



ICH E17 の理念に基づく国際共同治験成績を用いた CTD の構成

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 2-2
E17 サブチーム

2023 年 7 月 25 日

目 次

目 次	2
表 一 覧	2
図 一 覧	3
略号一覧	4
1 はじめに	5
1.1 ICH E17 に関する誤解	6
1.2 ICH E5 vs. ICH E17	7
2 ICH E17	9
2.1 Training Material Module 6	10
2.2 効果修飾因子	12
3 3-layer approach	13
3.1 概略	13
3.2 Layer 1: 全集団での評価	14
3.3 Layer 2: さまざまな視点による効果修飾因子の検討及び地域間での一貫性の評価	17
3.4 Layer 3: 申請国/地域でのベネフィット・リスクの評価	37
4 CTD	39
4.1 CTD で検討すべき内容	39
4.2 CTD の記載内容及び記載箇所	39
4.3 M2.5.6 (ベネフィットとリスクに関する結論) の各項及びそのソース文書で検討・記載すべき事項	43
5 Example	51
5.1 Example 1: チカグレロル	51
5.2 Example 2: 仮想事例 ABC123	57
5.3 Example 3: 仮想事例 GHI789	61
6 まとめ	68
7 参考資料	70

表 一 覧

Table 2-1	Training Material Module 1~7 の概要	9
Table 3-1	モデルに基づく国/地域別の治療効果の各手法の長所と短所	30
Table 3-2	頻度分布表: 安全性のさまざまな視点による効果修飾因子検討の事例 (フィネレノン)	32
Table 3-3	頻度分布表: 安全性の地域間での一貫性の評価 (ダビガトラン)	36
Table 4-1	治療効果に影響を与える内因性・外因性要因 (効果修飾因子) 及び国/地域間の差に関する考察の方向性	42

図 一 覧

Figure 1-1	ICH E5 vs. ICH E17	8
Figure 2-1	MRCT の結果の信憑性を高めるための多角的な評価	11
Figure 2-2	効果修飾因子を探索する際の構造化されたステップ	11
Figure 3-1	3-layer approach の概略	13
Figure 3-2	Layer 1 : 有効性の内的一貫性の評価事例 (ペルツズマブ)	15
Figure 3-3	Layer 1 : 内的一貫性の評価事例 (仮想事例)	15
Figure 3-4	Layer 1 : 有効性の外的一貫性の評価事例 (ブロールシズマブ)	16
Figure 3-5	Layer 2 : 有効性の内的一貫性の評価事例 (ペルツズマブ)	18
Figure 3-6	Layer 2 : 有効性の内的一貫性の評価事例 (仮想事例)	19
Figure 3-7	Layer 2 : 有効性の内的一貫性の評価事例 (チカグレロル)	20
Figure 3-8	Layer 2 : 有効性の統計学的な不確実性の評価事例 (チカグレロル)	21
Figure 3-9	Forest plot : 有効性のさまざまな効果修飾因子検討の事例 (ダビガトラン)	23
Figure 3-10	効果修飾因子の組み合わせでサブグループ解析を行う事例 (仮想事例)	24
Figure 3-11	効果修飾因子のクラスタリングで併合解析する事例 (仮想事例)	25
Figure 3-12	Q-Q plot の事例 (仮想事例)	26
Figure 3-13	CART 分析の事例 (仮想事例)	27
Figure 3-14	Forest plot : 地域間の一貫性評価の事例 (フィネレノン)	28
Figure 3-15	Funnel plot : 有効性の地域間の一貫性評価の事例 (ダビガトラン)	29
Figure 3-16	Forest plot : 安全性のさまざまな視点による効果修飾因子検討の事例 (ダビガトラン)	33
Figure 3-17	統計モデルを用いたアプローチ : 安全性のさまざまな視点による効果修飾因子検討の事例 (ダビガトラン)	35
Figure 3-18	Funnel plot : 安全性の地域間の一貫性評価の事例 (ダビガトラン)	36
Figure 4-1	ICH E17 の理念に基づいた CTD を作成するための基本的な構成要素	40

略号一覧

略語	略していない表現（英）	略していない表現（日）
ARR	attributable risk ratio	寄与危険度比
bid	bis in die	1日2回
BMI	body mass index	体格指数
CART	classification and regression tree	—
CI	confidence interval	信頼区間
CTD	Common Technical Document	国際共通化資料
DOR	duration of response	奏効期間
EDC	electronic data capture	電子的データ授受システム
EMA	European Medicines Agency	欧州医薬品庁
EU	European Union	欧州連合
FAS	full analysis set	最大の解析対象集団
FDA	Food and Drug Administration	米国食品医薬品局
GCP	good clinical practice	医薬品の臨床試験の実施の基準
HDI	human development index	—
HR	hazard ratio	ハザード比
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for human use	医薬品規制調和国際会議
ITT	intention-to-treat	—
KM	Kaplan-Meier	—
LS	least squares	最小二乗
MedDRA	Medical Dictionary for Regulatory Activities	ICH 国際医薬用語集
MRCT	Multi- regional Clinical Trial	国際共同治験
NMPA	National Medical Products Administration	—
ORR	overall response rate	全奏効率
OS	overall survival	全生存期間
PFS	progression free survival	無増悪生存期間
PhRMA	Pharmaceutical Research and Manufacturers of America	—
PK	pharmacokinetics	薬物動態（学）
PMDA	Pharmaceuticals and Medical Devices Agency	独立行政法人 医薬品医療機器総合機構
PT	preferred term	基本語
Q-Q Plot, Q-Q plot	quantile-quantile plot	—
RR	relative risk	相対危険度
SE	standard error	標準誤差

1 はじめに

1998年にE5^注がICHで合意され、同年に日本で「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の指針」（医薬審第672号）が実装された後、後追いブリッジング戦略に基づき、多くの新医薬品が製造販売承認を得た（Uyama et al. 2005）。その後、2006年にE5 Q&A No.11がICHで合意された。このQ&Aは多地域試験の計画、実施、報告に関する指針を「E5の枠組みの中で示したもの」であり、後にICH E17で論じられる国際共同治験（MRCT）の前段階として、一定の役割を果たした。続く2007年に、「国際共同試験に関する基本的考え方について」（薬食審査発第0928010号）が発出されたことで、日本も参加するMRCT、国際共同開発が加速した。2000年代半ばは、1996年にICHで合意されたE6（GCP）によって築かれた臨床試験データの国際的な相互利用の理念が結実し始めた時期、EDCに代表される電子的なデータ収集、国際共同開発を支えるIT環境が整い始めた時期と重なり、全世界的に日米欧以外の国々も参加するMRCTが急速に増加した時代でもあった。このような時代背景の中で、人種、民族的要因又は医療環境等の異なる他国のデータを受け入れることについて、日本を含め多くの規制当局が直面するようになった。このような、世界各国／地域の規制当局や医薬品開発を行う医薬品業界の経験蓄積に基づく問題意識から、世界各国／地域での承認申請におけるMRCTの受け入れ可能性を高めることを目的として、ICHでE17ガイドラインが検討され、2017年に合意に至った。ICH E17ガイドラインは日本でも「MRCTの計画及びデザインに関する一般原則に関するガイドライン」（薬生薬審発0612第1号）が発出され、実装に至った。

ICH E17はMRCTの計画及びデザインに焦点をあてたガイドラインであり、被験薬にとって重要な民族的要因を検討し、その要因に関する地域間の違いを踏まえてMRCTを適切にデザインするための留意事項が示されている。臨床試験の計画及びデザインは、事前にどのような知見が得られているかに基づき、どのような考え方で試験結果を評価しようとしているのかを見越し、熟慮した上で決定されるものである。そのため、ICH E17に基づき計画、実施されたMRCT全体の結果に基づき、民族的要因が被験薬の治療効果（有効性及び安全性）に与える影響を多角的かつ体系的に検討したうえで、申請国／地域での有効性及び安全性を評価するというGlobal firstに基づいた考えが重要である。さらに、MRCT全体の結果に基づく評価を充実させることで、その結果を世界各国で共通利用でき、特定の申請国／地域で個別に実施してきた多くの検討を減らすことで承認取得までの期間を短縮でき、世界中での医薬品アクセスを早期化することが可能となる。

