

第4回 DSRT  
テーマ4: ベイズ統計学の医薬品の  
臨床開発での活用について

【議論のまとめ】  
テーブル1：午前

# はじめに

- [Note]は、当日の議論に対して、公開用にファシリテータが補足した内容である

# ベイズチームの議題

本発表の対象

- 議題1：ベイズ統計学の導入とFDAの医療機器ガイダンスから学べること
- 議題2：p値についての議論
- 議題3：ベイズ統計学の検証試験における意思決定での利用について
- 議題4：事前分布の構成・ヒストリカルデータの利用

# 議題1：ベイズ統計学の導入とFDAの 医療機器ガイダンスから学べること

# 本議題の内容

- 本議題の内容

- FDAのガイダンス”Guidance for the Use of Bayesian Statistics in **Medical Device Clinical Trials** “（以下、FDAガイダンス）の中で「**どういう状況でベイズ統計学が使えるか**」「**ベイズ統計学を使うことのメリット**」について記載された箇所の内容を整理し、FDAガイダンスに基づいた**医薬品開発での使用可能性**や、**使用が難しい状況**について議論する

# 議題1：ベイズ統計学の導入とFDAの 医療機器ガイダンスから学べること [論点1の導入：ガイダンスの抜粋(2.1～2.2)]

# FDAの医療機器ガイダンス

## **Guidance for Industry and FDA Staff**

---

### **Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials**

Document issued on: February 5, 2010

The draft of this document was issued on 5/23/2006

# ガイダンス目次

## 1. Introduction

## 本議題の主な対象

### 2. Foreword

- 2.1 What is Bayesian statistics?
- 2.2 Why use Bayesian statistics for medical devices?
- 2.3 Why are Bayesian methods more commonly used now?
- 2.4 When should FDA participate in the planning of a Bayesian trial?
- 2.5 The Bayesian approach is not a substitute for sound science
- 2.6 What are potential benefits of using Bayesian methods?
- 2.7 What are potential challenges using the Bayesian approach?
- 2.8 What software programs are available that can perform Bayesian analyses?
- 2.9 What resources are available to learn more about Bayesian statistics?

### 3. Bayesian Statistics

- 3.1 Outcomes and Parameters
- 3.3 What is a prior distribution?.
- 3.4 What is the likelihood of the observed data?
- 3.5 What is the posterior distribution?
- 3.6 What is a predictive distribution?
- 3.7 What is exchangeability?
- 3.8 What is the Likelihood Principle?

## 4. Planning a Bayesian Clinical Trial

- 4.1 Bayesian trials start with a sound clinical trial design
- 4.2 Selecting the relevant endpoints
- 4.3 Collecting other important information: covariates
- 4.4 Choosing a comparison: controls
- 4.5 Initial information about the endpoints: prior distributions
- 4.6 Borrowing strength from other studies: hierarchical models
- 4.7 Determining the sample size
- 4.8 Assessing the operating characteristics of a Bayesian design

## 5. Analyzing a Bayesian Clinical Trial

- 5.1 Summaries of the posterior distribution
- 5.2 Hypothesis testing
- 5.3 Interval estimation
- 5.4 Predictive probabilities
- 5.5 Interim analyses
- 5.6 Model Checking
- 5.7 Sensitivity Analysis
- 5.8 Decision analysis



- **6. Post-Market Surveillance**
- **7. Technical Details**
  - 7.1 Suggested Information to Include in Your Protocol.
  - 7.2 Simulations to Obtain Operating Characteristics.
  - 7.3 Model Selection
  - 7.4 Checking Exchangeability using the Posterior Predictive Distribution
  - 7.5 Calculations
- **8. References**

## 2.1 What is Bayesian statistics?

- **Bayesian statistics** is an approach for learning from evidence as it accumulates.
  - In clinical trials, traditional (frequentist) statistical methods may use information from previous studies only at the design stage. Then, at the data analysis stage, the information from these studies is considered as a complement to, but not part of, the formal analysis.
  - In contrast, **the Bayesian approach uses Bayes' Theorem to formally combine prior information with current information on a quantity of interest**.
    - The Bayesian idea is to consider the prior information and the trial results as part of a continual data stream, in which inferences are being updated each time new data become available.

