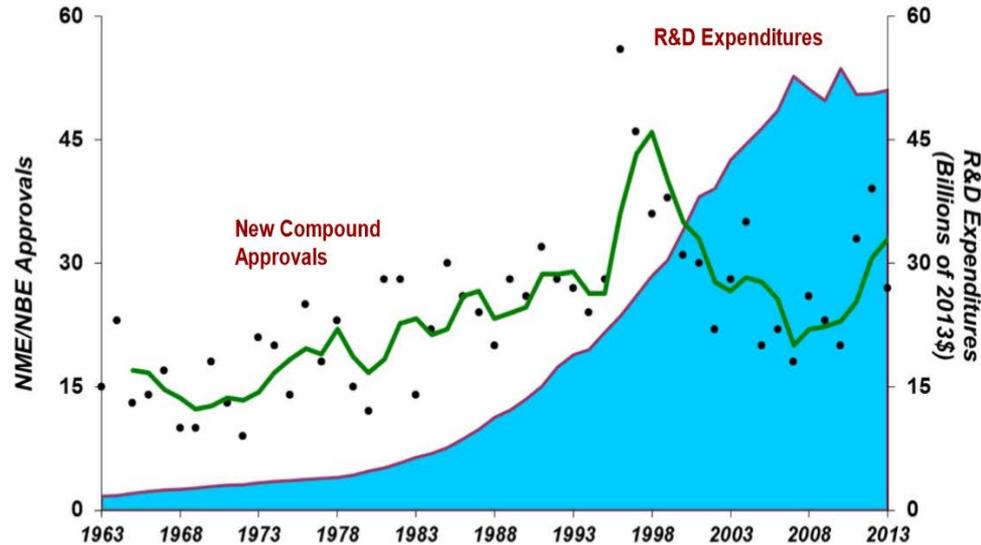


マスタープロトコル試験 バスケット型デザインとベイズ流アプローチ

平川晃弘

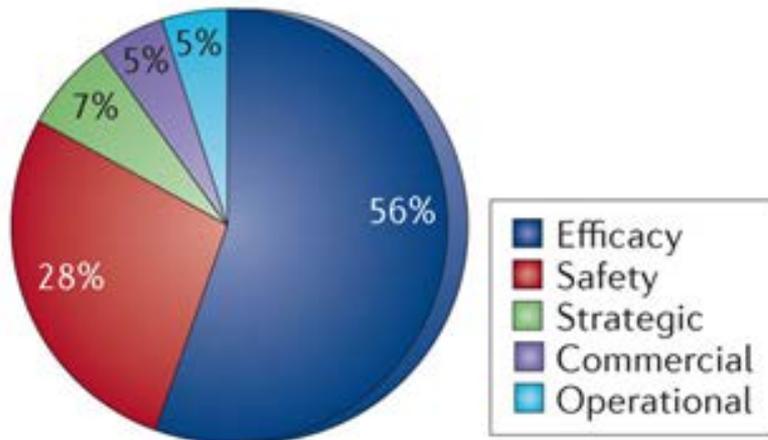
東京大学大学院 医学系研究科
生物統計情報学

医薬品開発の現状

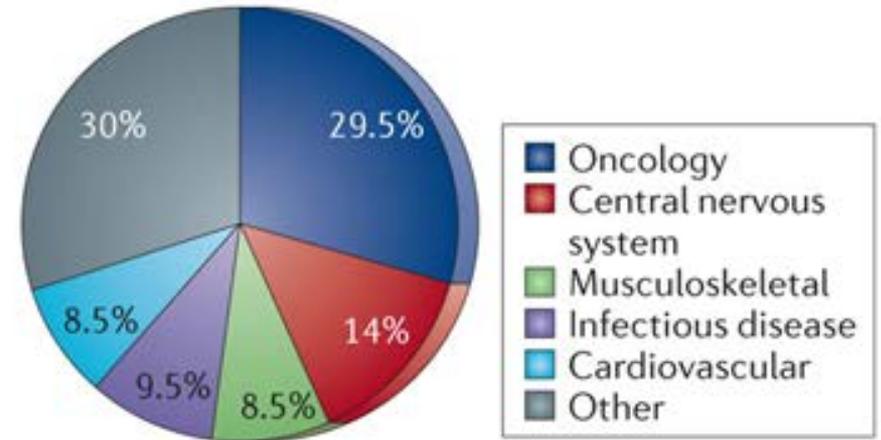


http://csdd.tufts.edu/files/uploads/Tufts_CSDD_briefing_on_RD_cost_study_-_Nov_18,_2014..pdf

a Causes of failure

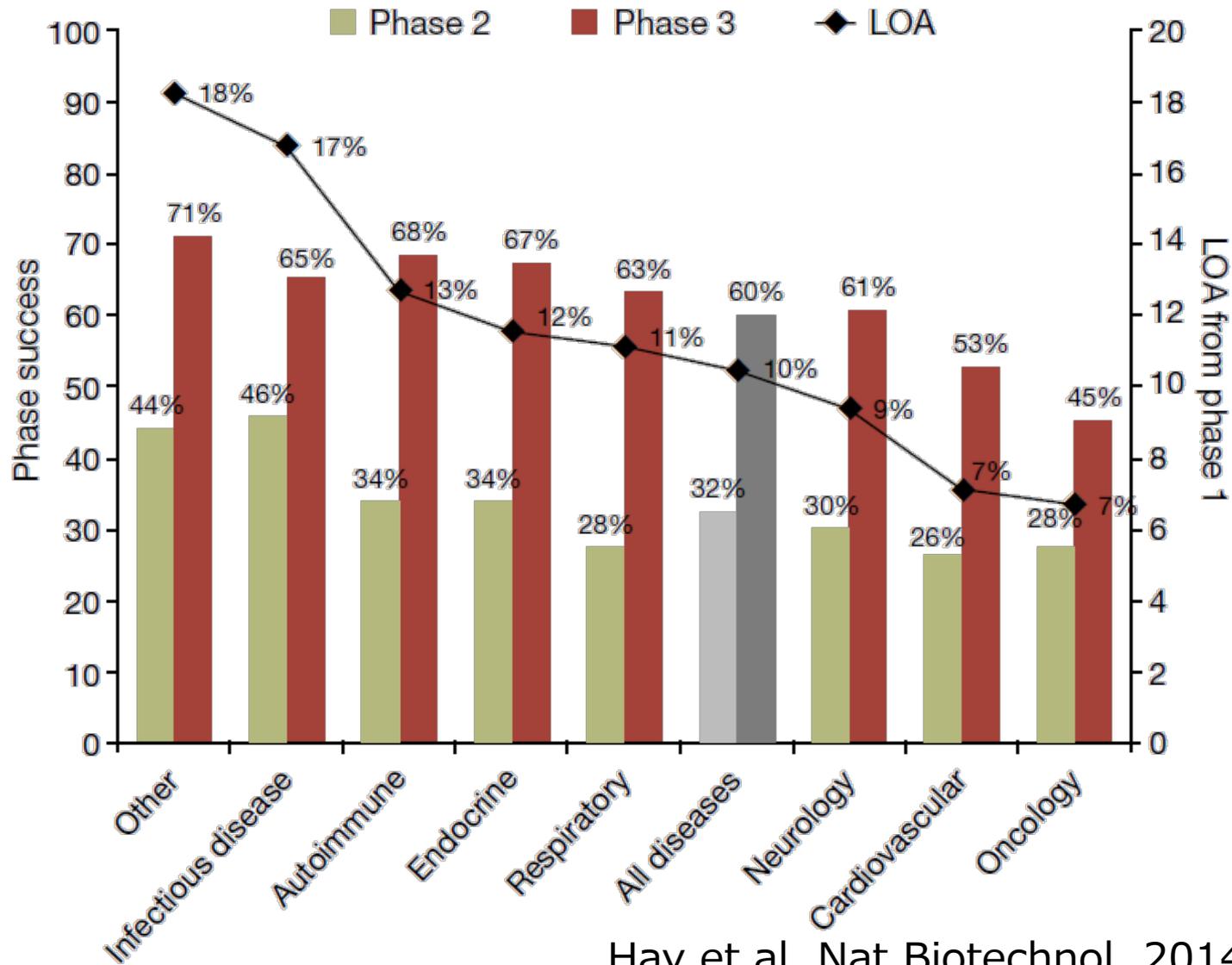


Failure by therapeutic area



Arrowsmith and Miller. Nat Rev Drug Discov 2013;12:569

臨床試験の成功確率



Hay et al. Nat Biotechnol. 2014 32:40-51



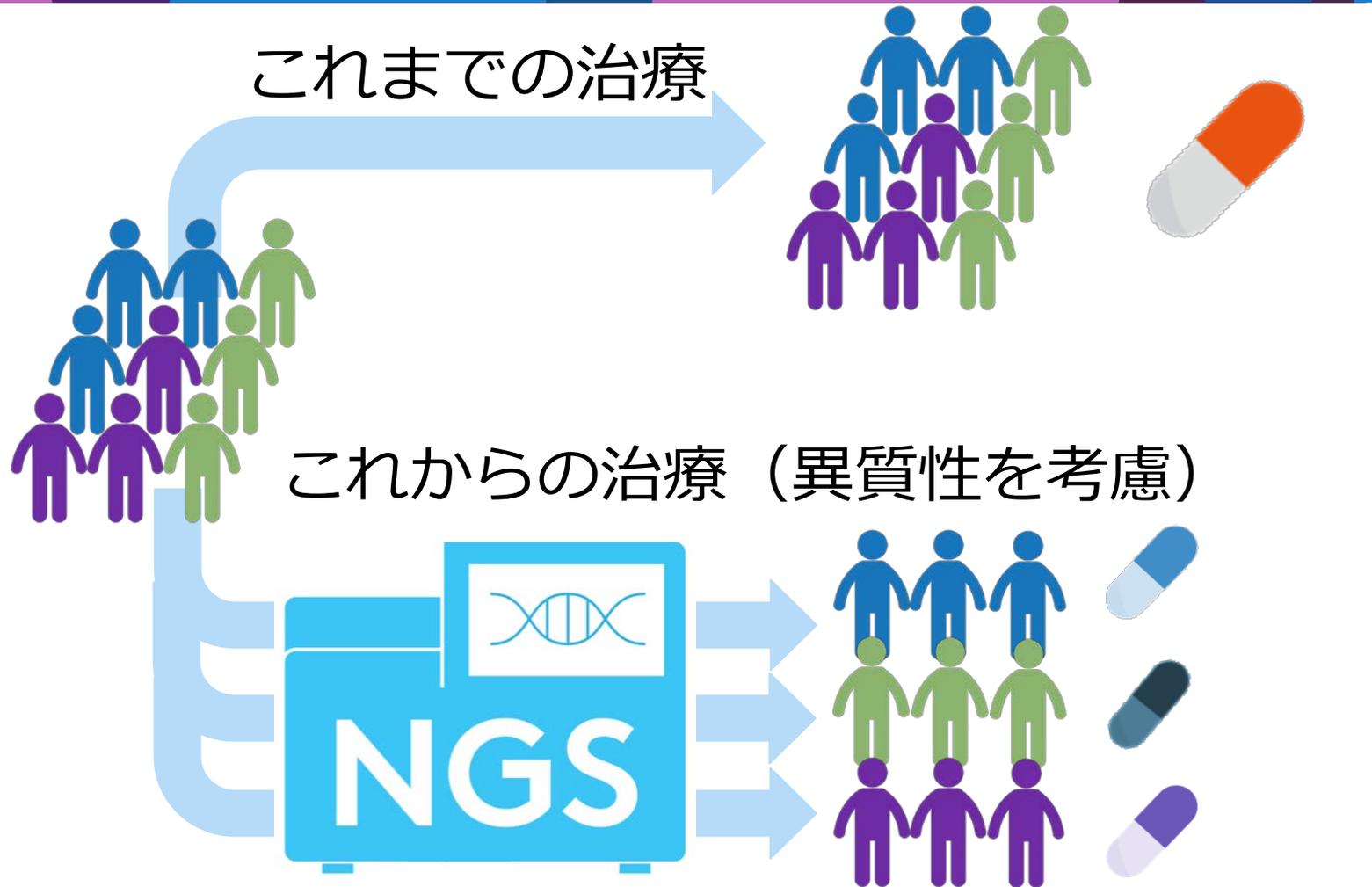
The 21st Century Cures Act

Subtitle C Modern Trial Design and Evidence Development	
Sec. 3021. Novel Clinical Trial Designs.	<ul style="list-style-type: none">Requires FDA to hold a public meeting and issue guidance documents that would assist sponsors in incorporating <u>adaptive designs and novel statistical modeling</u> into new drug applications.
Sec. 3022. Real World Evidence.	<ul style="list-style-type: none">Requires FDA to evaluate the use of <u>real world evidence</u> to help support the approval of a new indication for a previously approved drug and to help support or satisfy post-approval study requirements.
Sec. 3023. Protection of Human Research Subjects.	<ul style="list-style-type: none">Requires the Secretary of HHS to harmonize differences between the human subject regulations under the Common Rule and the Federal Food Drug and Cosmetic Act.Streamlines the institutional review board process for trials that are being conducted at multiple sites.
Sec. 3024. Informed Consent Waiver or Alteration for Clinical Investigations.	<ul style="list-style-type: none">Provides FDA the flexibility to waive or alter informed consent requirements for clinical trials with minimal risk, similar to existing flexibility for HHS and NIH under the Common Rule.

