希少疾患に対する治療効果の定量化についての一考察
~AMED研究で策定した指針の紹介~

平川晃弘

東京医科歯科大学大学院 臨床統計学分野

本発表について開示すべきCOI関係にある企業などはありません。

はじめに

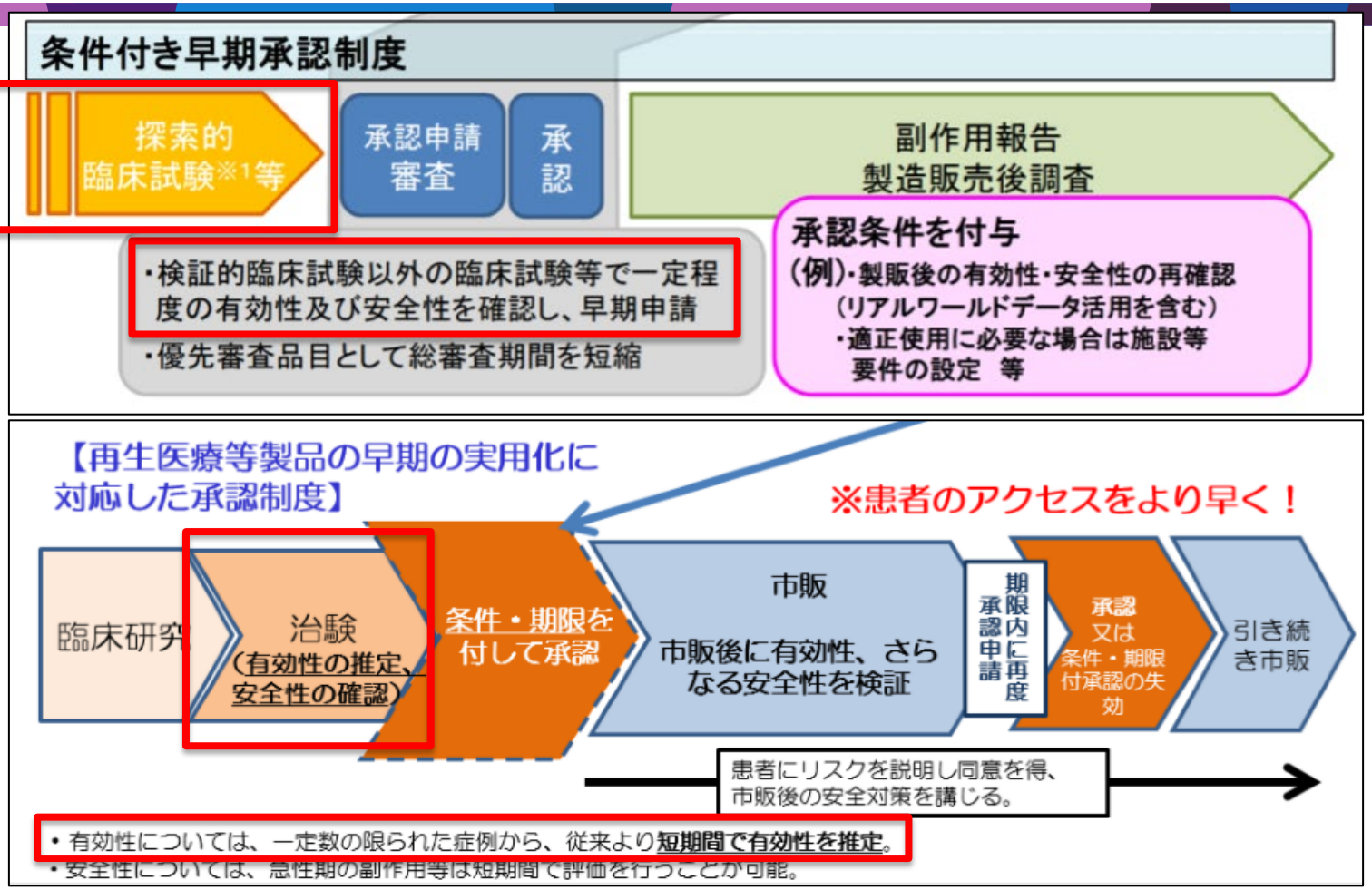
- 希少疾患領域の開発戦略
  1. 国際共同第3相試験をピボタル試験として承認申請
  2. 国内第2相試験を実施し，海外第3相試験と併せて承認申請
  3. **日本が先行して国内第2相試験を実施し，欧米に先駆けて承認申請**

