

【日本製薬工業協会シンポジウム】

Rare diseaseの臨床開発における治療効果推測の現状と課題

セッション2: Rare diseaseの治療効果の推測法 - CIDパイロットプログラムとその実例 -



2023年2月20日

医薬品評価委員会 データサイエンス部会 KT3

○ヤンセンファーマ株式会社 豊泉 樹一郎

中外製薬株式会社 松田 裕也

- 本演題の内容は、公開されている情報について、本タスクフォースにて集約・考察したものです。タスクフォースメンバーが所属する企業活動とは無関係であり、また企業の見解を示すものではありません。
- 演者は製薬会社の社員ですが、本演題は本タスクフォースの立場で発表しています。

Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products [1] – 1



➤ Complex Innovative Designs(CID)とは

- 複雑なアダプティブ, ベイズ流, 及びその他の新規の臨床試験デザインを指すと考えられているが, 革新的または新規と考えられるものは時間とともに変化する可能性があるため, 固定された定義はない
- CIDには有効性の証拠を提供する目的で, これまでほとんどまたは一度も用いられたことのない試験デザインが含まれ, 試験の動作特性を推定するために数式よりもシミュレーションが必要なことが特徴
- 目的, 実施, 動作特性(誤った結論を導く可能性など)を含むデザインの側面, 及び試験データをどのように解析し提示するかについて, 治験依頼者とFDAとの間で明確なコミュニケーションが求められる

Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products [1] – 2

- CID提案時に含めるべき内容
 - 試験デザインの選択, 及び医薬品開発計画全体にどのように適合するかの考察
 - 試験デザインの重要な側面(アダプテーションの計画, 中間解析の詳細, 決定基準を含む)に関する詳細な説明
 - CIDに関連した統計的な考察
 - 事前情報を用いる場合は, 事前情報の情報源及び選択方法の詳細, 妥当性及び関連がある全ての事前情報が考慮されていることを確認するための手順に関する説明
 - デザインの動作特性の詳細な評価(誤った結論を導く可能性, 及び治療効果の推定の信頼性を含む。ベイズ統計学の場合は, 事前分布の選択の影響)
 - シミュレーションが実施された場合は, シミュレーションに関する報告書
 - データモニタリング委員会, またはCIDの重要な側面を担う組織の役割や指示書
 - データへのアクセスを適切に制限し, 試験の完全性を維持するための包括的な計画
 - 患者の意見を試験デザインや解析に取り入れる場合には, 患者の意見を収集するための試験計画書

Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products [1] – 3



➤ ベイズ流のCID提案時に特に含めるべき内容

■ 事前分布

□ 事前分布に関する詳細な考察

□ 事前分布の作成に用いたデータまたはその他の外部情報を詳細に示し、FDAが外部情報の情報源、及び完全性、その妥当性、並びにデータの質、及び信頼性を理解できるようにする

□ 外部データの妥当性は、交換可能性の問題に関連しており、これはベイズ流のCIDの提案において言及されるべき

■ 決定基準

□ 添付文書に含めることを意図した主要評価項目、及び副次評価項目について、治験実施計画書で決定基準(例: $\Pr(\pi_A > \pi_B) > 0.99$)の提案が必要

□ 決定基準の設定根拠

CIDプログラムに採択された事例 [2]

対象疾患	CIDプログラム応募時に提案されたデザイン
全身性エリテマトーデス	動的な情報借用を伴うベイズ流階層モデルを用いて投与群間の奏効率を推定。Response-adaptive randomization及び複数のadaptationを行う中間解析を計画
アンメットメディカルニーズのある腫瘍領域	内部対照群と外部対照群の両方を含む。主要評価項目は無増悪生存期間とし、主要解析では内部対照データを使用。全生存期間などの主たる副次評価項目の解析には、ベイズ流の動的借用法を用いた外部同時対照を含める。
小児多発性硬化症	再発率に対する非劣性試験でベイズ流の負の二項モデルによる評価を行う際に、MAP priorを用いた評価も提案。
慢性疼痛	複数の慢性疼痛を対象に、複数の介入を行うマスタープロトコル。ベイズ流混合モデル反復測定解析で、ある疼痛でのプラセボ情報の借用、同じ薬剤での異なる疼痛間での情報借用を行う。
デュシェンヌ型筋ジストロフィー	複数の評価項目に対するベイズ流モデルに基づき、試験の様々な側面に対してアダプテーションを行う。 (成果物中で紹介)