ICH E17実装後、多くのMRCTが実施され、その結果に基づき、多くの新医薬品が製造販売承認を得ているものの、ICH E17の理念に基づき作成された申請資料（CTD）はほとんどなく、従来どおり、ICH E5に基づき、まず申請国／地域の結果に注目し、そのうえで他の地域又は試験全体との類似性を評価しているLocal firstの考えに基づいたCTDがほとんどである。この背景には、「ICH E17に対する誤解」と「ICH E5とICH E17の違いの理解不足」も影響していると我々は考えている。

注：ICH E5は、ICHで1998年2月にStep 4合意に至った後、軽微な修正が僅か1か月後にICHで承認され、1度改訂を経たもの（R1; Revision 1）として最終化されている。本書ではE5（R1）をE5と省略して記載する。

本書では、「ICH E17に対する誤解」と「ICH E5 と ICH E17 の基本的な考え方の違い」を説明し、2章で ICH E17、3章で 3-layer approach の framework を解説した上で、4章で ICH E17 の理念に基づいた CTD の構成及び記載内容の方針を提言する。なお、本書は、MRCT の結果の評価方法及び CTD でのまとめ方に焦点をあて、ICH E17 の理念に基づいた多角的・体系的な評価方法について理解を深め、結果の評価、解釈及び考察の方針を提案することを目的としている。本書が ICH E17 の理念に基づく MRCT の結果の解釈及び CTD 作成の普及につながる一助になれば幸いである。

1.1 ICH E17 に関する誤解

ICH E17 に関するよくある質問として、以下のようなものがある。

- 併合戦略が適用できないと ICH E17 を適用できない?
- ICH E17 を踏まえると、ICH E5 Q&A No.11 の枠組みの中で公表された 2007 年の「国際共同試験に関する基本的考え方について」（薬食審査発第 0928010 号）の方法 1、2 に準じた参加者数の設定は適切ではない?

これらの質問に対する回答はいずれも「No」である。

ICH E17 は、前述のとおり、MRCT の計画及びデザインの一般原則を示すガイドラインであり、世界各地域での承認申請で、MRCT の受け入れ可能性を高めることを目的としている。ICH E17 では、MRCT の試験デザインの多岐に渡る項目の一般原則を示しており、併合戦略はその 1 つにすぎず、「E17 の適用」とは「被験薬にとって重要な民族的要因を検討し、その要因に関する地域間の違いを踏まえて MRCT を適切にデザインすること」である。

しかしながら、ICH E17 策定時に新たに定義された「併合された地域 (pooled region)」及び「併合された属性別集団 (pooled subpopulation)」の印象が強かったためか、「ICH E17 の適用とは、併合戦略の適用である」と誤解されることが多い。また、参加者数の削減を目的として併合戦略への期待が高まったが、併合戦略を事前規定し、MRCT 計画時に規制当局と合意することは困難であるため、「MRCT の計画時に、併合戦略について規制当局と合意できなければ、ICH E17 は適用できない」という誤った認識が広まった。さらに、ICH E17 で新たに定義された併合戦略をとらず、従来通りの 2007 年の通知（薬食審査発第 0928010 号）の 6 で記載されている方法 1 又は 2 に準じて日本人参加者数を設定したため、ICH E17 が適用できないと誤解されることも多い。

このような「ICH E17 が適用できない」という誤解により、申請国／地域の結果に注目し (Local first) , そのうえで他の地域又は試験全体の類似性を評価するという、従来からの ICH E5 に基づく思考から抜け出せなくなったことが、ICH E17 に対する正しい理解及び運用が進まないことの本質ではないかと考える。

1.2 ICH E5 vs. ICH E17

ICH E17 以前に多地域試験の計画、実施、報告に関する指針であった ICH E5 Q&A No.11 は、軸足がブリッジング戦略にあって、まず申請国／地域の結果に注目し、その上で他の地域又は試験全体の類似性を評価するという **Local first** に基づく考え方を基礎としていた。ICH E17 が、ICH E5 や ICH E5 Q&A No.11 の改訂ではなく、別のガイドラインとして作成された意味を今一度理解した上で、ICH E17 の理念に基づいた多角的・体系的な評価方法について理解を深める必要がある。

ICH E5

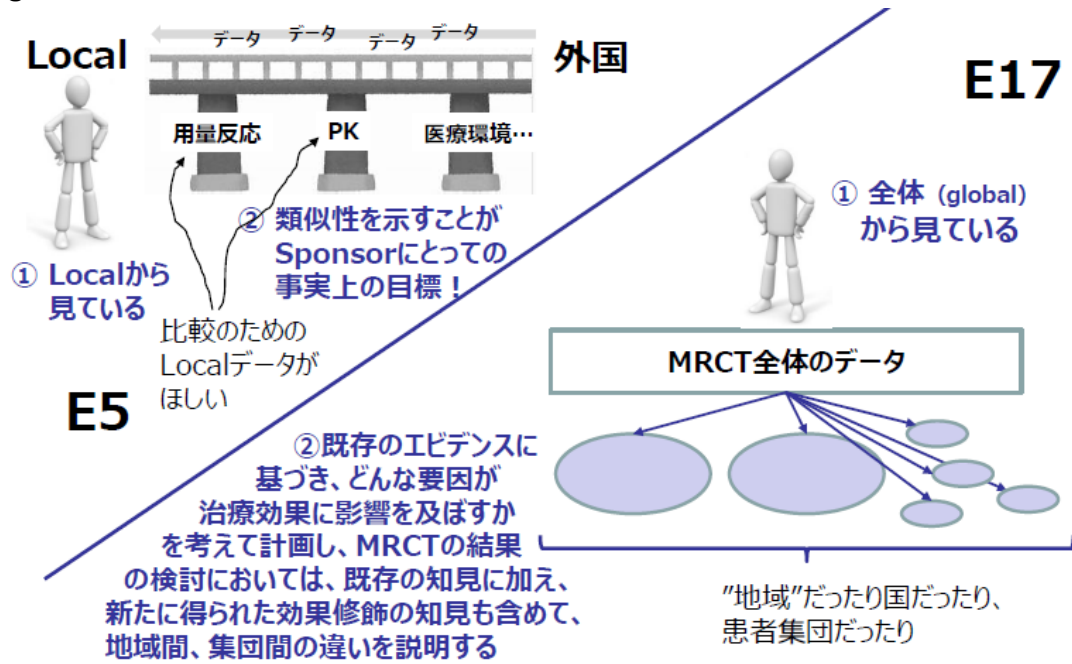
ICH E5 の功績は、治療効果に影響を及ぼす要因を外因性民族的要因と内因性民族的要因として整理したこと、ブリッジングという理念を導入し外国の臨床試験データを受け入れるための留意点を整理したことであろう。E5 が導入したブリッジング戦略は、基本的には申請国／地域 (**Local**) を主軸とした戦略である。たとえば、**Local** を日本とし、すでに外国で承認された臨床データパッケージがある場合に、外国で実施された臨床試験と同様のデザインのブリッジング試験を日本で実施し、その結果を比較し、有効性及び安全性、用量反応、薬物動態 (PK) が類似し、かつ医療環境等が外国と類似しているということが確認できれば、[Figure 1-1](#) のように橋の橋脚が出来上がり、外国の臨床データパッケージを日本の臨床データパッケージに外挿するという戦略である。すなわち、外国で既承認でありながら国内未承認の薬剤が多く存在する状況下で、既に存在する外国の検証試験結果を利用するために、国内外の臨床試験の類似性を示すことが主眼の戦略である。