- ベイズ流アプローチにより、臨床試験を合理化・効率化、試験中の柔軟な意思決定、臨床試験の成功確率の増大が期待される

<https://energycommerce.house.gov/cures>

プレシジョン・メディシン確立に向けて



次世代シーケンサー（Next Generation Sequencer）により数千万から数億のDNAの塩基配列を同時に測定

臨床試験方法論にも影響

分子マーカーA



分子マーカーB



分子マーカーC



複数の仮説を短期間で効率的に
評価する臨床試験方法論が必要

課題

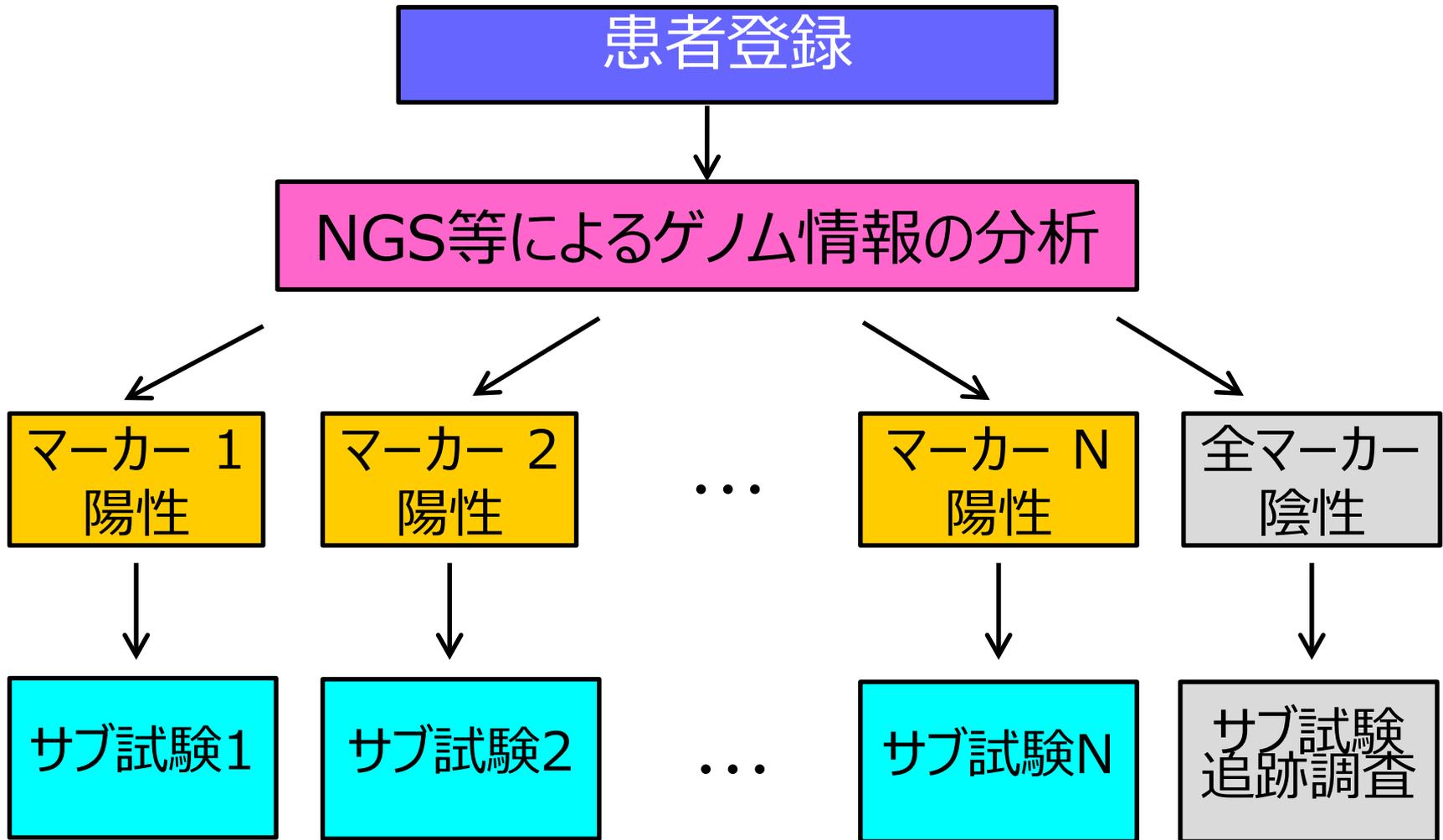
- がん種，ステージ，組織型等を規定した集団を対象に臨床試験を実施
- バイオマーカーごとの小集団で臨床試験を実施すると試験数の増加及び作業重複が多く効率的ではない
- 試験のプラットフォームが異なれば，試験間のデータを統合解析する際にその異質性への対処が必要
 - 分子マーカー測定法，患者選定・追跡の方法，データマネジメントシステムなど

マスタープロトコル

- 複数の仮説を評価することを目的に作成された包括的プロトコル
- 特徴
 - 共通のプラットフォーム（NGS, EDC, 安全性情報管理, SOP, SAP）
 - 分子マーカーで定義される特定集団を対象
 - 複数の癌種に対して, 1つの治療法
 - 単一疾患に対して, 複数の治療法

Renfro & Mandrekar, 2017

マスタープロトコル試験



マスタープロトコルを用いる試験



バスケット試験

- 複数の疾患やサブタイプに対して単一標的治療



アンブレラ試験

- 単一・共通化できる疾患に対して複数の標的治療



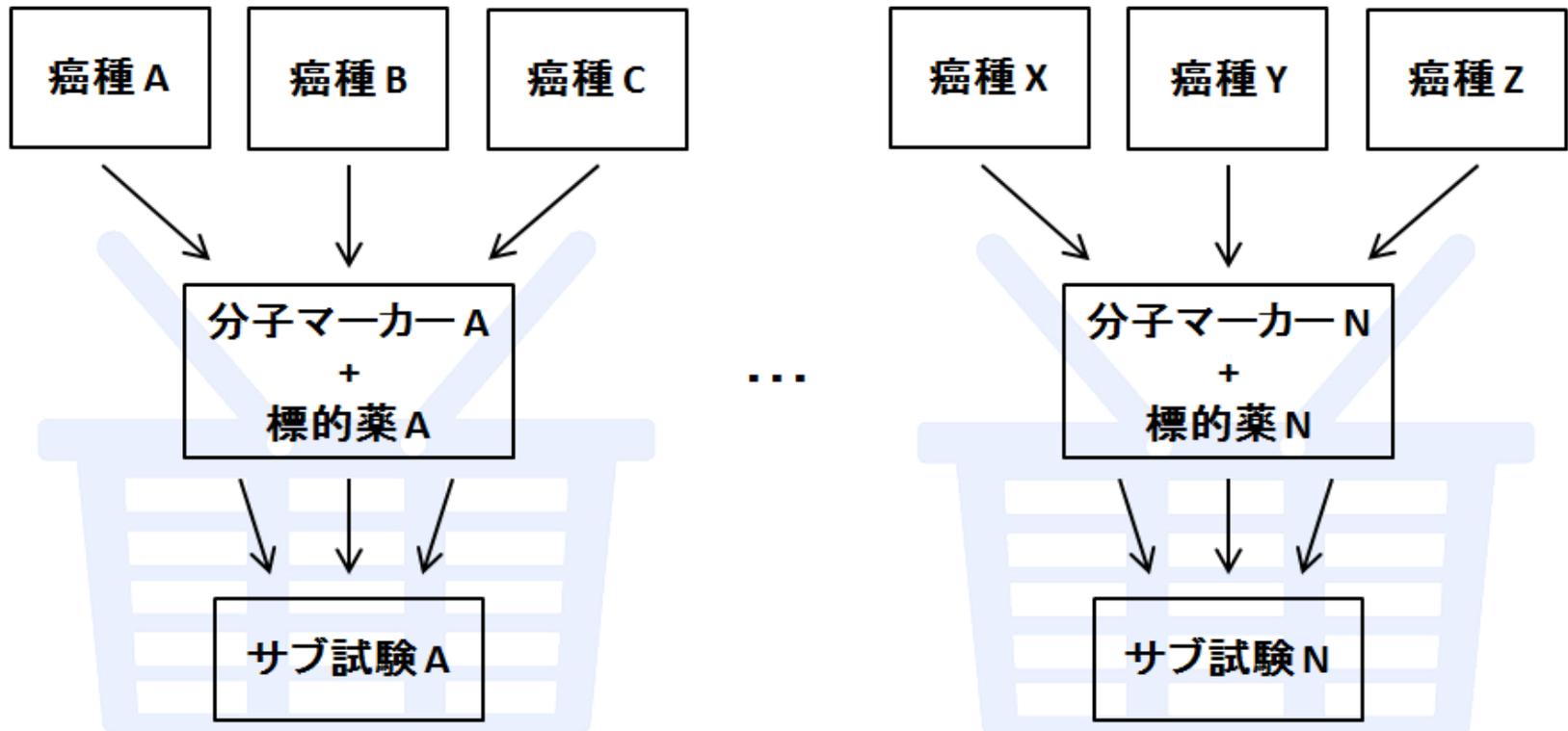
プラットフォーム試験

- 永続的に単一・複数疾患に対して複数の標的治療
- 試験中に新治療法や対象患者の追加・除外



バスケット試験

- 特定の分子マーカー（バイオマーカー）や遺伝子変異を有する複数のがん種に対して、それに対応する分子標的薬の治療効果をごん種別またはがん種横断的に評価



バスケット試験の特徴

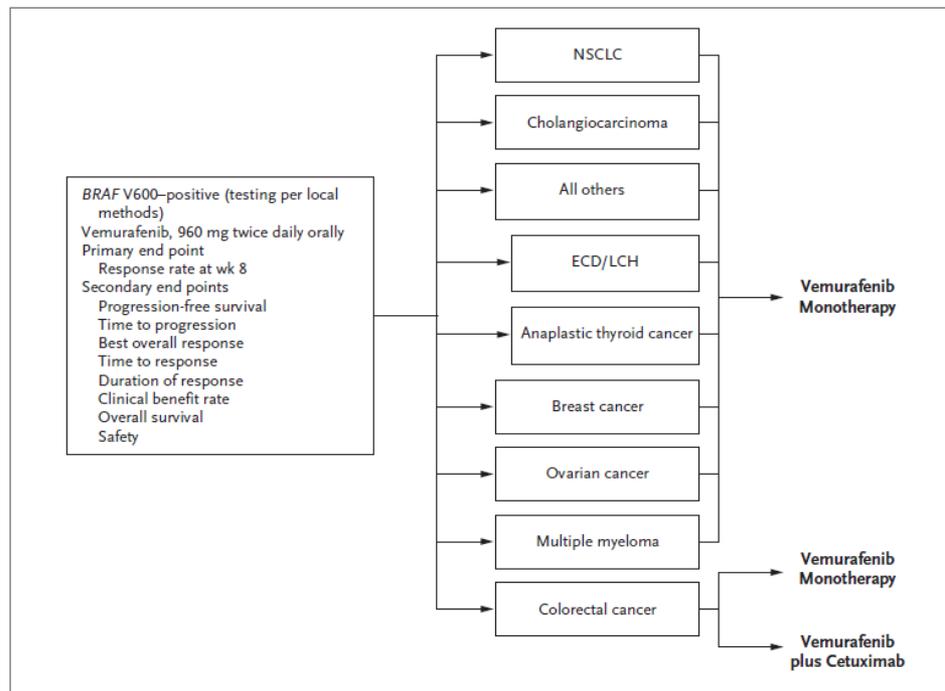
- 開発早期のPOC（proof-of-concept）の評価を目的とした探索的単相試験
- 各サブ試験の症例数は20-50例
 - 治療効果が大きい場合のみに仮説の有意性が示せるようにサンプルサイズ設定
- 早期無効中止を目的とした2段階・多段階デザインを利用

バスケット試験の特徴

- 腫瘍の分子特性がある標的治療に反応することを一定の確度で予測でき、かつそれが腫瘍の組織型とは無関係に成立することを前提とした試験
 - 疾患の潜在的連関性
- 各サブ試験では、複数のがん種を対象にすることから、生命予後の異質性が高く、時間イベント型の評価項目の評価は困難
- 腫瘍縮小効果を主要評価項目とすることが多い

BRAF-V600試験

- メラノーマ以外で*BRAF* V600変異を有する9つのがん種グループ
- ベムラフェニブの有効性を評価
- がん種（コホート）ごとにSimonの2段階デザインに基づいて症例数を決定



Hyman et al., N Engl J Med. 2015

NCI-MATCH (NCI Molecular Analysis for Therapeutic Choice) 試験

- 少なくとも1レジメンの治療を受けた固形癌とリンパ腫患者を対象
- 1薬剤に対して1バスケット試験を実施
- 主要評価項目は腫瘍縮小効果
- 重要な副次的評価項目は6か月PFS率
- 単群デザインにより, 試験ごとに35例を登録
 - Signal findingを目的としているので, どの試験も同じ統計的仮説