現段階でCIDパイロットプログラムを通じて承認された薬剤はない。

AMG 592 SLE Study [3]

➤ 全身性エリテマトーデスを対象とした第2相試験

Randomization	プラセボ + 標準治療	安全性 フォローアップ
	実薬低用量 + 標準治療	
	実薬中用量 + 標準治療	
	実薬高用量 + 標準治療	

➤ 試験のモチベーション

1. 用量選択
2. An adequate and well-controlled studyとされ、Phase III試験の数を減らすこと

➤ 症例数: 320例

➤ 主要エンドポイント: 52週時反応率

➤ 中間解析

■ 24週完了の患者数が規定に達した時点で計8回実施

□ 無益性中止

- ✓ 階層ベイズモデルを利用
- ✓ 実薬とプラセボとの差が、ターゲットより大きい事後確率がすべての用量で一定以下であれば試験中止

□ Response Adaptive Randomization

- ✓ 各実薬群において、52週での反応率が実薬群の中で最も高い事後確率を算出し、割付比に利用

□ Administrative Success

- ✓ すべての患者が24週に達した時点で解析を行い、将来の第III相試験が成功する予測確率が高い場合は第III相試験準備のトリガーとする

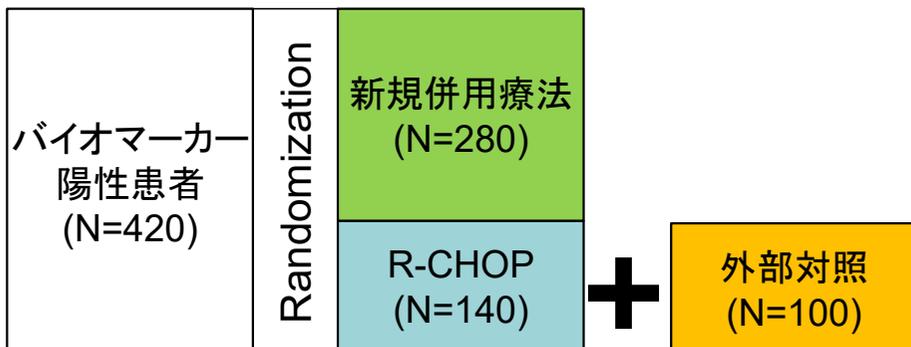
➤ 最終解析

■ すべての患者が52週完了時点で実施

- 階層ベイズモデルに基づき、有効性が用量群間で類似している場合は情報借用がされる。各用量群の反応率が、プラセボより高いことを、事後確率に基づき評価

1L DLBCLを対象とした第III相試験 [4]

- びまん性大細胞型B細胞リンパ腫(DLBCL)の一次療法を対象とした第3相試験
- CIDの目的: 全生存期間(OS)についてLabelに含めるための、試験期間の短縮とより効率的な試験実施



- 症例数: 420例
- 主要エンドポイント: 医師判定のPFS
 - 無作為化された患者のみで解析
- 重要な副次エンドポイント: 全生存期間(OS)
 - 無作為化された患者 + **外部対照**で解析

外部対照

■ 外部対照を用いるRationale

- 前もって現在進行中の社内のRCTを外部対照として用いることを計画
- 同じ適格・除外基準や類似した施設や医師の選択
- OSの臨床的意義やイベント判断の恣意性が最小限であること

■ 解析のスキーム

- 比較可能性の評価
 - ✓ 同一の適格・除外基準の適用
 - ✓ 内部対照と外部対照で大きく異なる因子の特定
- 傾向スコアマッチング
- **ベイズ流 commensurate prior**による動的な借用
 - ✓ 内部対照・外部対照の類似性によっては自動的に外部対照重みを減らし、ハイブリッドコントロールで推測を可能にする方法を選択

NEOS Study [5]