ICH E17

ICH E17 は、MRCT の計画及びデザイン的一般原則を示すガイドラインであり、MRCT の試験結果全体又は臨床データパッケージ全体、すなわち、**Global** を主軸とした戦略である。ICH E17 では、被験薬の治療効果に影響を及ぼす重要な民族的要因を検討し、その要因に関する地域間の違いを踏まえて MRCT を計画するとともに、MRCT の結果について、既存の知見に加え、新たに得られた治療効果に影響を及ぼす民族的要因 (効果修飾因子) も含めて、地域間、集団間の違いを検討するという考え方である。ICH E17 で示された一貫性評価 (**consistency evaluation**) は、地域間又は集団間に認められた違いを、効果修飾因子という視点で説明を試みることを含んでおり、地域間又は集団間の類似性を示すことがゴールではないことに御留意いただきたい。

Figure 1-1

ICH E5 vs. ICH E17



Source: 第7回 ICH E17 ワークショップ (2022)

ICH E17 の理念に基づく評価には、治療効果に影響を及ぼす民族的要因（効果修飾因子）を中心とした検討が重要な鍵となる。効果修飾因子、また、効果修飾因子を中心とした評価の詳細は、3章で説明する。

2 ICH E17

前述のとおり、ICH E17「国際共同治験の計画及びデザインに関する一般原則に関するガイドライン」は、世界各地域での承認申請でMRCTの受け入れ可能性を高めるために、MRCTの計画及びデザインの一般原則を示すことを目的として作成され、MRCTのデザイン及び治験実施計画を検討するうえで課題となる、参加者、投与量、評価項目や対照薬の選択、統計解析計画等に対するガイドが示されている。

また、本ガイドラインの重要な点の理解促進のため、Training Material が作成されている。Training Material は Module 1～7で構成されており、Module 1では Overview を、Module 2～7ではそれぞれ ICH E17の各セクションに焦点をあてた詳細な説明を提供している (Table 2-1)。

Table 2-1 Training Material Module 1～7の概要

Module	Title	Outline
Module 1	Basic principles and overview of training modules	<ul style="list-style-type: none"> • Training Material の目的及び概要
Module 2	Pre-considerations of regional variability when recruiting diverse populations in global drug development	<ul style="list-style-type: none"> • MRCTのデザインで、事前に治療効果に影響を及ぼす民族学的要因を検討し考慮することの重要性 • 治療効果に影響を及ぼす内因性又は外因性要因の特定方法とその対応策
Module 3	Selection of doses for use in confirmatory MRCTs	<ul style="list-style-type: none"> • 用量選択の際に考慮すべき事項 • 検証試験 (MRCT) で異なる投与レジメンが選択された事例紹介
Module 4	Overall sample size and allocation to regions	<ul style="list-style-type: none"> • 目標参加者数設定時の統計学的な検討事項 • 各地域の参加者数検討時に考慮すべき事項 • 地域や属性集団の併合の考え方 • ガイドラインで紹介されている5つの参加者数配分方法の紹介 (各方法の長所と短所)
Module 5	Pooling strategies	<ul style="list-style-type: none"> • 併合の意義 • 併合地域と併合属性別集団の定義 • 地域と属性別集団の併合方法 • 併合戦略の応用可能性
Module 6	Evaluation of consistency	<ul style="list-style-type: none"> • 一貫性の定義 • 一貫性評価の意義 • 一貫性評価の手順及び考慮すべき事項 • 地域別の一貫性評価の検討 • 事例紹介：PLATO 試験
Module 7	Selection of comparators	<ul style="list-style-type: none"> • 対照薬を選択する際の手順 • 特定の地域で承認されていない薬剤を選択する際の留意事項

Source: ICH E7 Training Material Module 1～7

ICH E17のタイトルに含まれ、ICH E17の基本的な適用対象となっている「計画とデザイン」と、本書が焦点をあてるMRCTの結果の評価（地域間及び属性別集団間の一貫性の検討をはじめ、結果の解釈及び評価）を結びつけるのがModule 6である。本書は、MRCTの結果の評価方法に焦点をあてようとしているので、ICH E17 Training Material Module 6をまずは解説する。

2.1 Training Material Module 6

Module 6では、Figure 2-1に示したとおり、MRCTの試験結果を評価する際の基本的な考え方が説明されている。この図の下の方に書かれたメッセージ「**These eventuality should be carefully considered at the planning stage.** (結果として何が起きるのか、どのような評価が行われるのか [these eventuality] を、計画段階で慎重に検討するべきである)」は、この Training Material を作成した ICH E17 Implementation Working Group がこだわった部分であり、MRCTの「計画とデザイン」が「結果の解釈及び評価」を見据えて行われるべきであることを伝えている。

一貫性は、「異なる地域間又は属性別集団間で、臨床的に重要な差がないこと (Absence of clinically relevant differences between treatment effects in different regions or subpopulations of an MRCT)」と定義されている。また、一貫性評価は、地域間又は部分集団間で結果が一致していることを説明することがゴールではなく、地域間又は属性別集団間での結果の差異の有無を多角的、体系的に評価し、差異が認められた場合、効果修飾因子 (2.2 項参照) による説明ができないかを検討することを意味する。すなわち、効果修飾因子を主軸に、全体集団の結果がすべての地域に適用できるかを評価することが重要である。

Module 6では、内因性要因と外因性要因を慎重に考慮した構造化されたステップと、一貫性を評価するための5つの視点が示されている (Figure 2-1)。構造化されたステップは、効果修飾因子の探索過程を段階的に示したものである。既知の因子がある場合は、試験計画時に事前に解析を規定した上で評価されるべきであり、予想外の地域間又は属性別集団間差が観察された場合は部分集団解析等による因子の探索が必要となり、探索しても説明できない地域間又は属性別集団間差については、更なる事後解析が必要になる可能性があること、場合によっては探索のための追加試験が必要となるかもしれないことが示されている。また、MRCTの結果の信憑性を高めるための多角的な評価の視点として、「生物学的合理性 (Biological)」、「内的一貫性 (Internal consistency)」、「外的一貫性 (External consistency)」、「統計学的な不確実性 (Statistical uncertainty)」、「臨床的意義 (Clinical relevance)」の5つが示されている (Figure 2-2)。

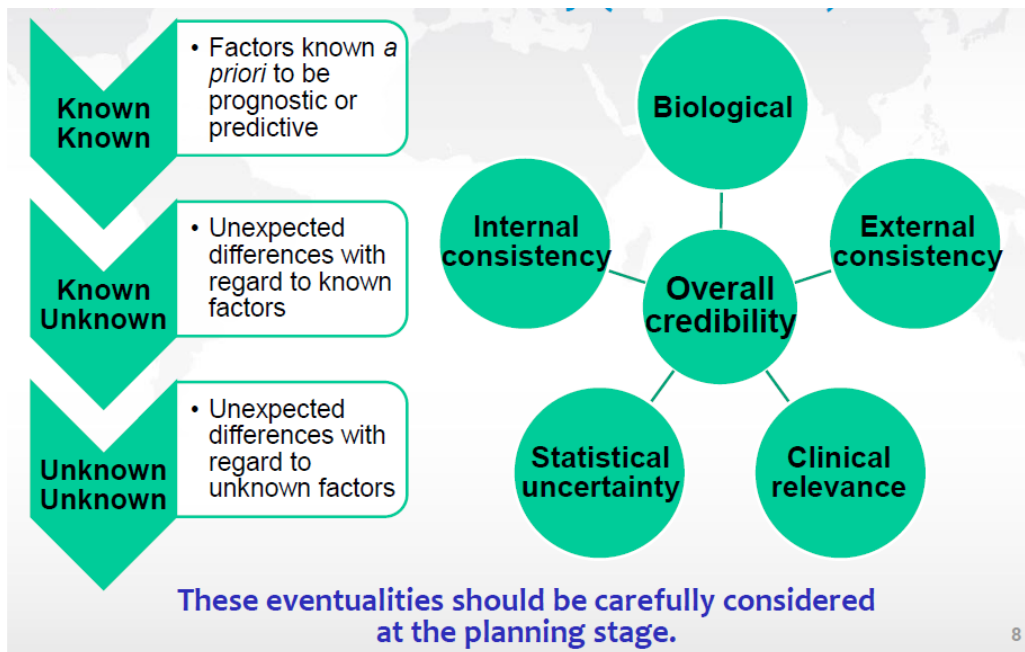
これら5つの視点は、以下のように説明されている。

生物学的合理性 (Biological)	生物学的に説得力のある説明が可能か
内的一貫性 (Internal consistency)	同一試験内で互いに支持する複数の所見が得られているか、たとえば生物学的あるいは医学的に関連のある複数の評価項目において、同じ結論 (命題) を支持する所見が得られているか
外的一貫性 (External consistency)	ある試験で得られた所見と、他の試験結果などの外部情報とを並べて見て、それらが同じ結論 (命題) を支持する所見であるか
統計学的な不確実性 (Statistical uncertainty)	得られた所見がどの程度確からしいか、言い換えれば、推定に無視できないバイアスは混入していないか、推定の精度はどの程度高いか
臨床的意義 (Clinical relevance)	臨床的な判断や治療方針の決定の根拠を与える重要な知見か。何らかの集団間で観察された差異は臨床的に注目すべきものか