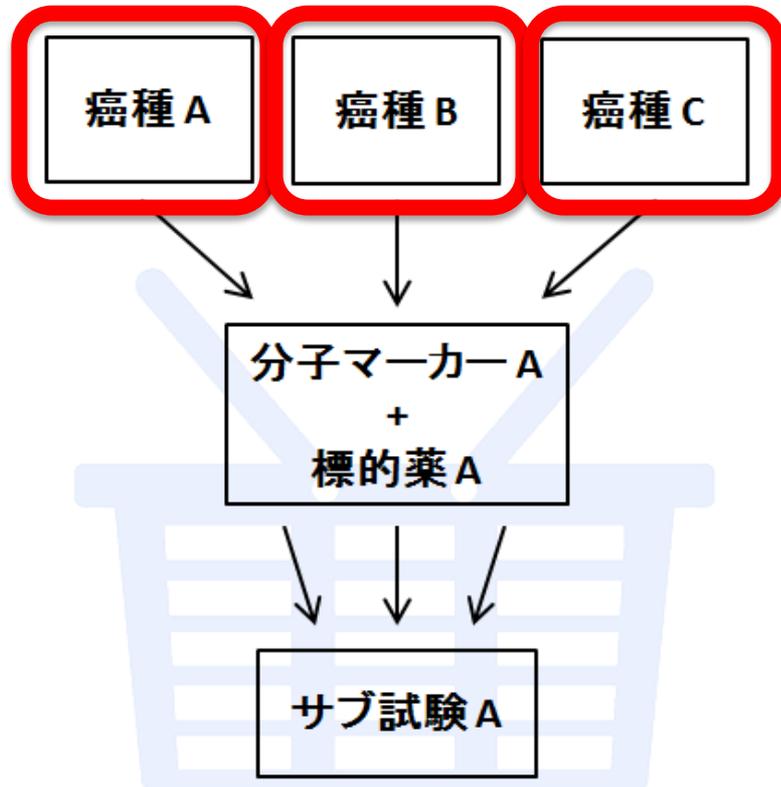
- One-stage design (no interim futility rule)
 - 31 evaluable patients per arm
 - N = 35 total (assuming 10% ineligibility)
 - Need ~8 months of follow-up after accrual is complete in a given arm
- Primary endpoint: Overall Response Rate $H_0: \leq 5\%$ vs $H_a: 25\%$
 - Reject H_0 if $\geq 5/31$ responses (~16%)
 - Type I error 1.8% / arm (one-sided)
 - Power 92%
- Secondary endpoints:
 - Progression Free Survival (PFS) at 6 months: Distinguish 15% (median PFS 2.2 m) vs 35% (median PFS 4 m)

<https://www.cancer.gov/about-cancer/treatment/clinical-trials/nci-supported/nci-match>, last accessed 1 October, 2018

登録中のサブ試験

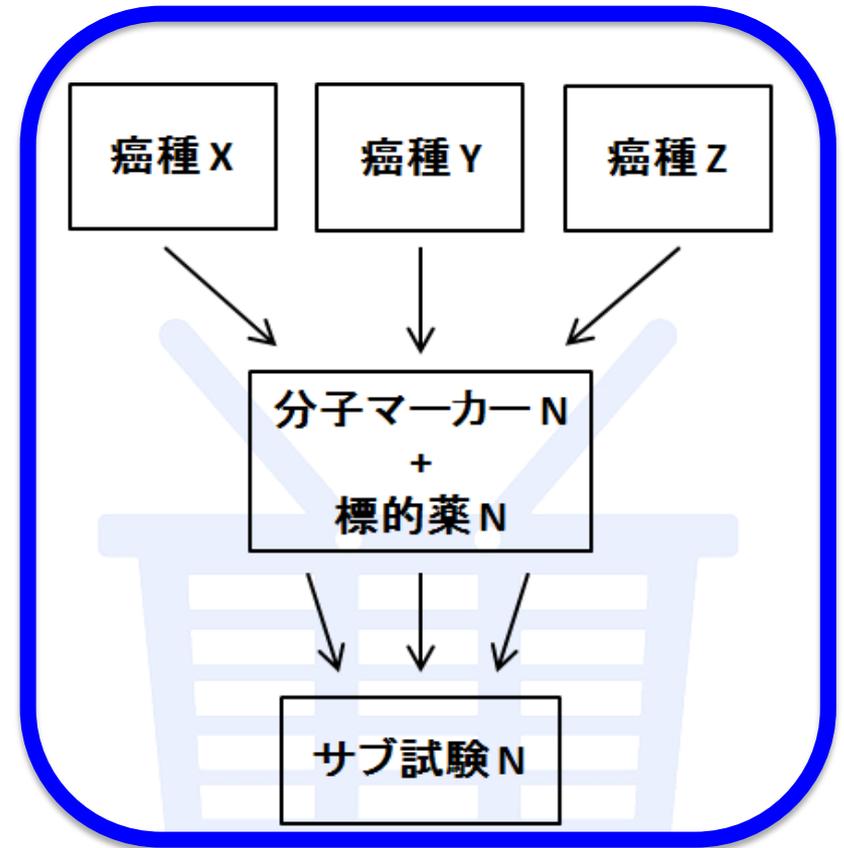
Arm	Targeted Genetic Change	Drug(s)
A	EGFR mut	Afatinib
C2	MET ex 14 sk	Crizotinib
E	EGFR T790M	AZD9291
F	ALK transloc	Crizotinib
G	ROS1 transloc	Crizotinib
J	HER2 amp	Trastuzumab, Pertuzumab
K1	FGFR amp	Erdafitinib
K2	FGFR mut or fusions	Erdafitinib
L	mTOR mut	TAK-228
M	TSC1 or TSC2 mut	TAK-228
S2	GNAQ/GNA11 mut	Trametinib
T	SMO/PTCH1 mut	Vismodegib
V	cKIT mut	Sunitinib
Z1C	CDK4 or CDK6 amp and Rb exp	Palbociclib
Z1E	NTRK fusions	Larotrectinib (LOXO-101)
Z1F	PIK3CA	Copanlisib
Z1G	PTEN loss without PIK3CA mut	Copanlisib
Z1H	PTEN (deleterious) seq result and PTEN exp	Copanlisib

注：2つの考え方



がん種ごとに例数設計
(BRAF V600試験)

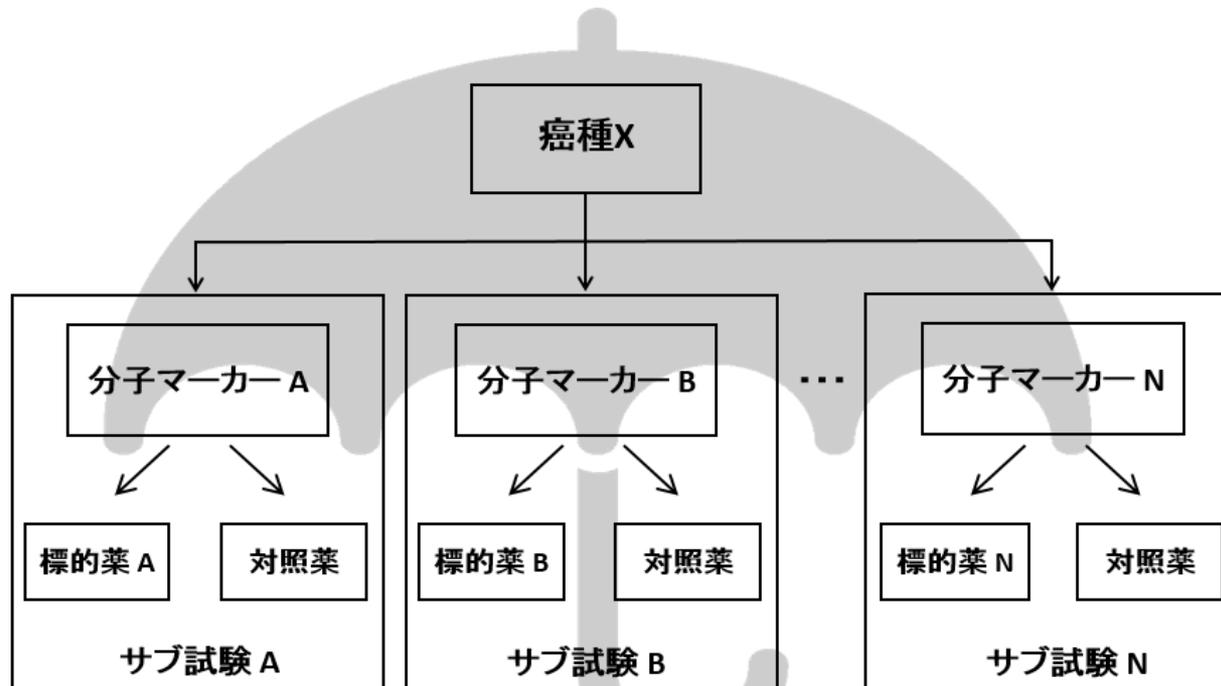
...



全がん種を1集団とみなして
例数設計
(NCI-MATCH)

アンブレラ試験

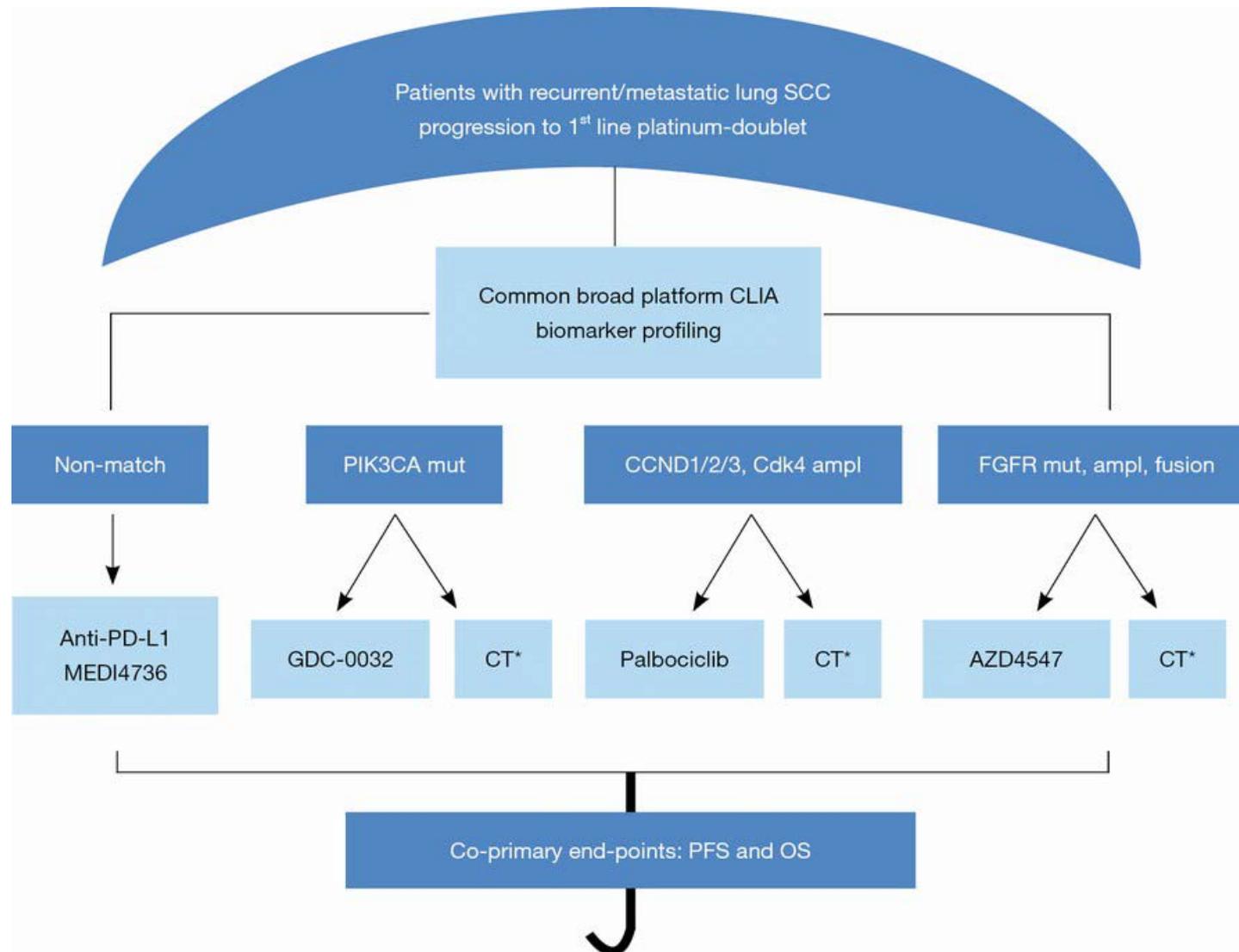
- 特定のがん種を対象にして、分子マーカーや遺伝子変異の別に基づいて、それぞれ対応する標的治療を評価するサブ試験を実施



アンブレラ試験の特徴

- サブ試験は、検証目的のランダム化試験であることが多い
- アンブレラ試験は大規模かつ長期的な試験
 - 期間中に標準治療が変更になると、当該標準治療を対照群としたサブ試験の臨床的意義が乏しくなる

Lung-MAP試験



<http://cco.amegroups.com/article/view/7845/8606>, last accessed 20 Nov, 2018

プラットフォーム試験

- 永続的に単一の疾患に対して複数の標的治療を評価し、さらに試験中に新たな治療法や対象患者の追加や除外を認める試験
 - 中間解析によって各標的治療の有効性及び無益性を評価
- バスケット試験やアンブレラ試験も、試験中に治療法の追加や除外を許容すれば、プラットフォーム試験として位置付けることも可能

FDA ドラフトガイダンス

- 2018年9月にリリース

Master Protocols: Efficient Clinical Trial Design Strategies to Expedite Development of Oncology Drugs and Biologics Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

統計的事項

329 VII. STATISTICAL CONSIDERATIONS

330

331 A. Nonrandomized, Activity-Estimating Design

332

333 In nonrandomized protocols, where the primary endpoint is ORR, the planned sample size
334 should be sufficient to rule out a clinically unimportant response rate based on the lower bound
335 of the 95 percent confidence interval around the observed response rate. The analysis plan should
336 describe the futility analyses to be conducted. FDA recommends designs such as the Simon two-
337 stage design²⁰ that limit exposure to an ineffective drug. If a sponsor anticipates that the results
338 would form the primary basis of an efficacy claim in a marketing application, the clinical
339 protocol and SAP should ensure that collected data are of adequate quality for this purpose.
340 Additionally, the SAP should prespecify the timing of the final analysis, ensure adequate data
341 collection and follow-up on all patients for efficacy and safety, and describe the plan for
342 independent review of confirmed ORR in solid tumors for each substudy. If preliminary results
343 from a substudy or substudies suggest a major advance over available therapy, the sponsor
344 should meet with the review division to discuss modifications to the protocol.