- 国内第2相試験は，**少数例の小規模(単群)試験**
  - 少数例データから臨床的有用性を評価する必要性

アダプティブデザイン・ベイズ流アプローチを利用することで，より科学的（定量的かつ客観的）な評価が可能

しかし，本邦ではその適正利用（統計的方法や結果解釈の仕方）に関する**検討が不十分**

これからの希少疾患領域の医薬品開発



希少疾患領域におけるアダプティブデザイン・
ベイズ流アプローチの利用に関するガイドラインが必要

https://www.pmda.go.jp/files/000205439.pdf, https://www.pmda.go.jp/files/000223145.pdf

AMED医薬品等規制調和・評価研究事業

希少疾患領域の医薬品開発を効率化するための小規模臨床試験のデザイン・統計解析法の研究開発とその適正利用のための基本的考え方の策定（研究代表者 平川晃弘）(2019-2021)

- AMED研究の目的
  1. 小規模臨床試験のためのアダプティブデザイン・ベイズ流アプローチの新規方法論開発
  2. 適正利用のための基本的考え方の策定

- 期待される成果
  - 希少疾患領域の医薬品評価に関するガイドライン策定に貢献
  - 治験相談・承認審査の内容深化
    - 早期承認制度等の客観的運用

シンポジウムの開催

- 産官学での意見交換を目的としたシンポジウムの開催
  - 2020年12月8日に柴田班との合同開催（参加者数：350名）

AMED医薬品等規制調和・評価研究事業合同シンポジウム「これからの臨床試験：新機軸の展開」の開催について

国立がん研究センター 柴田大朗
東京医科歯科大学 平川晃弘

【合同シンポジウムの開催趣旨】

医薬品等の有効性及び安全性を評価する新しい枠組みに関する検討が始まっています。患者レジストリデータやベイズ流アプローチの利活用はその一例であり、本邦においても日本医療研究開発機構（AMED）医薬品等規制調和・評価研究事業の一環で発足した研究班により、その利用可能性について検討が進められています。本合同シンポジウムは当該事業における2つの研究班の共催で実施されるものであり、午前の部では「希少疾患領域の治験におけるベイズ流アプローチの利用可能性」、午後の部では「患者レジストリデータを活用した臨床開発：生物統計学的側面からの検討」をテーマとして、研究班の活動を紹介すると共に、それぞれのテーマに関して産官学で意見交換を行います。

午前の部　希少疾患領域の治験におけるベイズ流アプローチの利用可能性

「希少疾患領域の医薬品開発を効率化するための小規模臨床試験のデザイン・統計解析法の研究開発とその適正利用のための基本的考え方の策定」班（研究代表者 東京医科歯科大学 平川晃弘）

本研究班では、希少疾患領域の臨床試験におけるベイズ流アプローチの適用に関する基本指針を策定しています。午前の部では、当該指針とベイズ流アプローチを用いた臨床試験の事例を紹介すると共に、パネルディスカッションをとおしてベイズ流アプローチの利用可能性について議論します。策定中の指針（案）は以下のURLからダウンロードできます。指針（案）に対するご意見・ご質問、またパネルディスカッションで取り上げて欲しいテーマ等がある場合は、こちらのフォームからご連絡ください（11月20日〆切）。