Source: これからの医薬品評価 – 国際共同開発での日本人データの意義って何? – (2022年日本製薬工業協会)

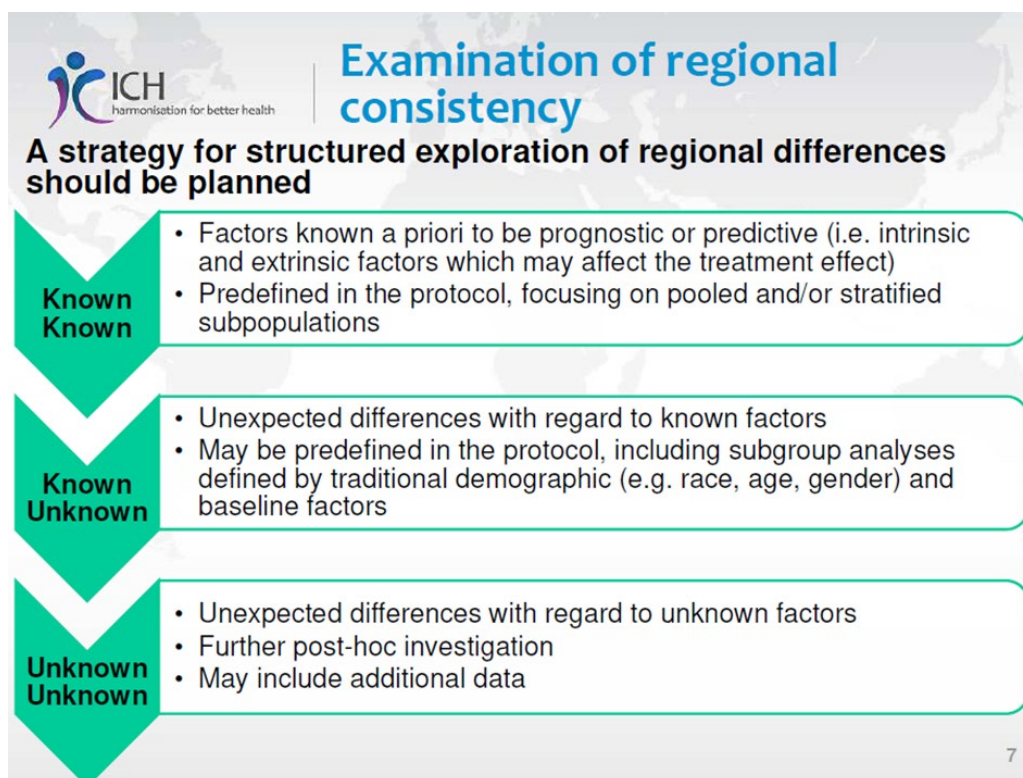
詳細は、体系的に MRCT の結果を評価するための framework である 3-layer approach とともに、3章で解説する。

Figure 2-1 MRCT の結果の信憑性を高めるための多角的な評価



Source: ICH E7 Training Material Module 6, 8 頁

Figure 2-2 効果修飾因子を探索する際の構造化されたステップ



Source: ICH E7 Training Material Module 6, 7 頁

2.2 効果修飾因子

本項では、効果修飾因子及びそのアプローチ方法について説明する。

医薬品の治療効果（有効性又は安全性）が異なる生物学的現象を効果修飾（Effect Modification）と呼び、効果修飾を引き起こす要因、つまり治療効果に影響を与える要因を効果修飾因子（Effect Modifier）という [「[これからの医薬品評価-国際共同開発での日本人データの意義って何？-](#)」（2022年日本製薬工業協会）]。効果修飾因子になり得る要因には、患者個人の特性を表す患者要因（内因性要因）、生活環境や医療環境の特性を反映する環境要因（外因性要因）があり、疾患の予後要因や治療の予測要因が効果修飾因子となることもある。また、同じ医薬品の安全性と有効性で、同じ効果修飾因子が存在する場合もあれば、異なる効果修飾因子が存在する場合もある。

効果修飾因子が特定された場合、臨床試験結果を解釈する際に、効果修飾因子の影響を考慮しながら治療効果を解釈することができる。たとえば、効果修飾因子に対して実施したサブグループ解析結果に基づき、各サブグループの治療効果を評価する。複数の効果修飾因子が存在し、それらの要因間に関連性がない場合は、それぞれ独立したサブグループ解析結果を評価することが可能である。また、別のアプローチとして、複数の要因を組み合わせたサブグループを構築すると、サブグループ解析結果から効果修飾因子を組み合わせた複合的な効果を説明することも可能となる。

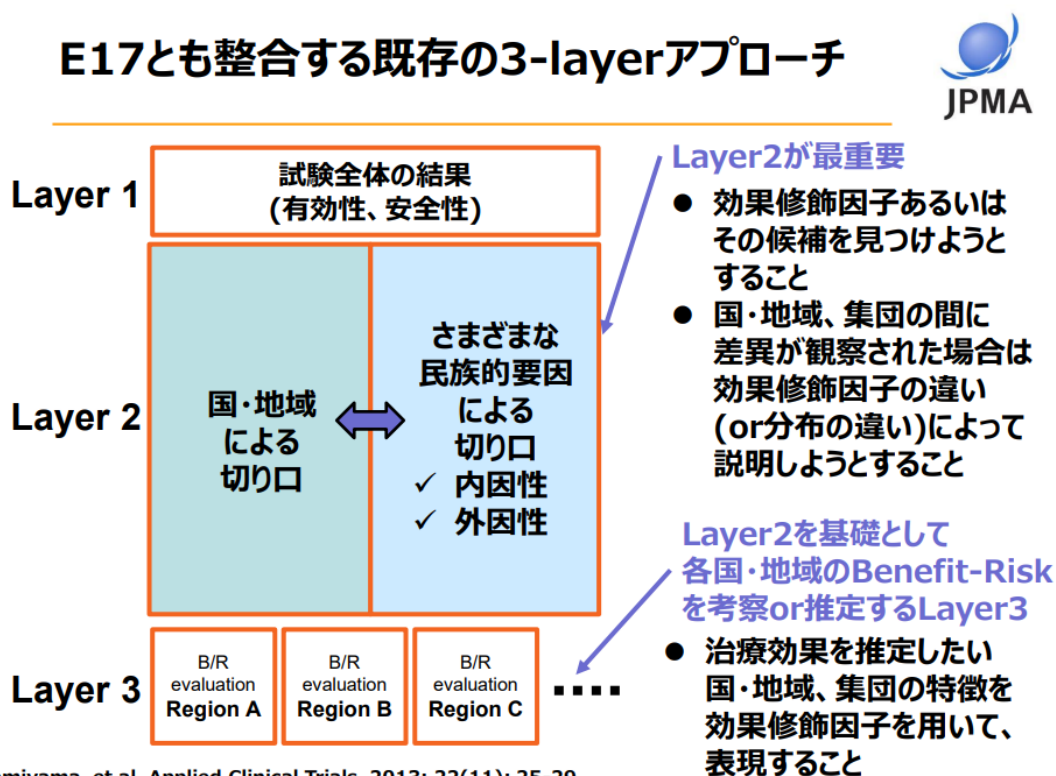
3 3-layer approach

3.1 概略

本項では、小宮山らが提案した 3-layer approach (Komiyama et al. 2013) の概略を説明する。3-layer approach は MRCT の結果を評価するための framework であり、MRCT 全体の結果を申請国／地域で最大限共有し、評価するための科学的なアプローチである。MRCT 全体の結果とは、主要評価項目の主要な解析結果が統計的に有意となり、試験目的が達成されたことのみではなく、その推定値や推定値に影響を及ぼす民族的要因の分析結果も含まれる。すなわち、申請国／地域のデータに焦点をあてずに、全体の結果を踏まえた上で、申請国／地域の結果を解釈するアプローチである。