## 2.2 Why use Bayesian statistics for medical devices?

### -With prior information

- **When good prior information on clinical use of a device exists**, the Bayesian approach may enable this information to be incorporated into the statistical analysis of a trial.
  - In some circumstances, the prior information for a device may be a justification for a smaller-sized or shorter-duration pivotal trial.

## 2.2 Why use Bayesian statistics for medical devices?

### -With prior information

- Good prior information is often available for medical devices because of their mechanism of action and evolutionary development.
  - The mechanism of action of medical devices is typically physical. As a result, device effects are typically local, not systemic.
    - Local effects can sometimes be predictable from prior information on the previous generations of a device when modifications to the device are minor.
- Good prior information can also be available from studies of the device overseas.
- In a randomized controlled trial, prior information on the control can be available from historical control data.

## 2.2 Why use Bayesian statistics for medical devices?

### -With prior information

- Our experience is that Bayesian methods are usually less controversial when the prior information is based on empirical evidence such as data from clinical trials.
- However, Bayesian methods can be controversial when **the prior information is based mainly on personal opinion** (often derived by elicitation from “experts”).

## 2.2 Why use Bayesian statistics for medical devices?

### -Without prior information

- **The Bayesian approach is also frequently useful in the absence of prior information.**
  - First, the approach can accommodate adaptive trials (e.g., interim analyses, change to sample size, or change to randomization scheme) and even some unplanned, but necessary trial modifications.
  - Second, the Bayesian approach can be useful for analysis of a complex model when a frequentist analysis is difficult to implement or does not exist.
  - Other potential uses include adjustment for missing data, sensitivity analysis, multiple comparisons, and optimal decision making (Bayesian decision theory).

## 2.2 Why use Bayesian statistics for medical devices?

### -Least burdensome

- **The Bayesian approach, when correctly employed, may be less burdensome than a frequentist approach.**
- Section 513(a)(3) of the Federal Food, Drug, and Cosmetic Act (FFDCA) mandates that FDA shall consider the least burdensome appropriate means of evaluating effectiveness of a device that would have a reasonable likelihood of resulting in approval (see 21 U.S.C. 360c(a)(3)).

# 議題1：ベイズ統計学の導入とFDAの 医療機器ガイダンスから学べること

## [論点1と議論の結果]



# 【論点1】 ベイズ統計学の利用可能性

- **FDAガイダンスの主張を一旦認めた上で**，以下の各論点について議論してください。
  - **論点1-1**
    - 医薬品・医療機器問わず，ベイズ統計学を用いる際に議論となりそうな点を挙げてください
    - 医薬品開発で，医療機器と比較してベイズ統計学の適用が難しいと考えられる理由を考えてください
  - **論点1-2**
    - 医薬品と医療機器の違いに注意しつつ，以下の状況の具体例をそれぞれ考えてください
      - 医薬品開発で，事前情報を用いたベイズ統計学が使えるような状況の例
      - 医薬品開発で，事前情報を用いたベイズ統計学が使えなさそうな状況の例
  - **論点1-3**
    - 事前情報がない場合の活用方法は，どういう場合に利用可能性が高そうでしょうか？

# 【Discussion Pointと議論の結果1】

- 医薬品・医療機器問わず、ベイズ統計学を用いる際に議論となりそうな点を挙げてください
  - 事前分布の妥当な設定と、周困および規制当局への説明
  - 規制当局との相談
  - MCMC法により生成されたマルコフ連鎖が事後分布に収束するか？
    - 特にパラメータが多いとき.
  - 結果の解釈に注意
    - 頻度論と同じように解釈できない.
      - 例. 信頼区間と信用区間
    - 頻度論とベイズ統計学で結果が異なる場合の結果の解釈
      - [Note] FDAガイダンス2.7節には以下の記載あり  
Bayesian and frequentist analyses approaches may differ in their conclusions  
two investigators, each with the same data and a different preplanned analysis (one Bayesian and one frequentist), could conceivably reach different conclusions that are both scientifically valid.
  - 医薬品の場合、背景の影響を受けやすい.
  - パラメータ $\theta$ がぶれることについて、議論を呼ぶ.
- **まとめ**  
**事前分布の設定、収束、結果の解釈が論点となりそう**