バスケット試験に対するFDAの考え方

■ 目的

- 腫瘍縮小効果の推定
 - Signal-finding trial

■ デザイン時の留意点

- 奏効率が低い場合に, (そのがん種又は薬剤) をrule out できるようなデザイン
- Simonの2段階デザインを推奨

ベイズ流デザインも活用

C. Master Protocols Employing Adaptive/Bayesian Design

In master protocols that incorporate adaptive designs, the SAP should provide all information described in the guidance for industry *Adaptive Design Clinical Trials for Drugs and Biologics* and the draft guidance for industry *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*²¹ and describe plans for futility analyses.²² Master protocols can use a Bayesian statistical method or other methods for planning or modifying the sample size, dropping an arm, or other adaptive strategies. The SAP should include details on implementation of Bayesian or other methods.

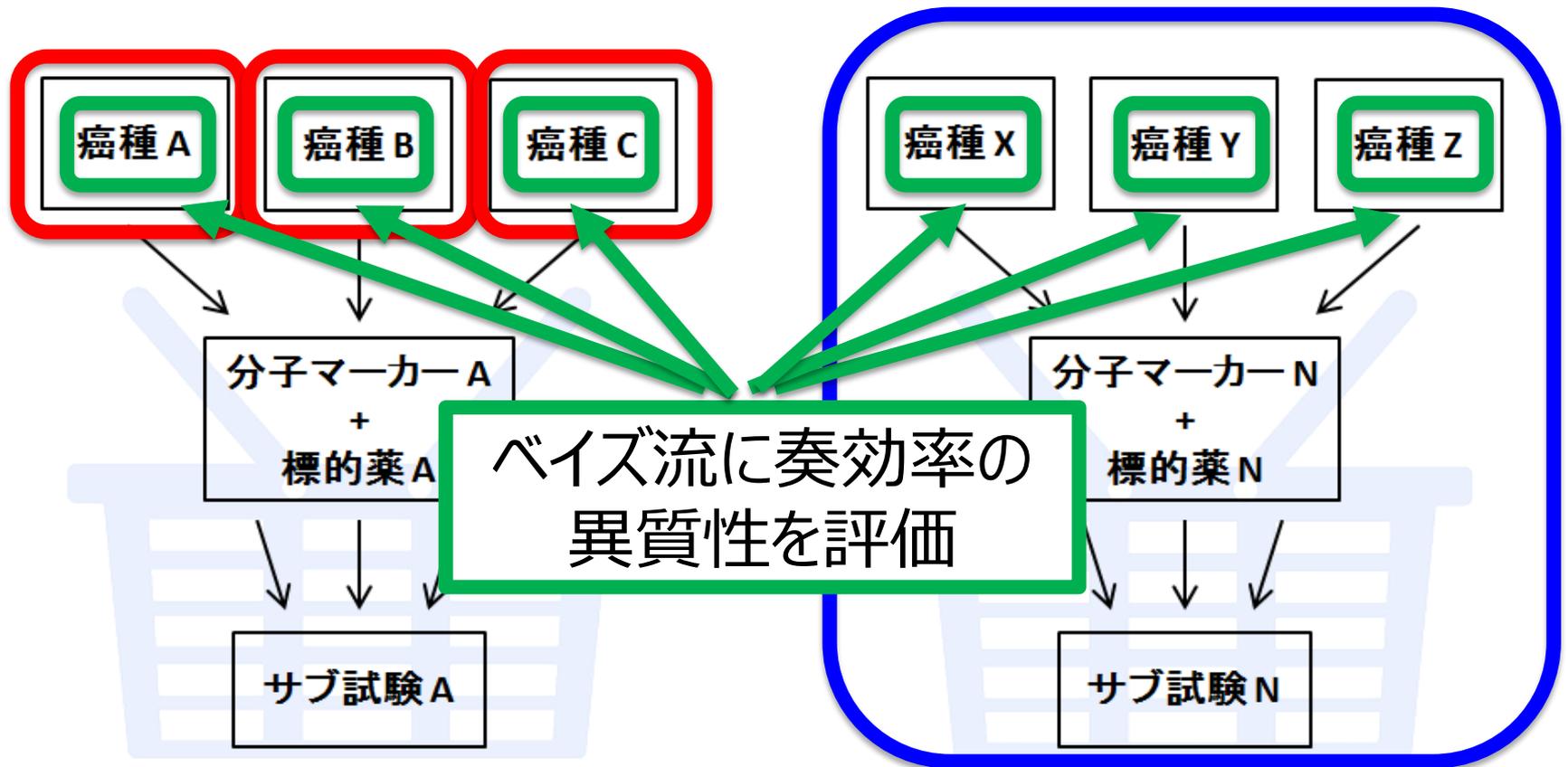
Adaptive Designs for Clinical Trials of Drugs and Biologics Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

- ベイズ流デザインにより、無効ながん種を評価

無効ながん種は？



がん種ごとに例数設計
(BRAF V600試験)

全がん種を1集団とみなして
例数設計
(NCI-MATCH)

比較試験の場合

D. Master Protocols With Biomarker-Defined Subgroups

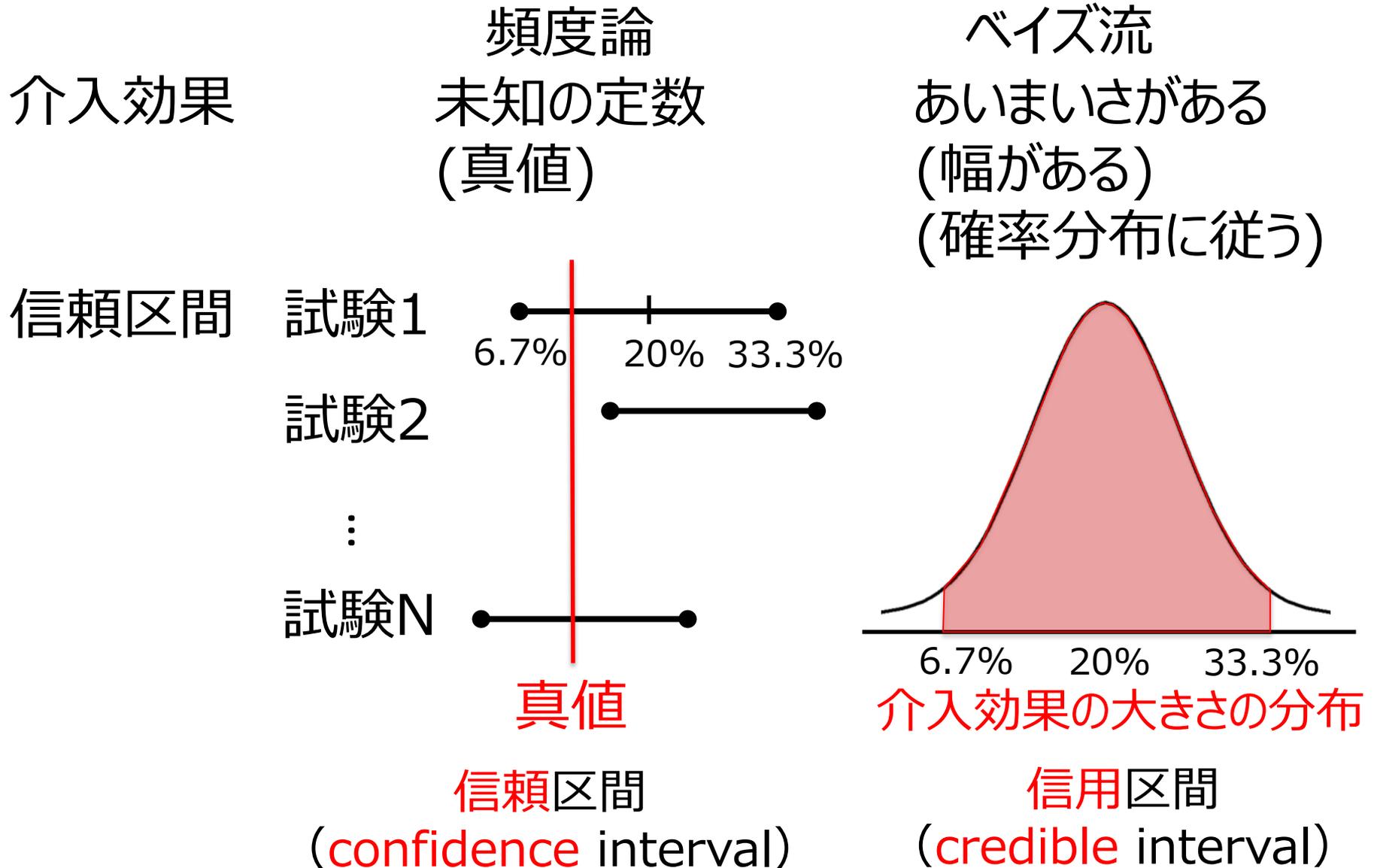
In master protocols with basket or complex design, where patient assignment to a treatment arm is based on the presence of a specific biomarker of interest, the protocol should clearly specify how patients with more than one biomarker of interest will be assigned to substudies. There are two approaches to making such assignments that FDA considers acceptable from a clinical trial design perspective, but other approaches may also be appropriate. One approach is to prioritize biomarkers or treatments. For example, in the BATTLE-1 trial,²³ investigators ranked the biomarker groups based on their predictive values and assigned patients with multiple biomarkers to the group for one of their biomarkers that has the highest predictive value. The other approach is based on a prespecified randomization ratio. For example, the Lung-MAP trial¹⁰ uses a reversed ratio of prevalence rates. Using reverse prevalence ratios, patients in the trial with tumors that have biomarkers with low prevalence have a greater likelihood to be assigned to a substudy for the lower prevalence population.²⁴

ベイズ流評価の考え方

統計学の流派

- 頻度論 (Frequentist) とベイズ流 (Bayesian) の二流派
 - “確率”をどう捉えるか, どう定式化するか, が異なる
- 臨床試験や観察研究における介入効果の評価に利用される統計学は主として頻度論
 - (点)推定値と95%信頼区間により介入効果の大きさを評価し, P値によって統計学的有意性を評価
 - 例
 - 介入群(N=100)と対照群(N=100)の有効例はそれぞれ50例と30例であった.
 - 有効率の差の点推定値 : 20%
 - 有効率の差の95%信頼区間 : [6.7%, 33.3%]
 - 有効率の差のp値 (カイ二乗検定) : $p = 0.0039$

頻度論とベイズ流の信頼区間

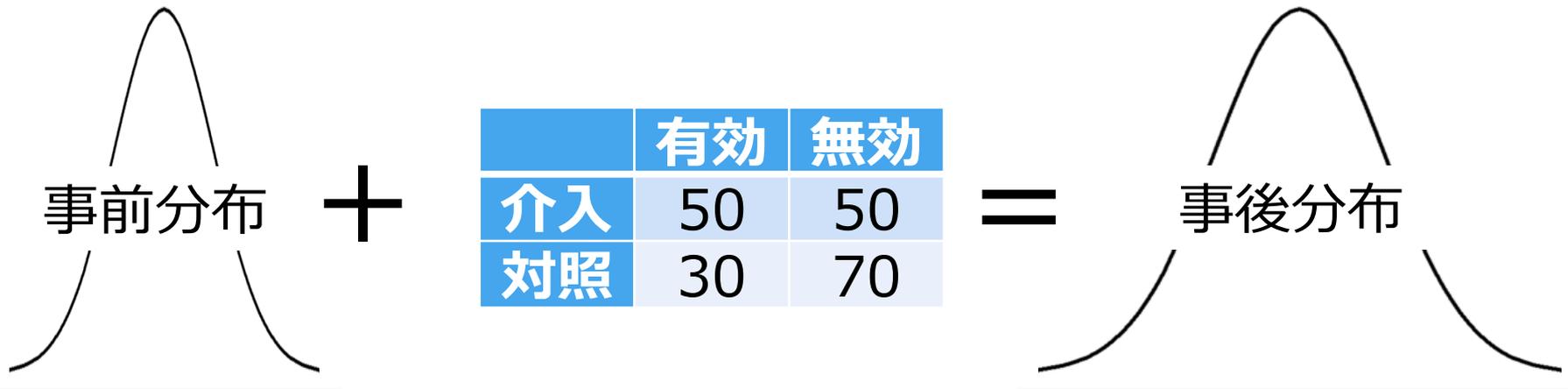


ベイズ流アプローチの用語

- 事前分布 (prior distribution)
 - 試験開始前に想定される介入効果の分布
 - ヒストリカルデータ等に基づき決定
 - 事前情報 (prior information) と呼ばれる場合もある
- 事後分布 (posterior distribution)
 - 試験で認められた介入効果 (データ) と介入効果の事前分布 (事前情報) を統計的に足し合わせて得られる介入効果の分布

ベイズ流アプローチの適用手順

1. 試験開始前に介入効果のあいまいさを事前分布として与える
2. 試験を実施し，介入効果のデータを収集
3. 介入効果のデータと事前分布を統計的に足し合わせて，事後分布を得る
4. 事後分布に基づいて介入効果を定量的に評価する
(※p値による評価は存在しない)



どのように利用されている？

実施
可能性

学術的
新規性

- | | | |
|-----------------------------------------------|---|-----|
| 1. 試験デザインと介入効果の評価を共にベイズ流で実施する | 難 | 高 |
| 2. 試験デザインは頻度論，解析はベイズ流 | 普 | 中～高 |
| 3. 既に終了した試験の介入効果をベイズ流アプローチにより再評価（レトロスペクティブ評価） | 易 | 中 |

事例紹介

1. 試験デザインと介入効果の評価を共にベイズ流で実施する

Park et al, N Engl J Med 2016;375:11-22 (I-SPY2)

2. 試験デザインは頻度論, 解析はベイズ流

Laptook et al. JAMA. 2017;318:1550-1560

3. 既に終了した試験の介入効果をベイズ流アプローチにより再評価 (メタアナリシスを除く)

