

第7回 データサイエンス ラウンドテーブル会議

がん第I相試験の用量探索デザインの実装及び実務上の問題

がん第I相試験について

- 古典的ながん第I相試験では、有効性や毒性が用量依存的に増加することを仮定し、毒性が許容できる範囲内の最大の用量と定義される最大耐用量（MTD、Maximum Tolerated Dose）において最大の有効性を期待するという考え方の下で、患者における忍容性及び安全性並びに薬物動態を確認しながらMTDを推定する。

がん第I相試験の用量漸増デザイン

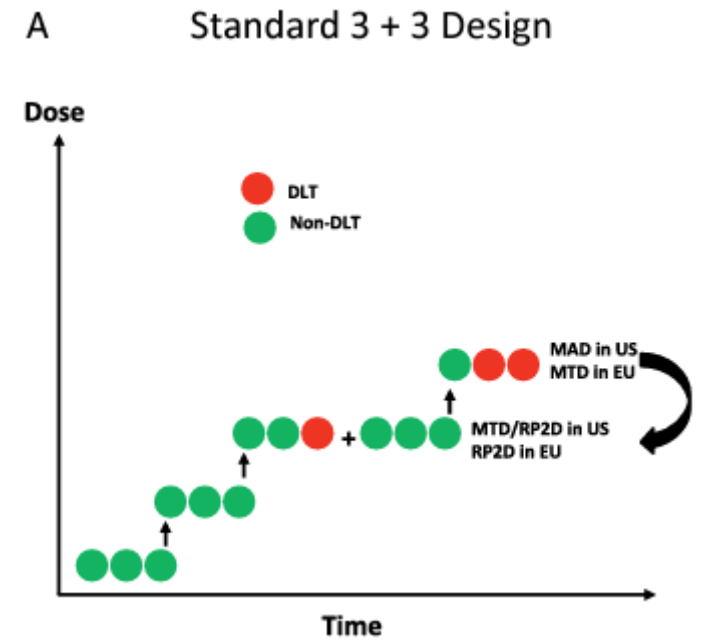
- 3+3デザインに代表されるRule-based（またはAlgorithm-based）のデザイン
- 連続再評価法（CRM、Continual Reassessment Method）等のModel-basedの試験デザイン
- Bayesian optimal interval（BOIN）デザイン等のModel-assistedデザイン

3+3デザイン

- 最初の用量レベルに被験者が3例登録され、用量選択のプロセスが以下のようにルール化されるのが典型である。
 - 最初の用量レベルの3例のうち0例がDLT発現であれば、次の用量レベルに増量し3例が登録される。
 - 3例のうち1例がDLT発現であれば、同じ用量に追加で3例が登録される。
 - 2例以上にDLTが発現した場合それ以上の増量はせず、1段階下の用量をMTDとする。

3+3デザイン

- MTD決定のプロセス



Hansen et al. (2014)

	用量レベル (低用量→高用量)				
累積登録例数	1	2	3	4	5
3	0/3 (Cohort 1)	0/3 (Cohort 2)	1/3 (Cohort 3)	2/3 (Cohort 5)	
6			0/3 (Cohort 4)		
			MTD		

CRM

CRMのイメージをつかむために、1パラメータの用量反応モデル（作業モデル）を用いてDLT発現確率を推定するための手順を示す。

$$\Pr(\text{用量水準 } j \text{ での毒性発現}) = \pi_j(\alpha) = p_j^{\exp(\alpha)}, \quad j = 1, \dots, J$$

用量に対して毒性が単調に増加することを仮定した下で、 $p_1 < \dots < p_J$ を興味の薬剤に対する各用量の事前に想定されたDLT発現確率とする。ここでは α が唯一の未知パラメータとなる。 n_j 例のうち y_j 例の被験者が用量 j でDLTを経験したと想定し、 D を観測データとしたとき α に関する尤度関数は以下のように与えられる。

$$L(D|\alpha) \propto \prod_{j=1}^J \left\{ p_j^{\exp(\alpha)} \right\}^{y_j} \left\{ 1 - p_j^{\exp(\alpha)} \right\}^{n_j - y_j}$$