3-layer approach (Figure 3-1) は、ICH E17 の理念とも一致する、国際共同試験の結果を評価するための framework である。MRCT の結果の解釈を 3 つの Layer (階層) に分け、Layer 1 及び Layer 2 では Global 共通で評価する。すなわち、試験全体の結果に基づいて、Layer 1 では被験薬の有効性及び安全性を評価し、Layer 2 でその結果に影響を与える要因、効果修飾因子あるいはその候補の有無を評価する。国・地域、集団の間に異なる反応が観察された場合は効果修飾因子の違い又はその分布の違いによって説明できないかを検討する。Layer 3 は、Layer 1 及び 2 の結果に基づき、治療効果を推定したい国・地域、集団の特徴を効果修飾因子で表現し、申請国／地域でのベネフィットとリスクを評価する。次項以降に、各 Layer の詳細を説明する。

Figure 3-1 3-layer approach の概略



Komiyama, et al. Applied Clinical Trials. 2013; 22(11): 25-29

Yoshida, et al. Therapeutic Innovation & Regulatory Science. 2015; 49(1): 175-180

16

Source: 第7回 ICH E17 ワークショップ (2022)

3.2 Layer 1: 全集団での評価

MRCT, 一地域で実施される Local の臨床試験等, 臨床試験の参加国にかかわらず, 臨床試験の主たる目的は, 特別な場合を除き, すべての参加者から得られたデータに基づいて治療効果を推定することである。そのため, 3-layer approach でも, 最初の Layer 1 では, MRCT を一つの臨床試験として評価し, 理解しようとする階層である。臨床試験には主要な目的があり, 目的を達成するために試験デザインの骨格 [組み入れ基準に反映される試験の対象となる集団, 治療群の構成, 主要評価項目とその評価方法 (統計的な仮説構造を含む), 参加者数等] が決められる。試験結果の評価の段階では, MRCT であっても特定の国や地域に焦点をあてずに, 主要な目的は達成されたのか, 他の有効性評価項目, 安全性に関わる結果はどうであったかを, まずは検討する。この検討が, 以降の階層 (Layer 2 及び Layer 3) の基礎を与え, 評価をさらに深掘りするための出発点となる。このような考え方は, ICH E17 の根底にある考え方とも整合している (たとえば, ICH E17 1.4 節の「基本的原則」, 第 2 章の「国際共同試験の計画及びデザインにおける一般的な推奨事項」等を参照)。

2.1 項で示した ICH E17 Training Material Module 6 の 5 つの視点は, MRCT に限らず試験全体の結果の信憑性を高める上での重要な視点であり, Layer 1 でも考慮すべきである。以下に, Layer 1 で考慮すべき 5 つの視点を, 一部, 有効性の事例とともに解説する。

生物学的合理性

試験全体の結果が, 臨床的, 薬理的又は作用機序等の複合的及び多角的観点から, 妥当な説明が可能であるかを検討する。生物学的合理性の検討は, 内的及び外の一貫性の検討とも関連する。一例として, 臨床的及び作用機序的な観点で検討した事例を以下に示す。

事例: フェソテロジンフマル酸塩

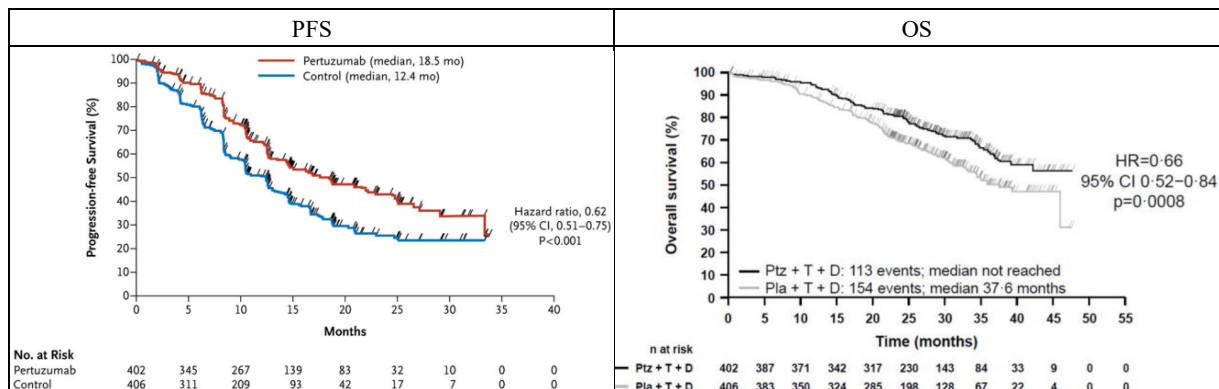
ムスカリン受容体拮抗薬の一つであるフェソテロジンフマル酸塩は, 膀胱平滑筋や膀胱知覚神経の神経終末に存在するムスカリン受容体に結合して, 膀胱収縮, 排尿反射の亢進を抑制する作用機序を有する。過活動膀胱患者を対象とした国際共同第 II 相試験では, フェソテロジンフマル酸塩の投与により, 24 時間あたりの平均切迫性尿失禁回数 (主要評価項目) でプラセボ群に比べ統計学的に有意な減少 (改善) が示された。同様に, 平均排尿回数及び平均尿意切迫感回数も, プラセボに比べ減少 (改善) した。また, 外国で実施された第 III 相試験でも, 同様の結果が得られた。このように, 臨床的及び作用機序的な観点から, 矛盾の無い結果であり, 生物学的合理性のある結果が得られている。

内的一貫性

同一試験内で, 互いに支持する複数の結果が得られているか, たとえば生物学的又は医学的に関連のある複数の評価項目の結果で, 同じ結論を支持する結果が得られているかを検討する。たとえば, Figure 3-2 に示すペルツズマブの事例のように, 無増悪生存期間 (PFS) と全生存期間

(OS)の結果が一貫しているかを評価するというように、医学的に関連のある主要評価項目と副次評価項目の結果が一貫しているかを評価することも、内的一貫性の評価の一例である。

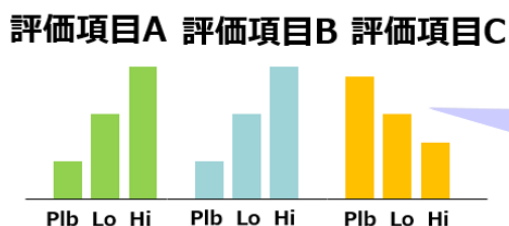
Figure 3-2 Layer 1 : 有効性の内的一貫性の評価事例 (ペルツズマブ)



Source: Baselga et al. 2012, Swain et al. 2013

また、Figure 3-3のように、複数の評価項目を群ごとの要約統計量によって評価する方法も有用である。

Figure 3-3 Layer 1 : 内的一貫性の評価事例 (仮想事例)



評価項目A, Bが+方向, Cが-方向に動くことが、「生物学的合理性」を持つならば、これらの結果は同じ結論を支持するもので、「内的一貫性」を確認したことになる。

Source: 第7回 ICH E17 ワークショップ (2022)