希少疾患領域の臨床試験におけるベイズ流アプローチの適用に関する基本指針（案）（PDF：443KB）

←指針（案）も同時公開

シンポジウムを終えて

- 普及に向けての課題
  - 事例の蓄積と共有（成功事例の存在）
  - 有用性の見える化
- 規制上の論点整理
  - ベイズ流アプローチを用いる理由
  - 事前情報の選択
  - 有効性の判断基準
  - 動作特性の検証

指針策定と論文化

AMED 医薬品等規制調和・評価研究事業

研究開発課題名：希少疾患領域の医薬品開発を効率化するための小規模臨床試験のデザイン・統計解析法の研究開発とその適正利用のための基本的考え方の策定

希少疾患領域の臨床試験における ベイズ流アプローチの適用に関する基本指針

（令和 3 年 7 月 16 日）

（注）本文書は、令和 1～3 年度の研究成果として研究班においてとりまとめたものである。

1

Received: 13 January 2021 | Revised: 20 October 2021 | Accepted: 28 December 2021

DOI: 10.1002/pst.2192

VIEWPOINT

WILEY

Regulatory issues and the potential use of Bayesian approaches for early drug approval systems in Japan

Akihiro Hirakawa[1] | Hiroyuki Sato[1] | Masataka Igeta[2] | Kei Fujikawa[3] | Takashi Daimon[2] | Satoshi Teramukai[3]

https://doi.org/10.1002/pst.2192

Jpn Pharmacol Ther（薬理と治療）vol. 49 suppl. 1 2021

［解析・技術動向］

希少疾患領域の治験における ベイズ流アプローチの利用可能性と留意事項

■

Bayesian Approaches in Rare Disease Drug Development

■

平川 晃弘*1 佐藤 宏征*1 井桁 正尭*2 藤川 桂*3
堀口 剛*3 大門 貴志*2 手良向 聡*3

https://www.pieronline.jp/content/article/0386-3603/49091/72

規制上の4つの論点

規制上の論点1：ベイズ流アプローチを用いる理由

- ベイズ流アプローチを用いる理由
  - ICH E9

> *臨床試験の計画と解析においては、頻度論的立場からの統計手法に基づく方法が主流になっていることから、本ガイドラインは仮説検定や信頼区間を議論する場合、主として頻度論的手法（用語集参照）を念頭に置いている。これは、他の方法が適切でないと主張するものではない。ベイズ流の手法や他の手法の使用も、それらの使用の理由が明らかであり、異なる仮定の下でも結果として得られる結論が十分に安定している場合には検討することができる。*

  - 特段の事情がない限りは，通常は頻度流アプローチを採用
  - ベイズ流アプローチは，その利用が有効性及び安全性評価において有用と考えられる際に，規制当局との合意を得た上で利用することが望ましい

規制上の論点2：事前情報の選択

- 事前情報として利用可能な情報：文献，過去の臨床試験成績，臨床経験，アンケート調査等
  - 実施予定の試験と対象集団，対照治療，試験治療又は対照治療の用法・用量，試験スケジュール（観察，検査，評価の方法・時期， 追跡期間等），評価項目等が同じであるかに留意
  - 実際には，これらの項目が全く同じである過去の臨床試験はほとんど存在しないため，試験デザインの類似性を精査し，臨床的，規制的に利用可能な事前情報であるのか規制当局と議論が必要

- 疾患レジストリデータの活用
  - 個々の患者データを利用できる場合も多いことから，実施予定の試験の対象患者と類似した集団を抽出し，事前情報を設定できるかもしれない
  - ただし，疾患レジストリのデータ品質管理は，臨床試験のそれとは異なることから，事前情報として利用するデータの品質が規制上受け入れ可能であるかが論点

- なお，既存情報の利用が困難な場合でも，ベイズ流アプローチの柔軟性の享受が試験目的の達成の上で有用である場合は，無情報事前分布の利用を検討するのがよいかもしれない