$f(\alpha)$ を α の事前分布とし、例えば $\alpha \sim N(0, \sigma^2)$ とする。ベイズの定理を用いることで用量 j でのDLT発現確率は事後平均

$$\hat{\pi}_j = \int p_j^{\exp(\alpha)} \frac{L(D|\alpha)f(\alpha)}{\int L(D|\alpha)f(\alpha) d\alpha} d\alpha$$

として推定できる。

CRMデザイン

さらに、推定されたDLT発現確率 $\hat{\pi}_j$ を用い、CRMの用量探索手順の例を以下に示す。

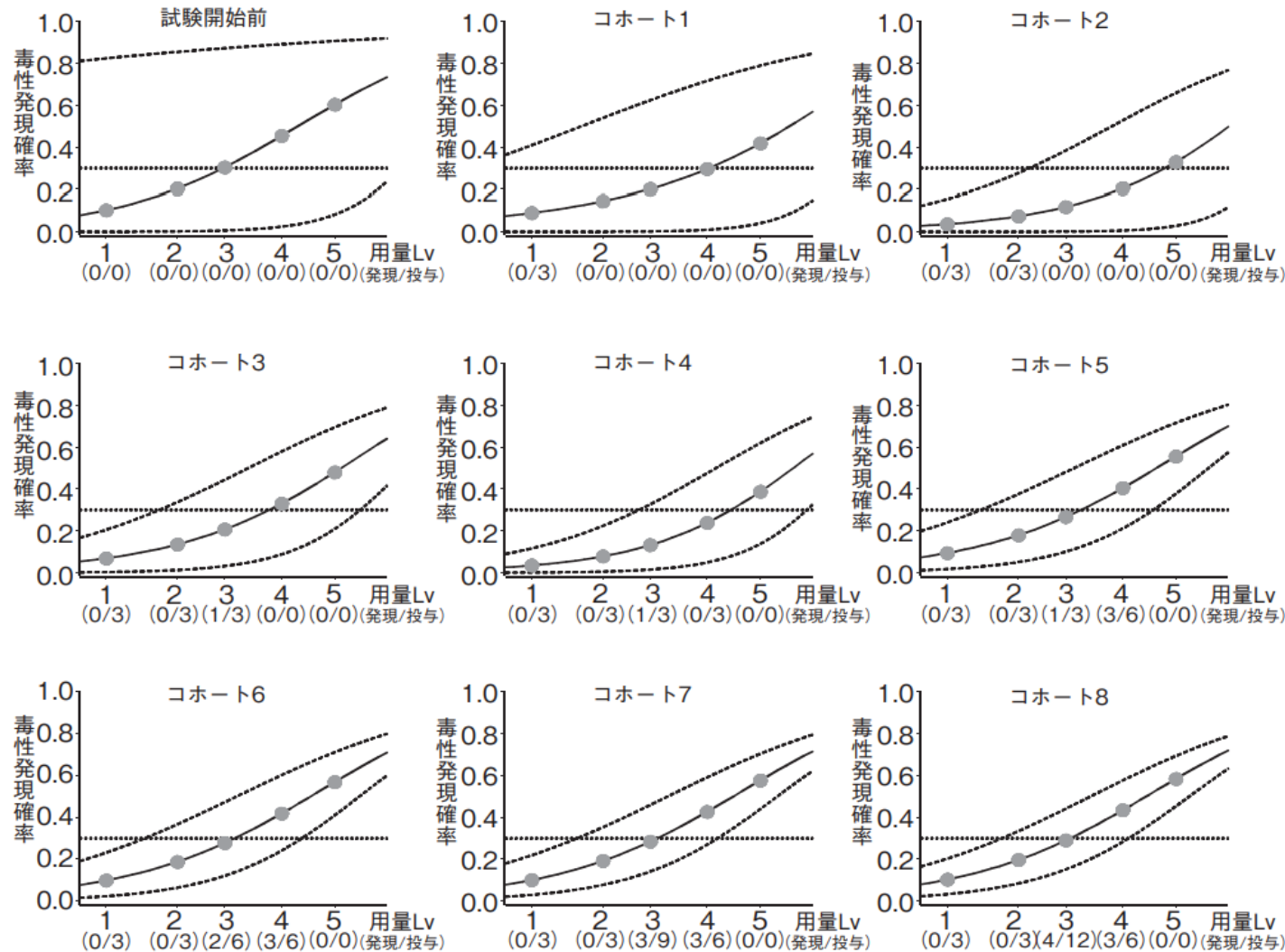
1. 最初のコホートを最低、または研究者が規定した用量で治療する。
2. 現時点での用量を j^{curr} とし、それまでに観測されたデータに基づいて推定されたDLT発現確率 $\hat{\pi}_1, \dots, \hat{\pi}_J$ を得る。目標DLT発現確率 ϕ に最も近いDLT発現確率 $\hat{\pi}_j$ をもつ用量

