



近年のがん第 I 相試験デザインとその選択

日本製薬工業協会

医薬品評価委員会 データサイエンス部会

2022 年度 タスクフォース 3

Ver 1.0

2023 年 6 月

目次

1.	はじめに	4
2.	主な用量漸増試験デザインの概説.....	7
2.1.	Rule-based (Algorithm-based) デザイン.....	7
2.1.1.	一般論	7
2.1.2.	単剤療法の開発.....	7
2.1.2.1.	有害事象の発現のみを指標としたデザイン.....	7
2.1.2.1.1.	3+3 デザイン	7
2.1.2.1.2.	Modified 3+3 デザイン	9
2.1.2.2.	有効性やバイオマーカーを同時考慮したデザイン.....	18
2.1.3.	併用療法の開発.....	19
2.1.3.1.	標準療法へ上乗せる併用療法.....	19
2.1.3.2.	新規療法を組合せる併用療法.....	20
2.2.	Model-based デザイン.....	23
2.2.1.	一般論 ^[14, 43]	23
2.2.2.	単剤療法の開発.....	23
2.2.2.1.	有害事象の発現のみを指標としたデザイン.....	23
2.2.2.1.1.	連続再評価法 (CRM, Continual Reassessment Method) ^[43]	23
2.2.2.1.2.	ベイズ流ロジスティック回帰モデル ^[55]	26
2.2.2.1.3.	イベント発生までの時間を考慮した CRM.....	29
2.2.2.2.	有効性やバイオマーカーを同時考慮したデザイン.....	30
2.2.2.2.1.	有効性情報を考慮に入れたデザイン (EffTox) ^[14, 68, 69]	30
2.2.2.3.	複数薬剤併用の用量漸増デザイン.....	32
2.2.2.3.1.	BLRM の併用療法への拡張	32
2.3.	Model-assisted デザイン	36
2.3.1.	一般論 ^[74]	36
2.3.2.	単剤療法の開発.....	36
2.3.2.1.	有害事象の発現のみを指標としたデザイン.....	36
2.3.2.1.1.	mTPI デザイン (mTPI, modified toxicity probability interval) ^[5, 75]	36
2.3.2.1.2.	キーボードデザイン (Keyboard) ^[75,77]	41
2.3.2.1.3.	ベイズ最適区間デザイン (BOIN, Bayesian optimal interval design ^[79]) ..	46
2.3.2.1.4.	Time-to-event BOIN	48
2.3.2.2.	有効性やバイオマーカーを同時考慮したデザイン.....	51
2.3.2.2.1.	BOIN-ET	51
2.3.2.2.2.	BOIN phase I/II (BOIN12).....	52

2.3.2.3.	その他のデザイン.....	55
2.3.2.3.1.	併用療法の開発.....	55
3.	適切なデザインの選択.....	59
3.1.	3つのクラスのデザインの特徴の整理.....	59
3.2.	フローチャート.....	61
3.2.1.	単剤療法かつ DLT に基づいて用量漸増する場合	63
3.2.2.	単剤療法かつ DLT に基づいて遅発毒性/登録速度を考慮して用量漸増する場合	64
3.2.3.	単剤療法かつ DLT と有効性または BM を同時考慮して用量漸増する場合	65
3.2.4.	併用療法かつ DLT に基づいて用量漸増する場合	65
3.2.5.	併用療法かつ DLT と有効性または BM を同時考慮して用量漸増する場合	66
4.	用量漸増デザイン選択に関わるその他の論点.....	67
4.1.	初回投与量、最高投与量、増量幅等の設定の一般的考え方.....	67
4.1.1.	初回投与量の設定.....	67
4.1.2.	増量計画と最高投与量の設定.....	68
4.2.	サンプルサイズ設定の考え方.....	69
4.3.	ソフトウェア.....	69
4.4.	総合判断という考え方.....	70
4.4.1.	次コホートの推奨用量の決定.....	70
4.4.2.	RP2D の決定.....	70
4.5.	動作特性の評価.....	71
4.5.1.	概要	71
4.5.2.	動作特性評価のためのシナリオ及び評価指標.....	72
4.6.	後追いがん第 I 相試験デザインの考え方	73
4.6.1.	用量漸増デザインの選択.....	74
4.6.2.	開始用量及び検討用量の設定.....	75
4.6.3.	被験者数・コホートサイズの設定及び忍容性確認の基準.....	76
4.7.	DLT が発現しないことが想定される場合のデザイン選択	76
5.	おわりに	78
6.	参考文献	79

1. はじめに

抗悪性腫瘍薬は、主に忍容性及び安全性並びに薬物動態を評価し用法・用量を検討・決定する第 I 相試験、有効性及び安全性を探索的に評価する第 II 相試験、そして従来の標準的治療と比較し主たる仮説を検証する第 III 相試験と大きく 3 段階に分けて段階的に開発することが一般的である^[1]。この一般的な流れは古典的な殺細胞性の抗悪性腫瘍薬の開発をモデルに確立されてきたものであり、ICH E8「臨床試験の一般指針」に定義される一般的な薬剤開発とはやや異なる開発モデルとなっている^[2]。即ち、抗悪性腫瘍薬以外の医薬品開発では、健康成人を対象に初期安全性や忍容性、薬物動態を確認する目的で第 I 相試験を実施し、続く第 II 相試験で興味のある疾患を対象に用法用量を探索するという流れをとるが、抗悪性腫瘍薬においてはがん患者を対象に第 I 相試験を実施し、当該試験において同時に用法用量を決定する。これは殺細胞性の抗悪性腫瘍薬では毒性の発現が不可避であり、健康成人を対象にした試験は倫理的な観点から許容されないこと、高い用量ほど毒性が強くなるとともに強い抗腫瘍効果が期待されること、これに関連して殺細胞性の抗悪性腫瘍薬の治療域が狭く、患者に投与可能な用量が毒性によって制限されてしまうという特徴に基づいている。こうした背景から、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。さらに、MTD を踏まえて第 II 相試験以降の推奨用量 (RP2D、Recommended Phase 2 Dose) を決定する。忍容性を検討する主たる評価項目として用量制限毒性 (DLT、Dose Limiting Toxicity) を事前に治験実施計画書で定義し、21 日～28 日程度の比較的短期間 (一般的な静脈注射の抗悪性腫瘍薬の 1 サイクル分の期間に相当) で定義される期間中に発現する DLT の頻度に基づいて MTD が推定される。

がん第 I 相試験の用量漸増デザインとしては、3+3 デザインに代表される Rule-based (または Algorithm-based) のデザインが採用されており、近年においても頻用されている^[3]。Rule-based のデザインに対して、より MTD の推定の正確性を高めること、MTD 付近の用量で治療される患者の割合を高めることなどが期待できる連続再評価法 (CRM、Continual Reassessment Method) 等の Model-based の試験デザインが提案されている^[4]。このクラスのデザインでは、DLT の発現確率と用量の関連に用量反応モデルを仮定し、それまでに得られているすべての用量で治療された被験者の用量と DLT の有無の結果を踏まえ、更新したモデルに基づいて次の被験者の用量や MTD を推定するのが一般的である。Model-based デザインは、MTD の推定の正確さをはじめとした動作特性において、Rule-based の試験より優れていることが報告されてきた。それでも Model-based の試験デザインの利用が進まない理由として、試験の計画や実施が Rule-based のデザインに比べて煩雑かつより多くの検討を要することや、統計担当者の強い関与の必要性などが挙げられている^[3,5]。このような実態を踏まえて、近年では Model-assisted デザインという新しいクラスのデザインが活発

に提案されてきている^[6]。このクラスのデザインでは、Model-based デザインのように DLT の発現確率と用量の関連に明示的な用量反応モデルを仮定しないものの、DLT の発現の有無に確率モデルを仮定し、実際に得られた各用量の DLT の有無の結果に基づいて、MTD 付近の用量により多くの患者が割り当てられるように統計的に最適化された用量調整ルールを用いる。このルールは、現在の用量での観察された DLT の発現数に応じて、用量の増減を示すことができ、事前に治験実施計画書に記載することができる。このルールに従い、各用量で治療された被験者の DLT の有無のすべての結果を利用して MTD を推定することで、Rule-based のように用量の増減の意思決定を事前にルール化しつつも、Model-based デザインのように MTD 付近で治療される被験者の増加やより正確な MTD の推定を可能としている。

一方、近年では技術革新とがんのバイオロジーの理解の進歩により、分子標的治療や免疫療法といった新しい作用機序の抗悪性腫瘍薬が台頭してきており、今後も遺伝子治療や再生医療など、がん治療のモダリティはさらに多様化していくことが想定される。これらの新しいクラスの薬剤は、殺細胞性の抗悪性腫瘍薬とは作用メカニズムが根本的に異なっており、一般に治療域がより広くかつ用量依存的な有効性の増加も保証されず、作用機序によっては DLT さえ想定されないことがあり得る。このような薬剤クラスのがん第 I 相試験においては、薬剤特性を踏まえて何を確認する試験とするのかを定義し、それにあった用量漸増デザインを選択することが肝要である。実際、これらのクラスの薬剤開発にあたっては第 I 相試験で拡大コホートを設定し、有効性及び安全性を探索的に評価して第 II 相試験を省略する場合や、第 II 相試験と第 III 相試験を一つの試験として実施する場合もある^[1]。少なくとも用量探索という側面においては、DLT を指標にした MTD の探索はもはや科学的な合理性を持たなくなるため、忍容性及び安全性を確認しつつ、生物学的に最適な用量（OBD、Optimal Biological Dose）を探索することが目的とされるべきであろう^[7]。

さらに、米国食品医薬品局（FDA、Food and Drug Administration）Oncology Center of Excellence は、2021 年からがん領域の医薬品開発における用量最適化（dose optimization）と用量探索のパラダイムを改革するイニシアチブ「Project Optimus」を開始し、2023 年にドラフトガイダンスが発出された^[8,9]。また、当該イニシアチブは有効性の最大化だけでなく忍容性及び安全性の観点でも最適化した用量選択を行うことを目的に産学官、専門団体と患者との協働や教育を推進することを目標としている。当該イニシアチブ設立の背景には、上述のとおり、分子標的薬のように新しいクラスの抗悪性腫瘍薬においては用量依存的に有効性が高くなるとも限らず、かつ短期間で定義される DLT では捉えられない遅発性の DLT の発現の可能性などから MTD に基づく用量設定が適切でない可能性があること、実際に FDA において承認を得た後に忍容性及び安全性上の理由で承認用量の変更を要する事案、または至適用量の再検討を求める事案が複数出てきていることなどが挙げられる^[10]。当該イニシアチブでは、分子標的薬や生物学的製剤に対して、MTD に基づく用量選択の考え方を再検討することを提言している。具体的には、がん第 I 相試験では、曝露量

やPD マーカーに基づいて検討を進めるべき2つ以上の用量（例えば用量の範囲の最小用量と最大用量）を選択し、選択した複数用量を用いて小規模なランダム化比較試験を実施し、得られた早期の安全性、有効性及び曝露反応データ等に基づき、検証的試験のための最適な用量を選択することを提案している。また、低用量では高用量に対して有効性が劣るというマインドセットから脱却するためのステークホルダーの教育の重要性についても説いている。このような新しいパラダイムへの移行を想定したとしても、既存の用量漸増デザインの特性を正しく理解し、適切に利用することで、後続の開発における最適な用量選択に資する情報が得られることが期待される。このような観点で、今後も適切な用量漸増デザインを選択することは大きな課題の一つになると考える。

本報告書では、上述のようにがん治療のモダリティが多様化する時代において、がん第I相試験の用量漸増デザインを適切に選択するための指針を与えることを目的としている。用量漸増デザインの選択にあたっては機能横断的な検討が必要になることから、本報告書の読者としては統計担当者だけでなく臨床担当者やオペレーション担当者、臨床薬理担当者など統計担当者でない関係者も想定している。このため、デザインの説明においては、機能横断的な議論に資する要点や特性に焦点をあてた概要として紹介し、詳細は特記しないかぎり大門らの教科書や総説、及びオリジナルの報告に譲る形をとった^[11,12,13]。本報告書では、2章でこれまでに提案されている主な用量漸増デザインを概説し、3章で想定される用量漸増デザインを大別し、大別したカテゴリごとに推奨される用量漸増デザインやその選択の考え方を説明する。4章では用量漸増デザインの選択時に考慮・検討すべきその他の話題について触れる。統計担当者以外の読者で、Rule-based/Model-based/Model-assisted デザインの概観を把握している（詳細は知らなくてもそれぞれのクラスの方法の一般的性質のイメージが持てる）読者は、3章や4章から読み始めて、必要に応じて2章に戻るといった読み方をすることも可能である。

本報告書が、統計担当者を含めた機能横断的ながん第I相試験の用量漸増デザインの検討に寄与し、本邦並びに世界のがん治療の開発の一助となれば幸いである。

2. 主な用量漸増試験デザインの概説

2.1. Rule-based (Algorithm-based) デザイン

2.1.1. 一般論

伝統的ながん第 I 相試験は、より高用量でより効果が強くなると同時に、より毒性も強くなるという仮定が置かれる殺細胞性薬剤の評価を目的に行われてきた。そのため、MTD の決定が目的となり、DLT 発現の有無をエンドポイントとして設定し、「3+3 デザイン」など Rule-based デザインを用いることが典型であった^[14]。DLT は、有害事象共通用語規準 (CTCAE、Common Terminology Criteria for Adverse Events) に基づく Grade3 (G3) 以上の非血液学的毒性や、Grade4 (G4) 以上の血液学的毒性などの有害事象として定義され、治験実施計画書などに事前規定される。

Rule-based デザインでは、複数名の被験者で構成されるコホート単位でこれから検討する用量に被験者を割り当て、DLT 評価期間中の DLT の有無を評価する。実際に得られた DLT の有無の結果と予め定められたルールに基づいて、次のコホートに割り当てる用量を現時点の用量から増量するのか、減量するのか、あるいは現状の用量を維持するのかを評価する。最終的な MTD も事前に決めたルールに基づいて判断するのが一般的である。このようなルールを治験実施計画書などに事前規定し、ルールに従って試験を実施することが特徴である。なお、これらのルールは一般に統計的な根拠に基づくものではない。

Rule-based デザインは殺細胞性の抗悪性腫瘍薬の開発において歴史的に多用されてきたが、Daniel ら (2021) ^[15]によると、2014 年から 2019 年までに公表された免疫治療または分子標的薬を対象とするがん第 I 相試験においても Rule-based デザインは 92% (788/856 試験) で使用されていた。

本項では、「3+3 デザイン」について概説した後に、数多く提案されている「modified 3+3 デザイン」と呼ばれる「3+3 デザイン」のルールを変更したものの中から、臨床試験での実例が存在するデザインをいくつか概説する。また、独自のルールを実装した臨床試験の実例もあるので、これらについて概説する。

2.1.2. 単剤療法の開発

2.1.2.1. 有害事象の発現のみを指標としたデザイン

2.1.2.1.1. 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。なお、コホートサイズ 6 とした場合も以下に示した^[11, 14]。

- 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。

- 3例のうち2または3例が DLT 発現であれば、その用量は MTD を上回ったと判断し、増量はそこで中止される。

6例に投与された場合、以下のルールで用量選択が行われる。

- 6例のうち1例以下が DLT 発現であれば、次の用量レベルに増量する。
- 6例のうち2例以上が DLT 発現であれば、その用量は MTD を上回ったとして増量はそこで中止される。

もしある用量が MTD を上回ったとされた場合、その一段階下の用量レベルで既に6例の被験者が投与されていれば、その用量が MTD であると決定される。一段階下の用量レベルでまだ3例しか投与されていなければ、一段階下の用量レベルに追加で3例を登録し、6例のうち DLT が1例以下の発現であれば、その用量が MTD とされるが、もし、DLT が2例以上に発現すれば、MTD はもう一段階下の用量レベルに下がる。このプロセスを、MTD が決定するまで繰り返す。最高用量に到達しても MTD が決定されなかった場合は、検討した用量レベルにおいては MTD に到達しなかったと判断される。

5段階の用量レベルを検討する状況下で MTD 決定のプロセスを表1に例示した。最初のコホートに3例が登録され、用量レベル1で治療され DLT 評価可能であった3例中 DLT が0例であったため、次のコホートは用量レベル2で治療された。このコホートで3例中1例に DLT が発現したため、次のコホートは用量レベル2に3例が追加され DLT が0例であったため、次のコホートは用量レベル3で治療された。用量レベル3でも合計6例が追加され DLT が1例であったため、次のコホートは用量レベル4で治療された。用量レベル4では、3例が追加され DLT が2例であったため、用量レベル3が MTD と決定された。

表 1 MTD 決定のプロセスの例示

	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	1/3	1/3	2/3	
6		0/3	0/3		
			MTD		

アルゴリズム化すると以下のようなになる。なお、k 番目の用量を d_k とし、初回用量は d_1 ($k=1$) と表すこととする。

- Step1: 3例の被験者を k 番目の用量 d_k で評価する。
- Step1-1: 3例の投与例のうち、DLT 発現が0ならば、 d_{k+1} に増量し Step1 に進む。
- Step1-2: 3例の投与例のうち、DLT 発現が1であれば、Step2 に進む。
- Step1-3: 3例の投与例のうち、DLT 発現が2以上であれば、Step3 に進む。
- Step 2: 3例追加し、6例の被験者を k 番目の用量 d_k で評価する。
- Step2-1: 6例の投与例のうち、DLT 発現が1ならば、 d_{k+1} に増量し Step1 に進む。
- Step2-2: 6例の投与例のうち、DLT 発現が2以上であれば、Step3 に進む。

- Step3: 用量増量の中止。
- Step3-1: 一段階下の用量レベルでまだ 3 例しか投与されていないならば、追加で 3 例を登録し、6 例のうち DLT が 1 例以下の発現であれば、その用量が MTD とされる

Bendell ら (2017) ^[16]の進行固形癌に対する Binimetinib の第 I 相試験の用量漸増パートでは、30 mg BID (N=4)、45 mg BID (N=4)、60 mg BID (N=7)、80 mg BID (N=4) の合計 19 例が Binimetinib の単剤投与を受けた。30 mg BID~60 mg BID はいずれも DLT が認められなかったが、80 mg BID で DLT 評価可能例 3 例のうち 2 例で DLT が認められた (G3 の網脈絡膜症及び G3 のざ瘡様皮膚炎) こと、60 mg BID の DLT 評価可能例が 6 例であったことから、最終的に 60 mg BID を MTD と決定し、拡大パートの開始用量とした。その他、Eyre ら (2019) ^[17]の進行性固形癌または再発・難治性リンパ腫に対する CXD101 の第 I 相試験、Evans ら (2019) ^[18]の HER2 陰性乳がんに対する Eribulin のリポソーム製剤の第 I 相試験でも用いられている。

2.1.2.1.2. Modified 3+3 デザイン

Modified 3+3 デザインとして、がん第 I 相試験をより速やかに進めるために 3+3 デザインを改良したデザインがいくつか提案されている^[14]。その中には、最終的に MTD を推定するときにロジスティックモデルや isotonic regression^[19]を利用するデザインも提案されている。なお、isotonic regression を利用した MTD 推定の詳細は Model-assisted デザインの 2.3.2.1.1.3 節に譲る。

1) Best of 5 デザイン

少なくとも 3 例の被験者が各用量レベルに登録されることは 3+3 デザインと同じであるが、3 例の投与例のうち DLT 発現が 1 例だけでなく 2 例でも被験者を追加し安全性を検討するのが特徴の 1 つであり、DLT 発現した場合のアルゴリズムは下記となる^[11]。4 例の投与例のうち DLT 発現が 1 例であれば増量し、5 例のうち DLT 発現が 2 例であれば増量するため、伝統的な 3+3 デザインと比べてより増量しやすく、応じて DLT 発現確率も高くなるデザインである。MTD は、用量漸増が中止された用量レベルより 1 段階低い用量レベルとして決定される。なお、k 番目の用量を d_k とし、初回用量は d_1 ($k=1$) と表すこととする。

- Step1: 3 例の被験者を k 番目の用量 d_k で評価す
- Step1-1: 3 例の投与例のうち、DLT 発現が 0 ならば、 d_{k+1} に増量し Step1 に進む。
- Step1-2: 3 例の投与例のうち、DLT 発現が 1 または 2 であれば、Step2 に進む。
- Step1-3: 3 例の投与例のうち、DLT 発現が 3 であれば、Step4 に進む。

- Step2: 1例追加し、4例の被験者をk番目の用量dkで評価する。
- Step2-1: 4例の投与例のうち、DLT発現が1ならば、dk+1に増量しStep1に進む。
- Step2-2: 4例の投与例のうち、DLT発現が2であれば、Step3に進む。
- Step2-3: 4例の投与例のうち、DLT発現が3であれば、Step4に進む。
- Step3: 1例追加し、5例の被験者をk番目の用量dkで評価する。
- Step3-1: 5例の投与例のうち、DLT発現が2ならば、dk+1に増量しStep1に進む。
- Step3-2: 5例の投与例のうち、DLT発現が3であれば、Step4に進む。
- Step4: 用量レベルがMTDを超えたと判断し、用量増量を中止する。

Forouzannia ら (2004) ^[20]は、このデザインを採用し胸部悪性腫瘍患者に対する Topotecan の第 I 相試験を行った。用量漸増ルールは表 2 に示すように、3 例を初回用量に登録し、3 例のうち DLT 発現が 0 例であれば (0/3)、増量する。DLT 発現が 1 例であれば (1/3)、被験者追加をしてその後 DLT 発現なければ (1/4) 増量し、DLT 発現があれば (2/4)、さらに被験者追加するというルールである。なお、初回用量から右側に行くほど増量される。なお、DLT は、G3 以上の非血液学的毒性 (脱毛症または悪心及び嘔吐を除く)、G4 の血液学的毒性、または毒性関連の理由により予定された Topotecan 及び放射線療法を 8 週間で完了できないことと定義した。

表 2 Topotecan の用量漸増

初回用量 DLT/N→Action		
0/3→増量	DLT/N→Action	
	1/4→増量	DLT/N→Action
1/3→1 例追加		2/5→増量
	2/4→1 例追加	
2/3→1 例追加		3/5→1 用量下が MTD
	3/4→1 用量下が MTD	
3/3→1 用量下が MTD		

結果は、表 3 に示す通り 4 用量が検討され、0.4 mg/m² per day (N=7) が MTD となった。0.5 mg/m² per day にて、G4 の血液毒性と下痢、G3 の食道炎の DLT が発現したためである。なお、各用量の DLT 発現件数は確認することができたが、DLT 発現例数が不明のため正確な用量漸増プロセスは不明であるが、7 例が登録された用量もあるので何らかの規定外の評価がされた可能性が示唆された。

表 3 Topotecan の第 I 相試験の DLT 発現件数

	Topotecan の用量 [mg/m ² per day]			
	0.2 mg/m ²	0.3 mg/m ²	0.4 mg/m ²	0.5 mg/m ²
被験者数	7	5	7	5
血液毒性 DLT 件数	0	0	0	5
非血液毒性 DLT 件数	3	3	2	2

2) 加速型漸増デザイン

Simon (1997) ^[21]は、有効性が期待できないような低用量が投与される被験者をより少なくし、かつ試験期間も短縮することを目的としたデザインを 4 通り検討・評価した。結論として、デザイン 2~4 のような各用量レベルに 1 例の被験者を割当て、毒性が出現するまで投与量を増加させ、DLT が 1 例に発現するか、G2 の薬剤に関連する毒性が用量を問わず累積で 2 例に発現した後は標準的な 3+3 で少しずつ増量を行うという方法（加速型漸増デザイン：accelerated titration design）を提案した。なお、デザイン 2~4 では同一被験者での増量を認め、投与された被験者が、前サイクルで G1 以下の有害事象であれば増量できる。G2 の有害事象であれば用量を維持し、G3 以上の有害事象が認められた場合は減量する。表 4 に各デザインの特徴をまとめた。

- デザイン 1: 各コホートに 3 例の被験者を登録する従来型の 3+3 デザインである。次コホートへの増量幅は 40%であり、同一被験者内での増量はしない。
- デザイン 2: 最初の投与サイクルで 1 例に DLT が発現または G2 の有害事象（種類を問わない）が累積 2 例に認められるまで、各コホートに 1 例の被験者を登録する。DLT または G2 の有害事象が認められた場合は、現在の用量レベルでコホートを拡大し、以降のすべてのコホートでは従来型の 3+3 デザインの仕様に変更する。次コホートへの増量幅は 40%である。
- デザイン 3: デザイン 2 と同じであるが、コホート拡大する前の初期加速段階では、次コホートへの増量幅を 96%にする。
- デザイン 4: デザイン 3 と同じであるが、いずれかの投与サイクルで DLT が発現または G2 の有害事象（種類を問わない）が 2 例に認められた場合は、現在の用量レベルでコホートを拡大する。以降のすべてのコホートでは従来型の 3+3 デザインの仕様に変更する。

表 4 各デザインの特徴

デザイン	1 (3+3)	2 (加速)	3 (加速)	4 (加速)
増量幅	40%	40%	96%	96%
コホート当たりの被験者数	3, 6	1	1	1
3+3 デザインへの移行基準	なし	1 サイクル中の DLT または累積 2 例の G2 有害事	1 サイクル中の DLT または累積 2 例の G2 有害事象	全サイクル中の DLT または累積 2 例の G2 有害事象

デザイン	1 (3+3)	2 (加速)	3 (加速)	4 (加速)
		象		
同一被験者での増量	なし	あり	あり	あり

加速型漸増デザインが提案された後、蓄積された経験を踏まえ、Heath ら (2009) [22] は、Wayne State University で行われ、1995 から 2005 年の間に公表された 9 つの化学療法の第 I 相試験に関する報告を行った。そのうち 4 試験は 3+3 デザインによって行われ、5 試験は加速型漸増デザインによって行われた。Green ら [14] は、加速型漸増デザインは、より多くの用量レベル (平均用量レベル 10.6 に対して 3+3 デザインでは 8.8) が検討できる点と、RP2D より低い用量で投与される被験者数が少なくなる (平均被験者数 23.8 に対して 3+3 デザインでは 34) という点では有望なデザインであるが、期待された試験期間が短くなる効果は認められなかったと結論付けた。

加速型漸増デザインではあるが、同一被験者内での増量を行わない変法デザインが、Mita ら (2017) [23] の進行性悪性腫瘍に対する Dinaciclib (MK-7965、旧名 SCH 727965) の第 I 相試験で使用された。Dinaciclib は、サイクリン依存性キナーゼ (CDK) 阻害剤であり、CDK2、CDK5、CDK1、CDK9 を阻害する。本試験のパート 1 (21 日おきに 2 時間点滴) のデザインは、開始用量を 1.85 mg/m^2 とし、被験者 2 例に G2 の有害事象が認められる、または DLT が発現するまで順次 2 倍に増量した。同一被験者内での増量は行わず、G2 の有害事象が認められた場合は 3 例以上、DLT が認められた場合は 6 例以上の被験者で評価した。毒性が認められた後の用量は、最大投与量 (Maximum Administered Dose (MAD)) に達するまで 1 コホート 3 例で 40% ずつ増量した。3 例の被験者で 2 例以上 DLT を経験した場合、前のコホートで少なくとも 6 例の被験者で評価した。 29.6 mg/m^2 で 1/6 例、次に 41.4 mg/m^2 で 1/6 例に DLT が認められ、 58 mg/m^2 で 2/4 例に DLT が認められたため MAD となった。次の用量は中間的な投与量レベルの 50 mg/m^2 とした。 50 mg/m^2 の最初の 6 例のうち 1 例で AST の上昇が認められたため、9 例追加し、最終的に 4/15 例 (<33%) に DLT が認められた。従って MAD 以下の最高用量の 50 mg/m^2 が RP2D となった。

2022 年 9 月時点で「accelerated titration design、Phase 1」の用語について PubMed で検索したところ、Yap ら (2013) [24] の進行性固形癌に対する OSI-930 の第 I 相試験など、約 20 試験での利用が確認できた。

3) Rolling-6

Skolnik ら (2008) [25] は、小児癌患者を対象とした第 I 相試験の試験期間を短縮することを目的としたデザインを提案した。Skolnik らはレビューにて、47 試験から報告された 1,066 人の小児癌患者の毒性死亡率は 0.5% であり、これは成人第 I 相試験で観察された毒性死亡率と異なることを確認した。小児癌患者は高リスク集団であると考えられた

が、実際の全体的な安全性プロファイルは良好であることが分かった。そこで、典型的な 3+3 デザインのように被験者 3 例ごとに登録を中断して DLT 評価するのではなく、DLT 発現例数や DLT 評価期間中の被験者数に基づいて新たに 2~6 例の被験者がある用量水準で同時に治療することを許容し、登録された被験者のうち 6 例が DLT 評価期間中である場合にのみ登録を中断することで、試験期間を短縮することを期待したデザインを提案した。提案の背景には、小児癌特有の課題があると推察する。小児癌は希少疾患であり治療選択肢が限られているため、がん第 I 相試験でも多施設共同試験が一般的で、登録の一時的な停止はマネジメントを複雑化させる。また、成人と比較して小児のがんの進行は早く、治療開始までの期間短縮が望まれる。

Rolling-6 デザインでは、最大 6 例の被験者を同時に試験に組み入れ、被験者 6 例の DLT 評価を待っている間のみ、試験への被験者組入れを中断する。新たな被験者へ割り付けられる用量は、当該被験者登録時に当該コホートに現在組み入れられている被験者数、認められた DLT の数及び DLT 発現有無を観察中（登録・治療されたが DLT 評価期間の観察を終えていない被験者）の被験者数に基づいて決定する。

- Step1: 3 例の被験者がある用量レベルに登録され、4 例目の被験者候補が現れた時点とする。
- Step1-1: 3 例すべてで DLT が認められなかった場合、4 例目の被験者を次の用量レベルに登録する。
- Step1-2: 1 例以上の DLT 評価が完了していない、または 1 例の DLT が認められた場合は、4 例目の被験者を同じ用量レベルに組み入れる。
- Step1-3: 2 例以上の DLT が認められた場合、1 つ下の用量レベルに組み入れる。
- Step2: 5 例目及び 6 例目の被験者に対しても Step1-1 から Step1-3 までのプロセスを繰り返す。

6 例中 1 例以下で DLT 発現が確認された最高用量または、2 例以上で DLT 発現が確認された用量以下、かつ DLT 発現が 1 例も確認されなかった用量を MTD とする。

