然のみによる共変量の不均衡の確率について具体的な記述が可能となる。妥当な大きさのサンプルサイズでは、無作為化を行うことにより、試験で観察されないものを含め、すべての共変量についてある程度の均衡を保つことができる。

Masking (also known as blinding) of physicians avoids bias that can be introduced by intended or unintended differences in patient care or evaluation of patient outcomes based on the treatment received during the course of the trial. Masking of patients minimizes biases due to the placebo effect.

医師のマスクング（盲検化ともいう）により、患者の治療における意図的または非意図的な違い、あるいは試験中に患者が受けた治療に基づく転帰の評価によって導入される可能性のあるバイアスが回避される。また、患者のマスクングにより、プラセボ効果によるバイアスが最小限に抑えられる。

We recommend you choose beforehand the type of analysis to be used (Bayesian or frequentist). Switching to a more favorable analysis method introduces a bias. It is difficult to justify a switch or account for such a bias in the analysis stage. In some cases, a Bayesian analysis of a new trial may salvage some information obtained in a previous non-Bayesian clinical trial that deviated from the original protocol. The information provided by such a trial may be represented by a prior distribution to be used in a prospective Bayesian clinical trial.

事前に使用する解析の種類（ベイズ法もしくは頻度論的手法）を選択しておくことを推奨する。より好ましい解析法への切り替えを行うとバイアスが生じる。解析段階で切り替えの根拠を示すことやそのようなバイアスを明確にすることは困難である。場合によっては、新たな試験のベイズ解析を行うことによって、以前ベイズ法を用いない臨床試験で得られたが当初の治験実施計画書から逸脱した一部の情報を回収できることもある。このような試験によって得られた情報は、ベイズ法を用いた前向き臨床試験で使用される事前分布によって示される。

For further information on planning a trial, see FDA's Statistical Guidance for Non-Diagnostic Medical Devices.<sup>5</sup>

臨床試験の計画に関する詳細は、FDAのStatistical Guidance for Non-Diagnostic Medical Devices<sup>5</sup>を参照のこと。

## **5.2 Selecting the relevant endpoints or parameters of interest**

### **5.2対象となる適切なエンドポイントまたはパラメータを選択する**

Endpoints (also called parameters in this document) are the measures of safety and effectiveness used to support a certain claim. Ideally, endpoints are:

- clinically relevant
- directly observable
- related to the claims for the device
- important to the patient.

エンドポイント（本書ではパラメータともいう）は、ある主張を支持するために用いる安全性および有効性の指標である。理想的なエンドポイントの条件は、以下の通りである。

- ・ 臨床的に関連性がある
- ・ 直接観察可能である
- ・ 試験機器の性能に関連している
- ・ 患者にとって重要である

For example, an endpoint may be a measure of the size of a treatment effect, or the difference between the effects in the treatment and control groups. The objective of a clinical trial is to gather information from the patients in the trial to make inferences about these unknown parameters.

たとえば、エンドポイントは、治療効果の程度あるいは治療群と対照群における効果の差を示す指標となる。臨床試験の目的は、試験に参加した患者から情報を収集し、このような不明確なパラメータについて推論することである。

### 5.3 Collecting other important information: covariates

#### 5.3 その他の重要な情報の収集：共変量

Covariates, also known as confounding factors, are characteristics of the study patients that can affect their outcome. There are many statistical techniques (Bayesian and frequentist) to adjust for covariates. Covariate adjustment is especially important in any situation where some degree of covariate balance is not assured through randomization, such as a Bayesian trial in which other trials are used as prior information. If adjustments are not made for differences in the covariate distribution between trials, the analysis can be biased. Covariate adjustment is also often used to reduce variation, which leads to a more powerful analysis.

共変量は交絡因子としても知られるが、これは被験者の転帰に影響を及ぼす可能性のある被験者の特性を示す。共変量の調整については多くの統計学的手法（ベイズ統計および頻度論的手法）がある。共変量の調整は、ベイズ統計を用いた臨床試験で他の試験を事前知識として用いる場合など、ある程度の共変量バランスが無作為化によって実現されない場合には特に重要である。試験間の共変量分布の差を調整しなければ、解析にバイアスがかかる可能性がある。また、共変量の調整はしばしばばらつきを抑えるために用いられるが、これによって解析の検出力がさらに高まる。

---

<sup>5</sup> <http://www.fda.gov/cdrh/ode/ot476.html>

<sup>5</sup> <http://www.fda.gov/cdrh/ode/ot476.html>

## 5.4 Choosing a comparison: controls

### 5.4 対照の選択：対照群

To facilitate evaluation of clinical trial results, we recommend you use a comparator, or control group, as a reference. Types of control groups you may use are:

- Concurrent controls
- self controls
- historical controls.

臨床試験結果の評価を容易にするため、対照あるいは対照群を参照として用いることを推奨する。用いることのできる対照群は以下の通りである。

- 同時対照群
- 自己対照群
- 歴史的対照群

We believe that self controls and historical controls are scientifically less rigorous than concurrent controls because of:

- potential problems with covariate adjustment
- placebo effect
- regression to the mean.

以下の理由から、自己対照群および歴史的対照群は同時対照群に比べて科学的正確性が低いと考えられる。

- 共変量の調整に関する潜在的な問題
- プラセボ効果
- 平均値への回帰

Another way to characterize the type of control is to distinguish between controls that are treated with an effective therapy (active controls) vs. controls that either receive no treatment (inactive controls) or are treated with a sham device (placebo controls). Bayesian methods are especially useful with active controlled trials seeking to demonstrate a new device is not only non-inferior to the active control but is also superior to no treatment or a sham control. A Bayesian trial can investigate this question by using previous studies comparing the active control to the inactive control. Bayesian methods for active control trials are discussed in Gould (1991) and Simon (1999).

対照の特性を示すもう1つの方法は、効果的な治療を受ける対照群（実対照）と治療を受けない対照群（無治療対照）あるいは偽の機器で治療を受ける対照群（プラセボ対照）とを区別する方法である。ベイズ法は、新たな機器が実対照に劣らないだけでなく無治療対照やプラセボ対照より優れているこ