$$j^* = \operatorname{argmin}_{j \in \{1, \dots, J\}} |\hat{\pi}_j - \phi|$$

を探索する。以下の規則に従い、次のコホートが治療される用量を決定する。

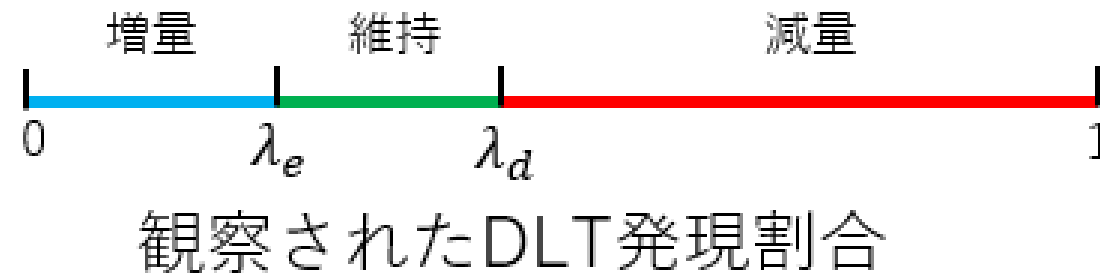
- $j^{curr} > j^*$ の場合、用量 $j^{curr} - 1$ に減量する
 - $j^{curr} < j^*$ の場合、用量 $j^{curr} + 1$ に増量する
 - その他の場合、用量は変更しない
3. 最大被験者数到達を含めた、事前規定した終了規準を満たした時点で目標DLT発現確率 ϕ に最も近い用量がMTDとして選択される。

CRMデザイン



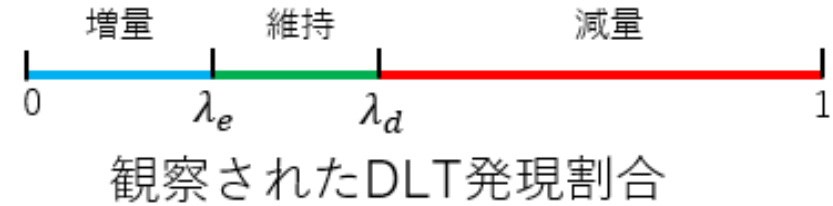
BOINデザイン

- 検討中の用量における安全性評価が完了した時点でDLT発現確率が下図のどの区間に含まれるかによって次のコホートの用量調整を行うデザインである。
- BOINデザインはFDAからfit for purposeとして指定されている。



λ_e : 減量的意思決定の閾値
 λ_d : 増量的意思決定の閾値

BOINデザイン



- 目標DLT割合0.3の場合：用量維持の区間は (0.236, 0.358)
- 観察されたDLT発現確率が
 - 0.167であった場合：次のコホートの用量は1つ上の用量
 - 0.333であった場合：次のコホートの用量は現在の用量と同じ
 - 0.500であった場合：次のコホートの用量は一つ下の用量
- DLT発現数に応じた用量調整表を事前に作成することができる。