Tan et al. Brith J Cancer 2002;86:843-850

新生児の低酸素性虚血性脳症に対する低体温療法

■ 背景

- 妊娠36週以降生誕の新生児低酸素性虚血性脳症に対し、生後6時間以内開始の低体温療法は死亡または障害発生を減少
 - 208例のランダム化比較試験 (Shankaran et al., N Engl J Med 2005; 353:1574-1584)

■ 仮説

- 6-24時間に当該療法を実施した場合、実施しなかった場合と比べて、死亡又は障害発生のリスクはどうか？

Research

JAMA | **Original Investigation**

**Effect of Therapeutic Hypothermia Initiated After
6 Hours of Age on Death or Disability Among Newborns
With Hypoxic-Ischemic Encephalopathy
A Randomized Clinical Trial**

ランダム化比較試験

■ 生後6時間以内開始の低体温療法の効果

- 標準治療に比べてリスク比(RR): 0.72
 - 95%CI[0.54, 0.95], $p = 0.01$

■ サンプルサイズ設計

- リスク比=0.72, 有意水準=5%, 検出力=80% => 243例/群
- 生後6-24時間以内開始の場合, 治療効果はもう少し小さい
- 患者登録期間は6年であり, 実施可能性から168例しか集積できない
- 頻度論に基づく仮説検定を実施する評価方法は現実的でない

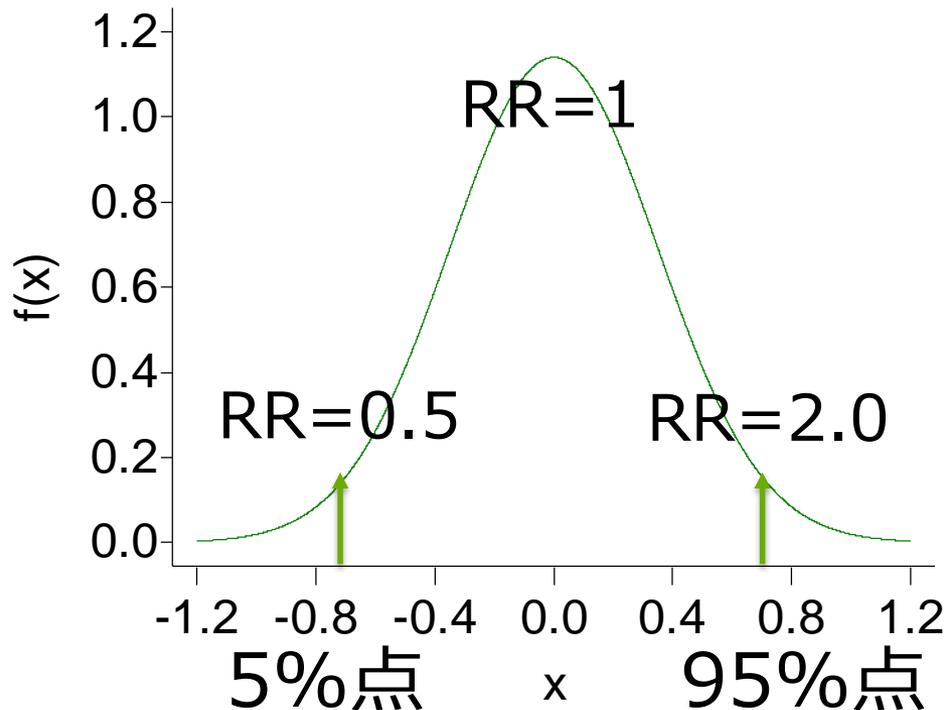
■ ベイズ流アプローチにより, 介入効果の事後分布を評価

事前分布の設定

- 3つの事前分布を設定
- Neutral prior (中立的)
 - 治療効果なし → $RR=1$
- Enthusiastic prior (熱狂的)
 - NEJMの比較試験と同じ効果 → $RR=0.72$
- Skeptical prior (懐疑的)
 - むしろ悪影響 → $RR=1.1$

中立的事前分布

- RR=1のあいまいさを事前分布で表現
 - RRは0.5~2.0と仮定
- 対数変換して正規分布を仮定
 - 平均： $\log(\text{RR}=1) = 0$
 - 標準偏差： $2 * \text{SD} = \log(\text{RR}=0.5) = -0.69 \rightarrow \text{SD} = 0.35$

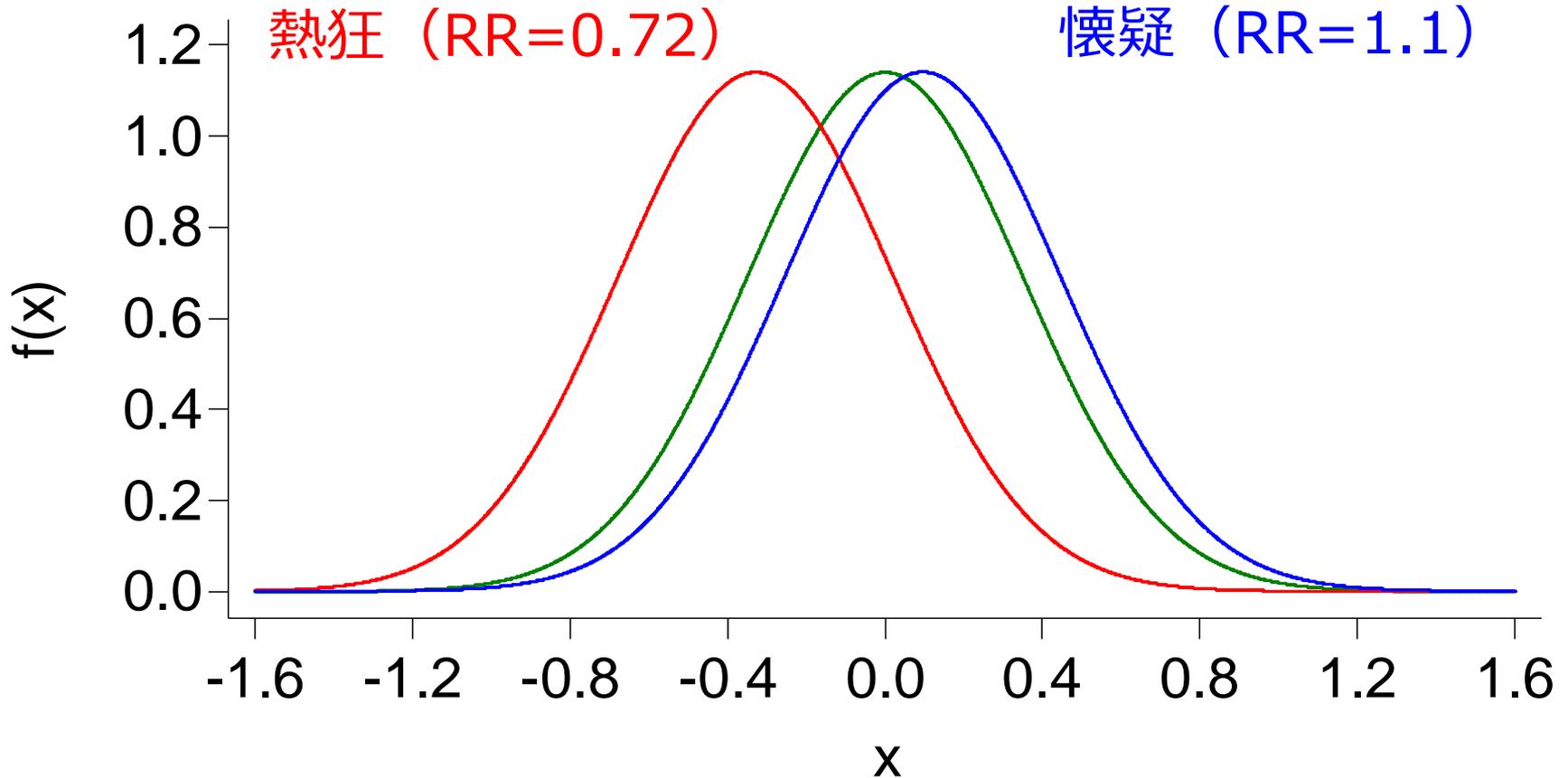


3つの事前分布

中立 (RR=1)

熱狂 (RR=0.72)

懐疑 (RR=1.1)



試験結果

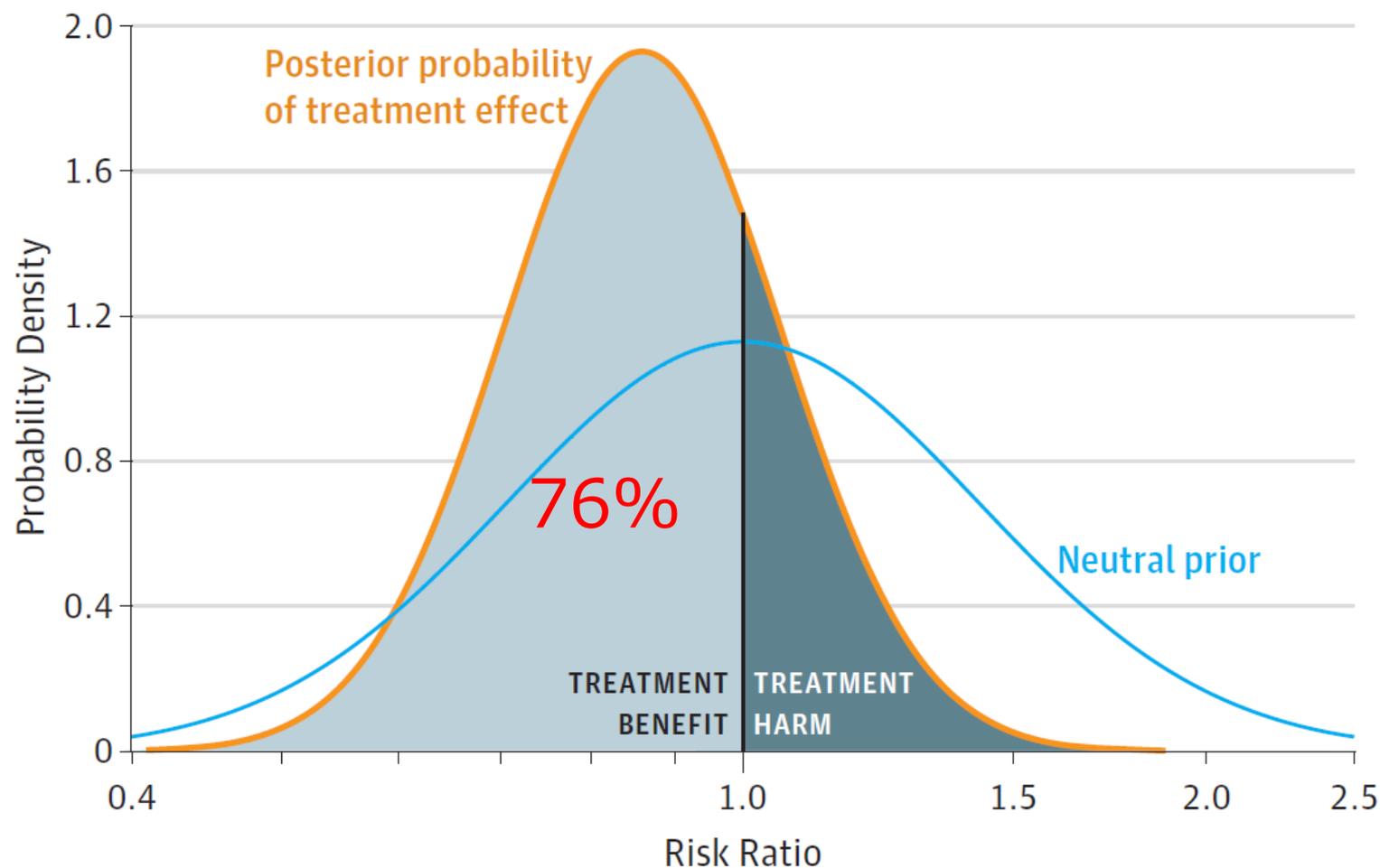
- P-TB: $RR < 1.0$ となる事後確率
 - $RR < 1.0$ の範囲での事後分布の面積

Table 3. Primary and Secondary Outcomes: aRRs and Posterior Probability of Treatment Effect^a

Outcome	No. (%)		Enthusiastic Prior (RR, 0.72)		Neutral Prior (RR, 1.0)		Skeptical Prior (RR, 1.10)	
	Hypothermia (n = 78)	Noncooled (n = 79)	aRR (95% Credible Interval)	P-TB, %	aRR (95% Credible Interval)	P-TB, %	aRR (95% Credible Interval)	P-TB, %
Primary Outcome								
Death or moderate-severe disability	19 (24.4)	22 (27.9)	0.78 (0.52-1.15)	90	0.86 (0.58-1.29)	76	0.89 (0.60-1.32)	73
Secondary Outcomes								
Death ^b	9 (11.5)	9 (11.4)	0.74 (0.45-1.21)	89	0.86 (0.54-1.44)	73	0.90 (0.56-1.52)	67
Moderate or severe disability ^c	10 (12.8)	13 (16.5)	0.74 (0.44-1.24)	87	0.89 (0.54-1.48)	68	0.93 (0.56-1.55)	61
Severe disability ^c	9 (11.5)	12 (15.2)	0.73 (0.43-1.23)	88	0.88 (0.53-1.50)	68	0.93 (0.55-1.55)	61
Moderate disability ^{c,d}	1 (1.3)	1 (1.3)						
Mild disability ^c	16 (20.5)	12 (15.2)	1.0 (0.62-1.62)	50	1.18 (0.73-1.91)	25	1.23 (0.76-2.0)	20

中立的事前分布を用いた際の事後分布

A Risk ratio



Abstract

Results Hypothermic and noncooled infants were term (mean [SD], 39 [2] and 39 [1] weeks' gestation, respectively), and 47 of 83 (57%) and 55 of 85 (65%) were male, respectively. Both groups were acidemic at birth, predominantly transferred to the treating center with moderate encephalopathy, and were randomized at a mean (SD) of 16 (5) and 15 (5) hours for hypothermic and noncooled groups, respectively. The primary outcome occurred in 19 of 78 hypothermic infants (24.4%) and 22 of 79 noncooled infants (27.9%) (absolute difference, 3.5%; 95% CI, -1% to 17%). Bayesian analysis using a neutral prior indicated a 76% posterior probability of reduced death or disability with hypothermia relative to the noncooled group (adjusted posterior risk ratio, 0.86; 95% credible interval, 0.58-1.29). The probability that death or disability in cooled infants was at least 1%, 2%, or 3% less than noncooled infants was 71%, 64%, and 56%, respectively.