➤ 小児多発性硬化症を対象とした第3相試験

Randomization	Ofatumumab (実薬1)
	Siponimod (実薬2)
	Fingolimod (実対照薬)

- 症例数: 180例 (1:1:1割付)
- 主要エンドポイント: 年間再発率(2年間追跡)
- 主要解析: 負の二項分布を用いたFingolimodに対する非劣性

➤ 外部データ

- Fingolimodで4試験 (成人3試験、小児1試験)、Ofatumumab及びSiponimodでそれぞれ成人2試験の結果が存在
- 成人から小児への情報借用するが交換可能性が成立しない可能性も加味して、robust mixture priorを利用
- 感度解析として事前分布の重みを変えたTipping point analysisを計画

➤ 中間解析

- すべての被験者で1年間追跡終了時点でOfatumumab及びSiponimodの両方で非劣性が示された場合には、二重盲検期を1年間短縮
- 新しい治療がFingolimodと同程度の治療効果で、小児において事前分布と同程度に有効な場合には早期に試験を終了できる可能性大

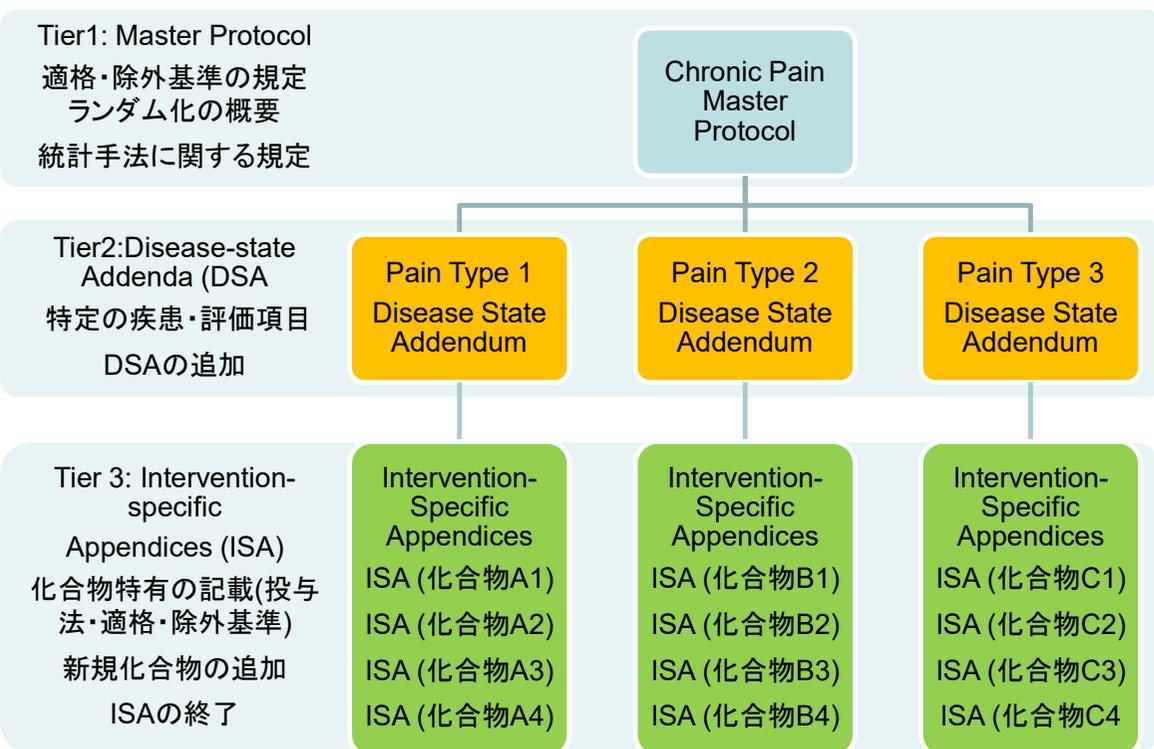
Chronic Pain Master Protocol [6]