	現在の投与量の投与被験者数 (例)									
	1	2	3	4	5	6	7	8	9	
DLT発現数が右の値以下の場合、増量	0	0	0	0	1	1	1	1	2	
DLT発現数が右の値以上の場合、減量	1	1	2	2	2	3	3	3	4	
DLT発現数が右の値以上の場合、投与中止	NA	NA	3	3	4	4	5	5	5	

議題

1. がんの第I相試験の用量探索デザインについて、どのような目的で、どのデザインをよく使用しているか。Algorithm-basedデザインの適用が多い場合、Model-basedデザインやModel-assistedデザインを使用しない理由は何か。
2. Model-basedデザインやModel-assistedデザインを利用して気づいた課題や難しさなどがありますか。利用経験がない場合は、利用にあたっての障壁や懸念などがありますか。
3. Model-assistedデザインの登場も踏まえて用量漸増デザインをどのように選択するのが適切と考えるか。
4. 規制当局と用量漸増デザインに関してどのような議論をしているか。どのような議論が重要と考えるか。
5. Model-basedデザインやModel-assistedデザインについて、例数設計や実際のモデル解析はどのソフトウェアを使用したか。モデル解析においては、解析結果のQCをどのように取るか。
6. 用量の増量/減量等の意思決定の際に、臨床担当者などの開発メンバーとどのような連携をしたか。上手くいっている点、課題に感じる点などは何か。
7. 海外で第I相用量探索試験が先行し、“後追い”の形で国内第I相用量探索試験を実施する場合、日本国外の開発担当者とのような連携をしているか。
8. FDAのProject Optimusを踏まえて、がん第I相用量探索試験の考え方の変化にどのように対応していくか。

第7回 データサイエンス ラウンドテーブル会議

がん第I相試験の用量探索デザインの実装及び実務上の問題

議論結果

3グループに分かれ、以下から議題を選択し、議論結果をまとめた

1. がんの第I相試験の用量探索デザインについて、どのような目的で、どのデザインをよく使用しているか。Algorithm-basedデザインの適用が多い場合、Model-basedデザインやModel-assistedデザインを使用しない理由は何か。
2. Model-basedデザインやModel-assistedデザインを利用して気づいた課題や難しさなどがあるか。利用経験がない場合は、利用にあたっての障壁や懸念などがあるか。
3. Model-assistedデザインの登場も踏まえて用量漸増デザインをどのように選択するのが適切と考えるか。
4. 規制当局と用量漸増デザインに関してどのような議論をしているか。どのような議論が重要と考えるか。
5. Model-basedデザインやModel-assistedデザインについて、例数設計や実際のモデル解析はどのソフトウェアを使用したか。モデル解析においては、解析結果のQCをどのように取るか。
6. 用量の増量/減量等の意思決定の際に、臨床担当者などの開発メンバーとどのような連携をしたか。上手くいっている点、課題に感じる点などは何か。
7. 海外で第I相用量探索試験が先行し、“後追い”の形で国内第I相用量探索試験を実施する場合、日本国外の開発担当者とのような連携をしているか。
8. FDAのProject Optimusを踏まえて、がん第I相用量探索試験の考え方の変化にどのように対応していくか。

Group A

①がんの第Ⅰ相試験の用量探索デザインについて、どのような目的で、どのデザインをよく使用しているか。Algorithm-basedデザインの適用が多い場合、Model-basedデザインやModel-assistedデザインを使用しない理由は何か。

- 採用される試験デザイン
 - Algorithm-basedデザインが基本。
(Model-assisted デザインの導入までは至っていない)
- Algorithm-basedデザインが採用される背景
 - 臨床開発部門に対して、方法論の導入提案に至れていない。
 - 失敗の事例があるために、挑戦ができていない。
 - 先行の海外試験ではAlgorithm-basedデザインを採用しているため...との意見により統計担当からの提案が難しい。
 - 先行の海外試験がある場合には、国内試験を簡単に試験を済ませたい。
 - 国内他社もAlgorithm-basedデザインを採用しているため、その他のデザインの検討に至らない。
- Model-basedデザインやModel-assistedデザインの採用背景
 - 経験値を有する担当者が多く、利用することの疑問が少ない。
 - 海外先行で試験を実施、その後日本で実施されることが多く、導入目的は不明確。

Model-basedデザインや
Model-assistedデザインに
チャレンジしたいけれど...