Cash ら (2021) ^[26]の Prexasertib の中枢神経系腫瘍を含む再発・難治性固形癌の小児患者を対象としたがん第 I 相試験で Rolling-6 が使用された。Prexasertib はチェックポイントキナーゼ (CHK) 阻害薬で CHK1 と CHK2 を選択的に阻害する。成人 RP2D (105 mg/m²、28 日おきに 1 日目と 15 日目に静脈内投与) の 80%である 80 mg/m²を開始用量とし、100 mg/m²、125 mg/m²、150 mg/m²の増量が計画され、同一被験者内での増量は認められなかった。MTD や RP2D が設定された時点で、代表的な 12 歳未満の被験者を対象に PK データを追加取得するため、PK 拡大コホートとして 6 例を追加登録するデザインであった。予定通り 80 mg/m²、100 mg/m²、125 mg/m²、150 mg/m²の用量レベルに 6 例ずつが結果的に登録され、DLT 評価可能例はそれぞれ 6 例、5 例、4 例、5 例であった。すべての用

量レベルで DLT は認められなかった。MTD には達しなかったものの、投与された最高用量である 150 mg/m² での忍容性が確認されたため、当該用量が RP2D と設定された。追加の PK 拡大コホートは RP2D である 150 mg/m² が 6 例に投与され、当該コホートにおいても DLT 評価可能な 5 例のうち DLT は認められなかった。

4) 増減デザイン

Storer (1989、2001) [27, 28] は、最初のうちは被験者 1 例に投与するたびに評価を行い、3+3 デザインと増量や減量のルール異なる増減デザイン (up and down design) を開発した。

- Step1: 最初の用量レベルに被験者 1 例を登録する。
- Step1-1: 最初の用量で DLT が発現しなかった場合は一つ上の用量に増量し、DLT が発現するまでより高用量に被験者 1 例を登録する。
- Step1-2: 最初の用量で DLT が発現した場合は一つ下の用量に減量し、DLT が発現しなくなるまでより低用量に被験者 1 例を登録する。
- Step2: 増量して DLT 発現した用量レベル、もしくは減量して DLT 発現しなくなった用量レベルに 3 例の被験者を登録する。
- Step2-1: 3 例とも DLT が発現しなければ次のレベルに増量し、高用量に被験者 3 例を登録する。
- Step2-2: 1 例に DLT が発現すれば用量は変更せず、同じ用量に被験者 3 例を登録する。
- Step2-3: 2 例以上に DLT が発現すれば減量し、低用量に被験者 3 例を登録する。

上記のアルゴリズムを予定された被験者数の DLT 評価期間が完了された時点で終了する。MTD については、すべての用量に登録された被験者の DLT 発現情報を用いてロジスティックモデルより推定された DLT 発現確率に基づいて決定する。シミュレーション研究によると、この増減デザインは伝統的な 3+3 デザインに比べて、許容できないほど高用量が投与される被験者を過度に増加させることなく、モデルを用いて MTD の推定精度を高め、低用量が投与される被験者を減らせることが示されている。なお、単剤療法でこのデザインを採用した事例を確認することはできなかったが、MTD の推定にロジスティックモデルを利用するという特徴があるので紹介した。

5) 薬理的誘導型増量 (pharmacologically guided dose escalation) デザイン

がん第 I 相試験において、MTD により早く到達するために用量漸増を加速したいが、これは被験者を高用量に曝露させるリスクを伴い、過度な毒性を示す可能性がある。問題となる治療が薬剤の形態である場合、用量漸増を安全に加速するという問題を解決する 1 つのアプローチは、用量-毒性関係ではなく、薬物動態と薬力学 (毒性力学) の関係を利用

することである。この概念に基づき、Collins ら (1986) ^[29]は pharmacologically guided dose escalation (PGDE) デザインを提案した。このデザインでは、薬物曝露時間にわたり測定される血漿中薬物濃度の曲線下面積 (AUC) に焦点をあて、AUC を薬物動態の指標とする。毒性は注目すべき薬力学的反応として、DLT を指標とする。このデザインでは、DLT は AUC の関数である可能性があり、動物モデルにおけるこの関係はヒトに外挿できると仮定している。PGDE の用量設定アルゴリズムを以下に示す。なお、マウスの致死量 10% (LD₁₀、10% lethal dose) における AUC とヒトの MTD での AUC が同じであると仮定している。

- Step1: マウスの LD₁₀ とその AUC を求める。
- Step2: 被験者を特定の用量 [開始用量は LD₁₀ の 1/10 相当の用量 (MELD₁₀) など] より開始し、AUC を測定し、各被験者が DLT を発現したか評価する。
- Step2-1: 各被験者の平均 AUC が MELD₁₀ に関連するヒトにおける事前に規定した目標 AUC に達しない場合は、目標 AUC までの距離に応じて用量を漸増し、Step2 を繰り返す。
- Step2-2: 各被験者の平均 AUC が目標 AUC に到達した場合、または DLT が発現した場合は、Step3 に進む。
- Step3: 試験を終了する。

Ames ら (1990) ^[30]は Pirozantone (Oxantraxole、NSC-349174) における切除不能な癌患者 (白血病を除く) に対するがん第 I 相試験で使用した。Pirozantone はアントラサイクリン系薬剤の心毒性を抑える目的で作成された薬剤である。開始用量は 7.5 mg/m² で、各用量レベルに 3 例ずつ登録する。新しい被験者は前の用量の 3 例の被験者が最低 3 週間観察された後により高い用量レベルに登録された。この試験では投与量の増加は、被験者の薬物曝露量 [濃度 (Concentration) × 時間 (Time)] が LD₁₀ 用量 (90 mg/m²) におけるマウス曝露量の 40% (59 /μg-min/mL) となるまで 2 倍で増量し、その後の用量漸増は前の用量レベルの 133%を超えないように計画した。

Pirozantone は 7.5、15、30、45、90、120、140、160 mg/m² と増量され、同一被験者内でも増量された。Step2-1 のプロセスにより上記以外にも 60、67、130、135 mg/m² の用量も投与された。7.5 及び 15 mg/m² の投与後に Pirozantone が検出されたのは血漿検体 1~2 本のみであったため、これらの用量では正確な薬物曝露量の計算を行うことができなかった。30 mg/m² で目標値 (target C×T value) をわずかに超えた被験者がいたので、次の用量は 1.5 倍の 45 mg/m² とした。45 mg/m² で target C×T value を超えた被験者がいなかったため、次の用量は 2 倍に増量した。90 mg/m² を投与した 3 例すべてが target C×T value を超えたため、次の用量から修正 Fibonacci 法を用いて設定した。その後順調に増量したが、160 mg/m² では、適切な程度まで白血球減少症が再現性よく認められたことと、同時期に

行われていたがん第1相試験で 190 mg/m² で許容できない重篤な骨髄抑制が2例の被験者に認められたため 160 mg/m² 以上の増量は行わなかった。その結果、160 mg/m² を RP2D とした。

6) 累積コホートデザイン (Cumulative Cohort Design)

Ivanova ら (2007) ^[31]は、MTD は事前に規定した目標 DLT 発現確率となるような用量とし、次の被験者の用量を減量、増量または維持するか判断は、現在の用量での DLT 発現確率が目標 DLT 発現確率から一定の範囲内に含まれるかどうかによって決める、新しいデザインを導入した。目標 DLT 発現確率からの乖離が許容 (目標 DLT 発現確率- δ_L 、目標 DLT 発現確率+ δ_U) の範囲内であれば次コホートの被験者を現在の用量で治療、範囲外であれば増 (減) 量を行うことが特徴である。なお、 δ_L 、 δ_U は目標 DLT 発現確率の許容される乖離の程度の下限と上限を示す。用量漸増完了後、MTD は治療された被験者全員の DLT 情報から isotonic regression を用いて推定された DLT 発現確率に基づいて決定される。被験者数が 20 で用量レベルが 6 のとき、いくつか目標 DLT 発現確率に対する、許容 δ の推奨値が示されている。なお、 $\delta_L = \delta_U = \delta$ と下限と上限を等しくした例示で、目標 DLT 発現確率が 0.3 の場合における δ の推奨は 0.1 である。

Prasanna ら (2022) ^[32]は、転移性乳癌患者を対象に Phenelzine と Nab-paclitaxel の併用のがん第 I 相試験 (Nab-paclitaxel は固定用量) で本デザインを使用した。Phenelzine は抗うつ薬及び抗不安薬として使用されるモノアミン酸化酵素阻害薬 (MAOI) である。本試験の主要目的は、Phenelzine と Nab-paclitaxel の併用療法の安全性と忍容性を決定し、MTD と OBD を決定することである。目標 DLT 発現確率は 30%を設定した。OBD は、最適なりジン特異的脱メチル化酵素 1 (LSD1) 阻害と循環系腫瘍細胞 (CTC) 減少が達成される用量として定義された。Nab-paclitaxel は、100 mg/m² を 3 週間投与/1 週間休薬する 1 サイクル (28 日) を 3 サイクル実施する。Phenelzine は用量漸増し、3 サイクルを通じて毎日経口投与する。5 つのコホートはそれぞれ Phenelzine の開始用量が徐々に高くなるように設定する。用量レベル A は 15 mg/day で開始し、2 週目には 30 mg/day に、3 週目には 45 mg/day に増量し、試験期間中これを維持する。用量レベル B は 45 mg/day で開始し、試験期間中、一定量とする。同様に、用量レベル C、D、E は、それぞれ 60、75、90 mg/day で開始し、試験期間中この用量を維持する。次の用量レベルの増量は、前の用量レベルで最初の 8 週間に観察された DLT 発現数に基づいて決定する。

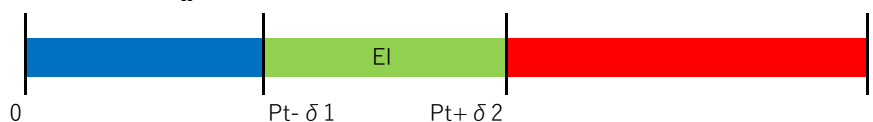
8 例の登録被験者の結果に基づき OBD が決定された。用量レベル A と B で 6 例が Phenelzine を最大 45 mg 投与し、用量レベル C で 2 例が最大 60 mg を投与した。G3 の事象は 2 例で報告され、Phenelzine 60 mg で好中球減少症と 45 mg で入院を伴う長期の下痢 (DLT) で、2 例とも試験を中止した。しかし、Phenelzine 60 mg は 45 mg の場合よりも良好な有効性を示したことから、試験は Phenelzine 60 mg で早期に中止された。その結果、MTD は未達であるが、治験実施計画書に従い、Phenelzine 60 mg を OBD 及び RP2D とし

た。また、Foster ら (2012) [33]も本デザインを使用した。目標 DLT 発現確率の範囲を(0.18 以上、0.34 以下)、MTD は DLT の推定発現率が 0.26 に最も近い用量と規定した。検討する用量レベルは 5 とされた。投与された 20 例の被験者での MTD 検討過程が追いやすく参考になる。

7) Interval 3+3 デザイン

Liu ら (2020) [34]は、伝統的な Rule-based の 3+3 デザインは、シミュレーションによる動作特性の比較を行うと、一般に Model-based デザインより MTD 推定精度が低いことが知られているため、単純だがより高度な規則に基づく interval 3+3 (i3+3) デザインを提案した。シミュレーションによると、i3+3 デザインは 3+3 デザインより MTD 推定精度が優れていて、Model-based デザイン及び Model-assisted デザインと比較しても同等の性能を示した。Interval の設定は、2.3.2.1.1 節の mTPI デザインと同様であり、目標 DLT 発現確率 P_t とその同等性区間 (EI, equivalence interval) として、 $EI = [P_t - \delta_1, P_t + \delta_2]$ を設定する必要がある。なお、 δ_1, δ_2 は目標 DLT 発現確率からの乖離の程度の下限と上限を示す。例えば、 $P_t = 0.3$ かつ $\delta_1 = \delta_2 = 0.05$ のとき、 $EI = [0.25, 0.35]$ のようになる。MTD は DLT 発現確率が EI の範囲内かつ P_t との乖離が一番小さい用量レベルとなる。事前規定した被験者数の DLT 評価が完了したとき、更に用量漸増が進む場合や DLT 発現確率が EI を超えて過大となった場合に試験終了となる。i3+3 では 3+3 デザインのようなルールに加え、観測された DLT 発現確率が P_t の同等性区間 EI に含まれているのかどうかを考慮したデザインである。具体的なアルゴリズムを次に示す。i3+3 では、用量 d に登録された被験者数を n 、用量 d で DLT が発現したに被験者数を x とし、DLT 発現確率 x/n が EI に含まれるか否かで次用量を以下のように推奨する。

- $\frac{x}{n} \leq EI$ の下限、次用量は $d+1$
- $\frac{x}{n}$ が EI 範囲内、次用量は d
- $\frac{x}{n} \geq EI$ の上限 and $\frac{x-1}{n} \leq EI$ の下限、次用量は d
- $\frac{x}{n} \geq EI$ の上限 and $\frac{x-1}{n}$ が EI 範囲内、次用量は $d-1$
- $\frac{x}{n} \geq EI$ の上限 and $\frac{x-1}{n} \geq EI$ の上限、次用量は $d-1$



例えば、 $P_t = 0.3$ 、 $EI = [0.25, 0.35]$ と設定した場合、登録された 3 例のうち 0 例の DLT 発現であれば $d+1$ に増量で次コホートへの移行が可能となる。3 例のうち 1 例の DLT 発現であれば $1/3$ (0.33) が EI 範囲内のため用量 d で被験者の追加登録が推奨される。3 例のうち 2 例の DLT 発現であれば $2/3$ (0.67) が EI 以上かつ $(2-1)/3=1/3$ (0.33) が EI 範囲内のため $d-1$ に減量となる。MTD の決定については、isotonic regression より推定された値を利用する。

Lin ら (2022) ^[35]は進行性固形癌患者を対象とした第 I 相試験で、始めの 2 用量は加速型漸増デザインを、それ以降は i3+3 デザインを使用した。なお、EI の設定等が不明であるため詳細な事例紹介は割愛する。

2.1.2.2. 有効性やバイオマーカーを同時考慮したデザイン

DLT の出現可能性が低く、かつその薬剤固有の生物学的な標的を持つ薬剤には、有害事象や DLT 発現の有無をエンドポイントとした 3+3 デザインが適切でない可能性がある。生物学的製剤では、用量-毒性関係が明らかでなく、増量すれば有効性も高まるという単純な仮定が正しくない可能性があり、用量を上げてても有効性は高まるのではなくプラトーに達するかもしれない。この種の薬剤に対するがん第 I 相試験の目的は、許容できる毒性の範囲内で、一定以上の生物学的効果が得られる用量レベルを見出すことが望ましい。

Hunsberger ら (2005) ^[36]は、さまざまな用量レベルにおいて毒性がほとんど生じないと想定される分子標的薬などにも適用できるデザインとして、増量は被験者ごとの奏効に基づいて行われる方法を 2 つ提案した。方法 1 は、例えば 6 例中 4 例以上の奏効が得られた用量について、次相の推奨用量とする。方法 2 は、奏効割合が一定に維持されてしまう場合は用量漸増を中止したいデザインである。用量を説明変数、各用量レベルの奏効割合を応答変数としてカーブを描き、回帰直線の傾きを考慮し用量漸増を決定する。傾きが 0 以下、少なくとも 1 例の奏効が観測できないと用量漸増を中止する。奏効割合が最も高い用量が RP2D となる。仮に DLT 発現した場合は、3+3 デザインと同様のルールに移行し、DLT が 2 例発現した用量の 1 つ下を MTD として決定を行う。3+3 デザインに移行するため、有効性を同時考慮したデザインとして完全に合致するものではないことにご留意されたい。

Christopher ら (2017) ^[37]は、治療歴のある転移性または局所進行固形癌に対する Avelumab 単剤のがん第 I 相試験を行った。当該試験は、3+3 デザインに基づいて用量漸増を行っているものの、有害事象以外の要素も考慮して RP2D を決定した事例である。本試験の用量漸増パートでは、Avelumab の 4 用量 (1、3、10 及び 20 mg/kg、2 週ごと) を評価した。結果は、Avelumab が投与された合計 18 例が DLT 評価可能例で、1、3 mg/kg コホートでは各 3 例、10、20 mg/kg コホートでは各 6 例が DLT 評価可能例であった。DLT は 20 mg/kg で 1 例のみ発現した。薬物動態解析では、3mg/kg と 20mg/kg の投与量の間で曝露量が比例することが示された。末梢循環血単核球上の PD-L1 の標的占有率を測定した結果、3mg/kg および 10mg/kg の用量で 90%以上となった。Avelumab による治療後、絶対リンパ球数または PD-L1 を発現するものを含む複数の免疫細胞サブセットに実質的な違いは認められなかった。Avelumab の毒性は 20mg/kg まで許容範囲であり、MTD には達していないが、薬物動態、標的占有率及び免疫学的分析に基づき、10 mg/kg の 2 週ごと投与を今後の開発用量として選択した。

2.1.3. 併用療法の開発

2.1.3.1. 標準療法へ上乗せる併用療法

標準療法に新規療法を上乗せする 2 剤併用療法を想定し、用量漸増は新規療法のみで、標準療法の用量は固定した状況であるとする。上乗せによる薬物相互作用の情報が限定的であると考えられる場合、重篤な有害事象など想定外の事象発生の可能性を考慮する必要があるであろうと考える。実際の臨床試験の事例を調査した結果、以下のような modified 3+3 デザインが適用されていた。

James ら (2020) ^[38]は、限局型小細胞肺癌に対する Pembrolizumab と化学放射線療法の併用第 I/II 相試験を行った。本試験の用量漸増パートでは、Pembrolizumab のみの用量漸増を行うこととし、Pembrolizumab の開始用量は 100 mg (単剤療法の MTD である 200 mg の半分) で、最高用量を 200 mg とした。3+3 デザインの用量漸増規則に従って用量を調節し、70 mg に減量するか、100 mg で継続するか、または 150 mg に増量可能とした。放射線及び併用化学療法 (Etoposide と Cisplatin または Carboplatin) について手順及び用法は固定されていた。Pembrolizumab について、開始用量から増量だけでなく減量もある 3+3 デザインであるのが本試験の特徴である。合計 40 例 (100、150、200 mg でそれぞれ 6、3、31 例) が DLT 評価可能例となり 1 例に DLT が認められたが、放射線療法に関連する事象であった。本試験の 3+3 用量漸増部分において Pembrolizumab の MTD を決定することができなかったため、Pembrolizumab 200 mg を RP2D として第 II 相試験パートに進んだ。

Catherine ら (2020) ^[39]は、再発または難治性ホジキンリンパ腫に対する Ipilimumab、Nivolumab、Brentuximab Vedotin の併用第 I/II 相試験を行った。用量漸増パートでは、合計 6 コホートを設定し、群ごとに新規療法に相当する薬剤の用量を探索した。3+3 デザインではあるものの、用量の組合せが多く、3 剤併用コホートも設定されているのが本試験の特徴である。

Ipilimumab 群 :

- コホート A: Ipilimumab 1 mg/kg + Brentuximab Vedotin 1.8 mg/kg (用量固定)
- コホート B: Ipilimumab 3 mg/kg + Brentuximab Vedotin 1.8 mg/kg (用量固定)

Nivolumab 群 :

- コホート D: Nivolumab 3 mg/kg (用量固定) + Brentuximab Vedotin 1.2 mg/kg
- コホート E: Nivolumab 3 mg/kg (用量固定) + Brentuximab Vedotin 1.8 mg/kg

三剤併用療法群 :

- コホート G: Nivolumab 3 mg/kg (用量固定) + Ipilimumab 1 mg/kg (用量固定) + Brentuximab Vedotin 1.2 mg/kg
- コホート H: Nivolumab 3 mg/kg (用量固定) + Ipilimumab 1 mg/kg (用量固定) + Brentuximab Vedotin 1.8 mg/kg

結果は、合計 61 例が DLT 評価可能例で DLT は 4 例に発現し 6 件報告された。コホート B、E 及び H が MTD とされ、これらの用量の組み合わせを用いて拡大コホート C、F 及び I に被験者登録が開始された。

Carlos ら (2022) ^[40]は、進行固形癌患者を対象に Emactuzumab と Atezolizumab の併用第 Ib 相試験を行った。本試験の用量漸増パートでは 3+3 デザインを用いて 1200 mg の Atezolizumab の固定用量に対して Emactuzumab の併用用量として 500、1000、1350 mg の合計 3 の用量レベルを設定した。合計 17 例 (500、1000、1350 mg でそれぞれ 5、6、6 例) が DLT 評価可能例となった。いずれのコホートでも DLT は発現せず MTD は未到達であったため、OBD の観点から Emactuzumab 1000 mg + Atezolizumab 1200 mg の併用療法を至適用量と決め、拡大パートに進んだ。

2.1.3.2. 新規療法を組合せる併用療法

2つの新規療法での2剤併用療法を想定し、用量漸増は2つの新規療法共に実施する状況であるとする。Ivanova ら (2004) ^[41]は、2療法での新しいノンパラメトリックデザインを提案した。各療法について、試験開始前に定められた用量において、他の療法の用量が固定されている場合、DLT は用量と共に単調増加すると仮定した。この仮定を用いて、すべての可能な用量レベルの組合せを二次元格子上で検討し、事前に規定した目標 DLT 発現確率 P_t に最も近い DLT 発現確率を有する用量の組み合わせを MTD として得ることが可能となる。

ある2療法 S と T があるとする、それぞれの用量レベルが療法 S は 1 から I まで (S_1, S_2, \dots, S_I)、療法 T は 1 から J まで (T_1, T_2, \dots, T_J) あるとする。コホートサイズを k とし、あるコホートでの併用療法の用量組合せを (S_i, T_j) と表すことにする。用量組合せ (S_i, T_j) での DLT 発現確率を P_{ij} 、目標 DLT 発現確率 $P_t = 1 - 0.5^{(1/k)}$ として、下表のような毒性判断をし、 P_{ij} と P_t の乖離が最も小さい用量組合せを MTD として決定する。なお、 $k = 1$ のとき、 $P_t = 0.5$ となり、 $k = 2$ で $P_t = 0.3$ 、 $k = 3$ で $P_t = 0.21$ 、 $k = 6$ だと $P_t = 0.11$ となる。

	DLT 発現確率 $P_{ij} <$ 目標 DLT 発現確率 P_t	DLT 発現確率 $P_{ij} >$ 目標 DLT 発現確率 P_t
DLT が 0 例 (0/k)	毒性増加可能	現状毒性を維持
DLT が 1 例以上 (1/k)	現状毒性を維持	毒性減量

現在の用量組合せが (S_i, T_j) であるとする次のような場合に各コホートへの移行が可能となる。

- 毒性増加、コホート (S_{i+1}, T_j) : DLT 発現確率 $P_{ij} < P_t$ かつ、直近コホートで DLT 無し
- 毒性減量、コホート (S_{i-1}, T_j) : DLT 発現確率 $P_{ij} > P_t$ かつ、直近コホートで DLT が 1 例以上

- 毒性維持、コホート (S_{i+1}, T_{j-1}) 、 (S_{i-1}, T_{j+1}) 、 (S_i, T_j) のいずれかを選択可能：DLT 発現確率 $P_{ij} < Pt$ かつ、直近コホートで DLT が 1 例以上
- 毒性維持、コホート (S_{i+1}, T_{j-1}) 、 (S_{i-1}, T_{j+1}) 、同用量の (S_i, T_j) のいずれかを選択可能：DLT 発現確率 $P_{ij} > Pt$ かつ、直近コホートで DLT 無し

T_j						
...						
T_{j+1}	○					●:現在の用量組合せ
T_j	▽	●	△			○:毒性維持
T_{j-1}			○			△:毒性増加
	S_{i-1}	S_i	S_{i+1}	...	S_i	▽:毒性減量

なお、用量 I について、療法 S が最大用量 S_i に到達したときは増量できないので、療法 T の増量 $j+1$ を行う。また、現状毒性を維持すると考えられるため、 (S_{i+1}, T_{j-1}) 、 (S_{i-1}, T_{j+1}) 、 (S_i, T_j) のいずれかの組合せを選択可能とする。これは、当該選択が MTD の決定に与える影響は大きくないことがシミュレーション結果より示唆されているためである。

Isakoff ら (2014) [42] の選択された進行性固形癌 (乳癌、膵臓癌、大腸癌、胆管癌またはグリオブラストーマ) に対する Bosutinib Hydrate と Capecitabine の併用におけるがん第 I 相試験では、増減デザイン (up and down design) という名称で Ivanova らのデザイン [37] を利用した。Bosutinib Hydrate と Capecitabine 両剤は既承認の薬剤ではあるものの、承認用法・用量を基にしない組合せを検討した点で、新規療法の組合せの併用療法の事例といえる。

Capecitabine は 625 mg/m^2 、 750 mg/m^2 、 1000 mg/m^2 の 1 日 2 回投与、Bosutinib Hydrate は 200 mg 、 300 mg 、 400 mg の 1 日 1 回投与の組み合わせを、各コホート $N=2$ で検討した。

開始用量 [Capecitabine (mg/m^2) / Bosutinib Hydrate (mg)] は "750/200 (コホート 1)" 及び "625/300 (コホート 2)" であった。コホート 1 では DLT 評価可能な 2 例のうち DLT が 0 例 (0/2 例) であったため、次用量として "750/300 (コホート 3)"、"1000/200 (コホート 4)" への被験者登録が可能となった。一方で "コホート 2 (625/300)" は当初 1/2 例で DLT が認められた (最終的には DLT と判定されなかった) ため次用量は減量 "625/200 (コホート 7)" となった。

コホート 3 (750/300) は DLT が 0/2 例であったため、次用量は 750/400 (コホート 5)、1000/300 (コホート 6) となった。コホート 4 (1000/200) は DLT が 0/2 例であったため次用量は増量するはずだが、コホート 6 が既に開始されていたため、同用量でさらに被験者を追加登録した (1000/200 (コホート 8))。コホート 7 は DLT が 0/2 例であったため次用量は増量し、750/200 (コホート 11) と 625/300 (コホート 12) となった。コホート 5

(750/400) は DLT が 1/2 例で認められたため、次用量は減量し、750/300 (コホート 9) となった。コホート 6 は DLT が 0/2 例であったため次用量は増量し、1000/400 (コホート 10) となった。

コホート 8 は DLT が 0/2 例であった。コホート 10 は DLT が 1/2 例で認められた。コホート 9、11、12 は DLT が 0/2 例であったが、コホート 8 の完了後に DLT 評価可能例が 24 例に達したため次用量が設定されなかった。なお、以上の用量の組み合わせと DLT 評価可能例数や DLT 発現例数等を表 5 にまとめた。本試験では、Capecitabine と Bosutinib Hydrate の併用療法の MTD は、DLT 発現確率が 33%未満の最高用量レベルと定義されていたため、コホート 6 [Capecitabine (1000 mg/m²) / Bosutinib Hydrate : 300 mg] が MTD と判断された。

表 5 用量の組み合わせと DLT 発現

登録順	用量レベル： Capecitabine (mg/m ²) /Bosutinib Hydrate (mg)	コホート 番号	DLT 評価可 能例数	DLT 発現 例数	次コホート
1	750/200	C1	2	None	C3/C4
2	625/300	C2	2	None	C7
3	750/300	C3	2	None	C5/C6
4	1000/200	C4	2	None	C8
5	750/400	C5	2	1	C9
6	1000/300	C6	2	None	C10
7	625/200	C7	2	None	C11/C12
8	750/300	C9	2	None	NA
9	1000/200	C8	2	None	NA
10	1000/400	C10	2	1	
11	750/200	C11	2	None	NA
12	625/300	C12	2	None	NA

2.2. Model-based デザイン

2.2.1. 一般論^[14, 43]

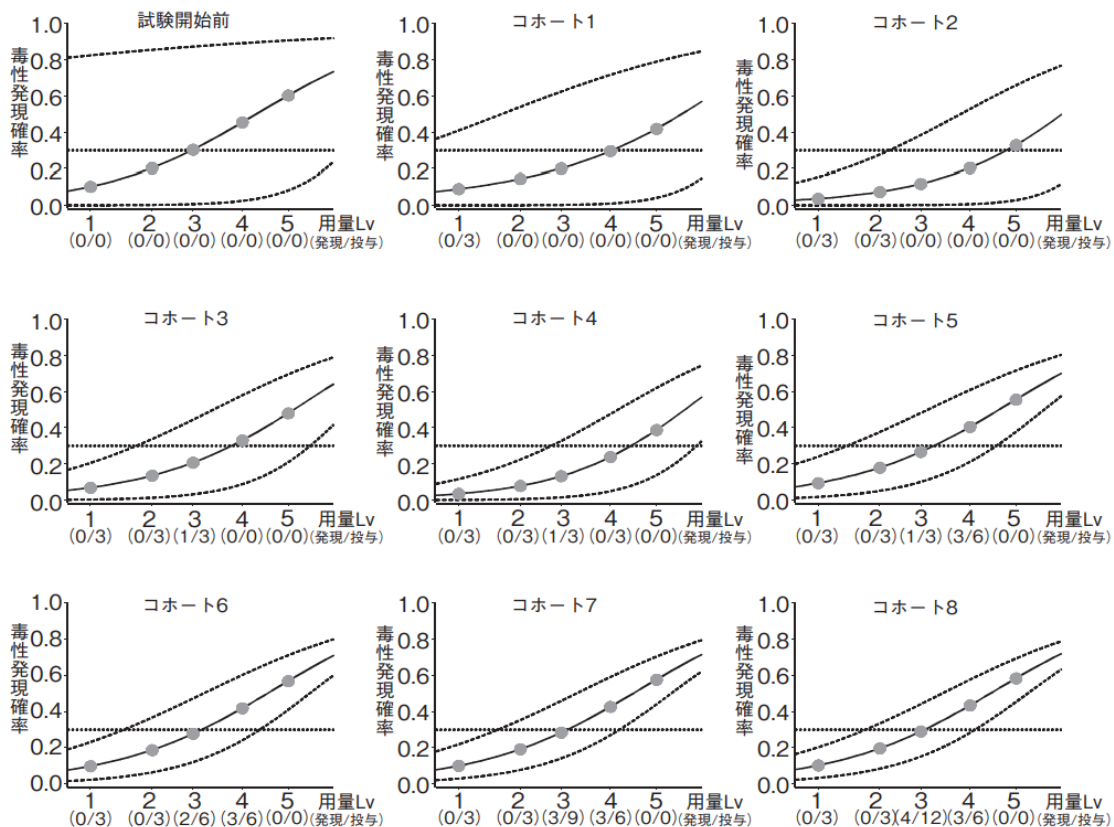
用量反応関係にパラメトリックな仮定をおかない Rule-based デザインに対し、Model-based デザインは異なる用量にわたって情報を併合するために一般にパラメトリックな用量反応関係を仮定する。真の用量反応関係は未知であり、Model-based デザインによって推定された MTD は用量反応モデルの誤特定の影響を受けやすい傾向にあるものの、すべての用量の情報を次用量の選択や MTD の推定に用いること、任意の目標 DLT 発現確率を指定できることから、3+3 デザインに代表される Rule-based デザインよりも高い MTD 推定性能を示すことが知られている。一般的ながん第 I 相試験では用量当たりの被験者数が少ないことから、用量調整の意思決定の精緻化という観点でもすべての用量の情報を利用する Model-based デザインは合理的であるといえる。さらに、ベイズ流のアプローチを用い、それまでに実施された試験で得られたデータや臨床担当者や臨床薬理担当者の意見をこれから実施する試験の解析に寄与させることで推定精度を高めることを試みることもできる。統計学的モデルに基づいて用量を選択する当該クラスの用量漸増デザインでは、FDA ガイダンス^[44, 45]等を参考に試験開始前にシミュレーション実験を用いることで動作特性を確認し、事前情報や用量反応モデルを含む各種設定の妥当性を担保することが必要となるだろう。