- 慢性疼痛を対象とした第2相試験
- CIDの目的: 複数の化合物で複数の適応について効率的にシグナルを検出

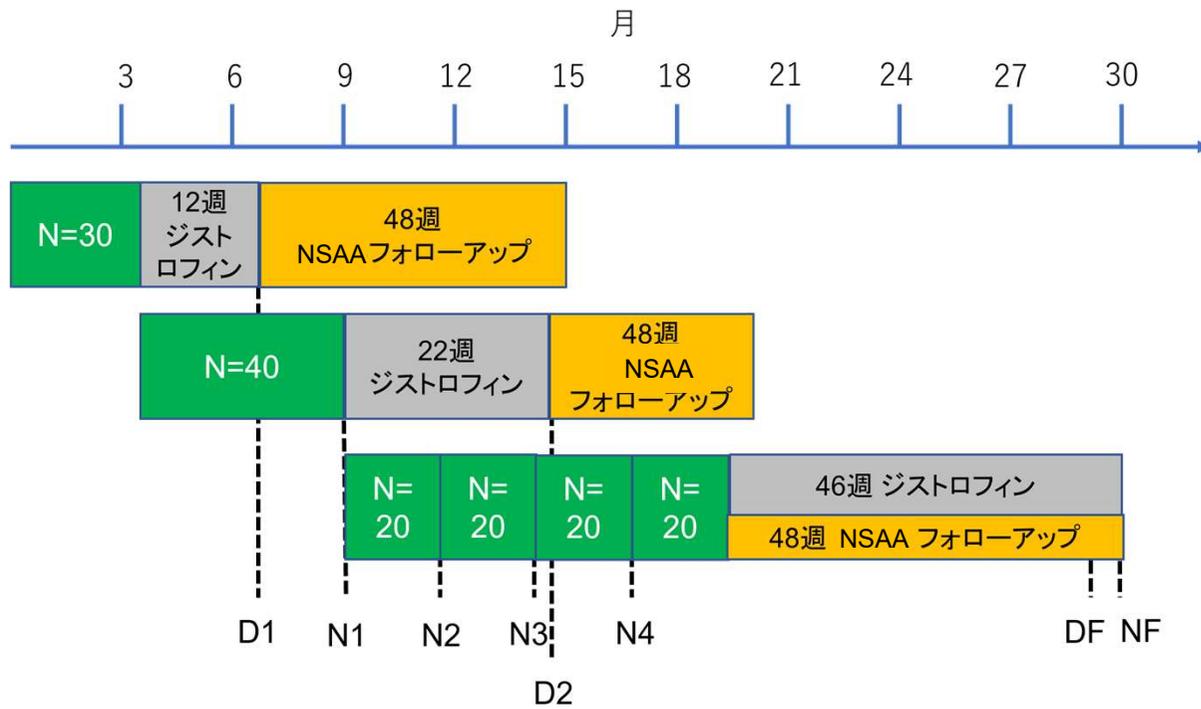
- 各ISA内で試験薬とプラセボで2:1割付
- 主要評価項目: Numerical Rating Scale
- 主要解析: ベイズ流MMRM

- 経時的な階層ベイズモデルを用い、同一の疼痛タイプを対象にしている他のISAからのプラセボ情報の借用
- 同一化合物で他の疼痛タイプでの治療効果について情報借用
- ベイズ流の成功基準を各ISAで設定(中間解析及び治療群追加に関しても利用)

- マスタープロトコルを用いる統計的利点
 - 疼痛タイプ内/外での化合物の比較
 - 標準化されたデータ収集
 - サンプルサイズの低減



DYSTANCE 51 study [7]



試験デザイン	ランダム化プラセボ対照二重盲検並行群間試験 (48週間の二重盲検期+48週間の非盲検期)
対象症例	歩行可能なデュシェンヌ型筋ジストロフィー (DMD) 患者
症例数	150例
投与群	プラセボ、suvodirsen低用量、高用量 (割付比1:1:1)
主要評価項目	<ul style="list-style-type: none"> FDA:筋生検で得られたジストロフィン変化量 (迅速承認に向けた代替エンドポイント) EMA & PMDA: 48週時のNSAA(臨床エンドポイント)