② Model-based デザインや Model-assisted デザインを利用して気づいた課題や難しさなどがありますか。利用経験がない場合は、利用にあたっての障壁や懸念などがありますか。

- どのように導入していく⇒臨床開発部門との連携が重要
 - Model-basedデザインの理解が必要だが、理解が困難なため導入が困難。
 - Model-assistedデザインなら利用可能性を受け入れてもらいやすいかもしれない。
- 導入のハードル
 - 統計担当者のリソースの問題
 - 準備のための時間が足りずに、提案に至れていない。
 - シミュレーションプログラムの作成など、時間がかかることが多い。
 - 提案できたとしても、新しいことのために試験実施中に何が起こるか分からず、踏み出しにくい。
 - 組織体制の問題
 - 海外本社から試験デザインが下りてくるため、提案できる状況ではない。
 - チーム、開発チームへの説明
 - 訴求ポイント：**MTDの推定精度**など
 - ただし、MTDの推定精度が高いことは臨床開発メンバーが利益を感じにくい部分かもしれない。
 - 臨床開発メンバーが負担に思わないレベルまでの統計担当者が準備することが必要。
 - 何から始めたらいいのかわかりにくい
 - 方法論や利点・欠点も理解はしているが、何から着手すると解決につながられるか。
- Model-based・Model-assistedデザインの利用時の課題
 - 臨床開発チームへの説明、シナリオ設定、30日調査への対応。

まとめ

- 気づいたこと
 - 国内企業（開発）においてはModel-based/Model-assistedデザインの導入に至れていない状況が多々ある。
 - いくつかの課題が考えられるが、統計担当者のリソースの問題が一番の原因ではないか。
 - 統計的側面の説明・解決だけでは、利用実現には結びつかない。
 - 外資系企業と内資系企業の差が原因である可能性も示唆された。
- 導入につなげるには
 - 国内企業での経験蓄積
 - 臨床開発メンバーに寄り添った提案
- 他グループへの質問
 - 日本FIH試験でどのようにModel-based/Model-assisted デザインの利用につなげているのか？

Group B

④規制当局と用量漸増デザインに関してどのような議論をしているか。どのような議論が重要と考えるか。

- PMDAから30日調査照会事項のチェックリストが用意されている。

<https://www.pmda.go.jp/files/000252155.pdf>

- 3+3 デザイン以外の用量漸増デザインを用いる場合（3+3 デザインと他の用量漸増デザインを組み合わせる場合も含む）には、用量漸増デザインの動作特性がシミュレーションにより検討され、当該検討結果に基づき、最大耐量（以下、「MTD」）を上回る用量で治験薬が投与される被験者が最小限となることを含め、用量漸増デザインの適切性が説明されている。なお、説明内容に以下の点が含まれることを確認すること。
 - MTD の定義、忍容性評価を終了する条件（想定する最大症例数）等。
 - 最も低い用量レベルで真の DLT 発現確率が過毒性と定義される確率をわずかに超える場合に、MTD としてどの用量が選択されるか（各用量の選択割合）、各用量における平均被験者数及び試験中止確率について、シミュレーションに基づき検討した結果。

- 安全性の観点が必要になることが想定される。例えば、
 - MTDより高い用量に曝露してしまう被験者の割合
 - MTDより高い用量を誤ってMTDとして選択してしまう割合
 - 最も低い用量レベルで真のDLT発現確率が過毒性と定義される確率をわずかに超える場合に、試験が早期中止する割合

④規制当局と用量漸増デザインに関してどのような議論をしているか。どのような議論が重要と考えるか。

- 試験の早期中止割合に試験横断的な基準はない。
 - シミュレーション実験を通じた各種動作特性に基づいて試験デザインの妥当性を評価すべき。
 - 3+3と同じぐらいの中止割合があればよいというわけではない。
- 3+3デザインではシミュレーション実験を求められていないのはなぜか。
 - 歴史的な背景・経験を踏まえて安全性が高いと考えられる。
- 標的毒性発現確率について
 - PMDA: 照会を受けた経験が少ない（30日調査照会事項のチェックリストに記載されている項目を含めた動作特性を総合的に評価していると予想される）。
 - FDA: 照会を受けた経験があった。