2.2.2. 単剤療法の開発

2.2.2.1. 有害事象の発現のみを指標としたデザイン

2.2.2.1.1. 連続再評価法 (CRM, Continual Reassessment Method) ^[43]

CRM は過去の知見に基づいて仮定される用量反応曲線を用いて任意の用量反応関係を記述し、任意の目標 DLT 発現確率の下での MTD の推定や次用量の選択を可能とする。図 1^[46]で示すように、試験中、DLT 発現情報が新たに蓄積されるにつれ、CRM は逐次的に用量反応曲線を更新し、すべての用量の DLT 発現確率の推定値を更新する。新規組み入れ被験者は更新された DLT 発現確率に基づいて最も適切な用量にコホート単位で逐次的に割り付けられ、事前に規定した最後の被験者の DLT 評価期間が完了したときに MTD が同定されることになる。また、用量反応曲線の推定に 1 パラメータの用量反応モデルを用いる CRM は、モデルを誤特定した場合でも MTD 推定は正しく行える傾向が報告されている^[47]。



実線：毒性発現確率の平均値；点線：90%信用区間；破線：標的毒性発現確率

図 1 用量反応曲線の遷移^[46]

O'Quigley ら (1990) ^[48]が提案したオリジナルの CRM は、各被験者の DLT 評価が完了した後、DLT の発現状況から MTD を推定し、推定された MTD に最も近い用量を次の検討用量とするというものであった。この方法の問題点として、被験者の不均一性が高い場合でも 1 例の被験者の結果に基づき用量レベルを飛ばして増量し得ることや、開始用量の決定に際して統計学的なモデルを用いており、必ずしも開始用量が臨床的に合理的と考えられている最小の用量ではない可能性があることから、安全性の懸念や MTD を過大推定する傾向があった。これらの懸念を解消するため、Faries (1994) ^[49]は、用量を飛ばして増量することを許容しない方法を、Goodman ら (1995) ^[50]は、必ず最初の用量から試験を開始する方法やコホートサイズを 1 例に限定しない方法をそれぞれ提案した。このような変法を用いることで多少 MTD の推定精度は低下するものの、より重要な点である安全性の改善が期待される。さらに、オリジナルの CRM では予定した被験者数に達した後に試験を終了することを規定しているものの、その他の終了規則として、Zohar ら (2001) ^[51]によるすべての用量で毒性が強い場合に中止する規則や、Korn ら (1994) ^[52]による 1 用量当たりの最大被験者数に基づく規則などが提案された。このような研究を経て、実際の臨床試験への適用の場面では、これらの変法を用いた modified CRM が利用されることが多い。

ここで、以下のようなオリジナルの CRM と同様の 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。

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

$f(\alpha)$ 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。

さらに、推定された DLT 発現確率 $\hat{\pi}_j$ を使い、上で触れたいくつかの変法を組み入れた 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。

Modified CRM が用いられた事例として、CD70 を標的した抗体薬物複合体である SGN-CD70A のがん第 I 相試験がある^[53]。本試験の腎細胞癌コホートは、CD70 陽性の転移性腎細胞癌患者を対象として SGN-CD70A (6 週間に 1 回静脈内投与) の安全性及び忍容性を評価し、MTD を決定することを目的として実施された。用量漸増デザインとして modified CRM が使用され、検討用量として 8 つの用量レベル (8、15、30、50、80、120、160 及び 200 mcg/kg) が設定された。本コホートにおける開始用量は、他のがん種のコホートから得られた結果を踏まえて 15 mcg/kg とされた。本試験では目標 DLT 発現確率は 30% と設定

され、DLT 発現確率が 30%未満となる確率が 50%以上の用量のうち最も高い用量がモデルに基づく推定 MTD とされた^[54]。最終的な MTD は、モデルから推定された DLT 発現確率及び集積された安全性情報に基づき安全性評価委員会により決定された。

本コホートには全体で 18 例が組み入れられ、15、30 及び 50 mcg/kg の 3 つの用量でそれぞれ 3 例、7 例及び 8 例が投与を受けた。30 mcg/kg 及び 50 mcg/kg コホートでそれぞれ 1 例及び 2 例に DLT が発現し、50 mcg/kg コホートで DLT 以外の毒性により 2 サイクル以上の投与が実施できなかった被験者が認められたことから 50 mcg/kg を超える用量の投与は行われなかった。各用量レベル（15、30 及び 50 mcg/kg）における最終的な DLT 発現確率はそれぞれ 10.2%、14.3%及び 23.0%、DLT 発現確率が 30%未満となる確率はそれぞれ 98.8%、96.8%及び 74.8%と推定され、50 mcg/kg は事前に定義したモデルに基づく MTD の条件を満たしていたが、DLT 以外の利用可能な安全性データも考慮され、MTD は 30 mcg/kg に決定された。

2.2.2.1.2. ベイズ流ロジスティック回帰モデル^[55]

ベイズ流ロジスティック回帰モデル（BLRM、Bayesian Logistic Regression Model）はモデルに基づくベイズ流用量漸増デザインの 1 つである。特徴として、2 つのパラメータを用いて柔軟な用量反応関係の表現が可能な点や DLT 発現確率の不確実性を考慮して次用量の選択が可能な点が挙げられる。

BLRM では用量-DLT 関係に以下のロジスティック回帰モデルを仮定する。

$$\text{logit}(\pi(j)) = \log \alpha + \beta \log(j/j^*)$$

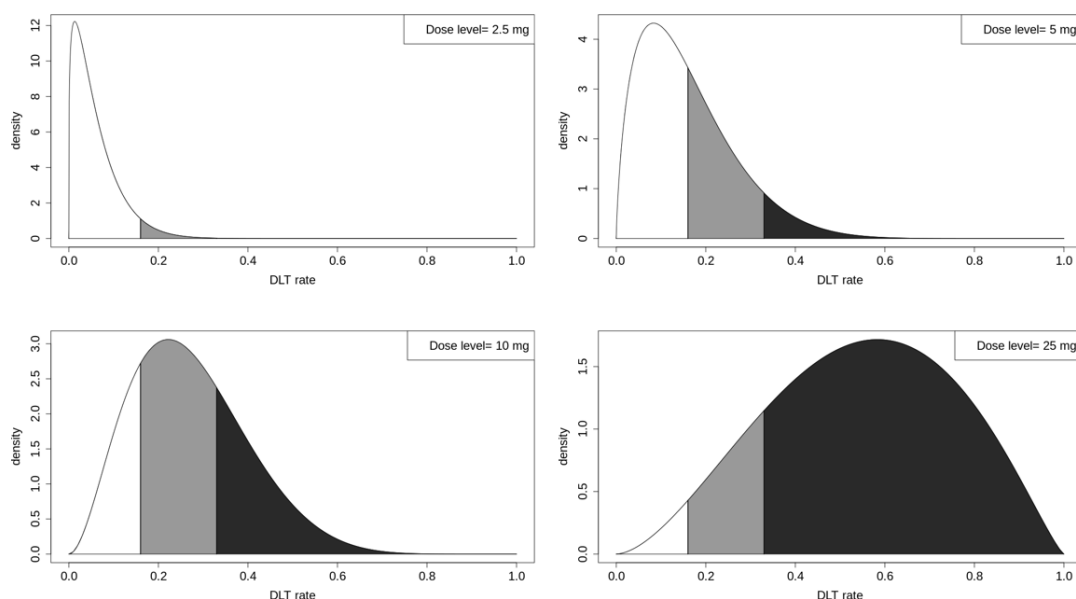
上式において、 $\pi(j)$ は用量 j での DLT 発現確率、 $\alpha, \beta (> 0)$ はモデルパラメータ、 j は用量、 j^* は参照用量（用量 $j = j^*$ の時 DLT 発現確率のオッズが α に対応）である。2 つのパラメータに対する事前分布は 2 変量正規分布を仮定した場合 $(\log(\alpha), \log(\beta)) \sim BVN(\mu, \Sigma)$ と表される。

次コホートの用量は他の Model-based デザイン同様それまでに得られた用量-DLT 発現状況に関するデータを利用し決定される。CRM では各用量における DLT 発現確率の事後分布における事後平均が目標 DLT 発現確率に最も近い用量を選択するため、事後分布の形状や不確実性は考慮していない。これに対し、BLRM では DLT 発現確率の事後分布全体を利用することでより不確実性を考慮した次用量選択方法を提案している。具体的には、各用量の DLT 発現確率の事後分布を、予め DLT 発現確率の区間で定義した各カテゴリ

（例：過少用量、標的用量、過大用量の 3 カテゴリを、それぞれ DLT 発現確率 $[0, 0.16]$ 、 $(0.33, 1.0]$ の区間として定義；図 2）に含まれる確率として要約（図 3 $(0.16, 0.33]$ ）し、これら区間に含まれる確率を次用量選択に利用する（例：次用量は標的用量区間の確率が最も高い用量を選択する等）。この際、次用量選択の基準には、過大用量の制御を伴う増量法（EWOC, escalation with overdose control）^[11, 56]の考え方を取り入れることも可能であ

る。EWOC は、MTD 以上の用量（過大用量）と推定される確率が一定未満（例：25%未満）となる用量のみを次用量の候補とすることで、過大用量へ投与する機会を制限しながら用量漸増を実施する方法である。EWOC の考え方を取り入れた次用量選択基準には、例えば、次用量は過大用量区間の確率が 25%未満の用量のうち標的用量区間の確率が最も高い用量を選択する等が考えられる。BLRM における次用量の選択方法については、その他の定量的な方法も提案されている^[55]。

図 2 は、DLT 発現確率の事後分布のイメージ図であり、各用量におけるその時点での情報に基づいた DLT 発現確率の分布を示している。図 3 は図 2 の各用量の DLT 発現確率の分布を特定の区間（過少用量、標的用量、過大用量）毎に要約したものである。次用量の選択方法について、EWOC の原則も取り入れ、「次用量は過大用量区間の確率が 25%未満の用量のうち標的用量区間の確率が最も高い用量を選択する」とした場合を例に考える。この場合、図 3 より、5 mg が EWOC の原則を満たす用量範囲内で最も高い標的用量区間の確率を示すため次用量として選択される（10 mg は標的用量区間の確率が最も高いが、過大用量である可能性が許容できない程高いため選択されない）。なお、次コホートの結果が得られた場合は、その情報も利用し事後分布（図 2、図 3）が更新され、新たな事後分布に基づきさらにその次の用量を選択していくこととなる。



過少用量区間（白色）：DLT 発現確率が $[0, 0.16]$ 、標的用量区間（灰色）： $(0.16, 0.33]$ 、過大用量区間（黒色）： $(0.33, 1.0]$ 、DLT 発現確率が各区間範囲である確率は各区間の曲線下面積に対応

図 2 各用量の DLT 発現確率の事後分布

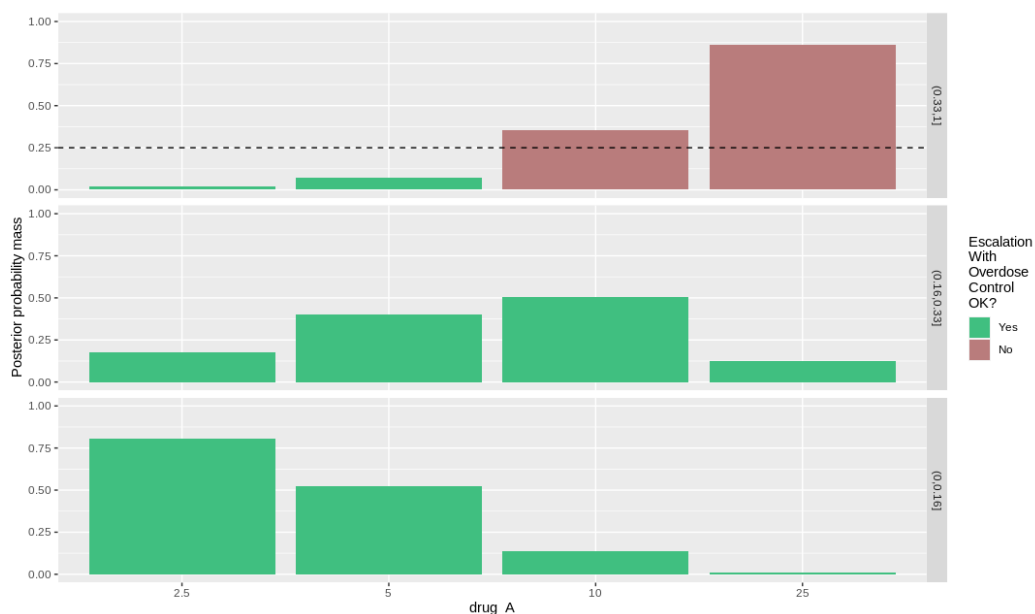


図 3 各用量の DLT 発現確率の事後分布の区分確率の要約

MTD は、次用量の選択方法同様、試験終了時に標的用量区間の確率が最も高い用量や EWOC の原則を満たす用量範囲における標的用量区間の確率が最も高い用量を選択するなどの考えに基づき選択される。

なお、DLT 発現確率のカテゴリの設定や事前分布の設定（ヒストリカルデータの利用可能性やその程度）、事後分布に基づく次用量や MTD の決定方法等については、興味の薬剤や疾患の特性等も考慮し、試験デザインを立案する際に調整されるべきである。なお、ヒストリカルデータが利用できない場合、情報量をほとんど持たない分布が事前分布として設定されるだろう^[55, 57]。

BLRM が用いられた実例として、Ceritinib の第 I 相試験がある。^[58]本試験の用量漸増パートは、ALK 遺伝子再構成を有する非小細胞肺癌患者を対象として Ceritinib の MTD 及び RP2D を決定することを目的として実施された。MTD の定義は、第 1 サイクルにおける DLT 発現確率の事後分布が過大用量区間 $[0.33, 1.00]$ に含まれる確率が 0.25 未満である用量のうち、標的用量区間 $[0.16, 0.33]$ に含まれる確率が最も高い用量とされた。BLRM に基づき、DLT に関して得られたデータ及びモデルパラメータの事前分布（無情報事前分布）から各用量レベルにおける DLT 発現確率の事後分布を算出し、3 つの区間（過少用量 $[0, 0.16]$ 、標的用量 $[0.16, 0.33]$ 、過大用量 $[0.33, 1.00]$ ）に含まれる確率をそれぞれ算出した上で、標的用量区間に含まれる事後確率が最大となる用量を次コホートの用量レベルとして推奨した。次コホートの用量レベルは、モデルに基づく DLT 発現の事後分布に加えて、その時点までに得られている安全性及び PK 薬物動態データも考慮され決定された。なお、EWOC の原則に基づき、事後分布が過大用量区間に含まれる確率が 25%未満である用量の

みを選択可能とした。次コホートの用量は、上記のモデルに基づく推奨に加えて、その時点までに得られている安全性及び薬物動態データも踏まえて総合的に決定した。

本試験の用量漸増パートには合計で 59 例（うち 54 例が DLT 評価対象例）が組み入れられ、50 mg QD から 750 mg QD までの 9 つの用量が検討された。本試験における開始用量は、非臨床試験の結果に基づき 50 mg QD と設定され、DLT は 400 mg QD の 14 例中 2 例、600 mg QD の 10 例中 2 例及び 750 mg QD の 8 例中 2 例に観察された。750 mg QD の評価が完了した時点ではモデルに基づいて 900 mg QD への増量が推奨されたが、DLT 以外の安全性データから増量は妥当ではないと判断されたため、当該用量の投与は行わず、MTD 及び RP2D は 750 mg QD とした。このとき、750 mg QD 及び 900 mg QD が標的用量区間に含まれる確率はそれぞれ 54.2%及び 60.1%であり、過大用量区間に含まれる確率はそれぞれ 3.3%及び 13.6%であった（過大用量区間に含まれる確率が 25%を超える用量レベルは、1050 mg QD の 26%であった）。

2.2.2.1.3. イベント発生までの時間を考慮した CRM

2.2.2.1.1 節で説明した CRM を含む多くの用量漸増デザインでは、被験者全員の DLT の有無を完全に追跡したのち、すべての用量の DLT の有無の結果を用いて用量反応モデルを更新し、新規組み入れ被験者の用量の選択を行う。しかしながら、放射線療法に伴う遅発性の DLT や予防的薬剤の DLT を評価の対象とする試験などでは、これらの DLT の有無の追跡が完了するまで長い期間を要するため、試験期間が非現実的なほど長くなる可能性がある。Cheung ら（2000）^[59]は、DLT 発現までの期間を考慮することで DLT 評価期間を完了している被験者の情報に加え、治験実施計画書で規定した DLT 評価期間を完了していない被験者の情報も用いて新規組み入れ被験者が割り当てられる用量を決定するイベント発生までの時間を考慮した CRM（TITE-CRM、time to event CRM）を開発した。具体的には、DLT 評価期間に対する実際に追跡された期間の割合を利用して DLT 観察期間中の DLT 未発現の被験者の情報を割り引いて用量反応モデルの推定に寄与させることを提案している。これにより、当該時点で登録されている被験者全例の DLT 評価完了を待たずして新たな被験者を試験に組み入れることが可能となり、観察中の被験者の情報も最大限に利用したうえで試験期間の短縮が期待される。しかしながら、治療された被験者の DLT 情報を追跡している間も被験者の登録を中断しないために、被験者の登録スピードが速い状況では、MTD を超える用量が被験者に割り当てられる可能性が高くなる恐れがある点については留意が必要である。

この問題に対し、Bekele ら（2008）^[60]は次の被験者が割り当てられる用量での DLT 発現の予測確率が許容できないほど高い場合に登録を一時的に停止する規則を導入した。追加の追跡により、DLT の予測確率が許容可能な水準まで低下した場合に被験者の登録が再開される。これにより、被験者全員の DLT 評価期間が完了するまで新しい被験者の登録を許容しない CRM の課題と、被験者全員の DLT 評価期間の完了を必要としないものの DLT 評

価が不十分な状況で新しい被験者の登録を許容する TITE-CRM の課題に対する妥協点を与えた。

また、Mauguen ら (2011) [61] は CRM を TITE-CRM と同じアイデアを用いてイベント発生までの時間を考慮した EWOC を開発している。

TITE-CRM が用いられた事例として、MKC-1 (cell cycle inhibitor) のがん第 I 相試験がある。[62] 本試験は、間欠経口投与での複数の第 I 相試験実施後、進行固形がん患者を対象に MKC-1 連日経口投与時 (1 サイクルは 4 週) の MTD 及び有効性を評価することを目的として実施された。DLT 評価期間は 12 週とし、MTD は DLT 発現確率が 33% を超えないと予測される用量のうち最高用量と定義した。MKC-1 の既存の毒性データが豊富にあり、急性 (サイクル 1) 及び遅発性 (サイクル 2 及び 3) の毒性を考慮する必要があるため、用量漸増には TITE-CRM が用いられた。DLT 発現確率の推定は新規の被験者が試験に参加するたびに、固定切片と未知の用量-DLT パラメータ (事前分布: $N(0, 1.2^2)$) による単一パラメータのロジスティック用量-DLT モデルを用いて継続的に更新した。DLT がみられない被験者は、予定した DLT 評価期間に対する実際に完了した DLT 評価期間の割合に等しい重みで確率計算に含めた。新規組み入れ被験者は、DLT 発現確率が 33% と期待される用量に割り当てた。過度に急激な用量増加を避けるために、用量の増加幅は最大 30 mg/day に制限し、最大用量は 460 mg/day を超えないこととした。本試験における開始用量は、既存の第 I 相試験の結果に基づき 60 mg BID (120 mg/day) と設定された。

本試験には事前規定した合計 24 例が組み入れられた。最初の 3 例には 120 mg/day、以降の被験者 21 例には TITE-CRM アルゴリズムに基づき、150 mg/day [n=1]、180 mg/day [n=2]、200 mg/day [n=1]、230 mg/day [n=1]、260 mg/day [n=5]、290 mg/day [n=6] 及び 320 mg/day [n=5] が投与された。DLT は 320 mg/day の 1 例でのみ観察されたが、事前規定した被験者数 24 例に到達し、4 週サイクルでの累積用量が間欠経口投与での累積用量 (150~275 mg/day) を超えたため、320 mg/day で用量漸増を中止した。

なお、TITE-CRM が用いられたその他の実例として、放射線療法及び放射線療法と全身療法との併用療法の適用が複数報告されている [63, 64, 65, 66, 67]。

2.2.2.2. 有効性やバイオマーカーを同時考慮したデザイン

2.2.2.2.1. 有効性情報を考慮に入れたデザイン (EffTox) [14, 68, 69]

Thall ら (2004) [68] は毒性と有効性に基づいて用量を選択する EffTox (Efficacy-Toxicity Tradeoff design) 法を提案した。この方法は、図 4 のような毒性と効果のトレードオフとなる曲線を定義することで、毒性発現確率だけでなく有効性が得られる確率を考慮して次用量の検討を行うことを可能としている。

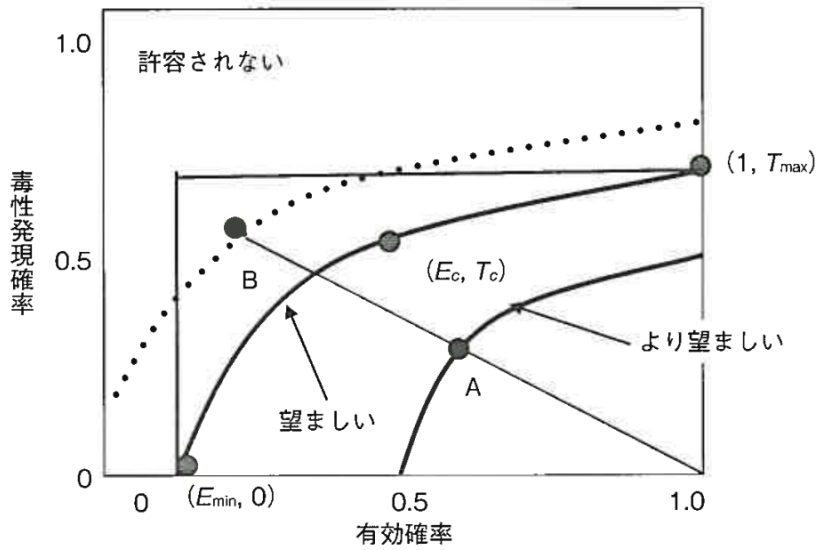


図 4 毒性と有効性のトレードオフ曲線のイメージ^[14]

毒性と有効性のトレードオフを示す曲線のセットを定義するにあたり、研究者はメディカルドクターの見解などに基づいて、許容される最小の有効確率 (E_{\min} : 毒性が全くないとき)、許容される最大の毒性発現確率 (T_{\max} : 有効確率が 100% のとき) T_{\max} : 有効確率が 100% のとき及びそれらの 2 つの極端の場合の間でかつ同程度に望ましい毒性発現確率と有効確率のペア (E_c, T_c) を設定する。この 3 つの目標値を用いて毒性と有効性に対する用量反応モデルをそれぞれ仮定し、さらに毒性確率と有効性が得られる確率の関連を表現するモデルを仮定する。基本的な用量選択手順は以下の通りである。

事前に定義された開始用量で最初の cohorts が治療される。次の cohorts では、安全性と有効性の用量反応曲線を更新し、有効確率が E_{\min} より大きく、かつ毒性発現確率が T_{\max} より小さいと推定された用量が次用量の候補となる。次用量の候補のうち、次の cohorts の被験者はまだ試されていない用量か、最も望ましい用量のうち低いほうの用量で治療される。図 4 では、用量 B に比べて用量 A は有効確率が高く、毒性発現確率が低いと推定されているため、用量 A の方が望ましい用量といえる。事前に規定した全員の被験者の毒性と有効性の観察が完了したときに残った次用量の候補のうち最も望ましい用量が次の試験の推奨用量として選択される。次用量の候補がなくなった場合は早期試験終了となり、どの用量も選択されない。

また、その他方法として、Braun (2002) ^[70] は CRM を毒性及び有効性を同時に考慮し、かつその各々の反応と用量との関係を説明可能な形式に拡張した bCRM (bivariate CRM) を提案している。

EffTox 法が用いられた事例として、Ponatinib Hydrochloride (Ponatinib, tyrosine-kinase inhibitor) のがん第 I/II 相試験 (MATCHPOINT) がある^[71]。本試験は急性転化期の慢性骨髄性白血病患者を対象に Fludarabine Phosphate、Cytarabine、darubicin Hydrochloride 及び顆

粒球コロニー刺激因子（FLAG-IDA）と Ponatinib を併用した際の Ponatinib の RD（事前規定した有効性を示す用量）及び安全性を評価することを目的として実施された。有効性は血液学的または軽度の細胞遺伝学的反応として定義される第 2 慢性期の誘発、安全性は DLT を指標とした。

EffTox 法を用いて、事前に規定した臨床的に重要な E_{\min} と T_{\max} の閾値（45%以上と 40%以下）を同時に考慮しながら、最適な Ponatinib の用量を決定した。EffTox 法を選択したのは、第 I 相（用量設定）と第 II 相（活性と毒性の推定）の両方の目的に対する回答を効率的に与え、比較的少数の被験者データより臨床的有用性を示す推奨用量を決定するためであった。本試験における Ponatinib の開始用量は、30 mg/day と設定した。次用量は候補用量（45 mg/day、30 mg/day、15 mg/day、15 mg/2day）から EffTox 法に基づいて選択された。本試験には合計 17 例が組み入れられた。有効性及び DLT を同時に考慮した EffTox 法より Ponatinib の RD は 30 mg/day と決定された。

2.2.2.3. 複数薬剤併用の用量漸増デザイン

2.2.2.3.1. BLRM の併用療法への拡張

Model-based デザインでは様々な併用時の用量漸増法が提案されている^[11, 72]。本項では、2つの薬剤を併用投与する治療法の用量探索時に適用可能な Model-based デザインとして BLRM を併用療法へと拡張したデザインを紹介する。なお、ここでは併用する薬剤それぞれの用量を漸増し MTD となる用量の組み合わせを探索する状況を想定する。

Wei ら（2014）の著書^[72]において Neuenschwander らは、2.2.2.1.2 節で説明した BLRM の併用療法への拡張として、DLT 発現確率のモデルについて「①それぞれの薬剤の単剤時の用量-DLT 関係、②それらの薬剤の交互作用」の 2つの要素を考慮する方法を提案している。即ち、興味のある併用時の DLT 発現確率の推定にあたって、単剤に対するモデルの変更は行わずに併用による DLT の増減をモデル化し明示的に考慮するというアイデアである。①に関しては、基本的な考え方は単剤における BLRM と同様であり、2.2.2.1.2 節に示したモデル式によって用量-DLT 関係をそれぞれの薬剤に対して表現する。

$$\logit(\pi_1(j)) = \log(\alpha_1) + \beta_1 \log(j_1/j_1^*)$$

$$\logit(\pi_2(j)) = \log(\alpha_2) + \beta_2 \log(j_2/j_2^*)$$

上式では 2つの薬剤 1、薬剤 2 に対してそれぞれ添え字 1、2 を用いた（その他の説明は 2.2.2.1.2 を参照）。②では、交互作用のパラメータ η を導入し、各薬剤がお互いの DLT 発現確率に影響を及ぼさない場合に対し、交互作用の大きさを表現可能とする交互作用モデルを提案した（ η の事前分布には正規分布を仮定）。交互作用モデルには例えば以下モデルが考えられる。

$$\text{odds}(\pi_{12}) = \text{odds}(\pi_{12}^0) \exp\left(\eta \frac{j_1 j_2}{j_1^* j_2^*}\right)$$

上式では、 π_{12}^0 は両剤の DLT 発現確率が独立であった場合の併用時 DLT 発現確率($\pi_{12}^0 = \pi_1 + \pi_2 - \pi_1\pi_2$)、 π_{12} は交互作用も考慮した場合の併用時 DLT 発現確率である。

用量漸増や MTD の選択に関する意思決定は、併用時の DLT 発現確率に関して、各薬剤の用量の組み合わせ毎に得られるため、単剤の用量漸増試験で BLRM を用いる場合と同様の対応が可能である。

図 5 は薬剤 A (2.5、5 mg) 及び薬剤 B (200、400、800、1200 mg) の併用時にそれぞれの用量組合せにおける DLT 発現確率の事後分布を特定の区間 (過少用量, 標的用量, 過大用量) 毎に要約したものである。次用量の選択方法について、EWOC の原則も取り入れ、「次用量は過大用量区間の確率が 25%未満の用量のうち標的用量区間の確率が最も高い用量を選択する」とした場合、図 5 では、薬剤 A=2.5 mg、薬剤 B=1200 mg の組合せ等はその時点では過大用量の可能性が許容できないほど高いため次用量としては選択されず (EWOC の原則を満たさず)、薬剤 A=2.5 mg、薬剤 B=800 mg の組合せが次用量として選択される。

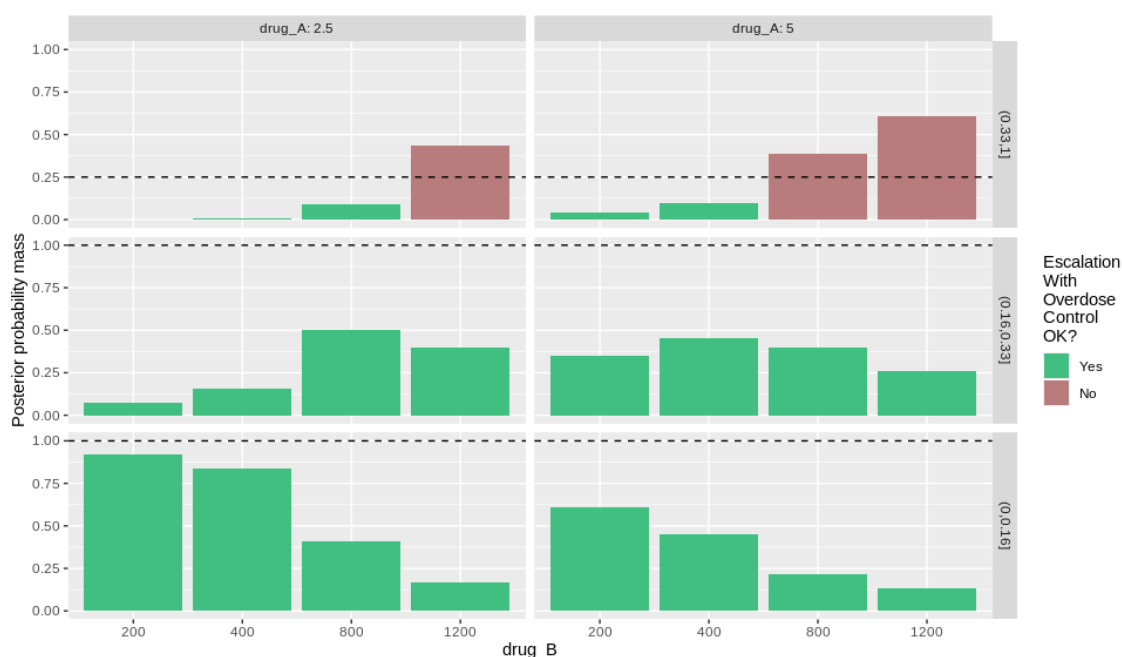


図 5 併用時における用量組合せ毎の DLT 発現確率の事後分布の区分確率の要約

併用治療の用量探索に BLRM が用いられた事例として、Buparlisib (Pan-PI3K Inhibitor) と Trametinib Dimethyl Sulfoxide (Trametinib, MEK1/2 Inhibitor) の第 Ib 相試験がある。^[73]本試験の用量漸増パートは、RAS または BRAF 遺伝子変異を有する非小細胞肺癌、卵巣癌または膵臓癌患者を対象として Trametinib と併用した際の Buparlisib の MTD 及び/または RP2D を決定することを目的として実施された。DLT 発現確率が 0.25 を超えないと推定される用量のうち最高用量を MTD と定義した。BLRM に基づき、Buparlisib 漸増後、