## 【Discussion Pointと議論の結果2】

- 医薬品開発で、事前情報を用いたベイズ統計学が使えるような状況
  - PMS
  - 治験でも対照薬にベイズを利用することができそう（対照薬は市販後でデータが多い）。
- 医薬品開発で、事前情報を用いたベイズ統計学が使えなさそうな状況の例
  - 大規模のRCTができる場合は、頻度論を用いる方がよいのでは？
  - 被験者数が十分に確保できるPIII試験において、過去試験の情報を借りて事前分布を構成し、データ解析するようなベイズ統計を正当化することは、ハードルが高い。
- [Note]事前分布に関して
  - FDAガイダンス3.3,4.4,4.5節に事前分布のことが記載されている

## 【Discussion Pointと議論の結果3】

- 事前情報がない場合の活用方法は、どういう場合に利用可能性が高そうでしょうか？
  - 頻度論による方法に加え、無情報事前分布を用いたベイズ流の方法で解析を行うことにより、結果を多方面から見ることができるかもしれない。

# 議題1：ベイズ統計学の導入とFDAの 医療機器ガイダンスから学べること [論点2の導入：ガイダンスの抜粋(2.6)]

## 2.6 What are potential benefits of using Bayesian methods?

### 2.6.1 More Information for Decision Making

- **The information from a current trial is augmented and the precision may be increased by the incorporation of prior information in a Bayesian analysis.**
- The Bayesian analysis brings to bear the extra, relevant, prior information, which can help FDA make a decision.

## 2.6 What are potential benefits of using Bayesian methods?

### 2.6.2 Sample size reduction via prior information

- In some instances, the use of prior information may alleviate the need for a larger sized trial.
- However, **a decrease in the sample size for the current trial may not be warranted by a Bayesian analysis incorporating prior information.**
  - See section 4.7 for further discussion on sample size issues in a Bayesian clinical trial.
  - Additionally, **if the prior information does not agree sufficiently with trial results, then the Bayesian analysis may actually be conservative relative to a frequentist or Bayesian analysis that does not incorporate the prior information.**

## 2.6 What are potential benefits of using Bayesian methods?

### 2.6.3 Sample size reduction via Adaptive Trial Design

- Adaptive designs use accumulating data to decide on how to modify certain aspects of a trial according to a pre-specified plan. They are often used to potentially reduce the size of a trial by stopping the trial early when conditions warrant. Adaptive trial designs can sometimes be easier to implement using Bayesian methods than frequentist methods. By adhering to the Likelihood Principle, a Bayesian approach can offer flexibility in the design and analysis of adaptive trials (see Sections 3.8 and 4.10).



## 2.6 What are potential benefits of using Bayesian methods?

### 2.6.4. Midcourse changes to the trial design

- 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 study 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.

## 2.6 What are potential benefits of using Bayesian methods?

### 2.6.5 Other Potential Benefits

- Exact analysis

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.

- Missing Data

Bayesian methods allow for great flexibility in dealing with missing data. See section 5.4 for a discussion of the use of these Bayesian methods.

- Multiplicity

Multiplicity is pervasive in clinical trials. For example, inferences on multiple endpoints or testing of multiple subgroups (e.g., race or sex) are examples of multiplicity. Bayesian approaches to multiplicity problems are different from frequentist ones, and may be advantageous. See section 4.9 for a discussion of Bayesian multiplicity adjustments.

# 議題1：ベイズ統計学の導入とFDAの 医療機器ガイダンスから学べること [論点2と議論の結果]

# 【論点2】 ベイズ統計学の利用可能性

- FDAガイダンス2.6節を踏まえ、以下の論点について議論してください。
  - **論点2-1**
    - ベイズ統計学をどういう状況で使用することに興味がありますか？ その場合に、どういう点が検討事項となりそうでしょうか。