④ 規制当局と用量漸増デザインに関してどのような議論をしているか。どのような議論が重要と考えるか。

- Project Optimusは最適な用量を決定するためのFDA主体のイニシアチブ
- Project Optimusに関連した議論が行われているのか？
 - 規制当局と明示的に議論はされていないものの、社内での検討は一部行われている。
 - 企業としてPhase 2以降の試験での用量の妥当性を説明できるような試験デザインである必要がある。
 - ランダム化試験での並行群間比較の他にも、拡大コホートやBackfillの活用が考えられる。

⑥用量の増量/減量等の意思決定の際に、臨床担当者などの開発メンバーとどのような連携をしたか。上手くいっている点、課題に感じる点などは何か。

- Model-assistedデザインを採用した時に感じた課題・工夫したこと
 - 会社上層部の理解を得るのが難しいため、根気強い説明が必要であった。
 - 上層部全体に話をもっていく前に、一部上層部へ試験デザインを説明して理解者を増やしておいた。
 - Model-assistedデザインの用量推移は比較的容易に描けるため、用量漸増のイメージを複数パターン用意して、視覚的に見せる。
 - シミュレーション結果の解釈を丁寧に説明する。
 - 用量増減表を理解してもらうために時間がかかった。
 - 3+3デザインと比べて被験者数の増加が課題となることがあるため、シミュレーションパラメータを調整したり、コホートサイズを小さくしたりする。
 - オペレーションの観点
 - Model-assistedデザインを採用できた会社では、統計担当者と距離が近かったため、理解は得られやすかった。
 - 試験期間の長さは気にしていた。
 - 実施医療機関側にBOINを知っている先生がおられたため受け入れられた。

Group C

①がんの第I相試験の用量探索デザインについて、どのような目的で、どのデザインをよく使用しているか。Algorithm-based デザインの適用が多い場合、Model-based デザインや Model-assisted デザインを使用しない理由は何か。

- Model-assistedデザイン (BOIN, mTPI2)
 - Model-assistedデザインは統計担当者以外のメンバーが理解・実施しやすい。
 - グローバルが主導してデザインしているのか？→日米欧チームで決定している。
 - BOINを選んだ理由→当局から受け入れられている実務実績があるため。(Fit for purpose、IND)
 - TITE-BOINも検討している。最近ではDose optimizationを受けてBackfillを含めて評価している。
 - Backfillに投与する用量の決定ルール→漸増の用量より低用量に組み入れている。
 - 用量漸増コホートの組み入れが止まっているときにBackfillに組み入れている。
- Model-basedデザイン (CRM, BLRM)
 - 各用量の毒性情報を利用でき、用量反応関係を明示できるため。
 - 社内でマニュアル化されているため。
- 3+3 デザインは使用していない。
 - 用量数が多いとき偶発的なDLT発現に引っ張られ、3+3デザインでは保守的なMTDとなるため。
 - Model-based/assistedデザインの方が性能が良い (MTD特定確率が高い) ため。海外でも認知されている。
 - Model-based/assistedデザインが社内標準となっている (マニュアル整備など)。
 - 3+3以外のデザインを導入するためには？→Model-based/assistedデザインの利点を啓発する。3+3デザインの増減ルールに論理的根拠はない。Model-assistedデザインには科学的根拠があり、性能の良さも示されている。
 - 3例コホートに縛られてしまうため。海外試験では3例を超過してしまうケースもある。Model-based/assistedデザインではコホートサイズを柔軟に運用できる。
- Model-based/assistedデザインを使用しない理由：なし

②Model-based デザインや Model-assisted デザインを利用して気づいた課題や難しさなどがありますか。利用経験がない場合は、利用にあたっての障壁や懸念などがありますか。