Trametinib の漸増を順次行った (図 6)。なお、漸増用量は各薬物単剤投与での MTD (1 日用量として Buparlisib : 100 mg、Trametinib : 3 mg) を超えないこととした。

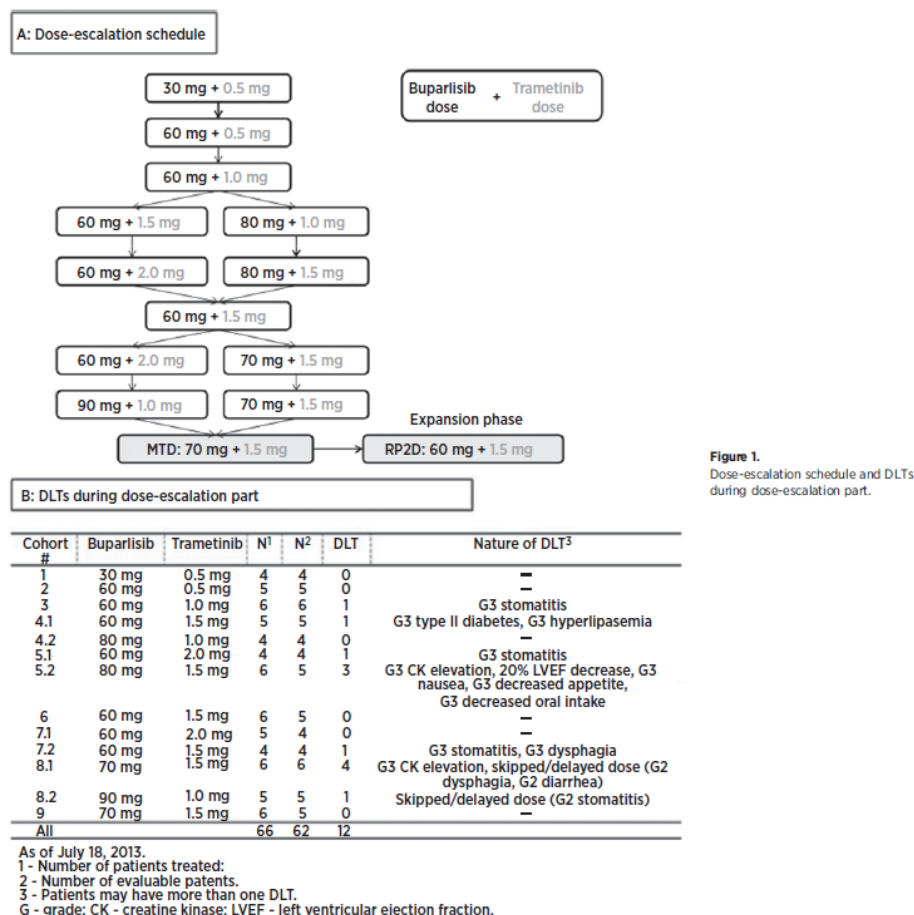


図 6 用量漸増スケジュール及び用量漸増パートでの DLT

BLRM に基づき、DLT に関して得られたデータ及びモデルパラメータの事前分布 (無情報事前分布) から各用量レベルにおける DLT 発現確率の事後分布を算出した。3つの区間 [過少用量[0, 0.16)、標的用量[0.16, 0.33)、過大用量[0.33, 1.00]] に含まれる確率をそれぞれ算出した上で、EWOC の原則に基づき、過大用量の確率が 25%未満となる用量を次用量の候補とした。次コホートの用量レベルは、BLRM に基づく推奨に加えて、その時点までに得られている安全性、PK、PD 及び有効性データを総合的に考慮して決定した。本試験における開始用量は、各薬物単剤投与の臨床試験の結果に基づき Buparlisib は 30 mg、Trametinib は 0.5 mg と設定された。

本試験の用量漸増パートには合計で 66 例 (うち 62 例が DLT 評価対象例) が組み入れられ、Buparlisib は 30 mg から 90 mg、Trametinib は 0.5 mg から 2 mg QD までの 13 のコホートで併用用量が検討された。図 6 に示すように、DLT は 62 例中 12 例に観察され、MTD

は Buparlisib で 70 mg、Trametinib では 1.5 mg に決定された。このとき、過大用量区間に含まれる確率はそれぞれ 0.1%未満であった。

当初は用量漸増パートで決定した MTD を用量拡大パートの用量に使用したが、最初の 10 例の被験者の忍容性を評価した結果、有害事象（主に口内炎と発疹）の発生率が高かったため、用量拡大パートにおける用量を変更することとした。当該時点の用量漸増パート及び用量拡大パートにおける DLT 情報を用いて、再度用量を推定し、Buparlisib と Trametinib の用量をそれぞれ 60 mg と 1.5 mg とし、当該用量を RP2D とした。

2.3. Model-assisted デザイン

2.3.1. 一般論^[74]

本節では、Rule-based デザインのように簡単な用量調整ルールに従って用量調整を行うことができる Model-assisted デザインを紹介する。Model-assisted デザインは Rule-based デザインと比較して真の MTD の特定割合が高くかつ MTD で或いはその付近で治療される被験者の割合が大きくなる^[74,75]。また、DLT 評価被験者数と DLT 発現被験者数の組み合わせごとに、次コホートの推奨用量（増量・維持・減量等）を試験計画段階で示すことができるため近年注目されているデザインである。

以降、これまで提案されてきた Model-assisted デザインを紹介する。Model-assisted デザインを紹介するにあたって、以下の 4 つの要素を中心にまとめた。

- a) 用量漸増規則
- b) MTD の選択方法
- c) 用量除外規則（事後確率に基づき毒性が高い用量（及びより高用量）は除外）
- d) 試験の終了条件

2.3.2. 単剤療法の開発

2.3.2.1. 有害事象の発現のみを指標としたデザイン

2.3.2.1.1. mTPI デザイン（mTPI, modified toxicity probability interval）^[5, 75]

mTPI デザインは、後述するベイズモデルの事後分布から得られた推定結果を基に用量漸増を行うデザインである。mTPI デザインでは、目標とする DLT 発現確率を基準に 3 つの区間、過少用量区間 [underdosing interval $(0, \delta_1)$]、適正用量区間 [proper dosing interval (δ_1, δ_2)]、過大用量区間 [overdosing interval $(\delta_2, 1)$] を設定する。 δ_1 と δ_2 はそれぞれ目標とする DLT 発現確率が存在する区間の下限と上限を示しており、検討している治療や対象とする疾患領域に応じて設定される値である。例えば、目標 DLT 発現確率が $\phi = 0.20$ 、適正用量区間を $\phi \pm 0.05$ とした場合、3 つの区間は、それぞれ $(0, 0.15)$ 、 $(0.15, 0.25)$ 、 $(0.25, 1)$ のように定義される。

mTPI デザインでは DLT 発現確率の事後分布の推定にベータ二項モデルを仮定している。すなわち、以下のように用量 j における総投与数 n_j と DLT 発現確率 P_j が与えられたもとでの DLT 発現数 y_j の分布に二項分布を、用量 j における DLT 発現確率 P_j の事前分布としてベータ分布を仮定している。

$$y_j | n_j, P_j \sim \text{Binominal}(n_j, P_j)$$

$$P_j \sim \text{beta}(a, b)$$

なお多くの場合、事前分布のベータ分布にはパラメータ $a = 1, b = 1$ が設定される。これは既知の通り、一様分布 $(0, 1)$ と等しくなる。

$$P_j \sim \text{beta}(1, 1) \equiv \text{Unif}(0, 1)$$

これより、観察データ (n_j, y_j) を与えた下で、 P_j の事後分布は以下のようなベータ分布に従う。

$$P_j | \text{data} \sim \text{beta}(1 + y_j, 1 + n_j - y_j)$$

以上のことから mTPI デザインは、DLT と用量の関係をモデル化する Model-based デザインとは異なり、用量間の関連を考慮せずに用量 j における DLT 発現確率のみをモデル化していることがわかる。

2.3.2.1.1.1. 用量漸増規則

次用量を決定するために、mTPI デザインでは 3 区間における unit probability mass (UPM) が使用される。UPM は P_j が該当区間に入る事後確率とその区間の長さの比として定義された以下のような式で表される。

$$\text{過少用量の UPM: } \text{UPM1} = \Pr(P_j \in (0, \delta_1) | \text{data}_j) / \delta_1$$

$$\text{適正用量の UPM: } \text{UPM2} = \Pr(P_j \in (\delta_1, \delta_2) | \text{data}_j) / (\delta_2 - \delta_1)$$

$$\text{過大用量の UPM: } \text{UPM3} = \Pr(P_j \in (\delta_2, 1) | \text{data}_j) / (1 - \delta_2)$$

図 7 のようにそれぞれの区間の長さが異なる場合があるため、各区間の長さで標準化した上でその区間に属する可能性の大きさを定量化した指標になっている。この UPM を用いて次のように次用量を決定する。

もし UPM1 が最も大きい場合、次のコホートは 1 用量分増量する。

もし UPM2 が最も大きい場合、次のコホートは現在の用量と同じとする。

もし UPM3 が最も大きい場合、次のコホートは 1 用量分減量する。

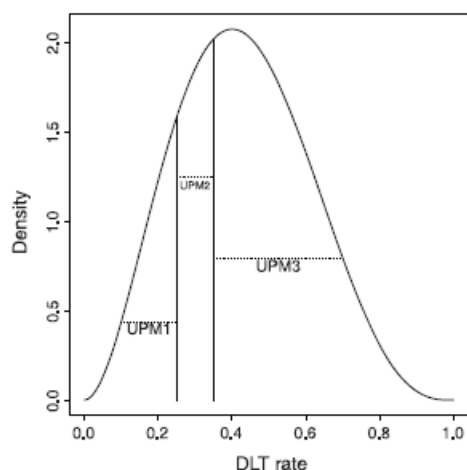


図 7 mTPI のイメージ

2.3.2.1.1.2. 用量除外規則及び試験の終了条件

mTPI デザインでは安全性を考慮し、もし現行用量 j における DLT 発現確率が目標 DLT 発現確率を超える事後確率がある閾値（例えば、95%）を超えている場合は、用量 j 及びそれよりも高い用量は試験の評価対象から除外されるという基準を設定している（以下、用量除外規則と呼ぶ）。そして、事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に、試験を終了する。

2.3.2.1.1.3. MTD の選択方法

試験終了後、各用量における DLT 発現確率の推定値を isotonic regression^[19] を用いて算出し、DLT 発現確率の推定値が目標 DLT 発現確率に最も近い用量を MTD として選択する。例えばある 5 つの用量レベルにおいて観察された DLT 発現確率（DLT を発現した被験者数/治療被験者数）がそれぞれ (0/3、1/3、0/3、3/15、2/4) であったとする。この場合、レベル 2 の用量の DLT 発現確率 (i.e. $1/3 \approx 33\%$) はレベル 3 の用量の DLT 発現確率 (i.e. $0/3 = 0\%$) より高くなっているため、用量毒性の単調性が成り立っていないことがわかる。このような現象は、たとえ真実としてレベル 2 の用量よりもレベル 3 の用量の毒性が高かったとしても、観測のランダム性により起こりえると考えられる。Isotonic regression は単調性が成り立っていない用量を平均化によって推定された DLT 発現確率に置き換えることにより、単調増加するように補正する方法であり、代表的なアルゴリズムとしては PAVA (pool adjacent violators algorithm) がある。すなわち、レベル 2 とレベル 3 の DLT 発現確率を $(1/3 + 0/3)/2 = 1/6$ と平均化した DLT 発現確率に置き換えることにより、各用量レベルの DLT 発現確率はそれぞれ $(0/3, 1/6, 1/6, 3/15, 2/4) = (0\%, 16.7\%, 16.7\%, 20\%, 50\%)$ と単調に増加するように補正される。ここで、想定している DLT 発現確率が 20% の場合、

レベル 4 の用量が MTD として選択されることになる^[75]。isotonic regression の SAS のサンプルコードは **Appendix** にて参照されたい。

2.3.2.1.1.4. mTPI デザインのケーススタディ

mTPI デザインの動作特性を説明するためにケーススタディを用いる。このケーススタディの設定を以下の通りとする。

- MTD 評価に用いる用量：4 用量（レベル 1～レベル 4）
- 目標とする DLT 発現確率: 0.3（30%）
- 過少用量、標的用量、過大用量: (0, 0.25)、(0.25, 0.35)、(0.35, 1)
- 最大評価可能被験者数: 30 例
- 1 コホートあたりの被験者数: 3 例
- 用量除外規則：DLT 発現確率が目標 DLT 発現確率を超える事後確率が 95%超となる
とき当該用量を評価対象から除外
- 終了条件：事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に、試験を終了

この設定における用量の増減規則を図 8 に示す。

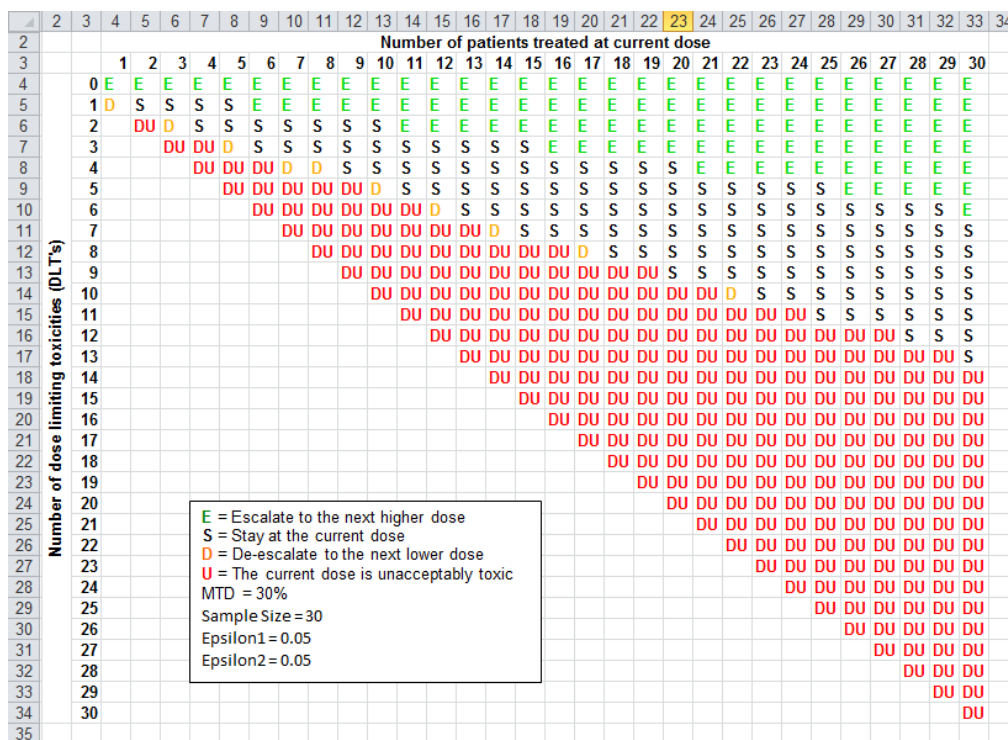


図 8 mTPI デザインの用量調整表

図 8 の横軸は用量ごとの投与被験者数を、縦軸は DLT 発現数を示す。また図中の“E”は増量、“S”は現用量維持、“D”は減量、“DU”は減量かつ現用量は評価対象から除外を意味す

る。例として2例中0例のDLT発現であった場合、図8の2例の漸増規則（Number of patients treated at current dose の“2”の列）に従い、DLT発現が0例では“E”になるので、次コホートの用量は1用量増量することが推奨される。

次に図8に基づいて、用量を増量、減量、維持する様子を示したケーススタディの結果を図9に示す。

このケーススタディによると、1番目のコホートでは3例中0例のDLT発現であったため、図8の3例の増減規則に従い、次コホートの用量は1用量増量したレベル2に決定した。2番目のコホートでは1例の被験者は評価不能であったため新たに1例を追加した。評価可能被験者数のうち3例中1例のDLT発現であったため、図8の3例の増減規則に従い、次コホートの用量は現用量維持のレベル2に決定した。3番目のコホートでは3例中1例のDLT発現であった。これによりレベル2は累積で6例中2例のDLT発現になるため、図8の6例の増減規則に従い、次コホートの用量は現用量維持のレベル2に決定した。4番目のコホートでは3例中0例のDLT発現であり、レベル2は累積で9例中2例のDLT発現になるため、図8の9例の増減規則に従い、次コホートの用量は現用量維持のレベル2に決定した。このように一連のDLT評価が最大評価可能数の30例に達するまで継続され、最終的にレベル4まで投与された。レベル1からレベル4までの各用量で観察されたDLT発現確率がそれぞれ(0/3, 2/12, 3/12, 2/3) = (0%, 16.7%, 25.0%, 66.7%)であることから、目標DLT発現確率(30%)に最も近い用量であるレベル3がMTDに選択された。

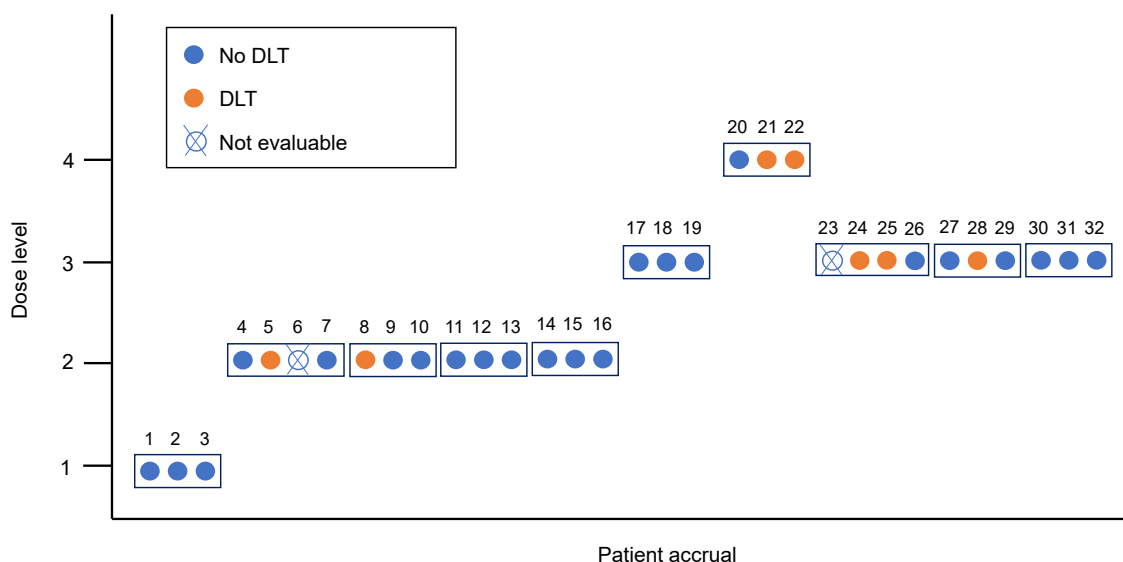


図9 ケーススタディによる mTPI デザインでの用量調整結果

2.3.2.1.1.5. 実際の試験例

mTPI が用いられた事例として、進行性非小細胞肺癌患者における HDAC 阻害剤の Mocetinostat と PD-L1 阻害剤の Durvalumab の併用に関する第 I/II 相試験 (Melissa et al. (2023))^[76]がある。当該試験の用量漸増パートは、根治目的の治療に適さない転移性または切除不能な局所進行性固形腫瘍を有し、かつ抗 PD-1 または抗 PD-L1 による前治療を受けた可能性のある患者を対象に、Mocetinostat の MTD 及び RP2D を決定することを目的に実施された。目標とする DLT 発現確率は 0.25 で、適正用量区間は (0.2, 0.3) として定義された。用量漸増パートで計画された Mocetinostat の評価用量は、50、70 及び 90 mg TIW で、必要に応じて 40 mg TIW まで減量が許容されていた。結果として用量漸増パートには合計で 20 例が組み入れられ、いずれの投与コホートにおいても DLT の発現は観察されなかった。本試験の MTD には最高用量の 90 mg TIW が選択されたが、MTD において Cycle 1 での減量または中断に至った有害事象が頻発したことを受けて、RP2D は 70 mg TIW が選択された。

2.3.2.1.1.6. mTPI デザインの問題点

mTPI を使用して用量調整を行うことの問題点として、次用量を決定するために用いる UPM という指標が直感的に解釈し難いことや過大な用量を推奨しやすい傾向にあることが指摘されている^[75,77]。後者に関して仮に、目標 DLT 発現確率を 0.20、過少用量、標的用量及び過大用量をそれぞれ(0, 0.17)、(0.17, 0.23)及び(0.23, 1)と設定した試験を想定し、試験のある時点で観察データが過少用量、標的用量及び過大用量の事後確率がそれぞれ 0.01、0.09 及び 0.9 を示したと仮定する。すなわち、この時点において現在の用量が過量投与にあたる確率が 90%で、適切な用量である確率は 9%程度であると解釈できる。しかしながら UPM を計算すると、過少用量は $0.01/0.17 = 0.06$ 、標的用量の UPM は $0.09/(0.23 - 0.17) = 1.5$ 、過大用量の UPM は $0.9/(1 - 0.23) = 1.17$ となり、標的用量の UPM が最大になることから、mTPI デザインは次コホートの用量として現在と同じ用量を推奨することになる。

2.3.2.1.2. キーボードデザイン (Keyboard)^[75,77]

Keyboard デザインとは、mTPI デザインの簡便さを保持しつつ、mTPI デザインで問題となった過剰な用量を推奨しやすい傾向をより抑制するように改良されたデザインである。mTPI デザインと同様に用量調整を簡便なベイズモデルに基づいて行う。mTPI デザインと大きく異なる点は、3つの用量区間の UPM を利用する代わりに真の用量毒性の位置を示すために DLT 確率の事後分布を等間隔に分割した区間を利用する。各区間は等間隔であるため UPM のように区間幅での標準化が不要となり、単純に最も事後確率が大きくなる区間を用いて次のコホートの用量を現在の用量よりも増量または減量するかを判断する。

この等間隔の区間のことを“Keys”とも呼ぶ。図 10 に Keyboard デザインのイメージを示す。

Keyboard デザインでは、まず“Target Key”と呼ぶ適正用量区間を特定する。例えば、目標 DLT 発現確率を 0.3 とした場合、“Target Key”はこの値を中心に(0.25, 0.35)のように等間隔の区間として設定される。その後、Target Key と同等な間隔を Target Key の両側に配置し、0 から 1 の範囲で同じ区間幅 Keys を形成する。例えば、先ほどの例のように Target Key として(0.25, 0.35)を与えた場合、区間幅は 0.1 となるので、左側に 2 つの Key (0.15, 0.25)及び (0.05, 0.15)を、右側に 6 つの Key (0.35, 0.45)、(0.45, 0.55)、(0.55, 0.65)、(0.65, 0.75)、(0.75, 0.85)及び (0.85, 0.95)を配置することになる。次に観察データを与えた下での DLT 発現確率の事後分布に基づいて Keys の事後確率を計算する。Keys の計算式は以下の通りである。

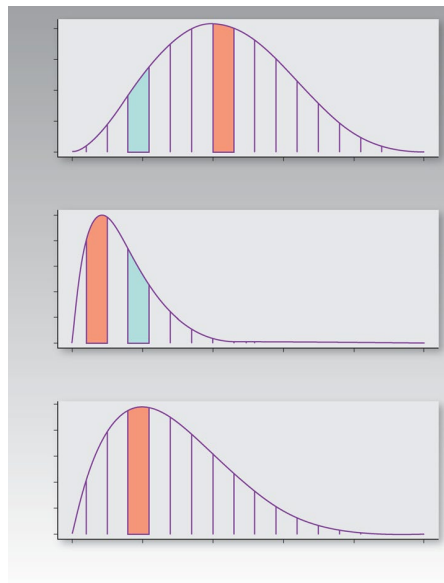
$$\Pr(p_j \in (c_1, c_2) | \text{data}) = \text{BETA}(c_2; a + y_j, b + n_j - y_j) - \text{BETA}(c_1; a + y_j, b + n_j - y_j)$$

c_1, c_2 はそれぞれ Key の区間の下限及び上限、 n_j は用量 j を投与された総被験者数、 y_j は用量 j を投与された被験者の中で DLT を発現した被験者数、 $\text{BETA}(c; a + y_j, b + n_j - y_j)$ はパラメータ $a + y_j$ 及び $b + n_j - y_j$ を持つベータ分布の c に対する累積分布関数 (DLT 発現確率が c 以下である確率) を表す (a, b は DLT 発現確率の事前分布 $\text{beta}(a, b)$ のパラメータであり、 $\text{beta}(1,1)$ を仮定した場合それぞれ 1 が与えられる)。

2.3.2.1.2.1. 用量漸増規則

次のコホートの用量を決定するために、最も事後確率が高くなる Key を特定する。この Key のことをその用量における真の DLT 発現確率を含む一番尤もらしい区間として“Strongest Key”と呼ぶ。この Strongest Key を用いて次のように次用量を決定する。

- もし Strongest Key が Target Key の左側に位置している場合、次のコホートは増量する。
- Strongest Key が Target Key の右側に位置している場合、次のコホートは減量する。
- もし Strongest Key が Target Key である場合、次のコホートは現在の用量と同じとする。



もし Strongest Key (赤色) が Target Key (青色) の右側に位置している場合は次のコホートは減量する (上図)。もし Strongest Key (赤色) が Target Key (青色) の左側に位置している場合は次のコホートは増量する (中央図)。もし Strongest Key が Target Key である場合、次のコホートは現在の用量と同じとする。(下図)

図 10 Keyboard デザインのイメージ^[75]

なお、用量除外規則、試験終了条件、及び MTD の選択方法は、mTPI と同様である (2.3.2.1.1.2 節、2.3.2.1.1.3 節参照)。

2.3.2.1.2.2. Keyboard デザインの特徴

mTPI デザインが過大用量で投与される傾向を招く原因は区間幅が異なることであるため、Keyboard デザインでは各 Key の区間幅が同じになるように定義している。実際、Yan (2017) ^[75]で、Keyboard デザインは mTPI デザインよりも過量投与するリスクを抑えられていることがシミュレーション結果から示されている。

また、同じような動機で mTPI デザインの問題点を改善した方法として mTPI 2 デザインがある。この方法は DLT 発現確率の区間幅を等間隔に分割している点で Keyboard デザインと同じであるが、現時点の用量における真の用量毒性の位置を示すために Keyboard デザインでは DLT 発現確率の事後分布を利用しているのに対して、mTPI 2 デザインでは UPM を利用している点で異なっている。最終的には Keyboard デザインと mTPI 2 デザインは同じデザインに分類されるが、Keyboard デザインの方がより単純に DLT 発現確率の事後分布のみで次コホートの用量調整を行うことができることから、直感的に解釈しやすいとされている ^[75,77]。なお、mTPI 2 の採用事例として進行性固形癌患者を対象とした CYH33 (PI3Kα inhibitor) の Ph1 試験がある ^[78]。

Keyboard デザインでは Target Key を中心に等間隔の Key を設定しているが、0 から 1 の範囲の両端が必ずしも等間隔になるとは限らない。上述の例のように Target Key が(0.25, 0.35)、区間幅が 0.1 の場合では(0, 0.05) 及び(0.95, 1) は Target Key と等間隔ではないので Key には含まれないことになるが、このことはつぎの理由により問題を引き起こすことはないとされている。事前分布に beta (1, 1) のベータ分布を与えた下では、DLT 発現確率 p_j の事後分布 $\text{beta}(1 + y_j, 1 + n_j - y_j)$ は単峰型か単調型のいずれかになる。加えて、目標とする DLT 発現確率 p_T は多くの場合 (0.1, 0.5) の範囲で設定され、Target Key の幅は p_T よりも小さくなることから、少なくとも 1 つの Key が Target Key の両端に配置可能である。もし、当該用量の真の DLT 発現確率 p_j が Key 以外の両端のいずれかに位置している場合 (e.g. $p_j \in (0, 0.05)$ or $(0.95, 1)$)、上述の通り事後分布は単峰型か単調型のいずれかになるため、 p_j に最も近い Key が Strongest Key になり、この Key は Target Key から見て p_j と Strongest Key は同側に位置することになる。よって結果として両端が Key に含まれていなくても、用量調整に影響を与えることはない。

2.3.2.1.2.3. Keyboard デザインのケーススタディ

Keyboard デザインを用いた臨床試験の結果が報告されていないことから、Yan (2017)^[75]で用いられているケーススタディを基に Keyboard デザインの動作特性を説明する。このケーススタディの設定は以下の通りである。

- MTD 評価に用いる用量：5 用量（レベル 1～レベル 5）
- 目標 DLT 発現確率: 0.3（30%）
- Key の区間幅: 0.1 [例：Target Key = (0.25, 0.35)]
- 最大評価可能被験者数: 30 例。ただし、1 用量あたりの最大評価可能被験者数は 18 例
- 1 コホートあたり被験者数: 3 例
- 用量除外規則：DLT 発現確率が目標 DLT 発現確率を超える事後確率が 95% 超となる時当該用量を評価対象から除外。
- 終了条件：事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に、試験を終了。

1 用量あたりの最大評価可能被験者数（18 例）の設定について、これはある用量で 18 例に達した場合、その時点で用量探索が収束したとみなすことを意図している。このような設定を入れることで、早期に試験を終了して MTD を決定できる可能性がある。

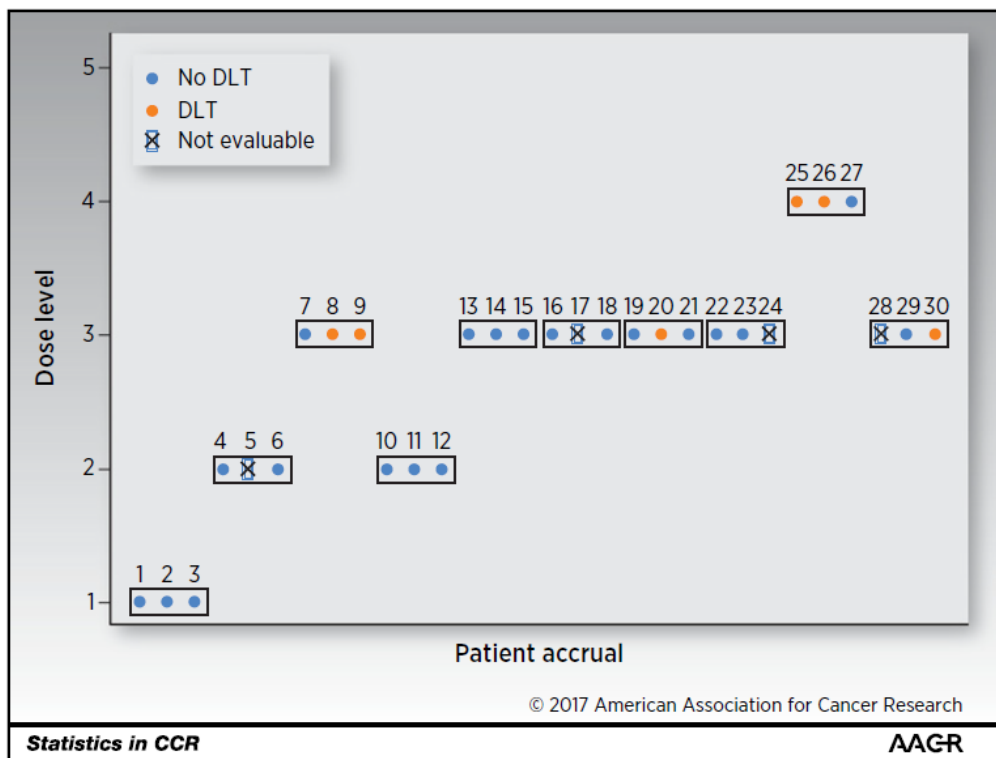
mTPI デザインと同様に、最大被験者数と DLT を発現した被験者数の全ての組み合わせから用量ごとの DLT 発現確率の事後確率を事前に計算することが可能であり、各組合せに対応する Strongest Key が一意に決定するため、試験開始前に用量の増減規則を作成することができる。このケーススタディの設定である 1 用量あたり最大被験者数を 18 例、目標とする DLT 発現確率を 30%、Key の区間幅を 0.1 の用量の増減規則を表 6 に示す。

表 6 Keyboard デザインの用量調整表

Number of patients treated	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Escalate if # of DLT <=	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3	4	4
Deescalate if # of DLT >=	1	1	2	2	2	3	3	3	4	4	4	5	5	5	6	6	6	7
	N	N																
Eliminate if # of DLT >=	A	A	3	3	4	4	5	5	5	6	6	7	7	8	8	8	9	9

Note. # of DLT is the number of patients with at least 1 DLT. When none of the actions (i.e., escalate, de-escalate or eliminate) is triggered, stay at the current dose for treating the next cohort of patients. "NA" means that a dose cannot be eliminated before treating 3 patients.