ジストロフィンの解析時点

- D1, D2: 中間解析, DF:最終解析

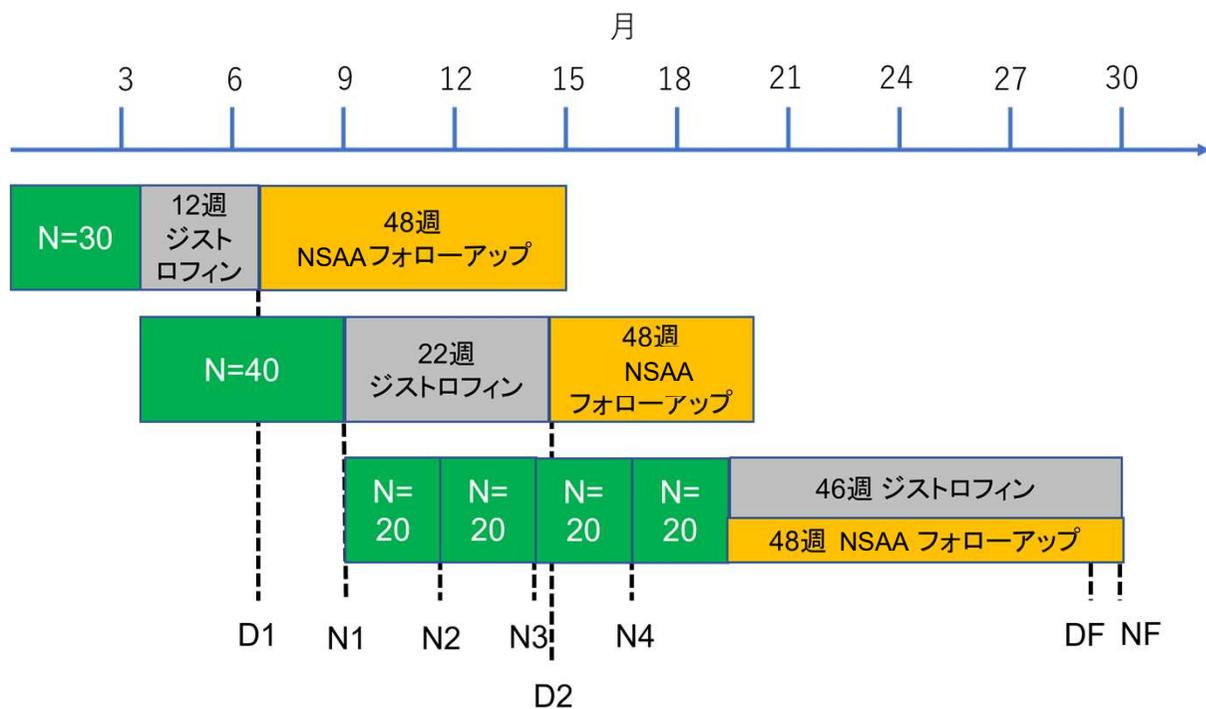
NSAAの解析時点

- N1, N2, N3, N4: 中間解析, NF:最終解析

NSAA: North star ambulatory assessment

・17個の各症状に0: Unable to achieve independently, 1: performs with modifications, 2: Normalで評価

ジストロフィンに関する解析計画 [7]



➤ 筋生検スケジュール (各患者2回)

- ベースライン (全員)
- Week 12 (最初の30例)
- Week 22 (次の40例)
- Week 46 (残りの最大80例)

➤ 解析計画

- 中間解析 (D1, D2)
 - Phase I試験の結果も加味し, Accelerate approvalを狙う
 - 低用量の中止
- 最終解析 (DF)

ジストロフィンに関する解析計画 —統計モデル・事前分布— [7]

➤ ベイズ流線形回帰モデルに基づく統計モデル

$$D_{ij} = \tau_j + \vartheta_{t(i),j} + \epsilon_{ij}$$

D_{ij} : 患者*i*の時点*j*(1: 12週, 2: 22週, 3: 46週)でのジストロフィンの変化量 τ_j : 時点*j*でのプラセボ群のジストロフィンの変化量
 $t(i)$: 患者*i*の治療群(0: プラセボ, 1: 低用量, 2: 高用量) $\vartheta_{T,j}$: 時点*j*での治療群*T*の治療効果(1: 低用量, 2: 高用量)
 $\epsilon_{ij} \sim N(0, \sigma_{Tj}^2)$: 誤差項(各時点、治療群で異なる想定)