- Model-assistedデザイン

- 3+3デザインと比較して性能が高いが試験終了までに時間を要する（最大30-40例など）。
 - Dose escalationを止める条件を使用すればある程度改善される。毒性評価をCommitteeにかけて評価するため時間を要する。
- 統計担当者以外の関係者から理解を得られているか？
 - 事前に用量決定ルールを表にできるため理解を得やすい。

- Model-basedデザイン

- 事前情報、BLRMならハイパーパラメータをどう設定するか？
 - 無情報事前分布を使用せざるを得ない。海外P1試験実施後の国内P1試験を計画するときには事前情報の議論となる。（海外情報を半分借りる？国内でAEが出やすいこともある）
 - 海外でP1試験実施後の国内P1について、用量の数は海外と同じか？（同じならBLRM、少ないなら3+3？）→一貫性を重視して海外P1と同じ手法を使う。用量数はプロジェクト次第であるが、海外の後追いであれば用量数を減らすケースが多い。1用量の場合もあるがBLRMは使用しない。DLTが発現したら用量追加しModel-basedデザインに戻す。
- 統計担当者以外の関係者が納得できるのか？
 - 実際、理解を得ているとは言い難いが、アウトプットは理解されている。
 - 仮想シナリオと次用量をいくつか示して、Protocolに記載している。現用量の推定DLT確率と次用量を示しているところもある。試験の進捗に合わせて、現状の被験者さんの結果別にシミュレーション結果を提示している。試験中は実際のデータを反映させた結果別に次用量を示して現場で運用しやすくしている。
 - がんP1試験における被験者さんの安全性担保が重要である。複雑なデザインであっても過量投与リスクを評価することが望ましい。最低用量がMTD超だったときに試験を止められるかを評価する。
 - BLRMの決定ルールに沿わない柔軟な用量決定をすることがある。
→想定する状況別のシミュレーション結果を求めることがある。

- 共通

- 過毒性シナリオ（Dose1>MTD）における動作特性がPMDAから問われる。

③Model-assisted デザインの登場も踏まえて用量漸増デザインをどのように選択するのが適切と考えるか。

• Model-assistedデザイン

- Algorithm-basedデザインのように運用できるため社内外の理解を得やすい。
- Model-basedデザインがマニュアル化されていない会社ではModel-assistedデザインを適用しやすい。
- GlobalでModel-basedデザインを使用しているが、用量数が少ないとき、MTDと一つ下の2用量を用いるときModel-assistedデザインを使用している。海外P1試験に日本人コホートを追加するときも適用している。
 - 日本人コホートの用量をどのように決めるか？→DLT発現した用量より下の用量から開始し、1用量ずつ評価している。（Dose 5でDLT発現したとき、Dose 3-4あたりから1用量ずつ日本人投与など）

• Model-based デザイン

- 併用に拡張したいとき、単剤情報も利用できるのでBLRMが第一選択となる。併用に使用できるBOIN拡張版もあるが使用していない。
- 用量判断は、漸増とBackfillを合わせて次用量を決めたいという要望がありModel-basedデザインを使用した経緯がある。
 - Backfillで毒性が出てしまったとき、Model-assistedで判断が難しいケースがあるが、Model-basedだと全被験者の情報を使用して判断できる。
- Model-basedデザインの導入を確立している会社ではModel-assistedデザインは選択肢とならない？→選択肢となる。
 - P1試験に日本が後追いとなるため議論にならない。（統計担当の好み?）

• 共通

- 標的療法など、毒性が出ないことが想定される薬剤のときもModel-assisted/basedデザインを適用している。
 - 用量反応曲線を確認した後に最適用量（有効性）を見るためにBackfillで低用量の投与が行われる。
- モデルで推奨されている用量を選ぶかどうか臨床担当者と議論している。必ずしもモデルによる推奨用量が選択されるわけではなく、統計担当者と臨床担当者の議論が重要となる。
 - 管理が難しい毒性のときに議論となる。
 - 事前に規定したモデルの性能を維持するためにモデル優先とすることはあるか？→Protocolでモデル重視など規定はせず、柔軟に運用したい。
- 実装するとき、複雑なシミュレーションプログラムが必要となる。