次に表 6 に基づいて、用量を増加、減少、維持する様子を示したケーススタディの結果を図 11 に示す。



図中の数字は、登録された被験者の番号を示す

図 11 ケーススタディによる Keyboard デザインでの用量調整結果^[75]

このケーススタディによると、1 番目のコホートでは 3 例中 0 例の DLT 発現であったため、表 6 の 3 例の増減規則に従い、次コホートの用量は 1 用量増量のレベル 2 に決定した。2 番目のコホートでは 1 例の被験者は評価不能であり、2 例中 0 例の DLT 発現であったため、表 6 の 2 例の増減規則に従い、次コホートの用量は 1 用量増量のレベル 3 に決定した。3 番目のコホートでは 3 例中 2 例の DLT 発現であったため、表 6 の 3 例の増減規則に従い、次コホートの用量は 1 用量減量のレベル 2 に決定した。4 番目のコホートでは再びレベル 2 が投与され、3 例中 0 例の DLT 発現であった。これによりレベル 2 は累積で 5 例中 0 例の DLT 発現になるため、表 6 の 5 例の増減規則に従い、次コホートの用量は 1

用量増量のレベル 3 に決定した。5 番目のコホートではレベル 3 が投与され、累積で 6 例中 2 例の DLT 発現であったため、表 6 の 6 例の増減規則に従い、次コホートの用量は現用量維持のレベル 3 に決定した。このような一連の DLT 評価は最大被験者数の 30 例に達するまで継続され、最終的にレベル 4 まで投与された。レベル 1 からレベル 4 までの各用量で観察された DLT 発現確率がそれぞれ $(0/3, 0/5, 4/15, 2/3) = (0\%, 0\%, 26.7\%, 66.7\%)$ であることから、目標 DLT 発現確率 (30%) に最も近い用量であるレベル 3 が MTD に選択された。

2.3.2.1.3. ベイズ最適区間デザイン (BOIN, Bayesian optimal interval design^[79])

BOIN デザインは、検討中の用量における安全性評価が完了した時点で DLT 発現確率が図 12 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^[80]。

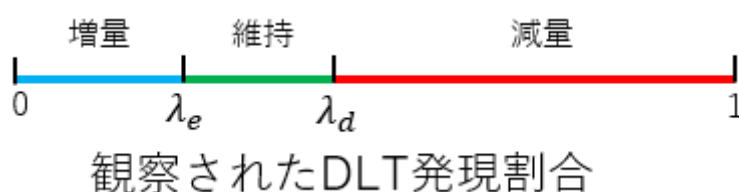


図 12 BOIN デザインの DLT 発現確率に基づく用量決定

用量維持の区間の境界 (λ_e, λ_d) は用量割り当ての誤りを最小化するように最適な値となるように計算される。増量が判断される目標とする DLT 発現確率 ϕ より低い毒性割合を ϕ_1 、減量が判断される ϕ より低い毒性割合を ϕ_2 とする。現在の投与量について、毒性割合が ϕ, ϕ_1, ϕ_2 となる確率がそれぞれ等しいというベイズの事前分布を与えた下で、用量維持の区間の境界は以下のように計算される。

$$\lambda_e = \frac{\log \frac{1 - \phi_1}{1 - \phi}}{\log \frac{\phi(1 - \phi_1)}{\phi_1(1 - \phi)}}, \lambda_d = \frac{\log \frac{1 - \phi}{1 - \phi_2}}{\log \frac{\phi_2(1 - \phi)}{\phi(1 - \phi_2)}}$$

上記式の詳細な計算方法は Liu and Yuan (2015)^[79]を参照。Liu and Yuan (2015)^[79]では、 ϕ_1 を $[0.5\phi, 0.7\phi]$ 、 ϕ_2 を $[1.3\phi, 1.5\phi]$ から選択することを提案し、 $\phi_1=0.6\phi$ 、 $\phi_2=1.4\phi$ をデフォルト値とすることを推奨している。本報告書では以降この推奨値¹を用いることにする。

¹ Liu and Yuan (2015) では、2 つの BOIN が提案されている。2 つの BOIN は λ_e と λ_d を固定した local BOIN と λ_e と λ_d が変動する global BOIN (投与例数に応じて用量維持の区間が変わる) である。global BOIN は真の MTD よりも大きな用

2.3.2.1.3.1. 用量漸増規則

例えば、 $\phi = 0.3$ の場合、用量維持の区間は $(0, 236, 0.358)$ になる。用量調整の例を示す。観察された DLT 発現確率が、0.167 だった場合は次のコホートの用量は 1 つ上の用量とする。観察された DLT 発現確率が、0.333 だった場合は次のコホートの用量は現在の用量と同じにする。観察された DLT 発現確率が、0.500 だった場合は次のコホートの用量は一つ下の用量とする。 ϕ を様々な値に変えた場合の λ_e , λ_d は以下になる。

表 7 ϕ_1 と ϕ_2 の推奨値を利用した場合の ϕ の値毎の用量維持区間

区間の境界	ϕ					
	0.15	0.20	0.25	0.30	0.35	0.40
λ_e	0.118	0.157	0.197	0.236	0.276	0.316
λ_d	0.179	0.238	0.298	0.358	0.419	0.479

毎回、DLT 発現確率を計算し、閾値と比較するのは手間になるため、以下のような用量調整表 ($\phi = 0.3$) を作成することができる。どの投与量でも同じ用量調整表を用いることができる。

表 8 用量調整表

	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

DLT 発現数が増量と判断される数を超え、減量と判断される数を下回る場合は次のコホートについて現在の投与量を維持する。

2.3.2.1.3.2. 用量除外規則及び試験の終了条件

用量除外規則及び試験の終了条件は mTPI デザインと同様である。

2.3.2.1.3.3. MTD の選択方法

MTD の選択方法は mTPI デザインと同様である。

2.3.2.1.3.4. 事例

BOIN が用いられた事例として、RET 阻害薬である Pralsetinib のがん第 I/II 相臨床試験 (Gainor et al. (2021))^[81]がある。本試験の用量漸増パートは、RET 融合遺伝子陽性進行

量への投与が多く local BOIN の方がシンプルであることから local BOIN の利用が Liu and Yuan (2015) より推奨されている。以降、本報告書では local BOIN を BOIN と呼ぶ。

非小細胞肺癌を対象として、Pralsetinib の MTD 及び RP2D を決定することを目的として実施された。MTD は、 $\phi=0.3$ に最も近い isotonic regression で調整された DLT 発現確率を持つ用量として定義された。また、本試験の用量漸増パートには合計で DLT 評価対象として 50 例が組み入れられ、30 mg QD から 600 mg QD までの 9 つの用量が検討された。DLT は 100 mg QD の 5 例中 1 例、300 mg QD は 11 例中 2 例、400 mg QD は 12 例中 2 例、600 mg QD は 3 例中 2 例に観察された。isotonic regression で調整された DLT 発現確率は掲載されていなかった。本試験の MTD は $\phi=0.3$ に最も近い 400 mg QD が選択された。また、安全性及び薬物動態の結果から RP2D としても 400 mg QD が選択された。

2.3.2.1.4. Time-to-event BOIN

分子標的薬、免疫療法、放射線治療などでは、DL の観察までに時間を要することがある。例えば、免疫チェックポイント阻害剤については、免疫関連毒性の一部が遅発性に発現する可能性があることがよく知られており、実際にイピリムマブやニボルマブなどの免疫チェックポイント阻害剤のがん第 I 相試験では、DLT 評価期間を 9 週間と通常の 28 日程度に比べて長期間に渡って評価をしている例がある^[82]。また、同様に放射線療法を評価する臨床試験では、DLT 評価期間を 6 か月と設定している例もある^[83]。このような場合においては、2.2.2.1.3 節で説明した様に、試験の長期化などの問題が生じる。Yuan et al. (2018)^[84]は、遅発毒性を考慮した Model-assisted デザインとして Time to event BOIN (以下 TITE BOIN) を提案した。

2.3.2.1.4.1. 用量漸増規則

TITE BOIN では、現コホートの DLT 評価が完了するのを待たずに新規コホートにおける用量を決定するために、DLT 評価が完了していない被験者の DLT 評価を補完する

((2.3.2.1.4.1)式参照) ことで DLT 発現確率を推定する。なお、補完は、観察期間中に一様に DLT が発現することを仮定し、一様分布に基づいて行う。

$$\hat{p}_j = \frac{\sum_{k \in O} y_{jk} + \sum_{k \in M} y_{jk}}{n_j} \quad (2.3.2.1.4.1)$$

ここで、 y_{jk} を用量 j に割り付けられた k 番目の被験者の DLT 発現の有無 (0: 発現無し、1: 発現有り) とする。 $k \in O$ は DLT 評価が完了している症例を表し、 $k \in M$ は DLT 評価が未完了の被験者を表すこととする。補完後の DLT 評価データを 2.3.2.1.3 節で説明をした BOIN の用量の増減基準に当てはめて、新規コホートの用量を決定する。

TITE BOIN では、現用量において DLT 評価が完了していない被験者の観察期間の総和を Total follow up time (TFT) とし、TFT を DLT 評価期間で割ることで標準化したものを Standardized TFT (STFT) として定義する。これは、未観測の観察期間中に一様に DLT が発現することに対応している。TFT は、例えば、DLT 評価期間が 3 か月の試験において、

現用量に 3 例の DLT 評価未完了例が存在し、それぞれ、1, 1.6, 2.5 か月観察されているとする。TFT=1+1.6+2.5=5.1、STFT=TFT/3=1.7 となる。(2.3.2.1.4.2)式は、STFT を用いて以下のように変換することが出来る。

$$\hat{p}_j = \frac{s + \frac{p_j}{1-p_j}(c-STFT)}{n_j} \quad (2.3.2.1.4.2)$$

(2.3.2.1.4.2)式中の s は、DLT 評価が完了している被験者の DLT 発現被験者数であり、 c は、DLT 評価未完了被験者数である。 p_j は、現用量 j における真の DLT 発現確率であるため、 $Beta(0.5\phi, 1 - 0.5\phi)$ を non-informative prior としたベータ二項モデルを用いて、DLT 評価が完了している被験者のデータから p_j の事後平均を求め置き換える。(2.3.2.1.4.2)式から推定される \hat{p}_j を 2.3.2.1.3 節で説明した用量維持の区間の境界 (λ_e, λ_d) と比較することで、BOIN と同様に用量の増減を決定する。より実際的には、 \hat{p}_j の推定に必要な当該用量の被験者数、DLT 発現被験者数、DLT 評価未完了被験者数、STFT の値を見て増減を決定され、安全性のために、現用量において 50%以上の症例が DLT 評価を完了していない場合には、次コホートへの割り付けを一時中断する。

2.3.2.1.4.2. 用量除外規則及び試験の終了条件

p_j の事前分布に $Beta(1, 1)$ を仮定し、用量 j に 3 例以上が組み入れられている条件の下で $Pr(p_j > \phi | n_j, s) > 0.95$ ($s = \sum_{k \in O} y_{jk}$) を満たした場合には、 $j' \geq j$ となる用量への組み入れを停止し、試験の評価対象から除外することを推奨している。そして、事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に、試験を終了する。

2.3.2.1.4.3. MTD の選択方法

予定していた被験者数のすべての DLT 評価を終えた際には、isotonic regression を用いて DLT 割合を単調増加に調整し、Target DLT rate に最も近い DLT 割合を持つ用量を MTD とする。より詳細は、2.1.2.1.1.2 節を参照とする。

2.3.2.1.4.4. TITE-BOIN のケーススタディ

TITE-BOIN の実際の適用例は、試験結果の詳細の報告がないため、ここでは、ケーススタディを用いて説明をする。

(2.3.2.1.4.2)式と 2.3.2.1.3 節で説明した用量維持の区間の境界 (λ_e, λ_d) と比較することで、用量調整表を求めることが出来る。表 9 にコホートサイズを 3、目標 DLT 発現確率を 0.2、DLT 評価期間を 3 か月とした際の用量調整表を示す。また、同様の設定における TITE-BOIN による用量調整の例を図 13 に示す。図 13 の例では、組み入れスピードを 2/月、目標症例数を 30 としている。最初のコホートは、最低用量から投与を開始する。Day

120 までには DLT の発現がなく、新たなコホートの組み入れは一時停止となっている。これは、表 9 に示された用量調整表において、DLT 発現がない場合には、新規コホートの組み入れに 2 例の DLT 評価完了が必要であることを反映している。Day 120 において、この 2 例の DLT 評価が完了し、次コホートの用量組み合わせの選択が可能になった。ここで、用量調整表に従って、2 コホート目は、用量が増量され、レベル 2 に組み入れられる。2 コホート目では、被験者 4 が、DLT を Day 150 に発現をしている。3 コホート目の最初の被験者である被験者 7 が組み入れられる Day 160 までには、被験者 4 が DLT 発現で DLT 評価完了、被験者 5 と被験者 6 は DLT の評価が未完了である。この時点で、用量調整表に基づいて、減量が確定されており、3 コホート目は用量レベル 1 に組み入れられる。また、Day 210 において、4 コホート目の最初の患者である被験者 10 が組み入れられるが、その時点で、用量レベル 1 に組み入れられている 6 例中に DLT の発現はなく、3 例が DLT 評価期間を完了している。この時、用量調整表に基づいて、増量し、5 コホート目は用量レベル 2 に割り付けられる。この様に割付を実施していき、試験は予定していた 30 例に達した 10 コホート目で終了となり、MTD として用量レベル 2 が選択された。ここで、TITE-BOIN を実施したこの例では、試験を 615(day)で終了することが出来る。もしも、TITE-BOIN を使用せず DLT 評価期間を完了する必要のある方法を用いていた場合には、少なくとも 1200 (day)が必要となる。

表 9 TITE-BOIN 用量調整表

Table 1. Dose-escalation and deescalation rule for TITE-BOIN with a target DLT rate of 0.2 and a cohort size of 3

No. treated	No. DLTs	No. data pending	STFT			No. treated	No. DLTs	No. data pending	STFT		
			Escalate	Stay	Deescalate				Escalate	Stay	Deescalate
3	0	<1	Y			12	1	6	≥1.24	<1.24	
3	0	≥2		Suspend accrual		12	1	≥7		Suspend accrual	
3	1	<2			Y	12	2	≤6		Y	
3	≥2	<1			Y&Elim	12	2	≥7		Suspend accrual	
6	0	<3	Y			12	3, 4	≤9			Y
6	0	≥4		Suspend accrual		12	≥5	≤7			Y&Elim
6	1	<3		Y		15	0	≤7	Y		
6	1	≥4		Suspend accrual		15	0	≥8		Suspend accrual	
6	2	<4			Y	15	1	≤7	Y		
6	≥3	<3			Y&Elim	15	1	≥8		Suspend accrual	
9	0	<4	Y			15	2	≤2	Y		
9	0	≥5		Suspend accrual		15	2	3	≥1.4	<1.4	
9	1	<2	Y			15	2	4	≥2.31	<2.31	
9	1	3	≥0.77	<0.77		15	2	5	≥3.48	<3.48	
9	1	4	≥2.15	<2.15		15	2	6	≥4.65	<4.65	
9	1	≥5		Suspend accrual		15	2	7	≥5.82	<5.82	
9	2	0		Y		15	2	≥8		Suspend accrual	
9	2	1		>0.52	<0.52	15	3	≤2		Y	
9	2	2		>1.59	≤1.59	15	3	3	>1.16	≤1.16	
9	2	3		>2.66	≤2.66	15	3	4	>2.34	≤2.34	
9	2	4		>3.73	≤3.73	15	3	5	>3.53	≤3.53	
9	2	≥5		Suspend accrual		15	3	6	>4.72	≤4.72	
9	3	<6		Y		15	3	7	>5.90	≤5.90	
9	≥4	<5			Y&Elim	15	3	≥8		Suspend accrual	
12	0	<6	Y			15	4, 5	≤11			Y
12	0	≥7		Suspend accrual		15	≥6	≤9			Y&Elim
12	1	<5	Y								

NOTE: "No. treated" is the total number of patients treated at the current dose level, "No. DLTs" is the number of patients who experienced DLT at the current dose level, "No. with data pending" denotes that number of patients whose DLT data are pending at the current dose level, "STFT" is the standardized total follow-up time for the patients with data pending, defined as the total follow-up time for the patients with data pending divided by the length of the DLT assessment window. "Y" represents "Yes," and "Y&Elim" represents "Yes & Eliminate." When a dose is eliminated, all higher doses should also be eliminated.

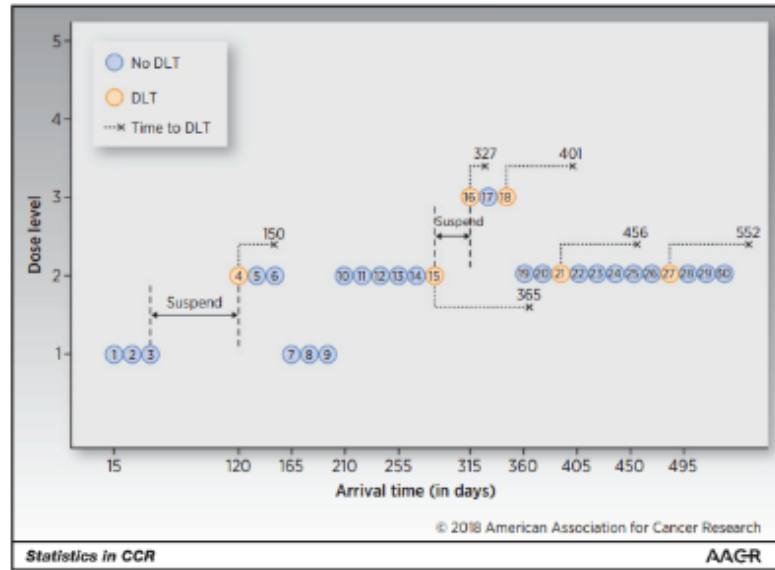


図 13 ケーススタディによる TITE-BOIN デザインでの用量調整結果

2.3.2.1.4.5. 実際の試験例

TITE BOIN は、実際に頭頸部癌の放射線療法の線量を探索する試験（Clinical trials.gov ID: NCT04870840）に適用されており、試験デザインが公開されている^[83]。2022 年 9 月現在進行中である。

2.3.2.2. 有効性やバイオマーカーを同時考慮したデザイン

2.3.2.2.1. BOIN-ET

Bayesian optimal interval design for dose finding based on both efficacy and toxicity outcomes (BOIN-ET)^[85]は安全性だけでなく有効性も考慮した OBD を選択するための用量調整デザインである。BOIN-ET は有効性について用量依存的な単調増加の仮定を必要としない、ノンパラメトリックな方法である。BOIN-ET では、BOIN の DLT 発現割合並びに有効割合に応じて用量調整を行う（図 14）。

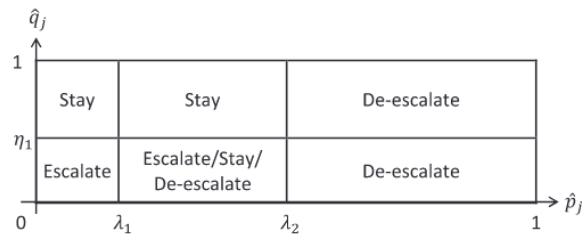


図 14 BOIN-ET の用量調整イメージ

2.3.2.2.1.1. 用量漸増規則

目標 DLT 発現確率を 0.3、ターゲットの有効性の割合を 0.6 とした場合に用量調整表は表 10 のようになる。E/S/D については有効性について単調性の仮定を与えないため、臨床試験の担当で事前にいずれかの設定を選ぶ必要がある。用量調整の例として、E/S/D はすべて S とした場合に、6 例の被験者が投与され、DLT 発現例数が 2 例、有効性の達成者数が 4 例の場合は次の用量は同じ用量とする。用量調整表を治験実施計画書に事前に定める場合はある一つの用量に対して想定される最大の投与例数だけ情報を用量調整表に入れる必要がある。

表 10 現在の用量の投与被験者 6 例に対する用量の調整表

	有効性達成者数							
		0	1	2	3	4	5	6
DLT 発現例数	0	E	E	E	S	S	S	S
	1	E/S/D	E/S/D	E/S/D	S	S	S	S
	2	E/S/D	E/S/D	E/S/D	S	S	S	S
	3	D	D	D	D	D	D	D
	4	D	D	D	D	D	D	D
	5	D	D	D	D	D	D	D
	6	D	D	D	D	D	D	D

E:Escalation, S:Stay, D:De-escalation

2.3.2.2.1.2. 用量除外規則及び試験の終了条件

安全性の観点からの用量除外規則は mTPI デザインと同様である。なお、有効性の観点からも有効性の達成割合が低い場合は用量を除外することができる。すべての用量が除外された場合は試験を終了する。そして、事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に試験を終了する。

2.3.2.2.1.3. OBD の選択方法

OBD は MTD 以下の用量の中から有効性が最大になると期待される用量を選ぶ。有効性は自由度 2 の fractional polynomial モデルを用いて調整する。

2.3.2.2.2. BOIN phase I/II (BOIN12)

BOIN12^[86]は安全性だけでなく有効性も考慮した OBD を選択するための用量調整デザインである。OBD を選択するために desirability (or mean utility) スコアを定義し、desirability スコアが最大となる用量を選択する。DLT の有無と有効性の有無の組み合わせで 4 通りのアウトカムとそのアウトカムのもとの utility を事前に定義し、それぞれのアウトカムが得られる確率の重み付きスコアとして各用量における desirability スコア $u(d)$ が算出される。

$$u(d) = p_1(d)u_1 + p_2(d)u_2 + p_3(d)u_3 + p_4(d)u_4$$

表 11 有効性の有無・毒性の有無別の確率

	有効性あり	有効性でない
毒性なし	$p_1(d)$	$p_2(d)$
毒性あり	$p_3(d)$	$p_4(d)$

u_1 、 u_2 、 u_3 及び u_4 を utility value と呼ぶ。utility value はメディカルドクター、臨床担当者や臨床薬理担当者と相談して将来的に薬剤が使用されるセグメントにおける UMN や標準治療を踏まえた臨床的位置づけを踏まえて設定する。例えば、原著では utility value は表 12 のような設定が提案されている。

表 12 有効性の有無・毒性の有無における utility values

	有効性あり	有効性なし
毒性なし	$u_1 = 100$	$u_2 = 40$
毒性あり	$u_3 = 60$	$u_4 = 0$

なお、 $p_1(d)$ 、 $p_2(d)$ 、 $p_3(d)$ 、 $p_4(d)$ は用量 d が投与された後の用量 d 内の観察された割合の値が入る。

BOIN12 は BOIN-ET と異なり、毒性や有効性について 2 カテゴリだけではなく、3 カテゴリ以上の場合でも適用できる。

2.3.2.2.2.1. 用量漸増規則

用量調整表が事前に作成可能であるため、用量調整時に desirability スコアを計算する必要はない。用量調整は現在の用量と前後 1 用量の中から行うとする。

表 13 表 12 の utility score のもとでの用量調整表

Num. Patients	Num. DLTs	Num. Effs	Desirability Score	Num. Patients	Num. DLTs	Num. Effs	Desirability Score
0	0	0	60	6	3	2	22
3	0	0	35	6	3	3	38
3	0	1	55	6	3	4	51
3	0	2	76	6	3	5	67
3	0	3	91	6	3	6	81
3	1	0	24	6	4	0	1
3	1	1	44	6	4	1	6
3	1	2	63	6	4	2	15
3	1	3	80	6	4	3	27
3	2	0	13	6	4	4	42
3	2	1	31	6	4	5	56
3	2	2	48	6	4	6	72
3	2	3	69	6	≥ 5	≥ 0	E
3	3	≥ 0	E	9	0	0	E
6	0	0	22	9	0	1	25
6	0	1	38	9	0	2	36
6	0	2	51	9	0	3	49

用量調整はまず DLT 発現被験者数のみに基づいて BOIN デザインに従い表 8 の用量調整表を用いて用量調整を行う。減量が判断されない場合は、表 13 に従って用量調整を行う。

用量調整の例を示す。例えば、2 つの用量があり、用量 1 の投与後の用量選択を示す。用量 1 は 3 例に投与し、DLT が発現せず有効性は 1 例に認められたとする。この場合は通常の BOIN に基づく推奨が増量となるため、desirability スコアに基づいて意思決定することになり、desirability スコアは 55 である。用量 2 はまだ一度も投与していないとする。この場合の desirability スコアは 60 である。desirability スコアがより大きい用量 2 を次のコホートの用量とする。BOIN12 デザインについて論文等の文献で紹介されている事例はなかった。

2.3.2.2.2.2. 用量除外規則及び試験の終了条件

安全性の観点からの用量除外規則は mTPI デザインと同様である。有効性の観点からの用量除外規則は、有効性の達成確率が期待された達成確率を超える確率が低い場合（閾値は例えば 0.9）に適用される。すべての用量が除外された場合に試験を終了する。そして、事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に試験を終了する。

2.3.2.2.2.3. OBD の選択方法

OBD は毒性に関して isotonic regression で MTD を決定してから MTD 以下の用量で最大の desiability スコアが期待されている用量である。

2.3.2.3. その他のデザイン

2.3.2.3.1. 併用療法の開発

Model-assisted デザインの枠組みで併用療法における用量探索を行うデザインは、BOIN for drug combination trials (以下 BOIN Combination) [87]と Keyboard combination design[88]が提案されている。どちらも単剤に対して提案された方法から、手法上の大きな変更はなく、単剤に対して用いられた閾値と同様のものを用いて用量の増量あるいは減量を決定する。用量の増減判断は、単剤に対して提案されたものと同様であるが、増量先あるいは減量先の併用療法には複数の用量組み合わせの候補が存在する。どの併用療法を選択するかについての決定が必要になることが単剤に対する手法とは異なる点であり、併用療法の選択までを含めた用量調整表の事前規定が出来なくなる。

2.3.2.3.1.1. BOIN for drug combination trials

Lin et al. (2017) は 2 剤併用療法の用量探索のための Model-assisted デザインとして BOIN Combination を提案している [87]。

2.3.2.3.1.1.1. 用量漸増規則

現在、最新の cohorts が割り付けられている用量組み合わせを $(i, j), i = 1, \dots, I, j = 1, \dots, J$ とすると、図 15 の様に現在の用量組み合わせにはその周りに 8 つの用量組み合わせが存在する。BOIN Combination では、2.3.2.1.3 節で説明した単剤用に開発された original の BOIN と同様の考え方、すなわち、現在の用量組み合わせでの DLT 発現確率と閾値を比較することで、用量の増減を判断する。そのため、用量の増減判断に関しては、original の BOIN と同じ用量調整表を使用することが出来る。BOIN Combination では、最初の cohort は最低の用量組み合わせから割り付けを開始し、増量時・減量時ともに、2 剤同時の用量変更を行わず、どちらか一方の用量のみを変更する。即ち、用量を増量と判断した場合には、図 15 の赤色で示した用量組み合わせである $(i + 1, j), (i, j + 1)$ の 2 つの候補用量組み合わせのどちらかに新規 cohort を割り付ける。また、減量と判断した場合には、図 15 の青色で示した用量組み合わせである $(i - 1, j), (i, j - 1)$ の 2 つの候補用量組み合わせのどちらかに新規 cohort を割り付ける。2 つの候補用量組み合わせから 1 つに選ぶ際には、BOIN Combination では、DLT 発現確率の事後分布に基づいて、当該用量組み合わせの DLT 発現確率が Target dose combination の範囲 (ϕ_1, ϕ_2) に存在する事後確率が最も高いとされる