- 事前分布: 無情報事前分布を仮定し, 分散は過去の臨床試験に基づく一様分布を設定
 - プラセボ群の各時点の変化量: $\tau_j \sim N(0, 10^2)$; $j = 1, 2, 3$
 - 各時点のSD: $\sigma_{Tj} \sim Unif(0, 10)$; $T = 0, 1, 2$ $j = 1, 2, 3$
 - 群間差: $\vartheta_{Tj} \sim N(0, 10^2)$; $T = 1, 2$ $j = 1, 2, 3$
- Phase I OLE試験のジストロフィンに関する結果の借用

ジストロフィンに関する解析計画 —中間解析と決定基準— [7]

➤ 中間解析での低用量群の中止

- 中間解析時にsuvodirsenの高用量群で効果があり、かつ低用量群での効果が高用量の半分以下である可能性が高い場合は低用量を中止。中止した場合、既に低用量群に割り付けられた患者は高用量群へ移行するが、解析には含めない

$$Pr(\vartheta_{2,j} > 0 \ \& \ \vartheta_{1,j} < 0.5\vartheta_{2,j} | d) > 0.95 \text{ for } j = 1, 2$$

■ 1回目の中間解析 (D1)

□ Week 12のデータを用いた低用量群の中止判断

✓ DYSTANCE 51試験及びPhase I OLE試験のデータを利用して判断

□ Week 22のデータを用いた低用量群の中止判断

✓ DYSTANCE 51試験ではWeek 22のデータはない状態だが、Phase I OLE試験(22週) vs DYSTANCE 51試験のプラセボ群(Week 12)で比較。プラセボ群ではWeek 12と22でジストロフィン値は同じ想定($\tau_1 = \tau_2$)。なおこの解析は低用量群の中止のみに関してで、有効性中止には用いない

■ 2回目の中間解析(D2)

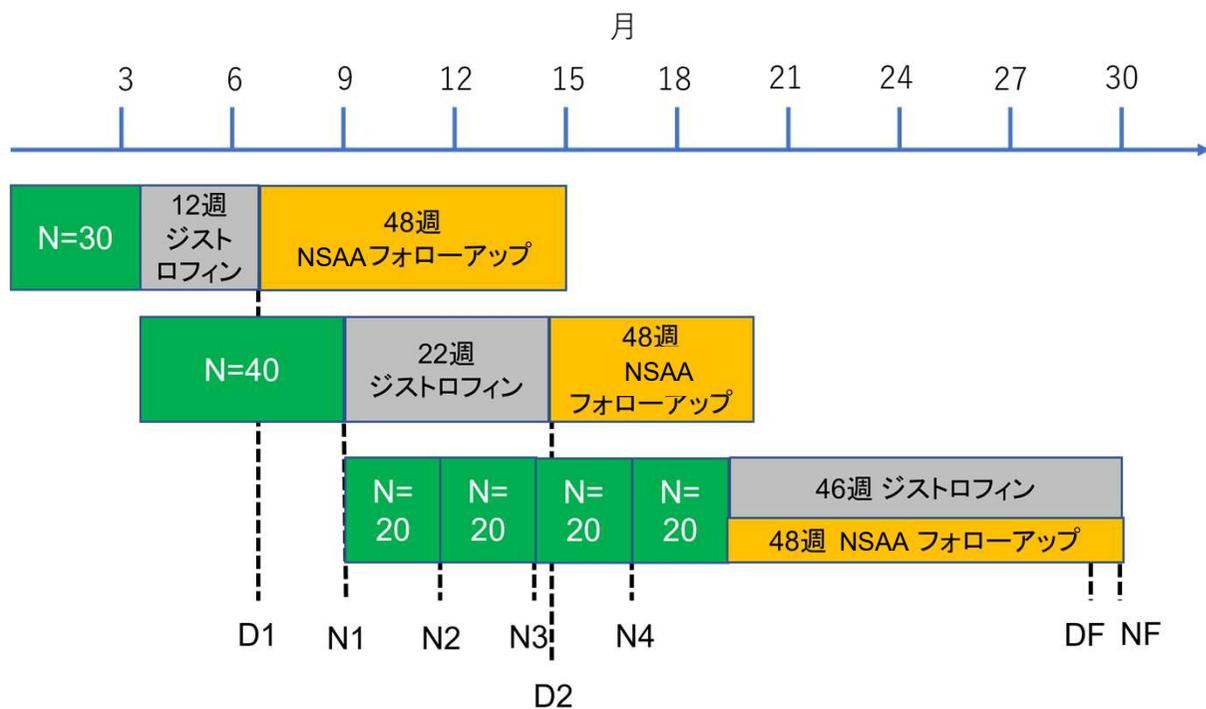
□ D1と同じくPhase I OLEデータを用い、Week 22の治療効果で低用量群の中止を判断