外的一貫性

ある試験で得られた結果と、他の試験結果等の外部情報が同じ結論を支持する結果であるかを検討する。同じ薬剤を同様の患者集団で評価した試験が複数ある場合には、試験間の一貫性を評価し、得られたエビデンスの頑健性を確認する。また、同じ薬剤で一貫性を評価可能な試験がない場合には、同種同効薬の試験結果との一貫性を評価することも1つの手段である。

安全性については、同じ薬剤を異なる適応症で試験したときのデータと比較して一貫性を評価することも有用である。

同じ薬剤を同様の患者集団で評価した試験が2試験ある場合の事例を以下に示す。

事例：ブロルシズマブ（遺伝子組換え）

Figure 3-4 は糖尿病黄斑浮腫（DME）の患者を対象としてブロルシズマブの有効性及び安全性を評価した2つの検証試験（KESTREL 及び KITE）を同じ論文内で評価した事例である（Brown et al. 2022）。この2試験は、アフリベルセプトを対照とした非劣性試験であり、ブロルシズマブ6 mg 群のアフリベルセプト群と比較した有効性及び安全性について、試験間の一貫性が評価されている。さらに、中心窩下脈絡膜新生血管を伴う加齢黄斑変性（nAMD）の患者を対象にブロルシズマブの有効性及び安全性を評価した HAWK 試験及び HARRIER 試験の結果と一貫していることも確認されている。

Figure 3-4 Layer 1：有効性の外的一貫性の評価事例（ブロルシズマブ）

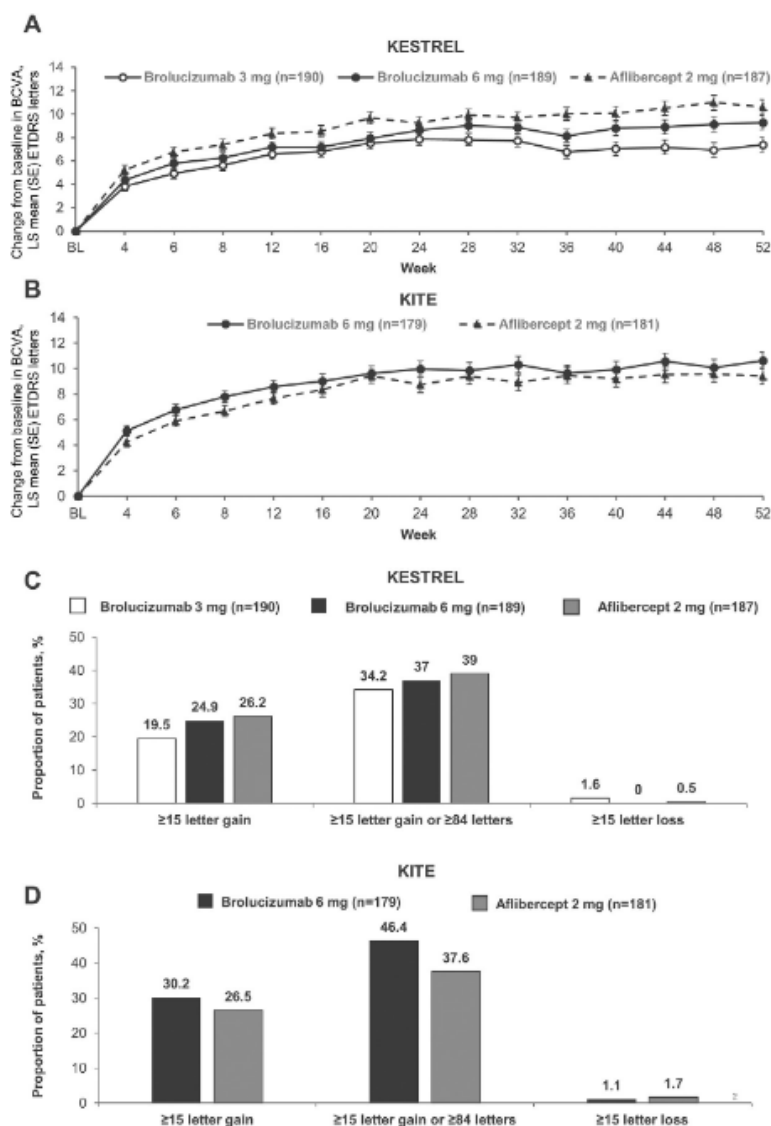


FIGURE 2. Mean change in BCVA from baseline to Week 52 in A) KESTREL and B) KITE. Proportion of subjects with 15-letter gain/loss in BCVA in C) KESTREL and D) KITE at Week 52. Full analysis set, last observation carried forward. BCVA = best corrected visual acuity, BL = baseline, ETDRS = Early Treatment Diabetic Retinopathy Study, LS = least squares, SE = standard error.

Source: Brown et al. 2022

統計学的な不確実性

不確実性の検討は、科学的なエビデンスを評価する上で、重要かつ不可欠な要素である。

Layer 1 では、試験デザイン（目標参加者数、盲検又は非盲検、等）、試験の実施状況（治験実施計画書からの逸脱状況、等）等を複合的に評価し、全参加者から得られた結果がどの程度確からしいか。推定に無視できないほどのバイアスは混入していないか。推定の精度はどの程度高いかを検討する。

臨床的意義

臨床的な判断や治療方針の決定の根拠を与える重要な知見かを検討する。

Layer 1 では、全体集団の結果で、得られた結果の臨床的重要性、日常診療への一般化可能性、各主要なベネフィット及びリスクに関連するエビデンスの強さ、限界及び不確実性の分析の結果が、臨床的な判断や治療方針の決定の根拠を与える重要な知見かを検討する。

3.3 Layer 2 : さまざまな視点による効果修飾因子の検討及び地域間での一貫性の評価

Layer 2 は、そのあとに続く Layer 3、すなわち各申請国／地域でのベネフィット・リスク評価の基礎となる重要な階層であり、Layer 2 も特別な場合を除き、MRCT の全参加者から得られたデータに基づいて治療効果に影響を与える要因を評価する。Figure 3-1 に示されたように、国・地域による切り口、さまざまな民族的要因（既知の効果修飾因子あるいはそれらの候補因子）による切り口から、治療効果に影響を及ぼす要因を探索していく。この探索の入り口となるのが、各因子によるサブグループ解析やそれらの結果を図示した Forest plot である。その後、以下で説明するようなさまざまな観点から探索が進められるであろう。2.1 項で示した ICH E17 Training Material Module 6 の 5 つの視点も考慮した上で、国や地域間の治療効果の違いが示唆された場合には、その差異を効果修飾因子の違い又はその分布の違いで説明できないかを検討する。

一方、多角的なデータ解析を経ても、国・地域間の差異が認められず、効果修飾因子の候補も示唆されない場合には注意が必要である。どのような患者集団においても治療効果に差がないこと（MRCT 全体の結果をどの申請国／地域に対しても外挿可能であること）の証明は、いわゆる“悪魔の証明”、“消極的事実の証明”であり、「証拠が無いことは、無いことの証明にならない [Absence of evidence is not evidence of absence]」という格言を念頭に置きながら、より多角的、より念入りな解析を行い、慎重なメッセージングを行う必要がある。

次項以降で、有効性評価、安全性評価別に、効果修飾因子又はその候補の有無を評価するためのアプローチである、さまざまな視点での効果修飾因子の検討（3.3.1.1 項、3.3.2.1 項）、地域での一貫性の評価（3.3.1.2 項、3.3.2.2 項）を説明する。

2.1 項で示した ICH E17 Training Material Module 6 の 5 つの視点は、MRCT の結果の信憑性を高める上での重要な視点であり、Layer 2 を評価する上で重要な視点である。以下に、Layer 2 での効果修飾因子の検討にあたり、考慮すべき事項を解説する。