用量組み合わせを選択する。ここで、 p_{ij} 、 n_{ij} 及び y_{ij} を用量組み合わせ (i, j) における DLT 発現割合、投与被験者数及び DLT 発現被験者数とすると、用量組み合わせ (i, j) が Target dose combination の範囲 (ϕ_1, ϕ_2) に存在する事後確率は以下の様に表すことが出来る。いずれの組み合わせでもこの事後確率が等しい場合は1つの組み合わせをランダムに選択する。

$$Pr(p_{ij} \in (\phi_1, \phi_2) | n_{ij}, y_{ij}) \quad (2.3.2.3.1.1.1)$$

ここで、 p_{ij} の事前分布には Beta (0.5, 0.5)を仮定する。

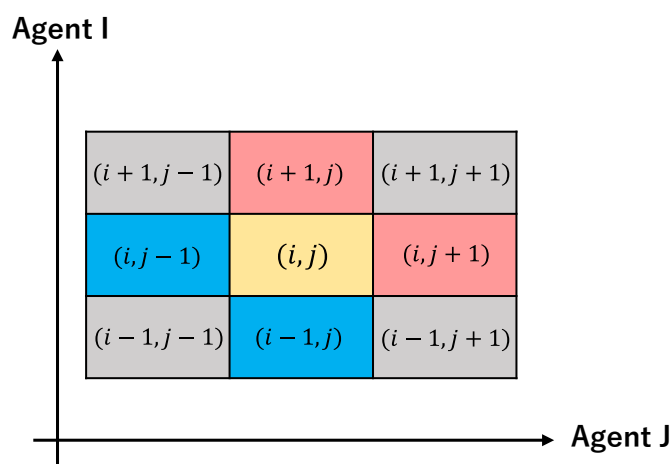


図 15 2 剤併用用量探索における用量組み合わせ

2.3.2.3.1.1.2. 用量除外規則及び試験の終了条件

BOIN Combination では、他の Model-assisted デザインと同様に安全性を考慮し、現在の用量組み合わせ (i, j) において、DLT 発現確率が目標 DLT 発現確率を超える事後確率がある閾値（例：95%）を超えている場合、その用量組み合わせへの組み入れを停止する。また、最低用量の組み合わせ $(1, 1)$ において、この基準に抵触した場合は、その時点で試験は中止される。そして、事前に計画した最大被験者数に到達した場合、あるいは用量除外規則に従って最低用量が除外された場合に、試験を終了する。

2.3.2.3.1.1.3. MTD の選択方法

予定していた被験者数のすべての DLT 評価を終えた後には、bi-variate isotonic regression を用いて DLT 発現確率を単調増加に調整し、目標 DLT 発現確率に最も近い DLT 割合を持つ用量の組み合わせを MTD とする。なお、 $(i, j-1)$ 、 (i, j) あるいは $(i-1, j)$ 、 $(i+1, j)$ の様に同一の用量*i*または用量*j*中に、目標 DLT 発現確率との距離が等しい用量組み合わせが存在する際には、 $p_{ij} < \phi$ を満たす中で最も高い用量、あるいは、 $p_{ij} > \phi$ を満たす中で最も低い用量の組み合わせを選択する。 $(i+1, j-1)$ 、 $(i-1, j+1)$ の様に異なる用量*i, j*中に Target

DLT rate との距離が等しい用量組み合わせが存在する際には、その中から、 $Pr(p_{i,j} \in (\phi_1, \phi_2) | n_{i,j}, y_{i,j})$ の値が最大となる用量組み合わせを選択する。

2.3.2.3.1.1.4. BOIN Combination のケーススタディ

BOIN Combination の実際の適用例は、2.3.2.3.1.1.5 節に示すが、試験の詳細の報告がないため、ここではケーススタディを用いて説明をする。

BOIN Combination の用量探索の例として、4 dose×4 dose の併用療法の用量探索の例を図16に示す。ここで、目標 DLT 発現確率を 0.33、目標被験者数を 60 例、cohort size を 2 とする。この時、BOIN による用量調整の閾値は、2.1.2.1.3 節の説明の様に、 $\phi_1=0.6TTL=0.6 \times 0.33$ 、 $\phi_2=1.4TTL=1.4 \times 0.33$ に基づいて、用量維持の区間の境界 (λ_e , λ_d) を (0.260, 0.395) と定めることが出来る。これに基づいて、用量の増減を判断する。まず、最初のコホートは、最低の用量組み合わせ(1, 1)に割り付けられ、ここで、2 例ともに DLT 発現が無く、用量維持の区間の境界 (0.260, 0.395) と照らし合わせて、次コホートでは、増量と判断される。そこで、増量可能な用量組み合わせである(1,2), (2,1)から増量後の用量組み合わせが選択される。この時、DLT 発現確率の事後分布に基づいて、Target dose combination の範囲(ϕ_1, ϕ_2)に存在する確率が最も高いとされる用量組み合わせを選択するが、まだ投与されたことのない用量組み合わせでは、選択される確率は等しくなる。そのため、(1,2), (2,1)からランダムに(1,2) が選択された。2 コホート目においても、DLT の発現は無く、ここでも増量が判断され、3 コホート目は同様のロジックに従い(2,2)に割り付けられる。6 コホート目まで増量が続き 6 コホート目は (4, 3) に割り付けられる。ここで、2 例中 1 例の DLT の発現が認められ、(4, 3)での DLT 発現確率は 0.5 となり、境界 (0.260, 0.395) と比較して次コホートでの減量が判断される。減量候補としては (3, 3) あるいは (4, 2) である。この時、(3, 3)と(4, 2)から、当該用量組み合わせの DLT 発現確率が Target dose combination の範囲 (ϕ_1, ϕ_2)に存在する確率が最も高いとされる用量組み合わせを選択する。ここで、(4, 2) が選択され、その後、60 例に達するまで BOIN Combination の dose escalation /de-escalation algorithm に基づいて割付が行われ、MTD として (4, 2) の用量組み合わせが選択される。

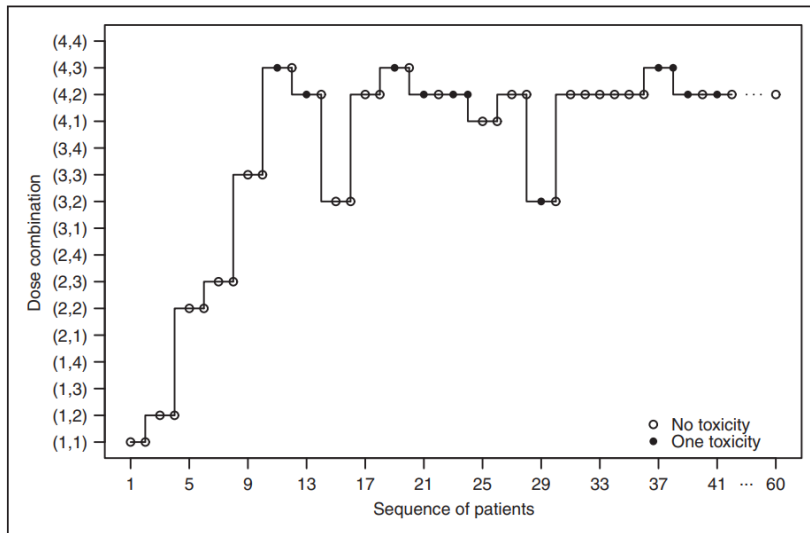


図 16 BOIN Combination 用量調整の例示

2.3.2.3.1.1.5. 実際の試験例

BOIN Combination は、実際に再発または難治性の急性骨髄性白血病試験において用いられている。CAR-T 細胞を抗原特異的ターゲットモジュールによって、腫瘍細胞に誘導する治療の試験であり、BOIN Combination を用いて CAR-T 細胞数及びターゲットモジュールの用量を探索している^[89]。

3. 適切なデザインの選択

3.1. 3つのクラスのデザインの特徴の整理

本節では2章で紹介してきた Rule-based/Model-based/Model-assisted の3つのクラスのデザインについて、Yuan 等 (2019) の報告^[6]でまとめられている表を参考に本タスクフォースでの検討を踏まえて追記・改編する形で、各クラスの方法の一般的な特徴を表 14 に整理した。

特に性能に関しては、3+3 デザインとの比較の文脈で model-based デザインや model-assisted デザインの性能が 3+3 デザインに優ることがシミュレーション研究を通じて様々に報告されてきている。例えば、Zhou 等 (2018) ^[90]では、modified CRM や BLRM に代表される model-based デザインや BOIN/Keyboard デザインに代表される model-assisted デザインが、3+3 デザインに比べて、正しい MTD を選択する確率や MTD で治療される確率が高いこと、真の MTD より高い用量を MTD として選択する確率や真の MTD より高い用量で治療される確率は 3+3 デザインに比べて同程度かやや高くなること、登録された被験者の 50%以上で MTD より高い用量で治療される確率は 3+3 デザインと同程度かやや低いことなどが示されている。また、Zhou 等 (2018) ^[90]では、modified CRM と mTPI /BOIN / Keyboard デザインの比較に焦点を当てて性能評価されており、上述と同様の評価指標において BOIN/Keyboard が modified CRM と同様の性能であることが示されている。同報告では、真の DLT 発現確率に関する用量反応関係が modified CRM の作業モデルで仮定した用量反応関係に近い場合は、modified CRM の方が BOIN/Keyboard よりも優れている傾向にある一方で、作業モデルで仮定した用量反応関係と真の用量反応関係が大きく異なるような場合は BOIN/Keyboard デザインの方が modified CRM よりも優れている傾向にあることも示されている。Liu 等 (2022) ^[91]では、2 剤併用療法の MTD を決定する状況において本稿で紹介した BLRM や BOIN Combination デザインの性能を 3+3 デザインとシミュレーション研究で比較しており、同様に Model-based/Model-assisted デザインの性能が優れることが示されている。Lin 等 (2020) ^[86]では、単剤療法の OBD を探索する状況において、3+3 デザインで決定した MTD のもとでの有効性コホート拡大パートで評価するという開発戦略に比べ、EffTox 法や BOIN12 のような Model-based/Model-assisted デザインの性能が平均的に優れることが示されている。

3.2 節では表 14 に整理した特徴を踏まえ、用量漸増デザインの選択の考え方を整理する。

表 14 3 つのクラスのデザインの特徴

特性	Rule-based	Model-based	Model-assisted
透明性/簡便性			
用量調整表を治験実施計画書で事前定義可能	○	×	○
推奨用量の推定に煩雑かつ繰り返しの計算を必要としない	○	△ ^{*1}	○
デザイン段階での事前検討の負荷の軽さ	○ ^{*2}	× ^{*2}	△ ^{*2}
統計担当者でない担当者の解釈の容易さ	○	×	△
柔軟性			
目標 DLT 発現確率や有効割合を明示的に設定可能	× ^{*3}	○	○
コホートサイズに依らない意思決定が可能	× ^{*3}	○	○
MTD や OBD で治療される被験者数が 6 例を超え得る ^{*4}	△ ^{*5}	○	○
望ましい動作特性を担保するためのサンプルサイズが設定可能	× ^{*3}	○	○
試験開始後の検討する用量レベルの追加に対する柔軟性	△ ^{*6}	○	○
用量増減時の用量選択肢の多さ ^{*7}	×	○	×
試験開始後の検討する用法の変更に対する柔軟性 ^{*8}	△	○	△
事前情報の利用可能性	× ^{*3}	○	○ ^{*9}
性能			
推定される MTD や OBD の正確性	△ ^{*10}	○	○
高い割合で MTD や OBD 近傍の用量に被験者を割り当てる	× ^{*3}	○	○
過量投与のリスクを制御可能	○	○	○ ^{*11}

*1 都度の推奨用量の推定は必要だがソフトウェアの利用や作成済みのコードの利用で必ずしも複雑にならないため

*2 デザインパラメータの検討に加え、一般に動作特性を評価するシミュレーション実験が必要となるが model-assisted デザインの方が model-based デザインに比べて検討するモデルパラメータが少なく負荷が軽いと考えられる。なお、本邦においては Rule-based のデザインであっても単純な 3+3 デザイン以外のデザインではシミュレーションが要求される^[92]。

*3 3+3 デザインなど例外はあるが一般には×

*4 RP2D 判断にあたって MTD や OBD と判定される用量でより多くの情報を得たいという需要があり得るためデザインの特徴の切り口として記述している

*5 選択するデザインによっては 6 例を超える被験者数を確保可能である

*6 減量を想定していない方法では低い用量を追加する場合に意思決定ルールが不透明になる

*7 Rule-based や Model-assisted デザインでは 1 段階ずつの増減量することが一般的であるが、Model-based デザインでは 2 段階以上の減量や、(安全性上の懸念が無い範囲で) 増量が可能なデザインを採用し得る

*8 前治療の追加や step-up dosing への変更、投与間隔や 1 回投与量の変更等があると、用量漸増デザインの抜本的な見直しを要するが、Model-based デザインは (強いモデルの仮定の下で) 変更前の情報を明示的に利用できる可能性がある

*9 事前情報を利用する場合の方法論も提案されている (model assisted design に関しては現在のところ single agent での方法論に限定されている)^[93]

*10 選択するデザインによっては isotonic regression など推定のための情報の有効活用がなされる

*11 mTPI は過量投与を推奨しやすい傾向にあることが報告されている

3.2. フローチャート

2章では Rule-based/Model-based/Model-assisted の3つのクラスで様々な方法を紹介した。実際のがん第I相試験で用いる方法は、表14に示すような特徴をはじめ、多くの要素を考慮して選択される。Revathiら^[94]がBOIN法とその拡張法の使い分けをフローチャートにまとめているものの、3つのクラス及びそれぞれに分類される方法を合わせ、方法選択の流れ及びその結果として推奨される方法をまとめたものはこれまで提示されていない。本節では、がん第I相試験の試験デザインの選択にあたっての試験デザインの大分類をフローチャート形式の図で整理し、大分類された試験デザインの中でどのような選択すべきかの推奨を可能な限り示すことを目的とする。フローチャート(図17)及び分岐点の解説を以下に示し、方法の推奨や複数の方法が推奨される場合はその使い分け、及び方法選択時の留意点などの図17の解釈を、図17にある大きな5つのフローに合わせて3.2.1-5節に示す。なお、本節に示す推奨は、これまでに得られているエビデンス等に基づく本タスクフォースとしての推奨であることに留意する。

図17に示した分岐点の解説

- ステップ1：単剤療法の開発であるか、併用療法の開発であるか。
 - 併用療法の開発であっても用量漸増する薬剤が1種である場合は、図17では単剤療法と同じとみなす。
- ステップ2：何をもって用量漸増の判断を行うか。DLTのみを考慮するのか、DLTと有効性またはバイオマーカーを同時に考慮するのか。
 - OBDの推定は、より明示的にDLTと同時に有効性やバイオマーカーを考慮したい場合に含まれると想定する。
- ステップ3：(Step2でDLTのみを考慮する場合) 用量漸増を早く進められる方法を特に選択したい状況であるか。

用量漸増を特に早く進めたい状況として、“注目しているDLTが遅発性である状況”、“患者登録のスピードが速く、次コホート開始までの待ち時間を少なくしたいと考えられる状況”などを想定する。

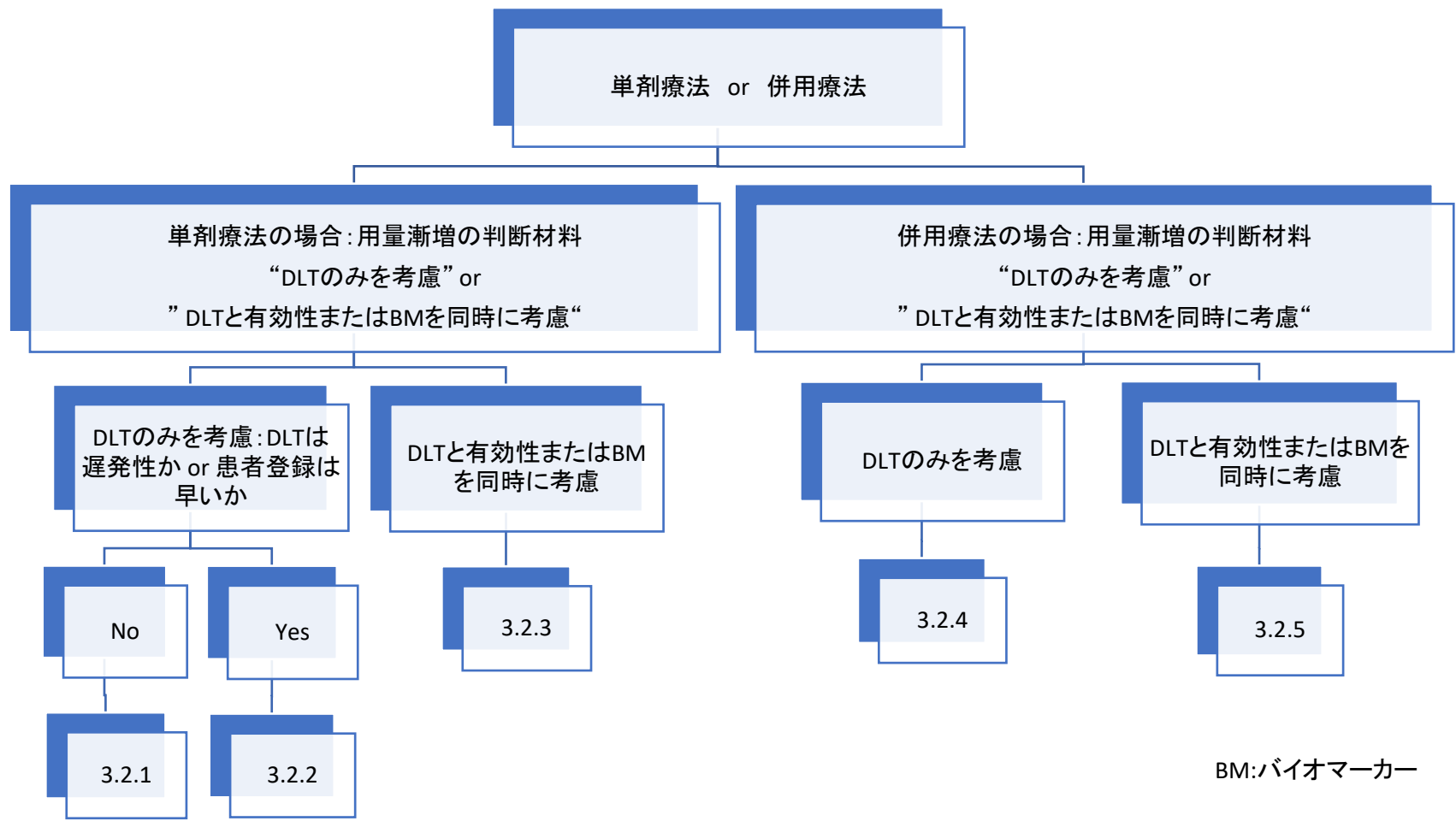


図 17 試験デザイン選択のためのフローチャート

3.2.1. 単剤療法かつ DLT に基づいて用量漸増する場合

DLT の発現頻度を指標に用量漸増の意思決定を行い、一定の終了基準を満たした時点で MTD を決定する。PK/PD や初期の有効性情報などを総合的に判断して RP2D を決定することになる。小規模な試験の中で MTD を正確に見極めることが RP2D を見極めるうえでも重要な課題であるため、表 14 にまとめた特性も踏まえて、Model-based デザインや Model-assisted デザインの利用が推奨される。用量と DLT 発現確率の関連についての事前情報が利用可能な場合にはそれらを明示的に利用することで意思決定の精度や正確さを向上し得るという観点でも Model-based デザインと Model-assisted デザインの利用が推奨できる。

Model-based デザインと Model-assisted デザインの使い分けにあたっては、用量反応関係に対する仮定の尤もらしさの検討が有用である。Model-based デザインでは、用量反応関係を明示的に仮定することによって、次コホートの推奨用量レベルの推定や、最終的な MTD の推定にすべての用量における DLT 発現の有無の情報を利用することができる。このように、用量反応関係を適切に表すモデルが利用可能であることは、Model-based デザインがより効果的に適用できる条件になりうるが、一般に用量反応モデルの適切性を検証することは困難である。このため、現実的には、事前のシミュレーション実験をとおして作業モデルが適切に MTD の選択に寄与できることが確認できる下で、Model-based デザインの利用は推奨できる。ただし、Model-based デザインでは推奨用量の推定や繰り返しの計算等、Model-assisted デザインに比べて相対的に手順が煩雑になり得る。また、多くのシミュレーション研究から、それらの研究で検討されている用量数や被験者数の範囲においては、Model-assisted デザインは Model-based デザインと同程度の動作特性を有することが示されている^[76]。このため、Model-based デザインの採用実績がなく実装に時間を要する場合や Rule-based デザインにより近い簡便な試験運営や意思決定手順が望まれる場合等では、Model-assisted デザインが選択肢となり得る。

Model-assisted デザインを採用する場合、本稿で紹介した BOIN 法と Keyboard 法（或いは mTPI2）の動作特性に大きな差異は報告されていないため^[76]、どちらのデザインを選んでも差し支えない。一方、mTPI はこれらの 2 つの方法に比べ過大な用量を推奨する傾向があるため推奨されない^[74]。

Model-based デザインを採用する場合、主に 1 パラメータで実装される mCRM と 2 パラメータ想定でかつ EWOC の考え方と親和性の高い BLRM を本稿では紹介した。一般に EWOC を採用した BLRM は、MTD を超える用量が投与される被験者の割合を減らすことが期待できるが、一方で正しい MTD を選択できる確率や MTD 付近で投与される被験者割合といった性能を犠牲にしている。これに対して、mCRM は MTD を超える用量が投与される被験者の割合を BLRM ほど厳密に制御できない代わりに正しい MTD を選択できる確率や MTD 付近で投与される被験者割合がより高くなることが期待される^[76]。これらの一

般的な特性を考慮したうえで、実際にケースごとに最適なデザインを選ぶためには複数のデザインの動作特性を評価するためのシミュレーション実験の実施が推奨される。

Rule-based デザインは、何らかの理由で Model-based/Model-assisted デザインが採用できなかった場合、または海外試験で RP2D が特定され日本では RP2D 付近の限られた用量での忍容性を確認する場合（例えば 2 用量以下など）に消極的に採用が検討され得る（4.6 節も参照すること）。この場合でも、DLT の発現が想定される限りは、MTD 推定の正確さを高めるという観点で、試験終了時に isotonic regression を適用する増減デザインや累積コホートデザインの採用を積極的に検討することが推奨される。また、加速型漸増デザインの single patient acceleration のアイデアは Model-based/Model-assisted デザインと組み合わせることが可能であり、開始用量が低く、検討用量レベル数が多い場合などでの積極的な検討が推奨される^[95]。

3.2.2. 単剤療法かつ DLT に基づいて遅発毒性/登録速度を考慮して用量漸増する場合

DLT の発現頻度を指標に用量漸増の意思決定を行い、一定の終了基準を満たした時点で MTD を決定する点、PK/PD や初期の有効性情報などを総合的に判断して RP2D を決定する点は 3.2.1 節の状況と同様である。ただし、DLT に遅発毒性を考慮する必要がある場合や患者登録のスピードが速いことが想定される場合に、DLT 発現の確認にタイムラグがあることで次コホート開始までの組み入れ停止期間が長くなることによる試験期間の長期化を回避できるデザインを検討したい状況である。

この状況で採用し得る試験デザインとして、2 章では Rolling-6、TITE-CRM、TITE-BOIN を紹介した。3.2.1 節と同様に、MTD をより正確に見極めることが RP2D の決定にあたって重要である点を考慮すると、ここでも Model-based/Model-assisted デザイン、即ち TITE-CRM や TITE-BOIN の選択が推奨される。これらのデザインの使い分けは 3.2.1 節と同様であり、事前のシミュレーション実験をとおして作業モデルが適切に MTD の選択に寄与できることが確認できる下では TITE-CRM が推奨される。一方、TITE-BOIN でも正しい MTD の選択割合や MTD 付近で投与される被験者割合といった動作特性が TITE-CRM と同等であり、かつ TITE-CRM より過大用量が投与される割合が小さいことが報告されている^[80]。このため、簡便な試験運営や意思決定手順が望まれる場合などは TITE-BOIN が推奨できる。

Rule-based デザインとして Rolling-6 デザインが提案されているが、何らかの理由で Model-based/Model-assisted デザインが採用できない場合に、試験期間の長期化を回避する目的で採用が検討され得る。

3.2.3. 単剤療法かつ DLT と有効性または BM を同時考慮して用量漸増する場合

MTD を優先的に決定したうえで RP2D を決定する 3.2.1 節や 3.2.2 節の状況とは異なり、DLT と同時に有効性や BM のアウトカムを直接的にモデルやルールに取り込んで用量調整の意思決定をする。そのうえで、MTD を考慮しつつ、有効性や BM のアウトカムを最大化する観点で OBD を決定し、RP2D 決定の参考とする状況である。このように、DLT だけでなく有効性や BM のアウトカムを同時考慮して用量調整し、OBD を推定するデザインは、Model-based デザインまたは Model-assisted デザインに限られる。これらのデザインの使い分けは 3.2.1.1 節と同様であり、事前のシミュレーション実験をとおして作業モデルが適切に OBD の選択に寄与できることが確認できる下では、Model-based デザインが推奨される。一方、Model-assisted デザインでも正しい OBD の選択割合や OBD 付近で投与される被験者割合といった動作特性が Model-based デザインと同等であることが報告されている^[85,86]。このため、簡便な試験運営や意思決定手順が望まれる場合などは Model-assisted デザインが推奨できる。

有効性や BM の指標として奏効の有無のような二値アウトカムを考慮する場合は、Model-based デザインとして 2 章で紹介した EffTox 法や bCRM の利用が考えられる。EffTox 法では毒性と効果のトレードオフを考慮した用量反応曲線を推定可能であるため、柔軟な追加用量の検討に資するかもしれない。また、よりシンプルに CRM を拡張した bCRM の利用も検討し得るが、いずれも実臨床での適用事例の報告は多くない。

Model-assisted デザインとしては、2 章で紹介した BOIN-ET デザインの利用や BOIN12 デザインの利用が考えられる。BOIN-ET は DLT 発現数に加えて有効性達成数を数えるだけで、BOIN デザインのように用量調整ができるデザインである。BOIN12 は DLT の有無或いは有効性/BM の反応の有無の組み合わせ毎に定義される utility とそれが得られる確率で定義される desirability score を最大化するように用量調整するデザインであり、有効性や BM が 3 カテゴリ以上のアウトカムにも応用できる。しかしながら、Model-based と同様にこれらの試験デザインの実臨床での適用事例は限定的である。

上述のとおり、Rule-based デザインとして同時考慮するために提案されている方法は確認されていない。ひとつの事例として Hunsberger ら^[36]は、毒性が無いと仮定した時に奏効例に基づき RP2D の決定をし、毒性が生じた場合は 3+3 デザインと同様のルールに移行するデザインを採用している。

3.2.4. 併用療法かつ DLT に基づいて用量漸増する場合

DLT の発現頻度を指標に用量漸増の意思決定を行い、一定の終了基準を満たした時点で MTD を決定する点、そのうえで PK/PD や初期の有効性情報などを総合的に判断して RP2D を決定する点は 3.2.1 節の状況と同様である。ただし、2 剤併用治療の用量漸増の検討を行う状況であり、用量調整を検討する範囲が 2 次元となる点が異なる。ここでも MTD

をより正確に見極めることが RP2D の決定にあたって重要である点を考慮すると Model based/Model assisted デザインが推奨される。これらのデザインの使い分けは 3.2.1.1 節と同様であり、事前のシミュレーション実験をとおして作業モデルが適切に MTD 推定に寄与できることが確認できている状況下では、Model-based デザインが推奨される。加えて、併用治療検討時には単剤での投与経験が利用可能なケースも想定されるが、そのようなケースでは Model-based デザインの事前分布の設定等に事前情報を利用しより効率的に MTD 推定が可能となることが期待される (Model-assisted デザインでは本報告時点では併用の文脈で事前情報を利用する方法は提案されていない^[93])。一方、Model-assisted デザインでも MTD の選択割合や MTD 付近で投与される被験者割合といった動作特性が Model-based デザインと同等であることが報告されている^[87,88]。このため、簡便な試験運営や意思決定手順が望まれる場合には、Model-assisted デザインが推奨される。

Rule-based デザインとして、実際に適用事例のある方法は特定できていないが、Ivanova らの方法が提案されている^[41]。しかしながら、この方法は目標 DLT 発現確率がコホートサイズに依存し、例えばコホートサイズが 2 の目標 DLT 発現確率は 0.3、コホートサイズが 6 の目標 DLT 発現確率は 0.11 となる性質があり、実務上の扱いが困難なため推奨できない。

なお、本文書執筆時点で 3 剤以上の併用療法を想定した用量漸増デザインに関する報告は確認できていない。ただし、技術的には本節でも引用している Model-based/Model-assisted デザインの方法を拡張することは可能と考えられる。一方、興味の併用療法が 3 剤以上の場合であっても、実際に用量を探索するのが 2 剤以下であれば、本節や 3.2.1 節と同様の考え方で試験デザインが検討できるため、用量漸増対象となる治療を限定することで試験を単純化するような検討は有用かもしれない。

3.2.5. 併用療法かつ DLT と有効性または BM を同時考慮して用量漸増する場合

併用療法を評価する状況下で、毒性だけでなく有効性或いは BM を同時に考慮して OBD を推定したいという状況があるかもしれない。理論的には 3.2.4 節や 3.2.4 節で触れたデザインを拡張することで実装可能と想定されるが、需要に限られるためかこのような状況を想定した試験デザインを確認することはできなかった。

4. 用量漸増デザイン選択に関わるその他の論点

4.1. 初回投与量、最高投与量、増量幅等の設定の一般的考え方

ヒトで初めて行う (FIH、First In Human) 試験での初回投与量、最高投与量、増量幅等の設定の一般的考え方は、令和 5 年 2 月 15 日時点で ICH S9 ガイドライン「抗悪性腫瘍薬の非臨床評価に関するガイドラインについて」^[96]、「「医薬品開発におけるヒト初回投与試験の安全性を確保するためのガイダンス」の改訂等について」^[97]及び「「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」について」^[98]等に示されており、以下に要約する。

4.1.1. 初回投与量の設定

ヒト初回投与量を慎重に設定することは、被験者の安全性を確保するために重要である。FIH 試験での初回投与量及び投与スケジュールは入手可能なすべての非臨床試験成績 (薬物動態、薬力学、毒性など) を考慮して科学的に裏付けられるべきである。