ジストロフィンに関する解析計画 —決定基準— [7]

➤ 治療効果の比較

- 複数用量への対応: Gatekeeping法を用い、高用量で有効性が示された時のみ、低用量での有効性を判断
- 複数解析時点への対応: Bonferroni法 (片側有意水準 $0.025/3$ 回 = 0.00833)
 - 1回目の中間解析: $Pr(\vartheta_{2,1} > 0|d) > 0.99167$
 - 2回目の中間解析: $Pr(\vartheta_{2,2} > 0|d) > 0.99167$
 - 最終解析: $Pr(\vartheta_{2,3} > 0|d) > 0.99167$

NSAAに関する解析計画 [7]



解析計画

- 中間解析 (N1-N4: 70例, 90例, 110例, 130例登録時点)
 - 事前に規定した有効性基準を満たした場合、症例登録を中止
- 最終解析 (NF)

NSAAに関する解析計画 —統計モデル— [7]

➤ ベイズ流反復測定疾患進行モデル

- 繰り返し測定値に対する疾患進行モデルを用いてプラセボと比較
- プラセボデータは後述するヒストリカルデータにベイズ流メタアナリシスを適用し補強する。データ間の異質性はベイズ流に評価し、結果が類似している場合はより多くの情報を借用
- 低用量群の中止がなかった場合は、高用量群と併合して解析
- ジストロフィンの解析とは独立

$$Y_{ij} = \gamma_i + \exp(\theta_{t(i)} + \eta_i + \alpha'X_i + \delta_{s(i)}) \sum_{k=1:j} \beta_k + \epsilon_{ij}$$

$\exp(\delta_s) \sim \text{Gamma}(1/\sigma_\delta^2, 1/\sigma_\delta^2)$ for $s = 1, \dots, S$
 $\sigma_\delta \sim \text{Unif}(0,1)$

被験者 $i: 1, \dots, n$

時点 $j: 0: \text{Baseline } 1: 12\text{週}, 2: 24\text{週}, 3: 36\text{週}, 4: 48\text{週}$

$s(i)$: 被験者 i の所属するデータソース (0: DYSTANCE 51試験, 1~S: ヒストリカルデータ)

Y_{ij} : 患者 i の時点 j での NSAA スコア

γ_i : ベースライン患者 i の変量効果

θ_T : 治療 T の治療効果

$t(i)$: 患者 i の治療群 (0: プラセボ, 1: 低用量, 2: 高用量)

η_i : 各患者での推移の違いに関する変量効果

X_i : 共変量 (C個)

$\delta_{s(i)}$: 背景情報の違いで説明できないデータソース間の変量効果

事前分布の設定については成果物参照

β_k : プラセボ群での各時点間の平均変化 ($\beta_0 = 0$)

$\epsilon_{ij} \sim N(0, \sigma_{t(i),s(i)}^2)$ 各治療群、データソースごとの標準偏差を仮定

NSAAに関する解析計画 –ヒストリカルデータ– [7]

- 以下の3試験の患者レベルのデータを利用
 - TadalafilとatalurenではCritical Path Institute's (C-Path's) Duchenne Regulatory Science Consortium (D-RSC)の枠組みも利用
- 交換可能性の過程に関して以下の4つを考慮
 - 客観的、一貫した、盲検下の評価
 - 一貫した適格・除外基準 (DYSTANCE試験の適格・除外基準も考慮)
 - 一貫した標準・背景治療
 - 対照群及び予後因子の可能性のあるベースライン因子の違いに対する調整の事前規定