JAMA Guide to Statistics and Methods

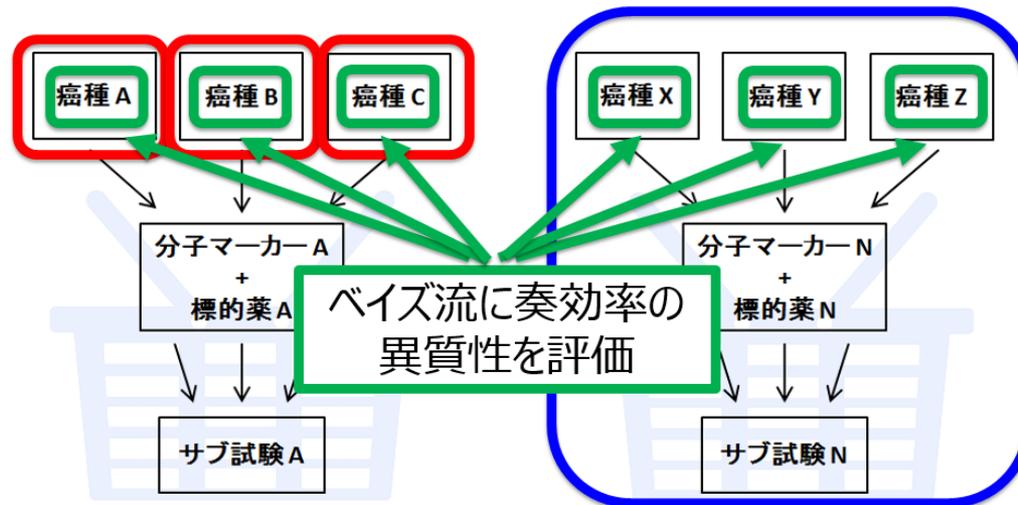
Bayesian Analysis: Using Prior Information to Interpret the Results of Clinical Trials

Melanie Quintana, PhD; Kert Viele, PhD; Roger J. Lewis, MD, PhD

バスケット試験における統計学的課題

奏効率の推定

- 各部分集団のサンプルサイズは数～数十
- 部分集団の奏効率を正しく解釈するために、**集団間でデータ（情報）を共有**することを考える
 - 情報借用（borrowing information）
 - 強度借用（borrowing strength）



情報借用への期待

■ 情報借用の目的は？

- （借用の可否や程度を検討した上で）他集団の結果を利用して部分集団の評価を行う

■ どのように情報借用するの？

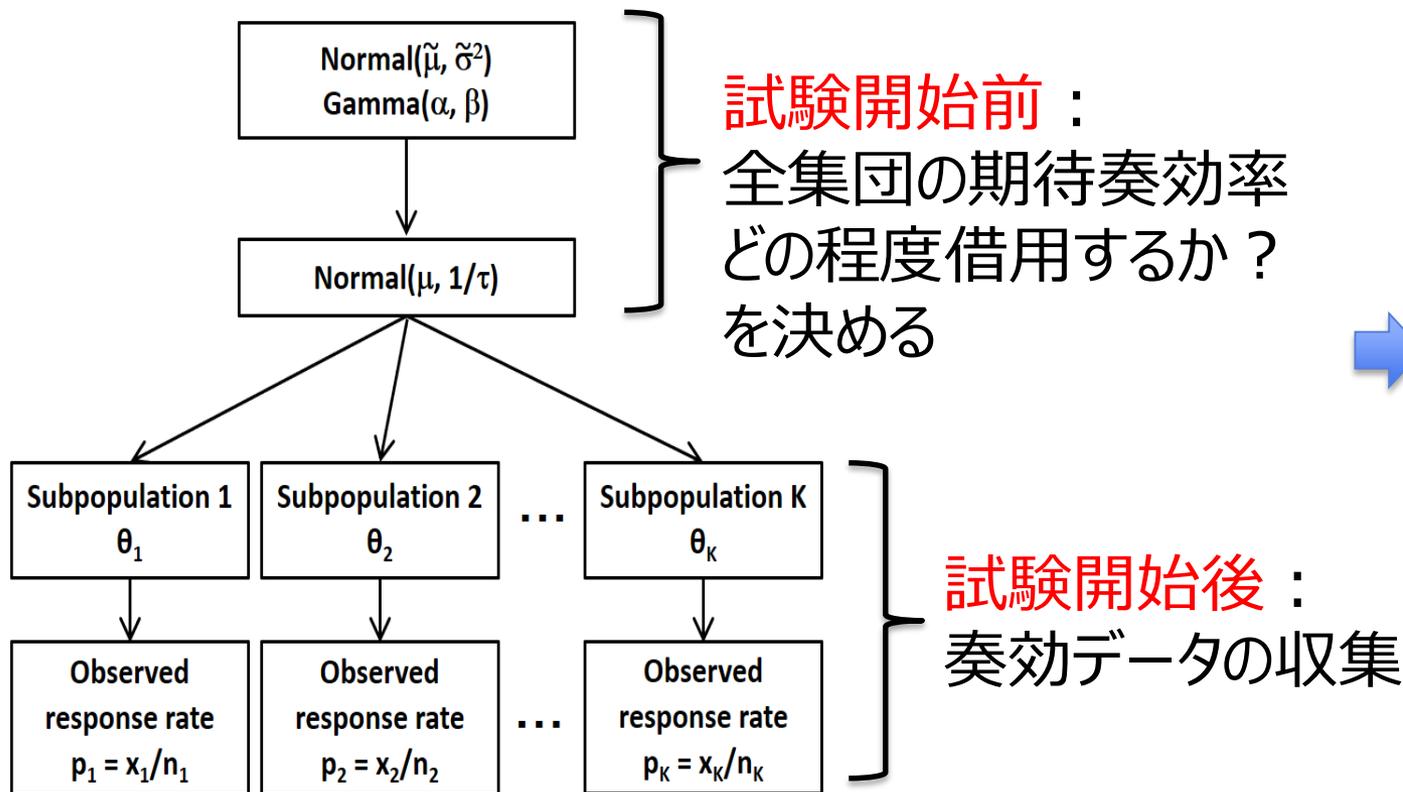
- 階層モデル, と呼ばれるベイズ流手法を利用

情報借用への期待

- 階層モデルから何を知りたい？
 - 各部分集団の奏効率を推定する際に，情報借用すると，**どの程度，奏効率が縮小するか？**
- 理論的には，事前情報と観察データから部分集団間の情報借用の程度を調整するアプローチ
 - ⇒ 『アウトカム-アダプティブ借用』という用語の方が分かり易い

階層ベイズモデル

- 各がん種の奏効率は、1つの正規分布に従っている、と仮定



奏効率を事前
規定した借用
の強さで調整

試験開始前：どの程度借用するか？

■ 期待奏効率：20%

■ $\mu = \log \left\{ \frac{0.2}{1-0.2} \right\} \rightarrow \mu \sim \text{Normal}(-1.386, 10)$

■ どの程度借用するか？

■ $\tau \sim \text{Gamma}(\alpha, \beta)$

(α, β)	(2, 2)	(2, 10)	(2, 20)
情報借用	強	中	弱
平均	1	0.2	0.1
分散	0.5	0.02	0.005
95%区間	(0.18, 2.38)	(0.035, 0.48)	(0.017, 0.24)

借用の程度 = 相関性

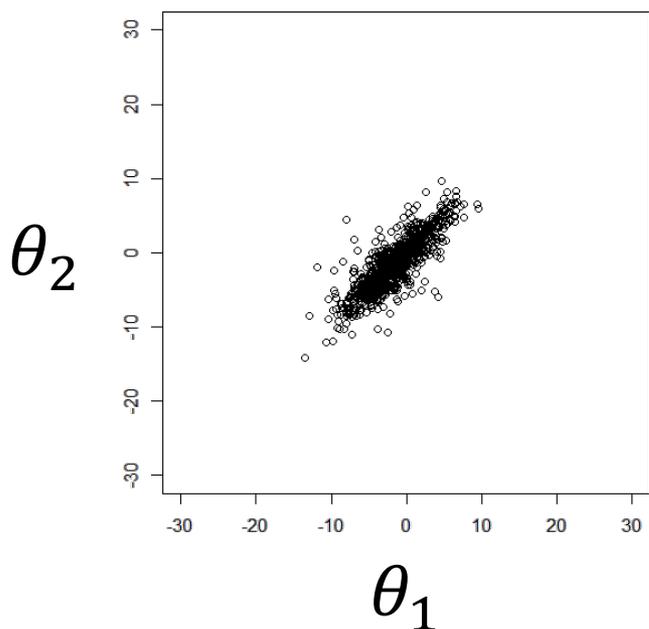
■ (θ_1, θ_2) の相関性

■ 1,000サンプル

$(\alpha, \beta) = (2, 2)$

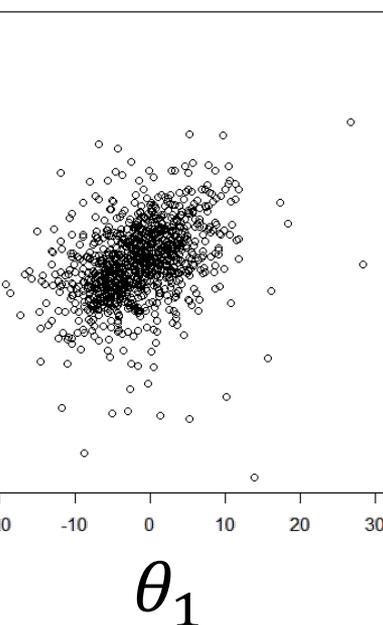
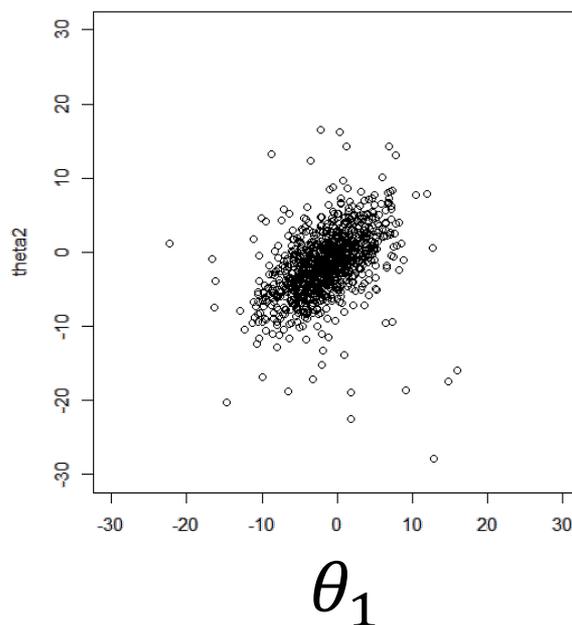
$(2, 10)$

$(2, 20)$



強相関

情報借用の程度は大きい



弱相関

情報借用の程度は低い

ベイズ流評価

- P値ではなく、事後確率と呼ばれる確率で奏効率の有意性を評価

- 部分集団*i*（がん種*i*）の奏効率に関する**事後確率**

$$Pr(\pi_i > \phi | data) > q$$

ϕ : 目標有効確率 (e.g., ≥ 0.05)

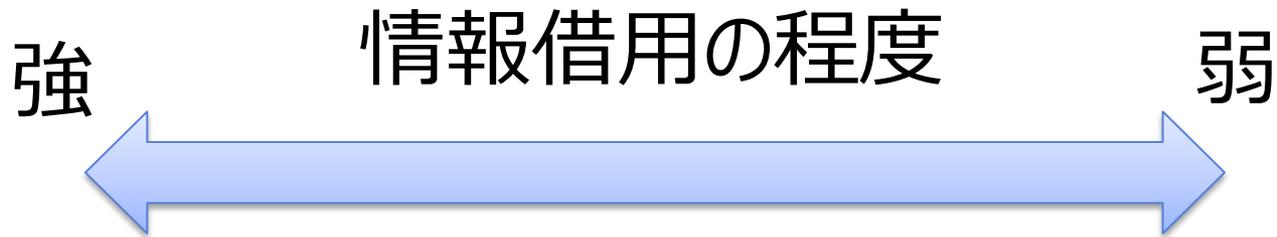
q : カットオフ値 (e.g., ≥ 0.7)