一般にヒト初回投与量は、最も感度の高い動物種を用いた非臨床毒性試験における無毒性量 (NOAEL、No Observed Adverse Effect Level) をもとに、アロメトリック補正、あるいは、薬物動態 (PK) 情報に基づいてヒト等価用量 (HED、Human Equivalent Dose) を算出し、さらに被験薬の特性や臨床試験デザインを踏まえた安全係数を考慮し設定される。

殺細胞性の抗悪性腫瘍薬では、原則、げっ歯類で供試動物の 10% に重篤な毒性が発現する投与量 (STD10、severely toxic dose in 10% of animals) の 1/10 量または非げっ歯類で重篤な毒性が発現しない最大投与量 (HNSTD、highest non-severely toxic dose) の 1/6 量に基づいて初回投与量を設定する。

また、抗悪性腫瘍薬に限らず、全身投与される低分子医薬品の多くにおいては、一般的に薬剤の投与経路、特性等に応じて、動物での投与量 (mg/kg) を体表面積 (mg/m²) に換算した上で、初回投与量を設定することを検討する。低分子及びバイオ医薬品のいずれにおいても、体重、AUC、その他の曝露量パラメータに基づいて投与量を外挿することが適切な場合もある。

免疫系に作用する薬物のうち、ヒトに対して高い特異性を示すバイオ医薬品では、薬理作用を示す動物種が存在せず、FIH 試験の初回投与量の設定が困難な場合も想定されるが、できる限り薬理評価と毒性評価を行っておくことが望ましい。難しい場合は、ヒト細胞を用いた *in vitro* における薬理作用の評価及びヒト組織交差反応性試験等に基づく毒性評価を実施しておくことが望ましい。免疫系に対しアゴニスト作用を有する可能性のあるバイオ医薬品については、非臨床試験の結果から予期できない有害作用が発現し得るリスクを考慮し、推定最小薬理作用量 (MABEL、Minimally Anticipated Biologic Effect Level) 等の科学的根拠に基づいて初回投与量を設定すべきである。

なお、ヒトへの初回投与量を設定する上で、NOAEL、MABEL等の設定根拠の違いにより異なる値が得られた場合は、科学的根拠に基づいてより慎重にFIH試験の初回投与量を決定する必要がある

4.1.2. 増量計画と最高投与量の設定

殺細胞性薬、分子標的薬のいずれであっても有害事象で用量が制限される場合には、DLT発現状況等に基づきMTD及びRDを決定する。MTD及びRDの探索にあたり、用量をどのように、どのレベルまで（増量計画及び最高投与量）増量するか、治験実施計画書に記載すべきである。なお、現時点において、最高投与量の決定に関する各局の指針はない。

通常、がん患者を対象とした臨床試験における増量計画または最高投与量は、非臨床試験で検討した最高投与量あるいは曝露量による制限を受けないが、一般的に殺細胞性の抗悪性腫瘍薬は有効性の期待できる治療域と中毒域が近接している。このことに十分留意し、治験薬の増量の方法を検討すべきである。一般的な増量法として伝統的方法であるFibonacciの変法を用いた3例コホート法（下表参照）等を用いることもあるが、科学の進歩に従って第2及び3章を参照し、適切な用量漸増デザインを採用することも可能である。いずれにせよ、治験薬の用量幅（増加計画）は、非臨床試験における用量-毒性曲線の勾配や薬理試験成績等に基づいて慎重に進めるべきである。その際には、用量-曝露-反応曲線の勾配や患者間の不均一性等も考慮し、既承認の類似薬がある場合はその臨床試験や非臨床試験成績等も参考とする。例えば、非臨床毒性試験において、重篤な毒性に関する急峻な用量あるいは曝露量反応曲線が得られる場合、またはあらかじめ重篤な毒性に関する適切なマーカーがない場合は、通常よりも小刻みな増量計画（例えば公比2以下）を考慮すべきである。

以上を踏まえた上で、投与量はMTDを決定するまで、または生物学的な効果が得られる用量レベルまで慎重に増量する。

なお、次コホートへの増量判定基準は事前に非臨床試験データ、類似医薬品データを基に試験デザインとともに検討し、治験実施計画書に記載すべきである。また、ヒトにおける情報がない場合での用量漸増幅の設定は不確実性を含んでおり、それまでに得られている各用量レベルのコホートから得られたPK/薬力学及び安全性の情報をもとに、用量レベルの追加など投与量、用量漸増手法を見直すこともありうる。このような場合のために、治験実施計画書に投与量変更の可能性とその手順を記載しておくべきである。

用量レベル	投与量（例）	増量率
1	X (15 mg)	
2	2X (30 mg)	100%
3	3.3X (50 mg)	67%
4	5X (75 mg)	50%

用量レベル	投与量 (例)	増量率
5	7X (105 mg)	40%
6	9.3X (140 mg)	33%
7	12X (180 mg)	29%

なお、原則として同一被験者での増量は行わない。ただし、ガイドラインに示された条件を満たす場合、同一被験者での増量が可能なこともあるが、増量した被験者の忍容性は安全性の評価には用いても増量後の用量における DLT 評価には用いない。

4.2. サンプルサイズ設定の考え方

Rule-based デザインと model-based デザイン並びに model-assisted デザインではサンプルサイズの設定方法について考え方が異なる。Rule-based デザインは実際の DLT 発現状況と試験終了のルールに依存してサンプルサイズが試験終了時に決定される。一方で、model-based デザイン並びに model-assisted デザインは MTD 特定割合が高くなるようなサンプルサイズを事前に設定することができる。サンプルサイズ設定はシミュレーションを通して行う必要がある。Optimal Selection of Adaptive Designs in Phase I Oncology Clinical Trials^[99]では、3+3 デザインをはじめとする Rule based デザインに加え、CRM や bCRM のような model-based デザイン、BOIN や mTPI のような model-assisted デザインなど幅広い方法が利用可能であり、MTD 特定割合に加え、MTD や MTD より高い用量で治療される被験者数などを含む様々な指標により、サンプルサイズを探索することができる。Model-assisted デザインでは Integrated Platform for Designing Clinical Trials^[100]を用いることで、BOIN、TITE-BOIN、BOIN12 などのサンプルサイズ設計を行うことができる。サンプルサイズ設定はシミュレーションを通して行う必要がある。MTD や OBD 特定の試験では、検出力 80%以上のような目安となる指標がないため、Iasonos et al. (2008) ^[101]では想定されるサンプルサイズの目安として、固定 8 用量の試験で、20 から 30 例としている。

4.3. ソフトウェア

前述の Optimal Selection of Adaptive Designs in Phase I Oncology Clinical Trials^[99]では Rule-based デザイン、model-based デザイン並びに model-assisted デザインの動作特性を一度に検討することができる。さらに以下で示すように多くの基本的なデザインについては実装可能なソフトウェアが公開されているものの、想定する臨床試験デザインが種々の modification を含む場合にはそのまま利用できないことがあるため、各種仕様の精査を行ったうえで状況によってはコードの修正や開発が必要となる点には留意する。

Model-based デザインについては、各手法に対応した R package (CRM, TITE-CRM: dfcrm package^[102] ; BLRM: blrm package^[103], Oncobayes2 package^[104] ; CRM, EWOC: bcrm package^[105]など) を利用することでシミュレーションによる動作特性の検討や実データに基づく次用量の選択及び MTD の推定が可能となる。さらに、MD Anderson Cancer Center

が提供しているデスクトップアプリケーション^[106]を利用することで、EffTox 法や bCRM についても容易に実装することができる。

Model-assisted デザインでは、BOIN、Keyboard、TITE-BOIN、TITE-Keyboard、BOIN12、BOIN Combination、Keyboard Combination デザインの用量調整を行うための用量調整表を Integrated Platform for Designing Clinical Trials^[100]より作成することができ、実データに基づく次用量の選択及び MTD (BOIN12 は OBD) の推定や動作特性の検討も可能である。

また、有料ソフトウェアとして EAST[®]、EAST Bayes[®]、ADDPLAN[®]などが幅広いデザインに対応しており、無料ソフトウェアに比して品質保証の観点で優れていることが多いため、これらの利用も併せて検討し得る。

4.4. 総合判断という考え方

4.4.1. 次コホートの推奨用量の決定

いずれのデザインを採用した場合でも次コホートの推奨用量の決定は、基本的に治験依頼者、治験責任医師及び外部の専門医等によって構成される安全性審査委員会によって実施される。単剤で DLT に基づいて用量漸増をするような試験において Model-based デザインや Model-assisted デザインを採用した場合、本委員会にその時点までの DLT 評価に基づく Model による推奨結果 (推奨用量や増量・維持・減量の判定) が提供される。本委員会では DLT の評価やモデルによる推奨結果だけでなく、安全性、PK、薬力学及びその他臨床効果に関するデータ等を考慮した上で推奨用量等が決定される。例えば、Model-based デザインにおいて使用したモデルが次コホートの推奨として増量を示したとしても、DLT 以外に注意すべき有害事象が発生していたり、PK 濃度が想定外の推移を示していたりする場合には、安全性審査委員会では維持または減量を選択する可能性がある。一方で、安全性、PK、薬力学及びその他臨床効果に関するデータ等に大きな懸念がなければ、事前評価している性能を担保するためにも基本的にはモデルによる推奨結果が採用されるべきである。

4.4.2. RP2D の決定

RP2D も基本的には安全性審査委員会の評価によって決定されるが、予め治験実施計画書に推奨用量の判定に係る基準を定める場合もある。その例を以下に示す。

	設定基準	設定理由
1	MTD は少なくとも x 例以上で評価する	MTD を決定する際の被験者数が極端に少なくならないように設定される。3+3 試験デザインを参考に 6 例と設定される場合もある。
2	DLT が 1 例に発現した場合、当該用量では少なくとも x 例以上で評価する	想定している DLT 発現確率を考慮し、DLT が 1 例に発現した際の必要最低限の評価被験者数として設定される（例：真の DLT 発現確率を 33%と想定している場合は 3 例以上で評価する）
3	PK データを評価するために 1 コホートあたり少なくとも x 例以上で評価する	RP2D を総合評価するにあたって PK データも考慮するために設定する。

このような基準は統計担当者が決定する訳ではなく、臨床担当者やメディカルドクターが中心になって決定する。

なお、一章でも述べた通り、FDA は Project Optimus を開始している。早期に用量の最適化を図るためにも、安全性、臨床活性、PK のデータに加え、バイオマーカーを用いる潜在的な有効性の評価や、有効性及び安全性の両方に対する用量/曝露（量）－反応関係性の評価が重要である。

RP2D を決定する際には、用量選択の強固な根拠を適切に説明できるようにしておくことが今後益々求められると考える。実臨床への応用例として、がん第 I 相試験で MTD 及び RP2D の候補用量を検討し、別途実施する第 II 相試験で RP2D の複数の候補用量を評価して用量最適化のためのエビデンスを集積するといった方策が考えられる。また用量漸増パートの後に続けてコホート拡大パートが設定されているようながん第 I 相試験では、コホート拡大パートで複数の RP2D の候補用量を評価する等の方策も考えられる。

4.5. 動作特性の評価

4.5.1. 概要

数ある用量探索試験のデザインから実際に試験に適用するデザインを選択する際には、その方法の動作特性を評価し、実際の試験の目的にかなう特性を持つデザインを選択する必要がある。動作特性の評価においては、まず想定される被験者数、用量反応関係、デザインによっては、組み入れスピードなどを仮定する。そして、それら条件の下で候補となる試験デザインがどのような特徴を持つのかを主にシミュレーションにより調査をする。これにより、研究者が目的に沿った試験デザインの選択をすることが可能となる。また、動作特性の評価結果は、治験依頼者が適用した試験デザインが倫理的及び科学的に妥当であることを対外的（規制当局や治験責任医師等）に示すためにも重要な役割を果たすと考えられるため、シミュレーションの設定や結果を治験実施計画書に記載することが推奨されるだろう。PMDA が公開している 30 日調査照会事項チェックリスト（抗悪性腫瘍剤分

野) [92]においても、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て用いる場合も含む) 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 MTD を上回る用量で治験薬が投与される被験者が最小限になることを含め、デザインの適切性を説明することが求められている。

動作特性の評価は多くの場合においてシミュレーションを伴うものとなるが、海外試験において RP2D が決定された後の後追い試験などで、国内で 1, 2 コホートの小規模な試験を行う際には、厳密な意味での MTD や OBD の推定を目的とはしておらず、動作特性の評価自体が必要とされない場合も考えられる。後追いが第 I 相試験についての詳細な議論は 4.6 節を参照すること。

4.5.2. 動作特性評価のためのシナリオ及び評価指標

シミュレーションにより動作特性を評価する際、想定する試験デザインに対し複数のシナリオを用い検討することが有用である。DLT の発現有無により単剤の MTD/RD の決定を目的とする試験の場合、シナリオとは各用量における真の DLT 発現確率の設定が該当する。例えば、全ての用量が MTD 未満 (すなわち検討予定の全用量レベルの真の DLT 発現確率が許容可能な範囲) のシナリオ、途中の用量で MTD を超えるシナリオ等が挙げられる。表 15 にシナリオ例を示す。また、安全性の観点から全ての検討予定用量が MTD を超えるシナリオにおいても検討し、適切に試験を中止することが出来るかを確かめることもあるだろう。日本が参加予定の試験の場合は、シナリオを検討するにあたり、PMDA が公開している 30 日調査照会事項チェックリスト (抗悪性腫瘍剤分野) [92]も参照されたい。

なお、Pseudo uniform algorithm など単調性などの仮定の下でランダムにシナリオを生成し、性能評価に用いる方法が提案されているが^[91,107]、治験実施計画書における検討に用いられる様な特定のシナリオの設定は、性能の評価及び解釈のしやすさの観点から、臨床担当者などの意見を取り入れ、統計担当者及び開発チームで設定をすることを推奨する。

表 15 単剤用量選択時のシナリオ例 (目標 DLT 発現確率 0.3)

シナリオ例	真の DLT 発現確率		
	用量レベル 1	用量レベル 2	用量レベル 3
全用量が MTD 以下	0.05	0.15	0.30
一部用量で MTD 超	0.15	0.35	0.55
全用量 MTD 超	0.35	0.45	0.50

上述の様に、デザインの検討時には、複数のシナリオに対し、想定する試験デザイン (採用予定の Rule-based/Model-based/Model-assisted の手法、用量漸増の基準や MTD/RD の決定基準、各コホートの被験者数や総被験者数、試験終了条件等) を反映したシミュレー

ションにより動作特性の評価を行うことで、試験デザインが目的にかなう特性を示すかを確認することになる。

なお、より正確に想定する試験デザインの性能を評価するという観点においては、シミュレーションで想定する試験デザインは MTD/RD の決定基準や試験終了条件などを可能な限り実際の治験実施計画書で定められた基準に一致させる。

動作特性の評価における評価指標としては、例えば、単剤の開発時においては、試験全体の被験者数の平均や MTD を超えるような過大用量の投与を受ける被験者数の平均、過少用量を MTD/RD として選択する割合、適正用量を MTD/RD として選択する割合、過大用量を MTD/RD として選択する割合等、試験の目的に応じて設定可能だろう。例えば、遅発毒性を考慮する試験デザインの検討時においては、試験終了までの期間なども考慮することが望ましい。試験デザインの動作特性を評価し、試験の目的にかなわず許容できない性能であればデザインを見直す必要があるだろう（例えば、いずれのシナリオにおいても過大用量を MTD/RD に選択してしまう割合が非常に高い場合には、MTD/RD の決定基準を見直す等）。

4.6. 後追いがん第 I 相試験デザインの考え方

近年、抗悪性腫瘍薬の開発においては、第 I 相試験実施後すぐに承認申請の主要な根拠となる試験（ピボタル試験）が国際共同治験という形で実施され、承認申請が行われるケースもよくみられる。このような状況で欧米から遅れず本邦で承認申請をするためには、日本が遅滞なくピボタル試験に参加し、海外と同時に開発を進めることが重要となってくる。

「国際共同治験に関する基本的考え方について」^[108]や「国際共同治験開始前の日本人での第 I 相試験の実施に関する基本的考え方について」^[109]では、特定の条件を満たす場合、日本人での忍容性の検討をスキップして国際共同治験に参加することも可能と記載されている。具体的には、海外でヒトでの忍容性が既に確認されており、民族的要因が安全性に大きな影響を及ぼさないと考えられ、国際共同治験に組み入れられる日本人の安全性が確保される場合などがあげられる。これらの根拠を示すことができた場合でも日本人の忍容性の検討が必要なのか、あるいはどのような根拠があれば日本人での忍容性の検討をスキップ可能であるかについては議論が多いところではあるが、抗悪性腫瘍薬については、比較的強い毒性が予想され安全性の確保に懸念がある等の理由から、多くの場合、ピボタル試験参加前に日本人での忍容性・安全性の検討・評価が求められているのが現状である。

抗悪性腫瘍薬の開発では、海外で FIH の第 I 相試験が進行し、日本での開発が海外から遅れて計画される状況がある。この時、日本がピボタル試験に参加するため、FIH の第 I 相試験を後追いするような国内第 I 相試験を実施することや、FIH の第 I 相試験に日本コホートを追加設定することが検討され（以下、これらを後追い第 I 相試験と呼ぶ）、日本人の忍容性確認が行われる。FIH の第 I 相試験に日本コホートを追加する場合は、海外 FIH の

試験と同一試験として実施でき、別試験として実施する場合と比較してドキュメントなどの管理統一や第一被験者の初回投与（FPD、First Patient Dosing）開始までの期間短縮、解析環境に関わるコスト削減という観点などで優れる可能性がある。なお、後追いのため、日本から試験参加する時期の問題で FIH の第 I 相試験のメインパートには参加できない状況を想定している。

後追い第 I 相試験での忍容性検討は、安全性や薬物動態/薬力学等の点から民族差がないことが予見できる場合、FIH の第 I 相試験で推定された MTD または決定された RP2D を日本でも共通で使用する想定で進められる。そのため、後追い第 I 相試験では、日本独自に MTD の推定や RP2D の決定をすることではなく、FIH の第 I 相試験で決定された RP2D が日本人でも忍容であるかどうかを確認することを主な目的として、少数の日本人患者を対象に忍容性評価を行うことが検討される。「後追い」のタイミングについて、RP2D が決定する前後の区分が考えられる。以下に後追い第 I 相試験をデザインする上で考慮すべき点について説明する。

4.6.1. 用量漸増デザインの選択

日本での後追い第 I 相試験のデザインは、海外で得られたデータが利用可能なため、検討用量レベルの範囲を限定することが可能であろう。海外で RP2D が特定された後であれば、日本での忍容性確認が試験の目的となる可能性が高い。また、RP2D が事前に分かっているということを事前情報として利用できるため、どのデザインを選択する場合でも検討用量レベルをさらに限定することが可能となる（4.6.2 節）。

以下に海外の開発進捗状況別（RP2D 決定前後）に、日本での後追い第 I 相試験について検討可能な主な用量漸増デザインを示す。RP2D 決定前の開始は、海外の進捗から可能な限り遅延しないように日本での開発をする場合、RP2D 決定後の開始は、次相への移行確度が高まった後に日本での開発をする場合などが考えられる。

<RP2D 決定前>

日本でがん第 I 相試験を計画実施する場合、計画時点までに FIH 試験で得られている事前情報を積極的に取り入れることが可能な Model-based や Model-assisted デザインを選択することが検討される。計画時には予定していない検討用量が追加された場合、3+3 デザインなどでは追加された用量に対するルールも追加しなければならないが、Model-based や Model-assisted デザインでは構築したモデルで対応可能な用量であれば追加対応は不要となる場合もあるからである。

FIH 試験に日本コホートを追加する場合、同一の治験実施計画書で試験を実施する一貫性の観点から FIH 試験で適用されているデザインと同様のものを選択することが検討される。また、日本で第 I 相試験を計画実施する場合と比較して、FIH 試験の治験実施計画書などのドキュメントが利用可能なため、計画段階から FPD までの期間を短くすることがで

きると考える。また、FIH 試験が実施中の場合、より最新の情報などを日本コホートの計画段階で反映できる可能性があり、RP2D に近い用量まで DLT が認められておらず、安全性が十分高いことが確認される場合も考えられる。このとき、RP2D 付近の限られた用量範囲で忍容性確認のために 3+3 デザインを選択する可能性も考えられる。

<RP2D 決定後>

日本でがん第 I 相試験を計画実施する場合、FIH 試験で RP2D が決定後なので、RP2D での忍容性確認が主目的となることが想定される。また、RP2D 付近の限られた用量範囲のみで忍容性確認を行うことが可能な場合などは、3+3 デザインを選択する可能性も考えられる。他国で承認済みであり、その後国際共同治験に参加する選択肢がない場合は、日本での忍容性を第 I 相試験で確認した直後に有効性を検証できるような第 II 相試験などのシームレスなデザインを検討することも考えられる。なお、他国での承認後については、開発ライセンスを別会社から取得し、日本でも開発を開始するような状況を想定している。

FIH 試験に日本コホートを追加する場合、同様に用量レベルを限定して RP2D が忍容可能であるかの評価が目的となることが想定される。しかし、RP2D 決定後であるため、現実的に FIH 試験に忍容性確認のための日本人コホートが設定可能なのか検討が必要である。

4.6.2. 開始用量及び検討用量の設定

日本での後追い第 I 相試験では、日本独自にフルスケールの用量漸増試験を実施する必要はなく、「抗悪性腫瘍薬の臨床評価方法に関するガイドライン（新ガイドライン）」^[98]や ICH E5 ガイドライン^[110]に沿って海外の試験成績（DLT を含む安全性データ等）を利用して開始用量を決定することが可能である。そのため、海外データがある場合は、開始用量は海外で検討済みの用量から選択することが多い。

検討用量の範囲については、海外で決定された RP2D での忍容性確認が主目的となるため、多くの場合、RP2D 以下となる。後追い第 I 相試験では、コストやスピードの観点から検討する用量レベルを少なくすることが望まれることも多いが、実施すべき検討用量レベルの範囲については海外データや類薬のデータなど利用可能な情報をもとに決定される。例えば、海外第 I 相試験が RP2D 決定前の状況であり、利用可能な情報から日本人での安全性確保に懸念がない場合、後追い第 I 相試験の開始用量として、海外第 I 相試験で忍容性が確認されている最高用量を選択することが検討される。また、RP2D 決定後の状況で海外での十分な投与経験に基づき、全体的な安全性プロファイルが良好であり（例：DLT が認められていない）、顕著な民族差がないと判断できる場合は、RP2D の 1 用量のみで忍容性の評価を実施することが検討される。一方で、海外データに基づき日本人に RP2D や海外第 I 相試験で忍容性が確認されている最高用量を投与する上で被験者の安全

性確保に懸念がある場合、当該用量よりも低い用量からの検討が求められ、複数用量レベルの検討が必要になる。例えば、海外第 I 相試験において RP2D よりも低い用量で DLT の発現が認められ、当該用量から検討する際に安全性が懸念される場合は DLT が認められていない最高用量が開始用量として選択され得る。

4.6.3. 被験者数・コホートサイズの設定及び忍容性確認の基準

日本での後追い第 I 相試験において、ある用量について忍容性の評価を行う場合、慣習的に 3+3 デザインの考え方にに基づき、1 用量あたり少なくとも 3 例で実施されることが多い。コホートサイズについては選択するデザインにより様々な設定が考えられる。例えば、3 例で評価中の DLT 発現時の追加症例数について、3+3 デザインでは 3 例を追加することになるが、Model-based や Model-assisted デザインでは比較的コホートサイズを柔軟に設定可能であるため、2 例を追加し DLT 発現が無ければ増量という場合もある。このように Model-based や Model-assisted デザインでは、DLT が発現し、同一用量レベルに症例を追加する際にコホートサイズが 3 例に縛られないことによる柔軟性を利用した事例もある [111]。

また、日本での後追い第 I 相試験の主目的は、日本人患者を対象に忍容性確認となる場合が多く、どのような基準で忍容性確認されるかなど忍容性評価を終了する条件や想定する最大症例数を治験実施計画書に規定する必要がある [92]。例えば 3+3 デザインでは、6 例のうち DLT が 1 例以下の発現であれば、増量可能または MTD となるので、このルールを参考に DLT 発現確率が 16.7%以下の場合を忍容性確認の基準とすることも考えられる。また、DLT が 2 例発現した場合、症例追加または MTD は 1 用量下となるので、DLT 発現確率が 33.3%以上の場合、忍容できない用量の基準と考えることもできる。このように各基準を 16.7%と 33.3%にした場合、乖離が大きいと感ずることもあるだろう。従って、忍容性確認の基準についてもある程度柔軟に規定できる可能性を考慮すると、Model-based や Model-assisted デザインを選択することを第一に検討することが考えられる。なお、個々の開発状況を考慮した時に 3+3 デザインの選択を否定するものでもない。

4.7. DLT が発現しないことが想定される場合のデザイン選択

近年開発されている多くの抗悪性腫瘍薬は分子標的薬や免疫療法薬であり、従来の殺細胞性の抗悪性腫瘍薬とは抗腫瘍効果のメカニズムも治療域も大きく異なる。実際このようなクラスの薬剤では、がん第 I 相試験の計画段階で得られている非臨床試験の結果やその作用メカニズムから、(中長期の毒性として定義する場合も含めて) DLT が発生しないことが想定されることも考えられる。このような状況では、DLT の発現が想定されないことの確からしさや、RP2D 選択にあたって追加でデータ取得する方策を組み合わせ、いくつかのデザイン選択方針が考えられる。

例えば、DLT の発生は想定しないものの、臨床で何が起こるかは実際には未知であることを踏まえ、DLT が出る前提で用量漸増していくことが考えられる（3.2 節のフローチャートで言えば、3.2.1 節、3.2.2 節、3.2.4 節に対応する方法を検討することになる）。この場合、想定どおり DLT が発現せず MTD に到達しなかった場合には、有効性や BM、PK などの情報に基づいて RP2D を決定するために別途被験者を追加して既存の用量レベルや追加の用量レベル（当初定義した最大用量レベルを超える用量の検討もあり得る）を評価することなどを事前想定したデザインを採用することになるかもしれない。用量漸増中であっても既に一定数の被験者の投与経験がある用量で被験者を追加し、総合判断のための情報を収集するバックフィルという考え方も検討され得る。また、用量漸増デザインとしても、DLT が発現しない想定がより確実と考えられる状況では、正しい MTD の選択確率といった動作特性などを考慮する必要がないため、より簡便性を重視して 3+3 デザインのような rule-based の方法を積極採用することもあり得るかもしれない。

また、DLT が出ないことが想定されるのであれば、有効性や BMなどを同時考慮して用量漸増し、OBD を推定しに行くというアプローチが考えられる（3.2 節のフローチャートで言えば、3.2.4 節、3.2.5 節に対応する方法を検討することになる）。この場合、DLT と有効性や BM とのバランスで次コホートの推奨用量や最終的な OBD/RP2D が決定されていくため、実際の DLT の発現の有無（想定の正しさ）をあまり気にせずにデザインできる点は利点といえる。一方、DLT が発現しない想定がより確実と考えられる状況では DLT を同時考慮する根拠に乏しくなるため、有効性や BM のみに基づいて RP2D を決定する試験デザインを採用することも考えられるかもしれない。

5. おわりに

近年の抗悪性腫瘍薬の開発は、治療の個別化や競争の激化、これらに伴う開発費の高騰やモダリティの多様化を受けて、開発計画の在り方の多様性を増している。このような状況にあって、3+3 デザインのような古典的ながん第 I 相試験を無条件に適用する時代は終わったと言えるのではないか。開発対象となる薬剤の有効性や安全性に関する情報量が限られる状況でも、背景や目的に合ったがん第 I 相試験デザインを選択することは、適切な RP2D の選択を通して、開発効率や開発成功確率の向上、上市に至った際の患者のベネフィットリスクバランスの最大化に資することが期待される。

本タスクフォースでは、統計家だけでなく開発計画や試験デザインを共に検討する開発担当者が、協働して適切ながん第 I 相試験のデザインを選択する際に参照頂くことを意識して本報告書を執筆した。この目的のために、Rule-based/Model-based/Model-assisted の 3 つのクラスごとに主な試験デザインを概説するだけでなく、実用事例やケーススタディを与えてデザインの理解の補助とした。また、3 つのクラスのデザインの特徴を表形式で整理するとともに、開発対象薬剤の背景や目的をフローチャートで大分類し、分類された状況ごとに用量漸増デザインの選択に関する推奨を与えた。またデザインの選択に関わるその他の論点をいくつかの視点で整理した。

本報告書では事例があり頻繁に利用され得るいくつかの代表的な試験デザインを紹介したが、他にも多くの試験デザインが提案されており網羅性は十分でないことに留意が必要である。より最新の優れた試験デザインの採用は常に検討されるべきであるが、その場合であっても、デザインが 3 つのクラスのいずれかに分類されるような状況であれば本報告書はデザイン選択時の議論のたたき台として活用可能と考える。

なお、本報告書は適切ながん第 I 相試験デザインの選択とその下での適切な RP2D の選択は今後も重要であることを想定して執筆している。一方、新しい潮流を踏まえてがん第 I 相試験への要件が大きく変わってきた場合には、試験デザインの選択の考え方も大きく変わり得ることに留意が必要である。その最たる例として Project Optimus が考えられる。すなわち、がん第 I 相試験で RP2D を決定することよりも、RP2D が存在することが想定される用量の範囲や RP2D の候補となる用量を効率的に絞り込むことに重きを置く場合、あるいは PK/PD のアウトカムを明示的に考慮する場合や、患者報告の有害事象重症度などを考慮する場合は、これらの目的に合った試験デザインの最適化が求められるかもしれない。