規制上の論点3：
治療効果に関する閾値と達成すべき事後確率の設定

- 単群試験において治療効果の存在を判定するためには，試験開始前に治療効果に関する閾値（又は，その不確実性を考慮した分布）の設定が必要
  - 閾値には，臨床的意義があると判断できる基準値を設定

- ベイズ流アプローチでは，治療効果が事前に設定したこの閾値を超える（事後）確率を求める
  - この事後確率が十分に高ければ，治療効果の存在の強い証拠
  - 閾値は，頻度流アプローチ（仮説検定）における帰無仮説の値に相当するため，ベイズ流アプローチの適用のために追加で必要となる情報ではない

- 達成すべき事後確率については，有意水準両側5%のようにコンセンサスのある数値はないが，十分に高いことが求められる

規制上の論点4：症例数設計と動作特性

- ベイズ流アプローチを採用する際の必要症例数
  1. 試験治療や対照治療の効果に関する事前分布（例えば，事前情報から仮定できる試験治療の効果の平均値やばらつき等）
  2. 計画中の臨床試験において期待される治療効果の想定値
  3. 試験治療が達成すべき治療効果に関する閾値
  4. その他計画中の試験の不確実性に係るパラメータ

  ⇒ どの程度の症例数を登録すれば，達成すべき事後確率が得られるか（ベイズ流検出力）を調べて，第一種の過誤確率も踏まえて決定

- コンピュータシミュレーション実験により動作特性を評価

- 規制当局を含め試験関係者がベイズ流アプローチを正しく理解・評価するためにも，試験開始前に動作特性を評価することは重要である

ベイズ流アプローチの利用可能性が高いケース

1. プラセボ効果を無視可能な主要評価項目を用いる単群試験

2. 奏効割合を主要評価項目としたバスケット型試験

3. プラセボ効果が認められる主要評価項目を設定した単群試験
   （プラセボ効果に対する事前情報）

4. 頻度流アプローチでは検出力が低くなるランダム化対照試験

5. 再生医療等製品の臨床開発

6. 小児領域の臨床開発

事例紹介

1. プラセボ効果を無視可能な主要評価項目を用いる単群試験

希少がん領域の医師主導治験

- 切除不能かつ化学療法歴を有するHER2：1+以上の進行・再発子宮癌肉腫に対するtrastuzumab deruxtecanの医師主導治験
  - STATICE 試験（UMIN000029506）

- 子宮癌肉腫
  - 10万人あたり0.5～1人程度（約450例/年）
  - III/IV期の5年生存率：9~26%
  - 化学療法
    - パクリタキセル＋カルボプラチン
    - ドセタキセル＋カルボプラチン
    - イホマイド＋シスプラチン

症例数設計

- 主要評価項目:奏効率
- 閾値奏効率 = 5%
- 期待奏効率 = 30%
- 有意水準 = 5%
- 検出力 = 90%

必要症例数 = 20

- 登録期間である2年で20例を集積することが困難かもしれない
- 登録開始後2年の時点で，一定の症例が集積されていれば，有効性評価を行いたい

  ⇒ ベイズ流デザインを活用

事後確率に基づく中間モニタリングデザイン

- 標準治療の奏効率：p_S
  - $p_S \sim Beta\,(\alpha_S, \beta_S)$
- 試験治療の奏効率：p_E
  - $p_E \sim Beta\,(\alpha_E, \beta_E)$
- 標準治療に対する上乗効果：δ