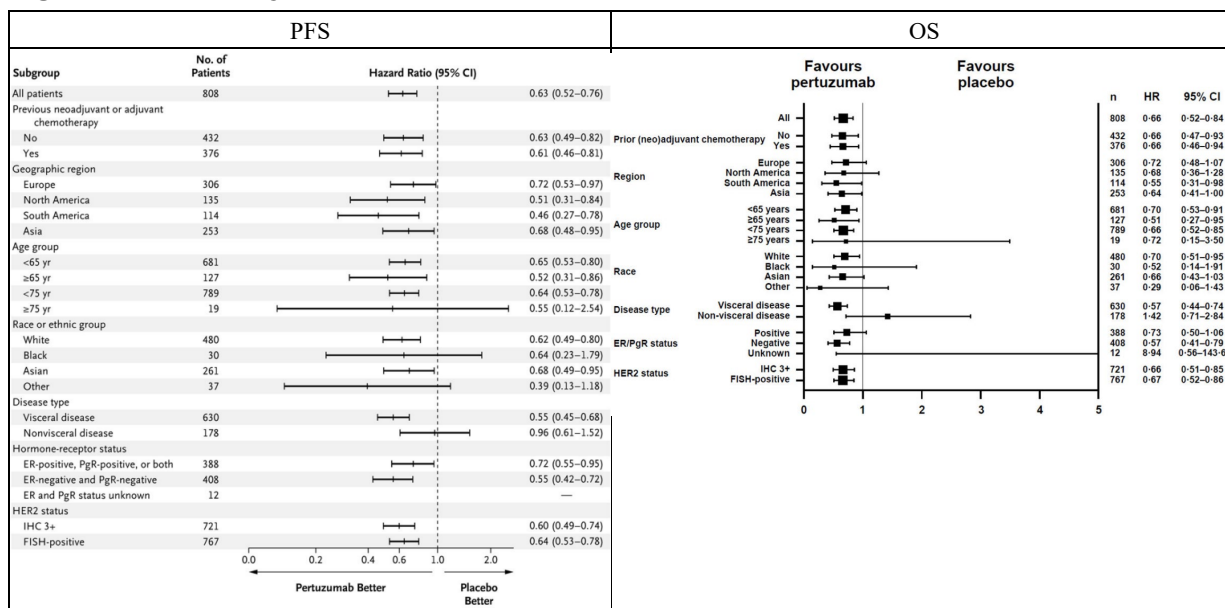
生物学的合理性

地域間又は集団間でみられた結果の差異、効果修飾因子の候補による治療効果への影響が、臨床的、薬理的又は作用機序等の複合的及び多角的観点から、妥当な説明が可能であるかを検討する。生物学的合理性の検討は、内的一貫性、外的一貫性の評価とも関連する。

内的一貫性

地域間又は集団間でみられた結果の差異、効果修飾因子の候補による治療効果への影響が、複数の生物学的又は医学的に関連のある複数の評価項目の結果、又は事前に規定したサブグループ間（地域又は属性的集団）で一貫してみられるかを検討する。たとえば、抗がん剤で、効果修飾因子の候補による治療効果への影響が PFS と OS で一貫しているかを、Figure 3-5 に示す forest plot をそれぞれ実施したうえで評価することも有用である。

Figure 3-5 Layer 2 : 有効性の内的一貫性の評価事例（ペルツズマブ）



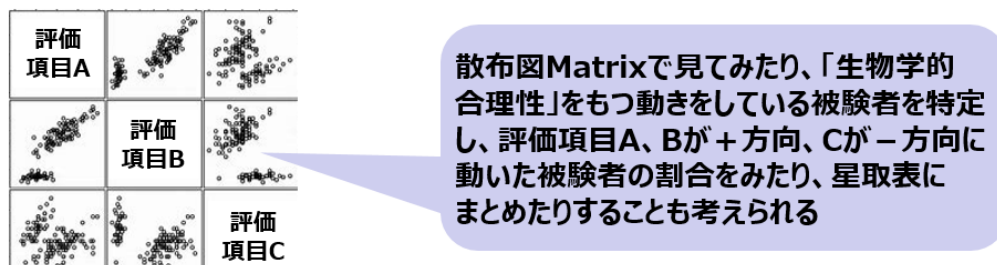
Source: Baselga et al. 2012, Swain et al. 2013

また、Figure 3-6 の事例のように、医学的に関連のある主要評価項目と副次評価項目の結果を、各参加者でプロットし、効果修飾因子の候補による治療効果への影響の一貫性を評価する方法も有用である。複数の評価項目がある場合に、すべての 2 変数同士の組み合わせで散布図を作成し、それを並べることで、臨床的及び作用機序的な観点から矛盾の無い「生物学的合理性」を示す参加者と示さない参加者を特定することが可能である。この散布図 Matrix を事前に規定したサブグ

ループ別に色分けして散布図を作成することで、「生物学的合理性」を示さない参加者で共通する因子、すなわち効果修飾因子の候補を特定可能な場合もある。

Figure 3-6 Layer 2 : 有効性の内的一貫性の評価事例 (仮想事例)

被験者レベルでも評価項目間の関連性



Source: 第7回 ICH E17 ワークショップ (2022)

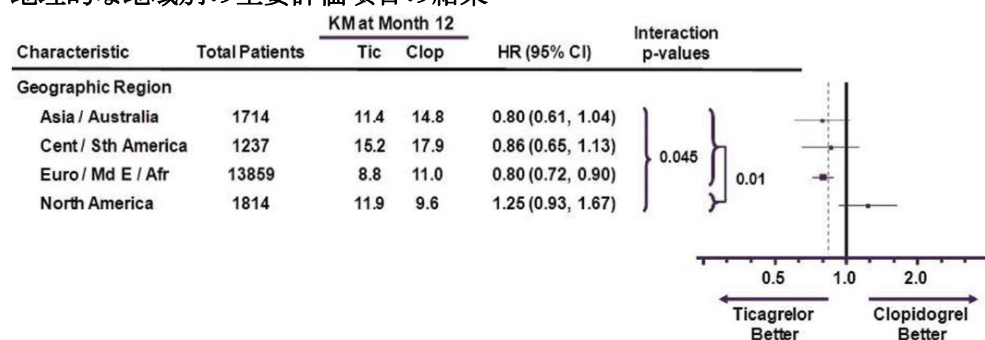
その他、サブグループ（地域又は属性的集団）間での効果修飾因子の候補による治療効果への影響を検討した事例として、ICH E17 Training Material Module 6 で、地域間の内的一貫性の事例として記載されている PLATO 試験を説明する。PLATO 試験の詳細は 5.1 項に示す。

事例：チカグレロル

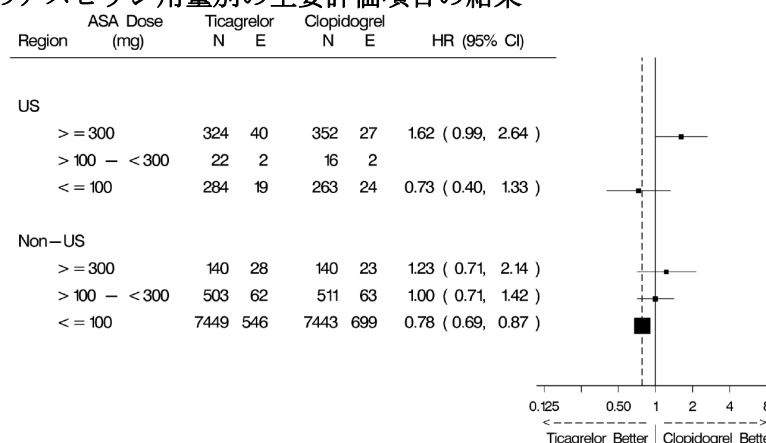
PLATO 試験では、主要評価項目である複合イベント（心血管死、心筋梗塞、及び脳卒中）のいずれかの初回発現までの期間で、地理的な地域別のサブグループ解析の結果から North America の治療効果がその他の地域と異なり（Figure 3-7 上図）、米国の結果の異質性が高いことが示唆された（5.1 項）。しかしながら、米国及び米国以外で効果修飾因子であるアスピリンの用量別のサブグループ解析を行った結果、米国、米国以外ともに、アスピリンの用量が高い集団で治療効果が逆転し、一貫した結果が得られた（Figure 3-7 下図）。