- 任意の時点で事後確率を評価し、試験の中止を検討することも可能

適用例

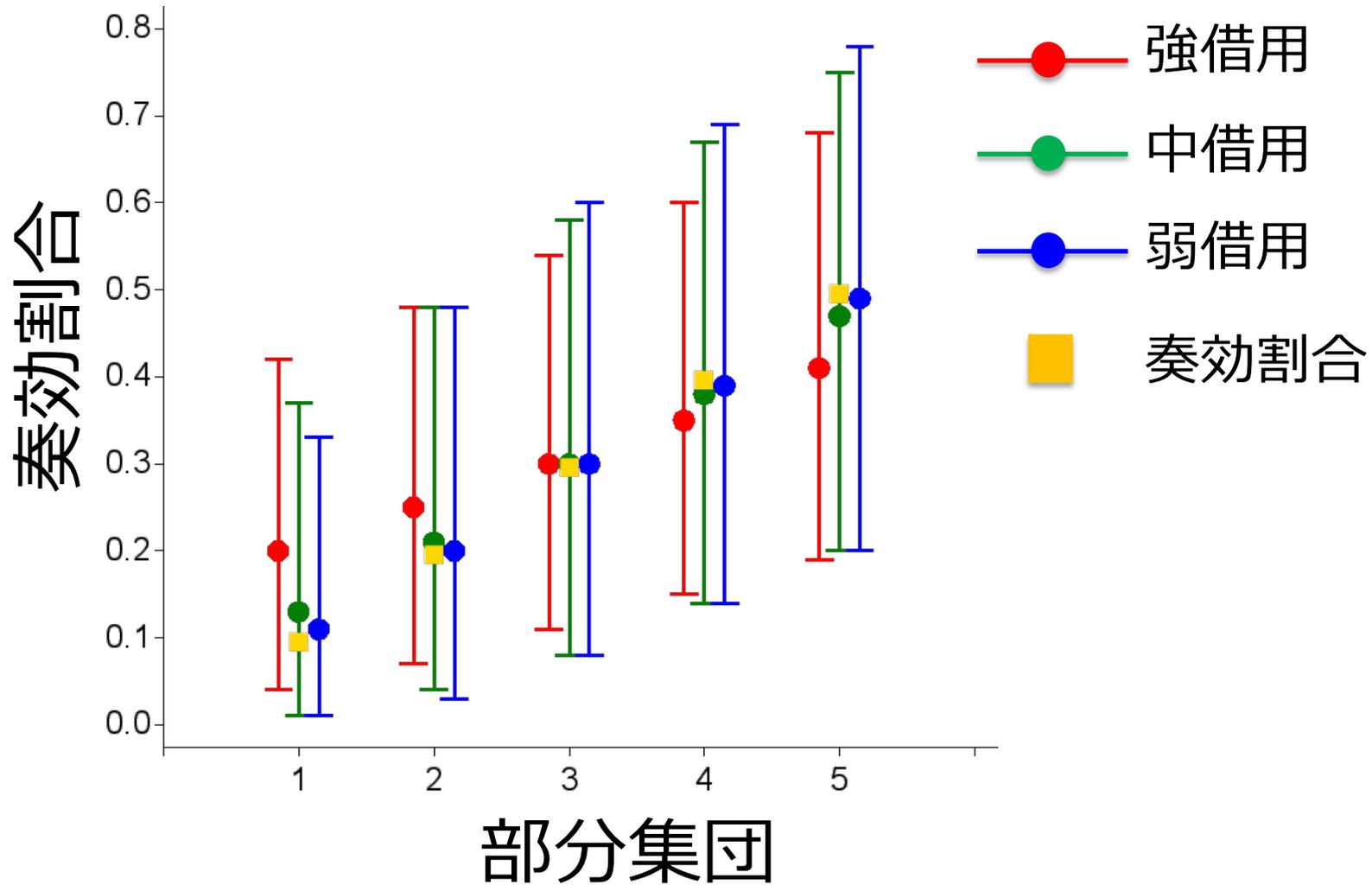
部分集団	症例数	奏効例数	奏効割合	95%CI
1	10	1	0.1	(0.00, 0.45)
2	10	2	0.2	(0.03, 0.56)
3	10	3	0.3	(0.07, 0.65)
4	10	4	0.4	(0.12, 0.74)
5	10	5	0.5	(0.19, 0.81)

事後平均と95%信用区間



部分 集団	Gamma(2, 2)		Gamma(2, 10)		Gamma(2, 20)	
	事後 平均	95%CI	事後 平均	95%CI	事後 平均	95%CI
1	0.20	(0.04, 0.42)	0.13	(0.01, 0.37)	0.11	(0.01, 0.33)
2	0.25	(0.07, 0.48)	0.21	(0.04, 0.48)	0.20	(0.03, 0.48)
3	0.30	(0.11, 0.54)	0.30	(0.08, 0.58)	0.30	(0.08, 0.60)
4	0.35	(0.15, 0.60)	0.38	(0.14, 0.67)	0.39	(0.14, 0.69)
5	0.41	(0.19, 0.68)	0.47	(0.20, 0.75)	0.49	(0.20, 0.78)

95%信用区間



$$Pr(\pi_i > 0.3 | \mathbf{n}, \mathbf{x}), q = 0.7$$

強 ← 情報借用の程度 → 弱

部分集団	(2, 2)	(2, 10)	(2, 20)
1	0.16	0.06	0.04
2	0.28	0.22	0.21
3	0.45	0.45	0.47
4	0.64	0.69	0.70
5	0.79	0.87	0.88

ベムラフェニブのバスケット試験

- *BRAF* V600 陽性の5癌種
 - 35%以上が望ましい

	症例数	奏効例数	奏効割合
肺癌	19	8	0.42
大腸癌	10	0	0.00
胆管癌	8	1	0.13
希少癌	14	6	0.43
甲状腺癌	7	2	0.29

奏効率の事後平均

	例数	奏効例数	奏効割合	強借用した結果	中借用した結果	弱借用した結果
肺癌	19	8	0.42	0.38	0.41	0.42
大腸癌	10	0	0.00	0.11	0.04	0.02
胆管癌	8	1	0.13	0.19	0.14	0.13
希少癌	14	6	0.43	0.38	0.42	0.42
甲状腺癌	7	2	0.29	0.27	0.27	0.28

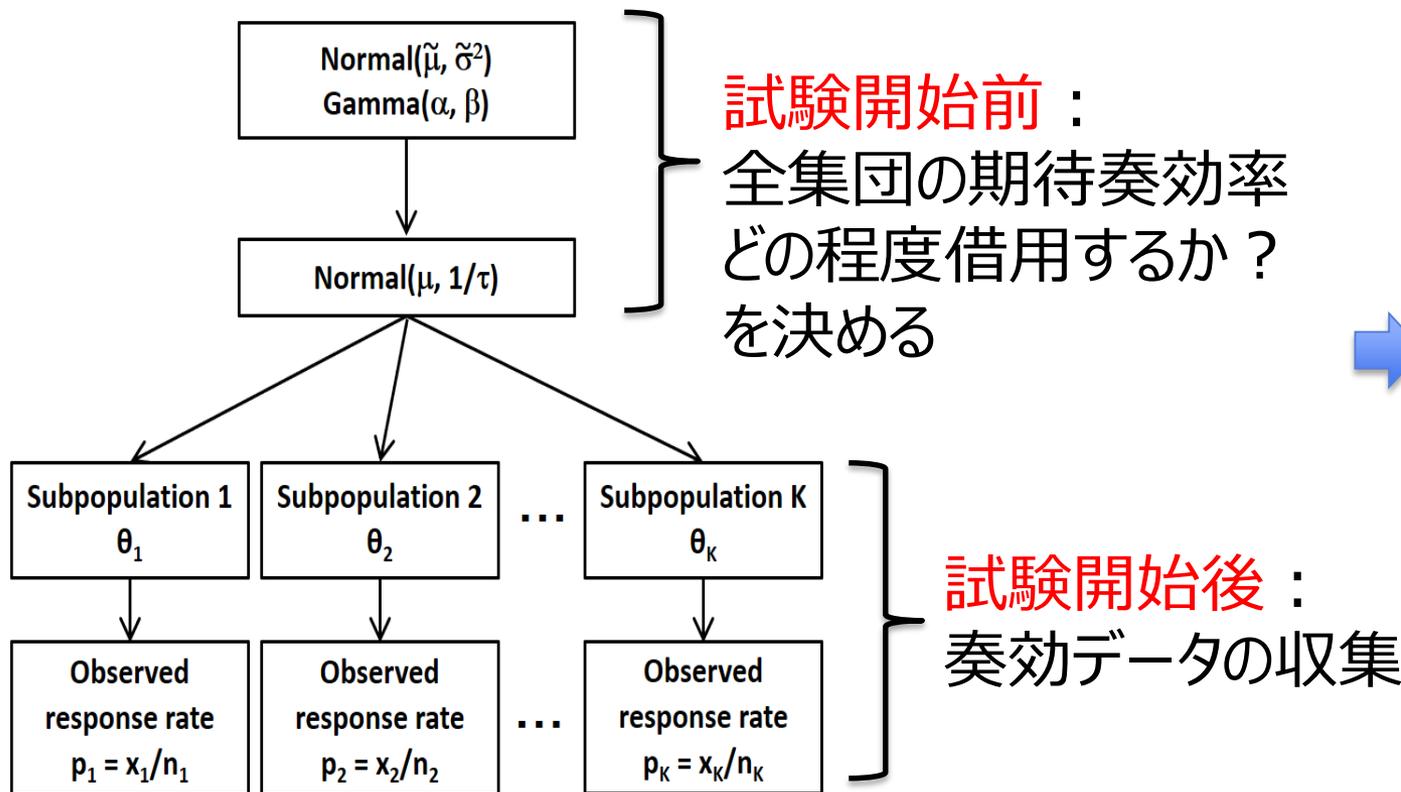
$$Pr(\pi_i > 0.35 | \mathbf{n}, \mathbf{x}), q = 0.7$$

情報借用する層を限定した方が合理的では？

	症例数	奏効 例数	奏効 割合	強借用 した結果	中借用し た結果	弱借用し た結果
肺癌	19	8	0.42	0.62	0.71	0.71
大腸癌	10	0	0.00	0.02	0	0
胆管癌	8	1	0.13	0.08	0.05	0.05
希少癌	14	6	0.43	0.57	0.68	0.70
甲状腺癌	7	2	0.29	0.26	0.29	0.31

発展的方法(EXNEXモデル)

- 各がん種の奏効率は、1つの正規分布に従っている、と仮定



~~奏効率を事前規定した借用の強さで調整~~

結果的に奏効率が類似しているがん種から情報を借用 (EXNEXモデル)

奏効率の事後平均

	例数	奏効 例数	奏効 割合	強借用 した結果	中借用 した結果	弱借用 した結果	EXNEX
肺癌	19	8	0.42	0.38	0.41	0.42	0.40
大腸癌	10	0	0.00	0.11	0.04	0.02	0.04
胆管癌	8	1	0.13	0.19	0.14	0.13	0.18
希少癌	14	6	0.43	0.38	0.42	0.42	0.40
甲状腺癌	7	2	0.29	0.27	0.27	0.28	0.31

$$Pr(\pi_i > 0.35 | n, x), q = 0.7$$

	症例数	奏効 例数	奏効 割合	強借用し た結果	中借用 した結果	弱借用し た結果	EXNEX
肺癌	19	8	0.42	0.62	0.71	0.71	0.67
大腸癌	10	0	0.00	0.02	0	0	0.003
胆管癌	8	1	0.13	0.08	0.05	0.05	0.11
希少癌	14	6	0.43	0.57	0.68	0.70	0.66
甲状腺癌	7	2	0.29	0.26	0.29	0.31	0.46

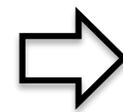
FDAインタビュー

■ 情報借用についてどう考えるか？

- 情報借用の正当性を説明する必要
 - 分子生物学的類似性
 - tumor growthの特徴

■ BRAF阻害剤

- メラノーマ：48%
- 大腸癌：<10%
 - 大腸癌でEGFRが高発現



借用の量を調整することが重要！

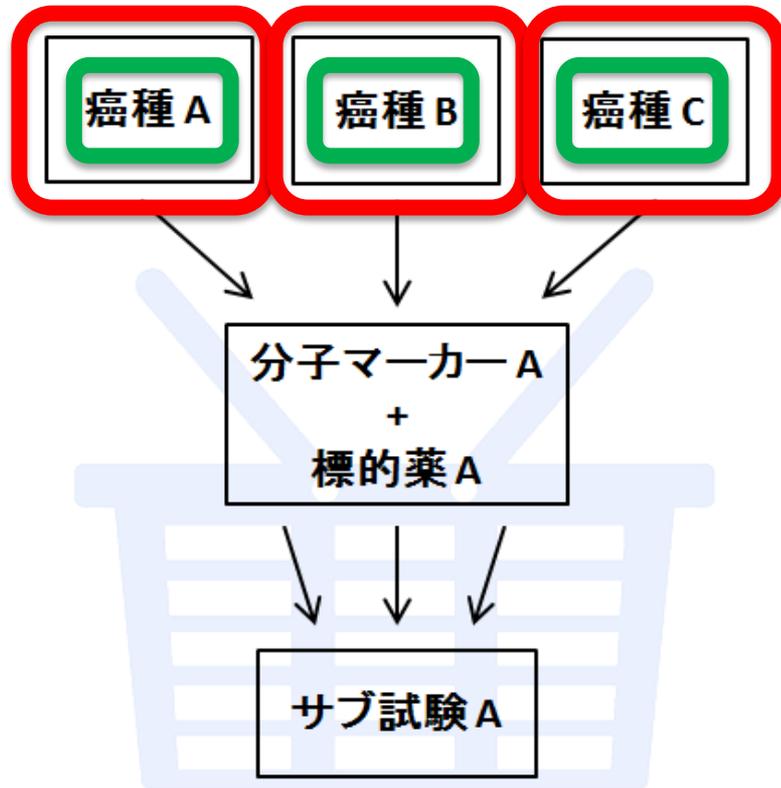
FDAインタビュー

■ 主解析で利用可能か？

- 情報借用の正当性を示した上で、事前規定しているのであれば可能
- 高い奏効率が低くなるのはそれほど問題にならないが、低い奏効率が高くなるのであれば、議論が必要