- $Y \sim Bin(n, p_E)$より，　$Y = y$ の下で以下を得る

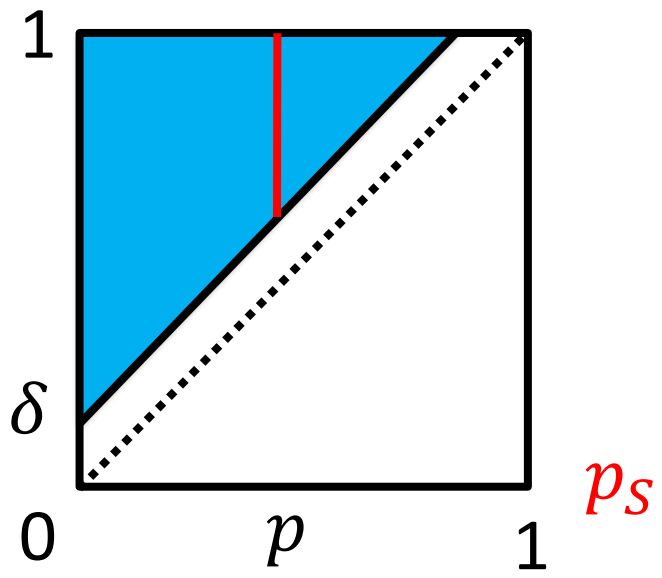
$$p_E \mid y \sim Beta\,(\alpha_E + \mathrm{y}, \beta_E + n - y)$$

Thall and Simon (1994 , Biometrics)

事後確率の計算

- n (=1, ⋯, N) 例時点で奏効がy例の場合のp_Eが$p_S + \delta$を上回る事後確率

$$\Pr(p_E > p_S + \delta | y) = \int_0^{1-\delta} \{1 - F(p + \delta; \alpha_E + y, \beta_E + n - y)\} f(p; \alpha_S, \beta_S) dp$$