# 【Discussion Pointと議論の結果1】

- ベイズ統計学をどういう状況で使用することに興味がありますか？その場合に、**どういう点が検討事項となりそうでしょうか。**
  - **興味があるのは、RCTができない状況（シングルアームのときなど）。ただし、事前分布の良し悪しは必ず議論するべきである。**
  - **日本でRCTできなくても、海外のデータを使い解析する。海外のデータはなるべく近いデータを事前情報に用いて解析する。**
  - **研究で出た結果を開発で利用する。**
  - **動物のデータから事前分布設定はよくないこともありうる。そのためP1から事前分布を用いたベイズを用いる際には注意が必要である**
  - **Adaptive designで使いたい。**
  - **検証的試験以外で使いたい。**
  - **社内でのdecisionの参考にベイズを使う。**
  - **PIIでP値を用いてどのように判定するか？ベイズなら頻度論と比べて判断が違つかもしいない。**
  - **ベイズは確率分布が得られるため、臨床的に意味のある最小の差よりも大きな差となる確率等を出してみたい。**

# 議題2： p値についての議論

# 本議題の内容

- 本議題の内容
  - 論文” **The ASA’s Statement on  $p$ -Values: Context, Process, and Purpose**”とその Supplemental material の一部を用いて、 $p$ 値の適切な使用方法を整理するとともに、「医薬品開発の意思決定に必要な条件は何か？」という観点からベイズ統計学の適用可能性を議論する

# 議題2： p値についての議論

## [議論の導入]



# 本議題の予習資料

- 必読

- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. The American Statistician.
  - <http://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108>

- 参考（必読資料の supplementary material. 以下のうちから1つ）

- Altman, Naomi: Ideas from multiple testing of high dimensional data provide insights about reproducibility and false discovery rates of hypothesis supported by p-values
- Benjamin, Daniel J, and Berger, James O: A simple alternative to p-values
- Berry, Donald A: P-values are not what they're cracked up to be
- Carlin, John B: Comment: Is reform possible without a paradigm shift?
- Gelman, Andrew: The problems with p-values are not just with p-values
- Goodman, Steven N: The next questions: Who, what, when, where, and why?
- Ioannidis, John P.A.: Fit-for-purpose inferential methods: abandoning/changing P-values versus abandoning/changing research
- Rothman, Kenneth J: Disengaging from statistical significance

# ASA's statement

1.  $P$ -values can indicate **how incompatible the data are with a specified statistical model.**
2.  $P$ -values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a  $p$ -value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. **A  $p$ -value, or statistical significance, does not measure the size of an effect or the importance of a result.**
6. **By itself, a  $p$ -value does not provide a good measure of evidence regarding a model or hypothesis.**

# 議題2： p値についての議論

## [論点と議論の結果]

# 論点1

- ASA論文の内容を踏まえ、下記論点について議論してください。
  - 論点1-0
    - Supplementary material を読んで来られた方は、大まかに内容をご説明ください
  - 論点1-1
    - 日常業務において、「p値の果たしている役割」を整理してください
    - 統計以外の部署の方による「p値に対する誤解」についても整理してください
  - 論点1-2
    - ベイズ統計学を用いることにより、頻度論より「よい」判断ができそうな状況はありますか？できるだけ具体的に検討してください

# 【Discussion Pointと議論の結果1】

- 統計以外の部署の方による「p値に対する誤解」についても整理してください
  - $p < 0.05$  が「差がある根拠」とされてしまう。臨床的に意味があるかを考える必要がある。
  - $p < 0.05 \rightarrow *$ ,  $p < 0.01 \rightarrow **$ で効果の大きさを示そうとする。

## 【Discussion Pointと議論の結果2】

- ベイズ統計学を用いることにより、頻度論より「よい」判断ができそうな状況はありますか？できるだけ具体的に検討してください
  - 信用区間は $\theta$ の変動範囲を示す物なので、 $\theta$ が○パーセントの確率で $x \sim x$ の範囲に入るという結論を下すことができるため、信頼区間よりと比べて人間の感覚に近いのではないか。
  - 社内の意思決定における資料作成に使えるのでは。
  - デザイン段階ではなく解析段階でも事前情報を使用し、推定精度を増すことができるのでは。
  - ベイズ統計の浸透度は低い。経験の蓄積が必要か。
  - P値は現在誤用されることが多い状況。これだけ誤って使われるのであれば、P値を使わずに他の指標を検討することも意味があるかもしれない。1つの指標としてベイズを使う方法を検討するのはどうか？
  - 頻度論・ベイズはどちらもメリット、デメリットがあることを認識するべきである。