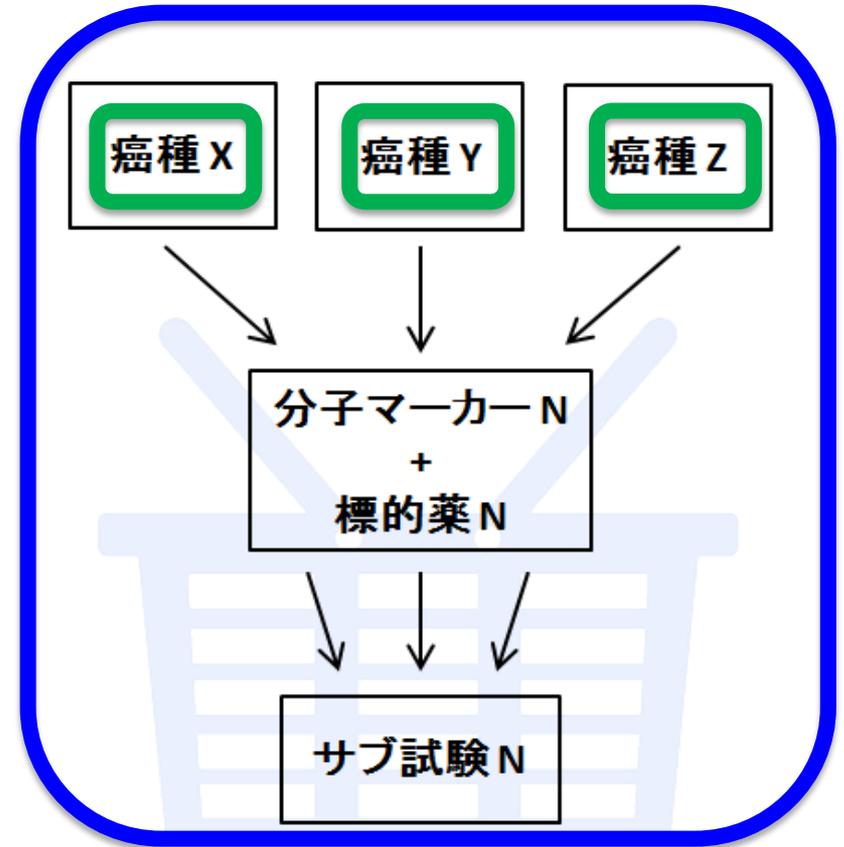
日本におけるバスケット試験の経験

どちらのタイプ？



がん種ごとに例数設計
(BRAF V600試験)

...



全がん種を1集団とみなして
例数設計
(NCI-MATCH)

希少がんバスケットデザイン

■ 課題

- 複数のがん種を対象とするものの、希少性により登録可能例数は非常に少ない
- アカデミアの場合、予算の関係もあり、数年で10-20例程度集積するのが限界

■ **ベイズ流デザインの利用により必要例数を合理化できる！**

単群試験：OSCAR trial

- 希少肉腫を対象とした医師主導第2相試験
 - A clinical trial of nivolumab (Opdivo®) in patients with unresectable clear cell sarcoma and Alveolar soft part sarcoma: OSCAR trial

■ 新しい試み

1. ベイズ流デザインの利用

Analytical Report

DIA

Utility of Bayesian Single-Arm Design in New Drug Application for Rare Cancers in Japan: A Case Study of Phase 2 Trial for Sarcoma

Therapeutic Innovation
& Regulatory Science
1-5
© The Author(s) 2017
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/2168479017728989
tirs.sagepub.com

Akihiro Hirakawa, PhD¹, Tadaaki Nishikawa, MD², Kan Yonemori, MD, PhD², Taro Shibata, MSc³, Kenichi Nakamura, MD, PhD⁴, Masashi Ando, MD, PhD⁵, Takafumi Ueda, MD, PhD⁶, Toshifumi Ozaki, MD, PhD⁷, Kenji Tamura, MD, PhD², Akira Kawai, MD, PhD⁸, and Yasuhiro Fujiwara, MD, PhD²

標準的な第2相試験

- 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 = 約20

奏効率の評価方法

- 20例を集積することが原則
 - 20例未満で終了することは許容されない
- 95%信頼区間(CI)に基づく評価

n	奏効例数	奏効率	95%CI の下限	95%CI の上限
20	3	0.15	0.032	0.379
20	4	0.20	0.057	0.437
20	5	0.25	0.087	0.491

懸念事項

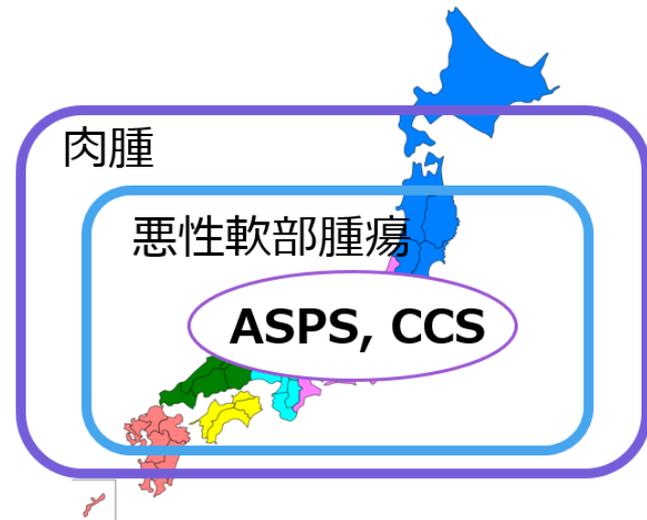
- 2年間で20例の集積は困難かもしれない

- 20例未達なら登録期間の延長

- 延長期間を見定めることすら困難

- 大幅な延長の可能性もある

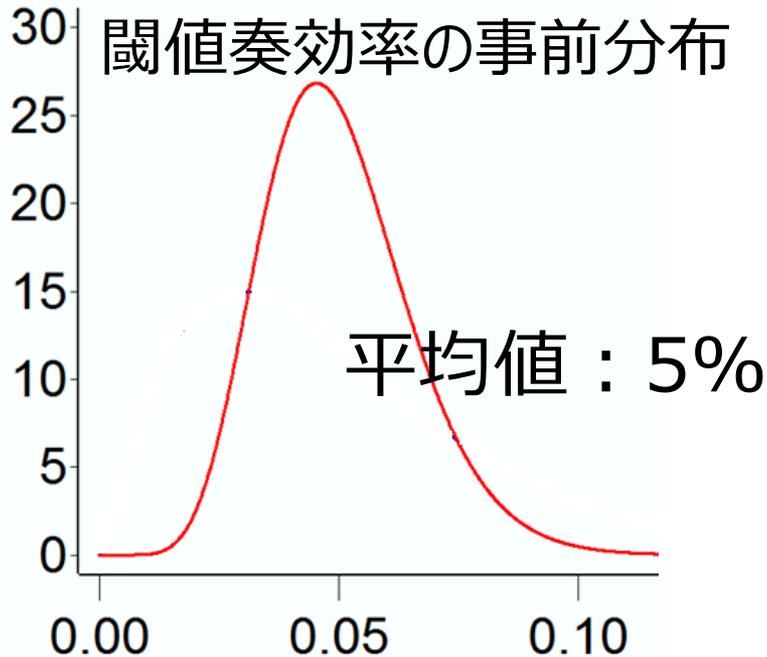
→ 2年時点での登録例数が一定数以上であれば有効性評価したい



ベイズ流デザインの活用

■ 症例数設定時に指定すること

■ 閾値奏効率と期待奏効率の事前分布



■ 事前分布の幅（ばらつき）は、平均値5%に対する**確か**から**しさ**で決める

- 100例中5例？
- 200例中10例？
- 1000例中50例？

症例数

15

16

17

18

19

20

21

22

23

24

25

必要奏効例数

3

3

3

3

3

3

3

4

4

4

4

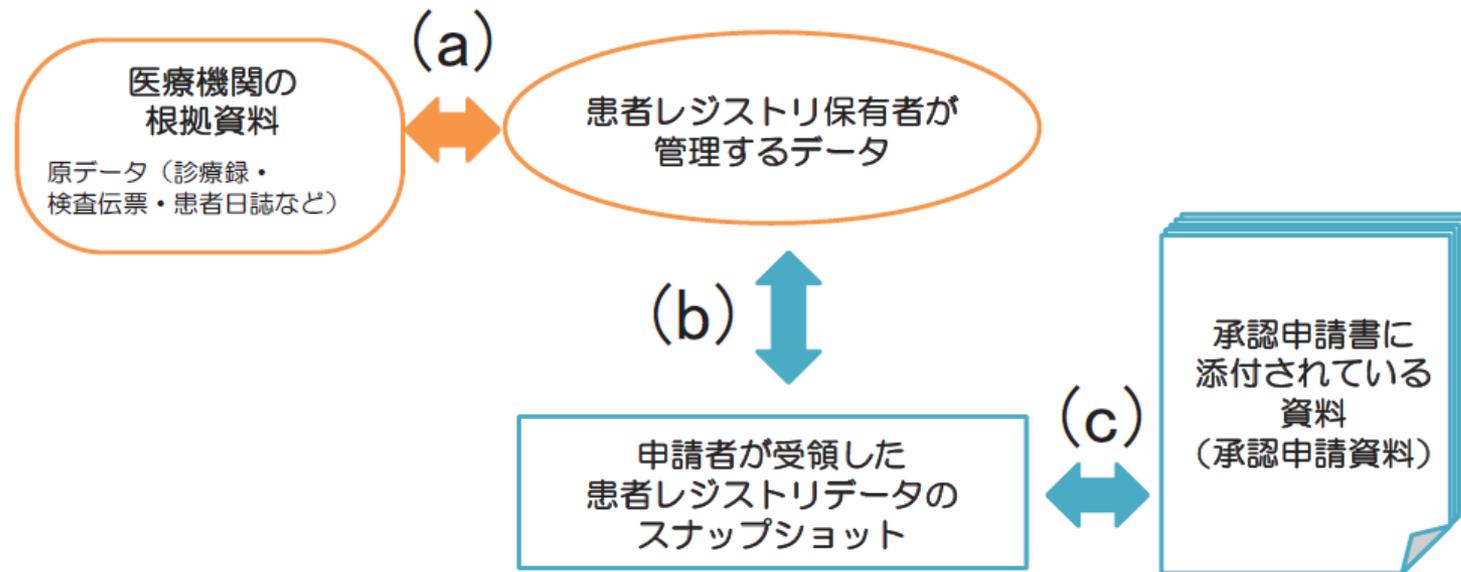
レジストリの利用

- ベイズ流デザインを利用する際、事前分布を決める情報・データは不足していることも多い
 - 事前分布決定の恣意性は、規制上の懸念
- データの信頼性が担保されたレジストリデータを利用することで、定量的な議論が可能
- 通常のデザインよりも、より効率的な意思決定が可能となり、コスト・時間も削減できる
 - 最小・最大例数の決定
 - 必要奏効例数：4例 → 3例

患者レジストリの信頼性の考え方

- 患者レジストリの設計・運用において担保すべきことは何か？
- 承認申請に利用するためには、どのようなプロセスで信頼性を担保すべきか？
 - ※信頼性の程度については言及が困難

医療機関・レジストリ保有者・申請者



- (a) 患者レジストリの設計・運用に関する事項
(医療機関と患者レジストリ保有者との関係)
- (b) 承認申請資料等に提示される情報がその利用目的に照らして十分な水準で信頼性が担保されているものと判断するために必要となる事項
(患者レジストリ保有者と申請者との関係)
- (c) 承認申請資料等に提示される情報がその利用目的に照らして十分な水準で信頼性が担保されているものと判断するために必要となる事項
(申請者と承認申請資料等との関係)

柴田大朗. AMED 林班 分担班による患者レジストリを薬事制度下で活用する際の「データの信頼性の考え方 (案)」より抜粋
AMEDシンポジウム「患者レジストリを薬事制度下で活用するためのデータの信頼性の考え方」2017年8月28日

目的	活用方法	(a) 医療機関 ～患者レジストリ 保有者間	(b) 患者レジストリ 保有者 ～申請者（ある いは企業等）間	(c) 申請者 ～承認申請資料間
承認申請等 以外	(1)市場調査・治験の実 施可能性調査			—
	(2)治験・製造販売後臨 床試験の実施計画作成	医療機関とレジストリ 保有者の2者間での取 り決めに従う	レジストリ保有 者と企業等の2 者間での取り決 めに従う	—
	(3)治験・製造販売後臨 床試験の候補患者のリ クルート			—
承認申請等	(4)希少疾病等の通常 の治験等による開発が困 難な領域における治験 等の対照群等	GCPに準じた医療機関 の原元データとの患者 レジストリのデータと 一致性を確認（困難な 場合理由を説明）	申請者による直 接確認	適合性調査
	(5)製造販売後調査	医療機関の原データが 適切に患者レジストリ に登録される設計・運 営体制となっているこ との確認（困難な場合 理由を説明）	申請者による直 接確認	適合性調査

柴田大朗. AMED 林班 分担班による患者レジストリを薬事制度下で活用する際の「データの信頼性の考え方（案）」より抜粋
シンポジウム「患者レジストリを薬事制度下で活用するためのデータの信頼性の考え方」2017年8月28日

まとめ

- 一定数集積できるがん種を対象にするのであれば、症例数はがん種ごとに頻度流に決定するのが理想
 - 一般的な閾値設定型単群デザイン
 - Simon の2段階デザイン
- 試験中又は試験終了後に、ベイズ流アプローチにより奏効率の異質性を評価
 - 階層ベイズ, EXNEX
- 希少癌等の少数例集団には、全がん種で症例数設定
 - 一般的な閾値設定型単群デザイン
 - Simon の2段階デザイン
 - **ベイズ流デザイン**
- ベイズ流アプローチによる奏効率の異質性評価は例数不足のために不安定になると思われる