$F(\cdot)$：分布関数

$f(\cdot)$：密度関数

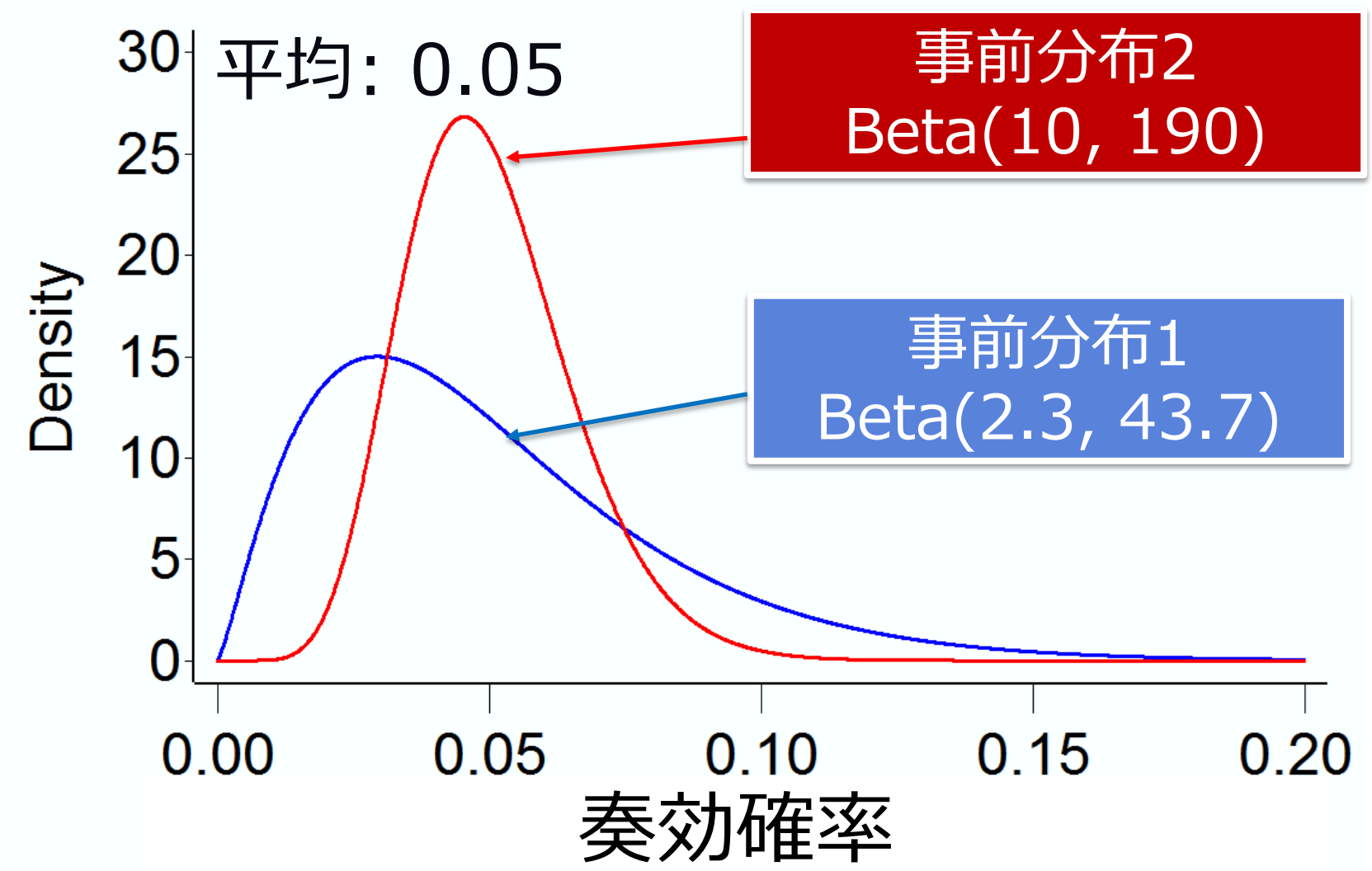
境界奏効例数の計算

$U_n = \Pr(p_E > p_S | y) \geq \theta_U$
を満たす最小のy

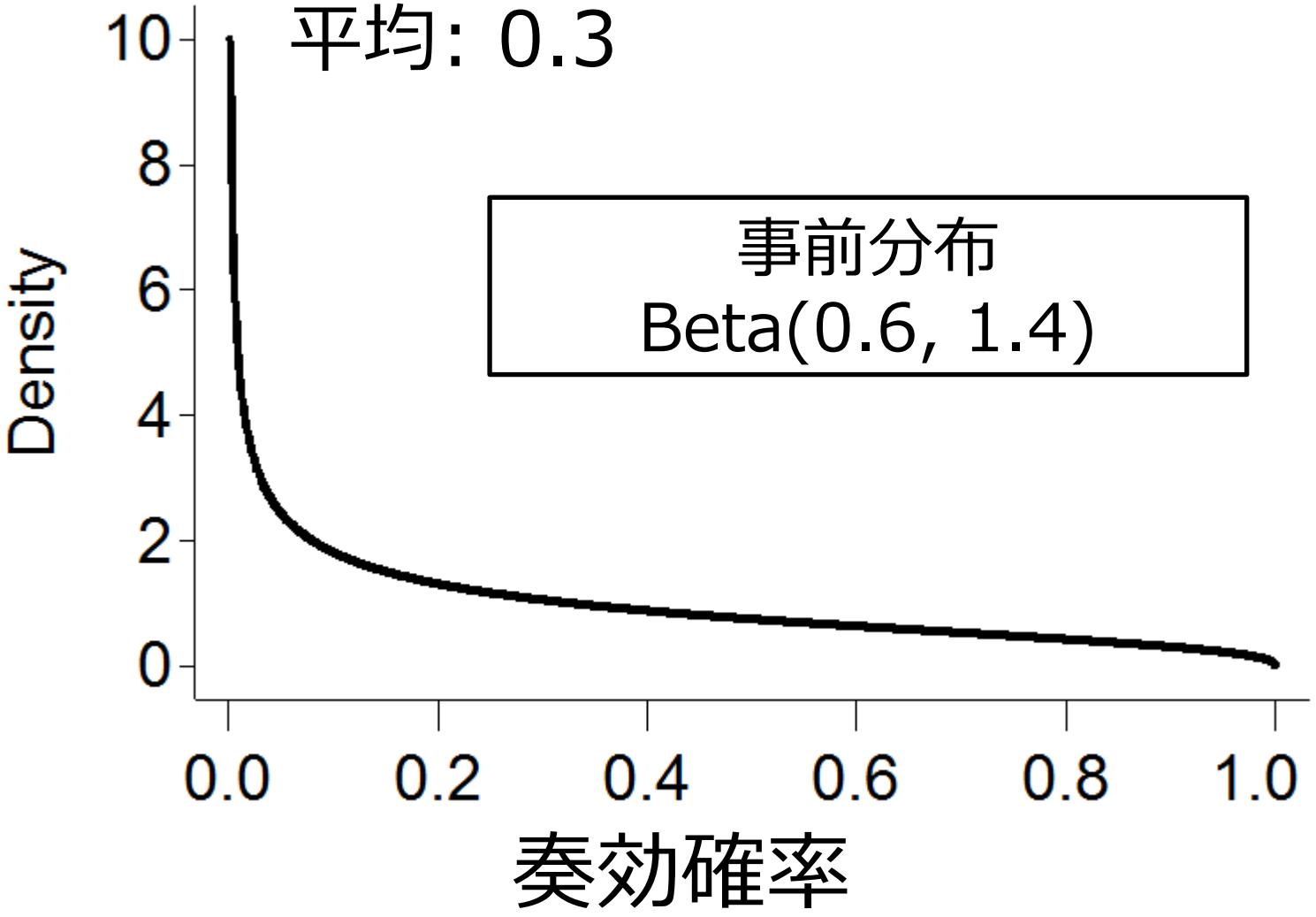
$L_n = \Pr(p_E > p_S + \delta | y) \leq \theta_L$
を満たす最大のy