データセット	例数	NSAAの ベースライン値 (SD)	NSAAのベースライン からの変化量 (SD)	残差誤差
Tadalafil DMD	90	22.5 (6.0)	-4.3 (3.7)	2.4
ACT-DMD	76	24.0 (5.8)	-3.5 (3.4)	2.2
B5161002	26	20.6 (5.9)	-4.6 (5.8)	3.2

NSAAに関する解析計画 ー中間・最終解析と決定基準ー [7]

➤ 中間解析

- 70, 90, 110, 130例登録時点で実施
- 事前にシミュレーションで規定した治療効果 DRR_N (70, 90, 110, 130例時点でそれぞれ0.50, 0.525, 0.55, 0.575)を満たす場合には, 追加の症例は登録しない
- DRR_N は中間解析時の症例数で有効性が確認できるために必要な治療効果の事後平均

$$Pr(\exp(\theta_T) < DRR_N | y) > 0.90$$

➤ 最終解析

- DRRが1未満である事後確率を評価
 - 低用量が中止していない場合には併合した集団($\theta_1 = \theta_2$ を仮定)で評価
 - 低用量が中止している場合には高用量のみ(θ_2)で評価

$$Pr(\exp(\theta_2) < 1 | y) > 0.975$$

- ・ 医薬品開発においては用量の選択を始め、様々な目的で試験が実施される。
- ・ 目的が複数に及ぶ場合、本来であれば探索的な臨床試験を別途実施したのちに、検証的な試験へと進む場合が多いが、希少疾病では臨床試験に登録可能な被験者数が限られているため、複数の試験を実施するのは難しい場合がある。
- ・ 様々な観点を組み合わせた複雑な臨床試験デザインが考えられ、効率的な医薬品開発につながる可能性がある。その場合、シミュレーションに基づく規制当局とのコミュニケーションが重要になる場合がある。

参考文献

- [1] Food and Drug Administration. "Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products- Guidance for Industry " (2020).
- [2] Price Dionne, and John Scott. "The US Food and Drug Administration's Complex Innovative Trial Design Pilot Meeting Program: Progress to date." *Clinical Trials* 18.6 (2021): 706-710.
- [3] Amy Xia, May Mo, and Tony Jiang. "Delivering Transformative Medicines to Patients by Leveraging Complex Innovative Designs AMG 592 SLE Case Study."
- Presentation http://www.bayesianscientific.org/wp-content/uploads/2022/08/CID_BSWG_AUG2022.pdf
- Youtube https://www.youtube.com/watch?v=0_fQsOXDHUw
- [4] Jiawen Zhu and Herb Pang. "Case Study of a Phase 3 Study Design with Hybrid Control in 1L DLBCL - FDA Complex Innovative Designs Pilot Program"
- Presentation <http://www.bayesianscientific.org/wp-content/uploads/2022/09/Bayesian-KOL-session.pptx.pdf>
- Youtube <https://www.youtube.com/watch?v=C9VNqcYVCvU>
- [5] Marius Thomas and Dieter A Häring. "NEOS: A Bayesian non-inferiority trial design in pediatric multiple sclerosis in the FDA Complex Innovative Designs Pilot Program"
- Presentation http://www.bayesianscientific.org/wp-content/uploads/2022/11/neos_dia_bswg_nov22.pptx
- Youtube https://www.youtube.com/watch?v=dvcSRh_WUpq
- [6] Saptarshi Chatterjee and Phebe Kemmer. "Learnings from Chronic Pain Master Protocol: FDA Complex Innovative Designs Pilot Program"
- Presentation http://www.bayesianscientific.org/wp-content/uploads/2022/10/DIA-BSWG-KOL-CPMP-21OCT2022_final.pdf
- Youtube <https://www.youtube.com/watch?v=ex5BgsBKjWo>
- [7] Lake, Stephen L., et al. "Bayesian adaptive design for clinical trials in Duchenne muscular dystrophy." *Statistics in Medicine* 40.19 (2021): 4167-4184.