6. 参考文献

1. 「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」について
<https://www.mhlw.go.jp/hourei/doc/tsuchi/T210401I0060.pdf>
2. ICH E8 臨床試験の一般指針 <https://www.pmda.go.jp/files/000156372.pdf>
3. Daniel V. Araujo, Marc Oliva, Kecheng Li, Rouhi Fazlzad, Zhihui Amy Liu, Lillian L. Siu. Contemporary dose-escalation methods for early phase studies in the immunotherapeutics era. *European Journal of Cancer*. 2021; 158:85-98.
4. O'Quigley, J., Pepe, M., and Fisher, L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics*. 1990; 46:33-48.
5. Ji Y, Liu P, Li Y, Bekele BN. A modified toxicity probability interval method for dose-finding trials. *Clin Trials*. 2010; 7(6):653-63.
6. Yuan Y, Lee JJ, Hilsenbeck SG. Model-Assisted Designs for Early-Phase Clinical Trials: Simplicity Meets Superiority. *JCO Precis Oncol*. 2019; 24:3.
7. Fraisse J, Dinart D, Tosi D, Bellera C, Mollevi C. Optimal biological dose: a systematic review in cancer phase I clinical trials. *BMC Cancer*. 2021 Jan 13;21(1):60.
8. Project Optimus <https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus>
9. Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases. FDA draft guidance. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/optimizing-dosage-human-prescription-drugs-and-biological-products-treatment-oncologic-diseases>
10. Shah M, Rahman A, Theoret MR, Pazdur R. The Drug-Dosing Conundrum in Oncology - When Less Is More. *N Engl J Med*. 2021;14;385(16):1445-1447.
11. Daimon T, Hirakawa A, Matsui S. Dose-Finding Designs for Early-Phase Cancer Clinical Trials. Springer, 2019.
12. 大門貴志. 連続再評価法とその周辺 (I) -連続再評価法とは-. 計量生物学. 2012;33(1):1-29.
13. 大門貴志. 連続再評価法とその周辺 (II) -連続再評価法の修正・拡張版と関連デザイン-. 計量生物学. 2012; 33(1):31-76.
14. S. Green, J. Benedetti, A Smith, J. Crowley[著]. JCOG データセンター[訳]. 米国 SWOG に学ぶ がん臨床試験の実践 -第2版-. 医学書院;2013.
15. Araujo DV, Oliva M, Li K, Fazlzad R, Liu ZA, Siu LL. Contemporary dose-escalation methods for early phase studies in the immunotherapeutics era. *Eur J Cancer*. 2021 Oct; 158:85-98.
16. Johanna C Bendell, Milind Javle, Tanios S Bekaii-Saab, Richard S Finn, Zev A Wainberg, Daniel A Laheru, Colin D Weekes, Benjamin R Tan, Gazala N Khan, Mark M Zalupski, Jeffrey R Infante, Suzanne Jones, Kyriakos P Papadopoulos, Anthony W Tolcher, Renae E

- Chavira, Janna L Christy-Bittel, Emma Barrett, Amita Patnaik. A phase 1 dose-escalation and expansion study of binimetinib (MEK162), a potent and selective oral MEK1/2 inhibitor. *British Journal of Cancer*. 2017; 116:575-583.
17. Toby A. Eyre, Graham P. Collins, Avinash Gupta, Nicholas Coupe, Semira Sheikh, John Whittaker, Lai Mun Wang, Leticia Campo, Elizabeth Soilleux, Finn Tysoe, Richard Cousins, Nick La Thangue, Lisa K. Folkes, Michael R. L. Stratford, David Kerr, Mark R Middleton. A Phase I Study to Assess the Safety, Tolerability, and Pharmacokinetics of CXD101 in Patients With Advanced Cancer. *Cancer*. 2019 Jan; 99-108.
 18. Evans TRJ, Dean E, Molife LR, Lopez J, Ranson M, El-Khouly F, Zubairi I, Savulsky C, Reyderman L, Jia Y, Sweeting L, Greystoke A, Barriuso J, Kristeleit R. Phase 1 dose-finding and pharmacokinetic study of eribulin-liposomal formulation in patients with solid tumours. *Br J Cancer*. 2019 Feb;120(4):379-386.
 19. Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. Wiley; 1972.
 20. Forouzannia A, Schiller J, Berlin J, Hutson P, Boothman D, Storer B, Wilding G, Mehta M. A phase I study of Topotecan, as a radiosensitizer, for thoracic malignancies. *Lung Cancer*. 2004 Apr; 44(1):111-9.
 21. Simon R, Freidlin B, Rubinstein L, Arbuck SG, Collins J, Christian MC. Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst*. 1997 Aug 6;89(15):1138-47.
 22. Heath E. I., LoRusso P. M., Ivy S. P., Rubinstein L., Christian M. C., and Heilbrun L. K. Theoretical and practical application of traditional and accelerated titration Phase I clinical trial designs: The Wayne State University experience. *Journal of Biopharmaceutical Statistics*. 2009; 19:414-423.
 23. Mita MM, Mita AC, Moseley JL, Poon J, Small KA, Jou YM, Kirschmeier P, Zhang D, Zhu Y, Statkevich P, Sankhala KK, Sarantopoulos J, Cleary JM, Chirieac LR, Rodig SJ, Bannerji R, Shapiro GI. Phase 1 safety, pharmacokinetic and pharmacodynamic study of the cyclin-dependent kinase inhibitor dinaciclib administered every three weeks in patients with advanced malignancies. *Br J Cancer*. 2017 Oct 24;117(9):1258-1268.
 24. Timothy A. Yap; Hendrik-Tobias Arkenau; D. Ross Camidge; Suzanne George; Natalie J. Serkova; Stephen J. Gwyther; Jennifer L. Spratlin; Rohit Lal; James Spicer; Nandita M. Desouza; Martin O. Leach; Jon Chick; Srinivasu Poondru; Ramesh Boinpally; Richard Gedrich; Katie Brock; Andrew Stephens; S. Gail Eckhardt; Stan B. Kaye; George Demetri; Michelle Scurr. First-in-Human Phase I Trial of Two Schedules of OSI-930, a Novel Multikinase Inhibitor, Incorporating Translational Proof-of-Mechanism Studies. *Clin Cancer Res*. 2013 Feb 15;19(4); 909-19.

25. Skolnik JM, Barrett JS, Jayaraman B, Patel D, Adamson PC. Shortening the timeline of pediatric phase I trials: the rolling six design. *J Clin Oncol*. 2008 Jan 10;26(2):190-5.
26. Thomas Cash, Elizabeth Fox, Xiaowei Liu, Charles G. Minard, Joel M. Reid, Adrienne C. Scheck, Brenda J. Weigel, Cynthia Wetmore. A phase I study of prexasertib (LY2606368), a CHK1/2 inhibitor, in pediatric patients with recurrent or refractory solid tumors, including CNS tumors: A report from the Children's Oncology Group Pediatric Early Phase Clinical Trials Network (ADV1515). *Pediatr Blood Cancer*. 2021 Sep;68(9):e 29065.
27. Storer BE. Design and analysis of phase I clinical trials. *Biometrics*. 1989 Sep;45(3):925-37.
28. Storer BE. An evaluation of phase I clinical trial designs in the continuous dose-response setting. *Stat Med*. 2001 Aug 30;20(16):2399-408.
29. Collins JM, Zaharko DS, Dedrick RL, Chabner BA. Potential roles for preclinical pharmacology in phase I clinical trials. *Cancer Treat Rep*. 1986 Jan;70(1):73-80.
30. Ames MM, Loprinzi CL, Collins JM, van Haelst-Pisani C, Richardson RL, Rubin J, Moertel CG. Phase I and clinical pharmacological evaluation of pirozantrone hydrochloride (oxantrazole). *Cancer Res*. 1990 Jul 1;50(13):3905-9.
31. Ivanova A, Flournoy N. Comparison of Isotonic Designs for Dose-Finding. *Stat Biopharm Res*. 2009 Feb 1;1(1):101.
32. Prasanna T, Malik L, McCuaig RD, Tu WJ, Wu F, Lim PS, Tan AHY, Dahlstrom JE, Clingan P, Moylan E, Chrisp J, Fuller D, Rao S, Yip D. A Phase I Proof of Concept Study Evaluating the Addition of an LSD1 Inhibitor to Nab-Paclitaxel in Advanced or Metastatic Breast Cancer (EPI-PRIMED). *Front Oncol*. 2022 Jun 3; 12:862427.
33. Matthew C. Foster, Chirag Amin, Peter M. Voorhees, Hendrik W. Van Deventer, Kristy L. Richards, Anastasia Ivanova, Jennifer Whitman, Wingkeung Michael Chiu, Nathan D. Barr & Thomas Shea. A phase I dose-escalation study of clofarabine in combination with fractionated gemtuzumab ozogamicin in patients with refractory or relapsed acute myeloid leukemia. *Leukemia & Lymphoma*. 2012 - Issue 7.
34. Liu M, Wang SJ, Ji Y. The i3+3 design for phase I clinical trials. *J Biopharm Stat*. 2020 Mar;30(2):294-304.
35. C.-C. Lin, J. Grewal, D. Sommerhalder, L.-Y. Bai, G.-L. Li, R. Schneider, L. Shen, Y.-M. Yeh, R. Yin, H.-C. Hsu, J. Baranda, C.-Y. Hsieh, S.X. Cai, Y.E. Tian, N. Ma, H. Xia, C.-C. Zhang, B.-Y. Li, M. Zhang, Z.-X. Yu. Safety, pharmacokinetic (PK), and pharmacodynamic profiles of the WEE1 inhibitor IMP7068 in patients (pts) with advanced solid tumors. European Society for Medical Oncology (ESMO) Congress 2022, Poster No. 482P.
36. Hunsberger S, Rubinstein LV, Dancey J, Korn EL. Dose escalation trial designs based on a molecularly targeted endpoint. *Stat Med*. 2005 Jul 30;24(14):2171-81.
37. Christopher R Heery, Geraldine O'Sullivan-Coyne, Ravi A Madan, Lisa Cordes, Arun Rajan,

- Myrna Rauckhorst, Elizabeth Lamping, RN, Israel Oyelakin, Jennifer L Marté, Lauren M Lepone, Renee N Donahue, Italia Grenga, Jean-Marie Cuillerot, Berend Neuteboom, Anja von Heydebreck, Kevin Chin, Jeffrey Schlom, Dr James L Gulley. Avelumab for metastatic or locally advanced previously treated solid tumours (JAVELIN Solid Tumor): a phase 1a, multicohort, dose-escalation trial. *Lancet Oncology*. 2017 May;18(5):587-598.
38. James W. Welsh, John V. Heymach, Chunxiao Guo, Hari Menon, Katherine Klein, Taylor R. Cushman, Vivek Verma, Kenneth R. Hess, Girish Shroff, Chad Tang, Ferdinandos Skoulidis, Melenda Jeter, Nathan Comeaux, Roshal R. Patel, Dawei Chen, Tugce Ozgen, Quynh-Nhu Nguyen, Joe Y. Chang, Mehmet Altan, Jianjun Zhang, Vassiliki A. Papadimitrakopoulou, George R. Simon, Lauren A. Byers, Bonnie Glisson. Phase 1/2 Trial of Pembrolizumab and Concurrent Chemoradiation Therapy for Limited-Stage SCLC. *Journal of Thoracic Oncology*. Aug 2020;15 (12):1919-27.
 39. Catherine S Diefenbach, Fangxin Hong, Richard F Ambinder, Jonathon B Cohen, Michael J Robertson, Kevin A David, Ranjana H Advani, Timothy S Fenske, Stefan K Barta, Neil D Palmisiano, Jakub Svoboda, David S Morgan, Reem Karmali, Elad Sharon, Howard Streicher, Brad S Kahl, Stephen M Ansel. Ipilimumab, nivolumab, and brentuximab vedotin combination therapies in patients with relapsed or refractory Hodgkin lymphoma: phase 1 results of an open-label, multicentre, phase 1/2 trial. *Lancet Haemato*. 2020 Sep; 7: e660-70.
 40. Carlos Gomez-Roca, Philippe Cassier, Dmitriy Zamarin, Jean-Pascal Machiels, Jose Luis Perez Gracia, F Stephen Hodi, Alvaro Taus, Maria Martinez Garcia, Valentina Boni, Joseph P Eder, Navid Hafez, Ryan Sullivan, David Mcdermott, Stephane Champiat, Sandrine Aspeslagh, Catherine Terret, Anna-Maria Jegg, Wolfgang Jacob, Michael A Cannarile, Carola Ries, Konstanty Korski, Francesca Michielin, Randolph Christen, Galina Babitzki, Carl Watson, Georgina Meneses-Lorente, Martin Weisser, Dominik Rüttinger, Jean-Pierre Delord, Aurelien Marabelle. Anti-CSF-1R emactuzumab in combination with anti-PD-L1 atezolizumab in advanced solid tumor patients naïve or experienced for immune checkpoint blockade. *Journal for ImmunoTherapy of Cancer*. 2022; 10 (5): e004076.
 41. Ivanova A, Wang K. A non-parametric approach to the design and analysis of two-dimensional dose-finding trials. *Stat Med*. 2004 Jun 30;23(12):1861-70.
 42. Isakoff SJ, Wang D, Campone M, Calles A, Leip E, Turnbull K, Bardy-Bouxin N, Duvillié L, Calvo E. Bosutinib plus capecitabine for selected advanced solid tumours: results of a phase 1 dose-escalation study. *Br J Cancer*. 2014 Nov 25;111(11):2058-66.
 43. Guosheng Yin [著], 寺良向 聡, 大門 貴志 [訳]. 臨床試験デザイン ベイズ流・頻度流の適応的方法 *Clinical Trial Design Bayesian and Frequentist Adaptive Methods*. メディカルパブリケーションズ. 2014.
 44. Adaptive Designs for Clinical Trials of Drugs and Biologics. <https://www.fda.gov/regulatory->

- [information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry](https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry) [Accessed 27 Apr 2023]
45. Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-use-bayesian-statistics-medical-device-clinical-trials> [Accessed 27 Apr 2023]
 46. 矢田真城, 浜田知久馬. 抗がん剤の第 I 相試験における MTD 推定に有用な CRM の歴史的発展と将来の展望. *Jpn J Clin Pharmacol Ther.* 2015;46(1).
 47. Paoletti X, Kramar A. A comparison of model choices for the Continual Reassessment Method in phase I cancer trials. *Stat Med.* 2009 Oct 30;28(24):3012-28.
 48. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics.* 1990 Mar;46(1):33-48.
 49. Faries D. Practical modifications of the continual reassessment method for phase I cancer clinical trials. *J Biopharm Stat.* 1994 Jul;4(2):147-64.
 50. Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for phase I studies. *Stat Med.* 1995;14(11):1149-1161.
 51. Zohar S, Chevret S. The continual reassessment method: comparison of Bayesian stopping rules for dose-ranging studies. *Stat Med.* 2001 Oct 15;20(19):2827-43.
 52. Korn EL, Midthune D, Chen TT, Rubinstein LV, Christian MC, Simon RM. A comparison of two phase I trial designs. *Stat Med.* 1994 Sep 30;13(18):1799-806.
 53. Pal SK, Forero-Torres A, Thompson JA, Morris JC, Chhabra S, Hoimes CJ, Vogelzang NJ, Boyd T, Bergerot PG, Adashek JJ, Li H, Yu X, Gartner EM, Carret AS, Smith DC. A phase I trial of SGN-CD70A in patients with CD70-positive, metastatic renal cell carcinoma. *Cancer.* 2019 Apr 1;125(7):1124-1132.
 54. Phillips T, Barr PM, Park SI, Kolibaba K, Caimi PF, Chhabra S, Kingsley EC, Boyd T, Chen R, Carret AS, Gartner EM, Li H, Yu C, Smith DC. A phase I trial of SGN-CD70A in patients with CD70-positive diffuse large B cell lymphoma and mantle cell lymphoma. *Invest New Drugs.* 2019 Apr;37(2):297-306.
 55. Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. *Stat Med.* 2008 Jun 15;27(13):2420-39.
 56. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Stat Med.* 1998 May 30;17(10):1103-20.
 57. Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. Robust exchangeability designs for early phase clinical trials with multiple strata. *Pharm Stat.* 2016 Mar-Apr;15(2):123-34.
 58. Shaw AT, Kim DW, Mehra R, Tan DS, Felip E, Chow LQ, Camidge DR, Vansteenkiste J, Sharma S, De Pas T, Riely GJ, Solomon BJ, Wolf J, Thomas M, Schuler M, Liu G, Santoro

- A, Lau YY, Goldwasser M, Boral AL, Engelman JA. Ceritinib in ALK-rearranged non-small-cell lung cancer. *N Engl J Med*. 2014 Mar 27;370(13):1189-97.
59. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics*. 2000 Dec;56(4):1177-82.
 60. Bekele BN, Ji Y, Shen Y, Thall PF. Monitoring late-onset toxicities in phase I trials using predicted risks. *Biostatistics*. 2008 Jul;9(3):442-57.
 61. Mauguen A, Le Deley MC, Zohar S. Dose-finding approach for dose escalation with overdose control considering incomplete observations. *Stat Med*. 2011 Jun 15;30(13):1584-94.
 62. Tevaarwerk A, Wilding G, Eickhoff J, Chappell R, Sidor C, Arnott J, Bailey H, Schelman W, Liu G. Phase I study of continuous MKC-1 in patients with advanced or metastatic solid malignancies using the modified Time-to-Event Continual Reassessment Method (TITE-CRM) dose escalation design. *Invest New Drugs*. 2012 Jun;30(3):1039-45.
 63. Giuliani M, Mathew AS, Bahig H, Bratman SV, Filion E, Glick D, Louie AV, Raman S, Swaminath A, Warner A, Yau V, Palma D. SUNSET: Stereotactic Radiation for Ultracentral Non-Small-Cell Lung Cancer-A Safety and Efficacy Trial. *Clin Lung Cancer*. 2018 Jul;19(4):e529-e532.
 64. Regnery S, Ristau J, Weykamp F, Hoegen P, Sprengel SD, Paul KM, Buchele C, Klüter S, Rippke C, Renkamp CK, Pohl M, Meis J, Welzel T, Adeberg S, Koerber SA, Debus J, Hörner-Rieber J. Magnetic resonance guided adaptive stereotactic body radiotherapy for lung tumors in ultracentral location: the MAGELLAN trial (ARO 2021-3). *Radiat Oncol*. 2022 May 25;17(1):102.
 65. Doyen J, Poudenx M, Gal J, Otto J, Guerder C, Naghavi AO, Gérard A, Leysalle A, Cohen C, Padovani B, Ianessi A, Schiappa R, Chamorey E, Bondiau PY. Stereotactic ablative radiotherapy after concomitant chemoradiotherapy in non-small cell lung cancer: A TITE-CRM phase 1 trial. *Radiother Oncol*. 2018 May;127(2):239-245.
 66. Biau J, Thivat E, Chautard E, Stefan D, Boone M, Chauffert B, Bourgne C, Richard D, Molnar I, Levesque S, Bellini R, Kwiatkowski F, Karayan-Tapon L, Verrelle P, Godfraind C, Durando X. Phase 1 trial of ralimetinib (LY2228820) with radiotherapy plus concomitant temozolomide in the treatment of newly diagnosed glioblastoma. *Radiother Oncol*. 2021 Jan; 154:227-234.
 67. Lao CD, Friedman J, Tsien CI, Normolle DP, Chapman C, Cao Y, Lee O, Schipper M, Van Poznak C, Hamstra D, Lawrence T, Hayman J, Redman BG. Concurrent whole brain radiotherapy and bortezomib for brain metastasis. *Radiat Oncol*. 2013 Aug 21; 8:204.
 68. Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade-offs. *Biometrics*. 2004 Sep;60(3):684-93.

69. Razelle Kurzrock, Chia-Chi Lin, Tsung-Che Wu, Brian P. Hobbs, Roberto Carmagnani Pestana, MD, and David S. Hong. Moving Beyond 3+3: The Future of Clinical Trial Design. *American Society of Clinical Oncology Educational Book* 2021:41; e133-e144.
70. Braun TM. The bivariate continual reassessment method. extending the CRM to phase I trials of two competing outcomes. *Control Clin Trials*. 2002 Jun;23(3):240-56.
71. Copland M, Slade D, McIlroy G, Horne G, Byrne JL, Rothwell K, Brock K, De Lavallade H, Craddock C, Clark RE, Smith ML, Fletcher R, Bishop R, Milojkovic D, Yap C. Ponatinib with fludarabine, cytarabine, idarubicin, and granulocyte colony-stimulating factor chemotherapy for patients with blast-phase chronic myeloid leukaemia (MATCHPOINT): a single-arm, multicentre, phase 1/2 trial. *Lancet Haematol*. 2022 Feb;9(2): e121-e132.
72. Wei Zhao, Harry Yang. *Statistical Methods in Drug Combination Studies*. New York: Chapman and Hall/CRC. 2014.
73. Bedard PL, Taberero J, Janku F, Wainberg ZA, Paz-Ares L, Vansteenkiste J, Van Cutsem E, Pérez-García J, Stathis A, Britten CD, Le N, Carter K, Demanse D, Csonka D, Peters M, Zubel A, Nauwelaerts H, Sessa C. A phase Ib dose-escalation study of the oral pan-PI3K inhibitor buparlisib (BKM120) in combination with the oral MEK1/2 inhibitor trametinib (GSK1120212) in patients with selected advanced solid tumors. *Clin Cancer Res*. 2015 Feb 15;21(4):730-8.
74. Yuan et al. Bayesian Optimal Interval Design A Simple and Well-Performing Design for Phase I Oncology Trials. *Clin Cancer Res*. 2016 Sep 1;22(17):4291-301.
75. Yan et al. Keyboard A Novel Bayesian Toxicity Probability Interval Design for Phase I Clinical Trials. *Clin Cancer Res*. 2017 Aug 1;23(15):3994-4003.
76. Melissa L. Johnson, James Strauss, Manish R. Patel, Edward B. Garon, Keith D. Eaton, Tavette Neskoriak, Josée Morin, Richard Chao, Balazs Halmos. Mocetinostat in Combination With Durvalumab for Patients With Advanced NSCLC: Results From a Phase I/II Study. *Clinical Lung Cancer*. 2023.
77. Zhou, H., Murray, T. A., Pan, H., & Yuan, Y. Comparative review of novel model-assisted designs for phase I clinical trials. *Statistics in Medicine*, 2018;37(14):2208-2222.
78. Xiao-Li Wei, et al. First-in-human phase Ia study of the PI3K α inhibitor CYH33 in patients with solid tumors. *Nature communications*. 2022;13(1):7012.
79. Liu, S., & Yuan, Y. Bayesian optimal interval designs for phase I clinical trials. *Journal of the Royal Statistical Society: Series C: Applied Statistics*. 2015 Apr;64(3):507-523.
80. FDA NEWS. <https://www.ascpt.org/Resources/ASCPT-News/View/ArticleId/27591/FDA-News-Issue-2-January-2022>
81. Gainor JF, et al. Pralsetinib for RET fusion-positive non-small-cell lung cancer (ARROW): a multi-cohort, open-label, phase 1/2 study. *Lancet Oncol*. 2021; 22:959-969.

82. Gainor JF, Schneider JG, Gutierrez M, et al. Nivolumab (NIVO) plus ipilimumab (IPI) with two cycles of chemotherapy (chemo) in first-line metastatic non-small cell lung cancer. *Journal of Clinical Oncology*. 2020 May;38(15), 9560-9560.
83. van Dijk, LV, rank SJ, Yuan Y, et al. Phase I safety and feasibility trial of MRI-guided adaptive particle radiotherapy. *Clinical and Translational Radiation Oncology*. 2022 Jan; 32:35-40.
84. Yuan Y, Lin R, Li D, et al. Time-to-Event Bayesian Optimal Interval Design to Accelerate Phase I Trials. *Clinical Cancer Research*. 2018 Oct; 24(20):4921-30.
85. Takeda, K., Taguri, M., & Morita, S. BOIN-ET: Bayesian optimal interval design for dose finding based on both efficacy and toxicity outcomes. *Pharmaceutical statistics*. 2018;17(4): 383-395.
86. Lin, R., Zhou, Y., Yan, F., Li, D., & Yuan, Y. BOIN12: Bayesian optimal interval phase I/II trial design for utility-based dose finding in immunotherapy and targeted therapies. *JCO precision oncology*. 2020; 4:1393-1402.
87. Lin R, Ying G. Bayesian Optimal Interval Design for Dose Finding in Drug-Combination Trials. *Statistical Methods in Medical Research*. 2017 Jul; 26(5), 2155-2167.
88. Pan H, Lin R. Zhou Y et al. Keyboard Design for Phase I Drug-Combination Trials. *Contemporary Clinical Trials*. 2020 May; 92(1-13):1-11.
89. Wermke M, Kraus S, Ehninger A, et al. Proof of concept for a rapidly switchable universal CAR-T platform with UniCAR-T-CD123 in relapsed/refractory AML. *Blood*. 2021 Jun; 137(22), 3145-3148.
90. Zhou H, Yuan Y, Nie L. Accuracy, Safety, and Reliability of Novel Phase I Trial Designs. *Clin Cancer Res*. 2018; 24(18):4357-64.
91. Liu R, Yuan Y, Sen S, Yang X, et al. Accuracy and Safety of Novel Designs for Phase I Drug-Combination Oncology Trials. *Statistics in Biopharmaceutical Research*. 2022;14(3):270-82.
92. 30 日調査照会事項チェックリスト (抗悪性腫瘍剤分野)
93. Zhou Y, Lee JJ, Wang S, Bailey S, Yuan Y. Incorporating historical information to improve phase I clinical trials. *Pharm Stat*. 2021;20(6):1017-34.
94. Ananthakrishnan R, Lin R, He C, Chen Y, Li D, LaValley M. An overview of the BOIN design and its current extensions for novel early-phase oncology trials. *Contemp Clin Trials Commun*. 2022 Jun 13;28.
95. Hughes TP, Mauro MJ, Cortes JE, Minami H, Rea D, DeAngelo DJ, Breccia M, Goh YT, Talpaz M, Hochhaus A, le Coutre P, Ottmann O, Heinrich MC, Steegmann JL, Deininger MWN, Janssen JJWM, Mahon FX, Minami Y, Yeung D, Ross DM, Tallman MS, Park JH, Druker BJ, Hynds D, Duan Y, Meille C, Hourcade-Potelleret F, Vanasse KG, Lang F, Kim DW. Asciminib in Chronic Myeloid Leukemia after ABL Kinase Inhibitor Failure. *N Engl J*

- Med. 2019; 381(24):2315-26.
96. ICH S9 ガイドライン「抗悪性腫瘍薬の非臨床評価に関するガイドラインについて」
(平成 22 年 6 月 4 日付け薬食審査発 0604 第 1 号) .
 97. 「「医薬品開発におけるヒト初回投与試験の安全性を確保するためのガイダンス」
の改訂等について」(令和元年 12 月 25 日付け薬食審査発 1225 第 1 号) .
 98. 「「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」について」(令和 3 年 3 月
3 1 日付け薬生薬審発 0331 第 1 号) .
 99. Sheau CC, Derek S, Assaf PO, Yu S. Optimal Selection of Adaptive Designs in Phase I
Oncology Clinical Trials. <https://cqs.app.vumc.org/shiny/AdaptiveDesignS/> [Accessed 17
Mar 2023].
 100. Ying Y, J. Jack Lee, et.al. Integrated Platform for Designing Clinical Trials.
<https://www.trialdesign.org/> [Accessed 17 Mar 2023].
 101. Iasonos A, Wilton AS, Riedel ER, Seshan VE, Spriggs DR. A comprehensive comparison of
the continual reassessment method to the standard 3+3 dose escalation scheme in Phase I
dose-finding studies. Clin Trials. 2008;5(5):465-77.
 102. Ken C. Package ‘dfcrm’. <https://cran.r-project.org/web/packages/dfcrm/dfcrm.pdf> [Accessed
17 Mar 2023].
 103. Furong S, Zhonggai L. Package ‘blrm’. [https://cran.r-
project.org/web/packages/blrm/blrm.pdf](https://cran.r-project.org/web/packages/blrm/blrm.pdf) [Accessed 17 Mar 2023].
 104. Sebastian W, Lukas AW. Andrew B. Package ‘OncoBayes2’. [https://cran.r-
project.org/web/packages/OncoBayes2/OncoBayes2.pdf](https://cran.r-project.org/web/packages/OncoBayes2/OncoBayes2.pdf) [Accessed 17 Mar 2023].
 105. Michael S, Graham W. Package ‘bcmr’. [https://cran.r-
project.org/web/packages/bcmr/bcmr.pdf](https://cran.r-project.org/web/packages/bcmr/bcmr.pdf) [Accessed 17 Mar 2023].
 106. MD Anderson Cancer Center. Biostatistics Software --- Desktop / Cloud.
<https://biostatistics.mdanderson.org/SoftwareDownload/SingleSoftware/Index/2> [Accessed
17 Mar 2023].
 107. Clertant M, O'Quigley J. Semiparametric dose finding methods. Journal of the Royal
Statistical Society. Series B. 2017 Nov;79(5):1487-1508.
 108. 国際共同治験に関する基本的考え方について (平成 19 年 9 月 28 日薬食審査発第
0928010 号厚生労働省医薬食品局審査管理課長通知) .
 109. 国際共同治験開始前の日本人での第I相試験の実施に関する基本的考え方について
(平成 26 年 10 月 27 日事務連絡厚生省医薬安全局審査管理課長通知) .
 110. ICH-E5 ガイドライン (「外国臨床データを受け入れる際に考慮すべき民族的要因に
ついて」平成 10 年 8 月 11 日医薬審第 672 号厚生省医薬安全局審査管理課長通知) .
 111. Takamatsu H, Iida S, Shibayama H, Shibayama K, Yamazaki H, Suzuki K. Daratumumab,
lenalidomide, and dexamethasone in Japanese patients with transplant-ineligible newly

diagnosed multiple myeloma: a phase 1b study. *Int J Hematol.* 2020; 111:692-701.

2023年6月

日本製薬工業協会 データサイエンス部会 タスクフォース3

井上 晃一 (日本イーライリリー株式会社)
大野 綾子 (ファイザーR&D 合同会社)
小出 裕一郎 (ファイザーR&D 合同会社) *
小島 将裕 (協和キリン株式会社)
柴山 和弘 (ヤンセンファーマ株式会社)
中川 雄貴 (中外製薬株式会社)
野村 怜史 (ノバルティスファーマ株式会社)
畑山 知慶 (ヤンセンファーマ株式会社)
兵頭 洋平 (住友ファーマ株式会社)

タスクフォースリーダー兼推進委員

浅川 誉 (中外製薬株式会社)
渡部 恵 (住友ファーマ株式会社)

担当副部長

角野 修司 (武田薬品工業株式会社)

*2023年5月31日時点の所属

Appendix

isotonic regression (PAVA)のサンプル SAS コード

```
%macro _pava(_dat,_len,_order="ascending");
data res;
    set &_dat.;
    array p(&_len.);
    array ap(&_len.);
    ap1=p1;
    i=1;
    l=1;
    do k=2 to &_len.;
        %if &_order.="ascending" %then %do;
        if ap[k-1] > p[k] then do;
            if i^=l then do;
                l=k;
            end;
            else if i=l then do;
                i=k-1;
                l=k;
            end;
            count=0;
            sum=0;
            do j=i to l;
                count=count+1;
                sum=sum+p[j];
            end;
            do j=i to l;
                ap[j]=sum/count;
            end;
        end;
    %end;
        %if &_order.="descending" %then %do;
        if ap[k-1] < p[k] then do;
            if i^=l then do;
                l=k;
            end;
        end;
    %end;
end;
```

```

        end;
        else if i=1 then do;
            i=k-1;
            l=k;
        end;
        count=0;
        sum=0;
        do j=i to l;
            count=count+1;
            sum=sum+p[j];
        end;
        do j=i to l;
            ap[j]=sum/count;
        end;
    end;
%end;
else do;
    ap[k]=p[k];
    i=k;
    l=k;
end;
end;
end;
output;

run;
%mend _pava;
/*sample code*/
/*
data test;
    p1=0;
    p2=0.2;
    p3=0.1;
    p4=0.05;
    output;

run;
%_pava(test,4,_order="descending");
*/

```