- 中止・継続ルール（n人中y例が有効）
  - $y \geq U_n \Rightarrow$ 中止 (有望)
  - $y \leq L_n \Rightarrow$ 中止 (非有望)
  - $L_n < y < U_n$ & $n < N \Rightarrow$ 継続

閾値奏効率の事前分布


平均: 0.05
事前分布2
Beta(10, 190)
事前分布1
Beta(2.3, 43.7)
Density
30
25
20
15
10
5
0
0.00
0.05
0.10
0.15
0.20
奏効確率

期待奏効率の事前分布



必要最小奏効例数

	事前分布1 Beta(2.3, 43.7) α : 0.0559			事前分布2 Beta(10, 190) α : 0.0757	
	L	U		L	U
n_{min} = 15	.	3		.	3
16	.	3		.	3
17	.	3		.	3
18	.	4		.	3
19	.	4		.	3
20	.	4		.	3
21	.	4		.	3
22	.	4		.	4
23	.	4		.	4
24	.	4		.	4
n_{max}=25	.	4		.	4

登録期間に関する規定

- 予定登録期間（2年）中は，20例を超えても25例まで症例登録を継続する
- 2年の時点で15~25例であれば，有効性の主たる解析を自実施する
- 15例未満であれば，登録継続又は中止のいずれかを検討する

Clopper-Pearson 95%信頼区間

n	奏効例	奏効率	95%信頼区間
15	3	0.200	(0.043, 0.481)
16	3	0.188	(0.040, 0.456)
17	3	0.176	(0.038, 0.434)
18	3	0.167	(0.036, 0.414)
19	3	0.158	(0.034, 0.396)
20	3	0.150	(0.032, 0.379)
21	3	0.143	(0.030, 0.363)
22	4	0.182	(0.052, 0.403)
23	4	0.174	(0.050, 0.388)
24	4	0.167	(0.047, 0.374)
25	4	0.160	(0.045, 0.361)

レフリーコメント

Why is the Bayesian design included among the limitations of the study?

While this type of design is not used as often as, for example, Simon's two-stage design, it is a perfectly reasonable approach.

参考事例

4. 頻度流アプローチでは検出力が低くなるランダム化対照試験

- 未熟児網膜症を対象としたアフリベルセプトのランダム化対照試験
  - 奏効割合の群間差の事後分布に基づき非劣性検証

Table 2. Primary Outcome of Treatment Success Rate Based on Bayesian Analyses Overall and in Prespecified Subgroups

	0.4 mg of intravitreal aflibercept			Laser photocoagulation			
	Treated bilaterally[a]	Treated unilaterally[b]	Estimated success rate, % (90% CrI)[c]	Treated bilaterally[a]	Treated unilaterally[b]	Estimated success rate, % (90% CrI)[c]	Between-group difference, % (1-sided 95% CrI)[d]
Primary outcome: proportion of infants without active ROP and unfavorable structural outcomes at 24 wk[e]							
Overall population	59/71 bilateral responders[f]	4/4 unilateral responders	85.5 (78.0 to 91.3)	28/34 bilateral responders[g]	4/4 unilateral responders	82.1 (70.5 to 90.8)	3.4 (−8.0 to ∞)
Zone I retinopathy of prematurity[h]	11/15 bilateral responders[g]	NA	70.8 (50.7 to 86.6)	4/6 bilateral responders	1/1 unilateral responder	64.4 (36.7 to 87.1)	6.2 (−23.5 to ∞)
Zone II retinopathy of prematurity[h]	38/42 bilateral responders[i]	4/4 unilateral responders	92.0 (84.1 to 96.9)	20/23 bilateral responders[g]	3/3 unilateral responders	84.9 (71.5 to 94.0)	6.9 (−4.8 to ∞)
Aggressive posterior retinopathy of prematurity	10/14 bilateral responders[j]	NA	73.3 (52.1 to 89.1)	4/5 bilateral responders	NA	72.2 (40.0 to 93.0)	1.1 (−28.6 to ∞)

Abbreviations: CrI, credible interval; NA, not applicable.

[a] Both eyes needed to respond to treatment for the primary outcome to be met.

[b] The response in the second eye was considered missing information in the analysis model.

[c] Reflects the median of posterior distribution based on the primary (bayesian) analysis model.

[d] The per-protocol–defined noninferiority margin was -5%.

[e] Assessed by investigators. Unfavorable structural outcomes included retinal detachment, macular dragging, macular fold, and retrolental opacity. The eyes were considered to be nonresponsive if rescue treatment was required and administered.

[f] In addition, 4 patients responded unilaterally only and 2 patients did not have data at week 16 or later and were handled by multiple imputations.

[g] In addition, 1 patient responded unilaterally only.

[h] Excludes infants with aggressive posterior retinopathy of prematurity.

[i] In addition, 3 patients responded unilaterally only and 1 patient did not have data at week 16 or later and was handled by multiple imputations.

[j] In addition, 1 patient did not have data at week 16 or later and was handled by multiple imputations.

Stahl et al. (JAMA, 2022)

参考事例

5. 再生医療等製品の臨床開発

- 探索的試験の成績を事前情報として検証的試験を実施することができないか検討中
  - 探索的試験 (RCT) で有効性を強く示唆する成績が得られた
  - 検証的試験 (RCT) では，探索的試験の情報を有効利用して，症例数と試験期間の観点で効率性を高めたい

- 探索的試験の成績は，Pocock の6条件を満たすヒストリカルデータ
  1. 対照群が同じ
  2. 選択・除外基準が同じ
  3. 主要評価項目が同じ
  4. ほぼ同じ施設で実施
  5. 患者背景が同じ（※検証試験終了後に考察）
  6. 各試験の成績の異質性を疑う他の要因がない（※検証試験終了後に考察）

対面助言で，何をどのように評価するかをPMDAと相談することが望ましい!

ご清聴ありがとうございました。