Figure 3-7 Layer 2 : 有効性の内的一貫性の評価事例 (チカグレロル)

地理的な地域別の主要評価項目の結果



米国及び米国外のアスピリン用量別の主要評価項目の結果



Source: ICH E7 Training Material Module 6

外的一貫性

地域間又は集団間でみられた結果の差異，効果修飾因子の候補による治療効果への影響が，複数のデータソースでみられるかどうかを検討する。具体的には，同じ薬剤を同様の患者集団で評価した試験が複数ある場合には，Layer 2 の内的一貫性で評価した効果修飾因子の候補による治療効果への影響が，試験間で一貫しているかを検討する。第 II 相試験で治療効果に影響を与える要因，すなわち，効果修飾因子の候補が特定された場合には，第 III 相試験で割付の層別因子にする，事前にサブグループ解析を計画する等の方法で，治療効果への影響を確認することも外的一貫性の評価の一例である。その他，同種同効薬の試験結果との一貫性を評価することも 1 つの手段である。

統計学的な不確実性

不確実性の検討は，科学的なエビデンスを評価するうえで，重要かつ不可欠な要素である。

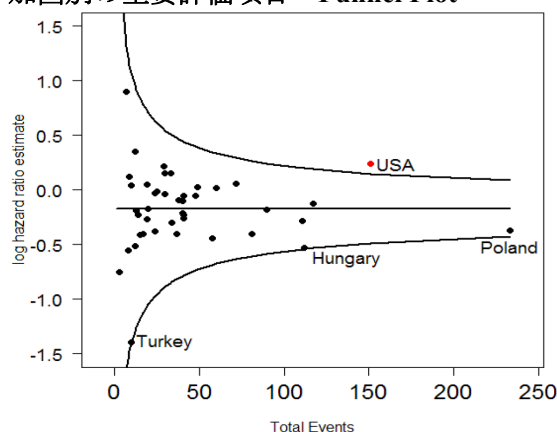
Layer 2 では，地域間又は集団間でみられた結果の差異，効果修飾因子の候補による治療効果への影響が，どの程度確からしいか，推定の精度を検討する。

事例：チカグレロル

下図は、ICH E17 Training Material Module 6 で地域間の一貫性での統計学的な不確実性の事例として記載されている PLATO 試験である。Funnel plot (Figure 3-8) で、米国での主要評価項目の結果が、95%信頼区間から乖離し、また、投与群と地域（米国又は米国以外）の交互作用項の p 値が 0.0095 であったこと、また、主要評価項目のハザード比が全体集団で 0.84 のときに米国で 1.27 以上となる確率は 0.006 未満であったことから、米国の結果には何らかの系統的な要因が存在することが示唆された。しかしながら、米国のみで検証できるだけの検出力はないため、米国での結果が偶然の結果である可能性も否定できない。

Figure 3-8 Layer 2：有効性の統計学的な不確実性の評価事例（チカグレロル）

参加国別の主要評価項目—Funnel Plot



Source: ICH E7 Training Material Module 6

臨床的意義

Layer 2 では、地域間又は集団間でみられた結果の差異、効果修飾因子の候補による治療効果への影響が、臨床的な判断や治療方針の決定の根拠を与える重要な知見かを検討する

事例：チカグレロル

ICH E17 Training Material Module 6 に記載されている PLATO 試験では、アスピリンの維持用量が効果修飾因子として特定され、アスピリンの維持用量が高い集団でチカグレロルの治療効果が低下することが明らかとなった。内的一貫性の事例で示すとおり、維持用量 300 mg 超の高用量のアスピリンと併用した場合には、対照薬と比較したチカグレロルのハザード比は米国及び米国外ともに 1 を超え、チカグレロルの対照薬に対する有効性のベネフィットは得られないことが示唆された。

3.3.1 有効性評価

3.3.1.1 さまざまな視点による効果修飾因子の検討

薬剤の有効性及び安全性に影響する可能性のある内因性及び外因性要因は多く存在し、効果修飾因子の検討に一律の方法はなく、薬剤の特性に応じて方法を検討する必要がある。次項以降に内因性及び外因性要因を踏まえた効果修飾を説明するためのアプローチの例を示す。

(1) One factor at a time アプローチ

効果修飾の存在を知るための単純で典型的な方法は、内因性又は外因性民族的要因の何らかの要因によって集団をサブグループに分け、治療効果の違いを検討するサブグループ解析である。単純なサブグループ解析だけでなく、Figure 3-9に例示のようなグラフィカルなアプローチを併用して検討することも有用である。治療効果が部分集団間で異なる場合には、その要因が効果修飾因子である可能性を示唆する。ただし、この方法は、それぞれの要因と治療効果の関係を検討することができるが、複数の要因による複合的な効果を説明することはできない。

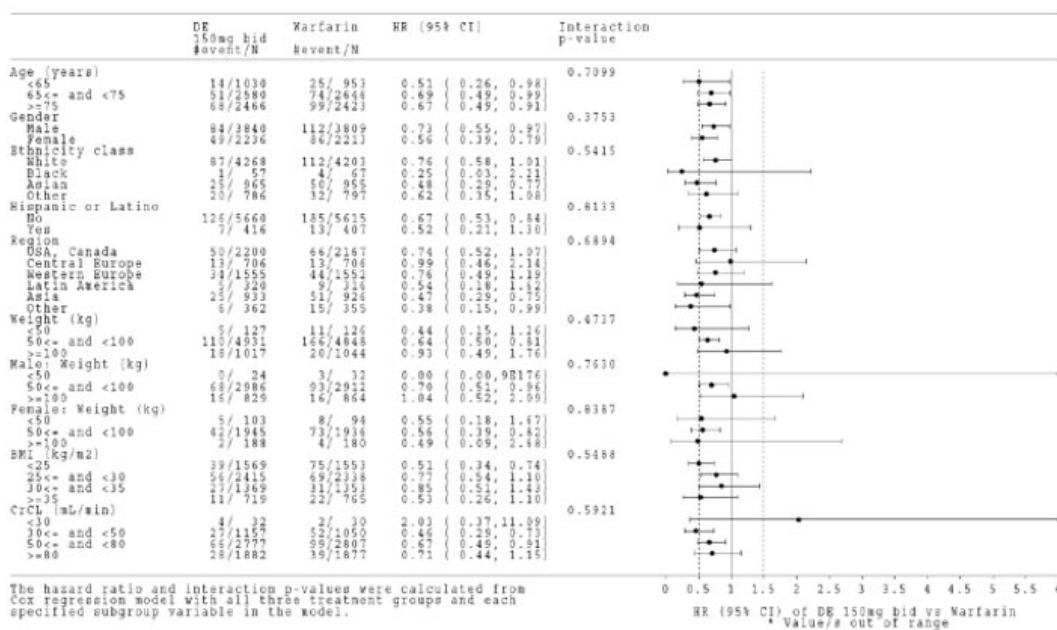
1) Forest plot

サブグループ解析のグラフィカルアプローチの1つに forest plot がある。サブグループ解析は交互作用の有無の検討を目的として実施されるため、各部分集団の要約統計量に加え、ハザード比（95%信頼区間）、交互作用の p 値が示されることがある。(3) 統計モデルを用いたアプローチ
1) 交互作用の検討に記載のとおり、交互作用の p 値はあくまで参考値として取り扱い、p 値が小さい場合には、その要因が効果修飾因子の可能性を考慮し、更なる検討が必要となる。

事例：ダビガトランエテキシラートメタンサルホン酸塩（ダビガトラン）

Figure 3-9は、ダビガトランの第III相MRCT（RE-LY試験）の脳卒中又は全身性塞栓症のワルファリンに対するハザード比の forest plot を示す。この結果では、いずれの項目も投与群との間に有意な交互作用は認められなかった。ダビガトラン群では脳卒中又は全身性塞栓症の発現割合は加齢及びクレアチニンクリアランス増加とともに増加した。その傾向は対照群でも同様であった。

Figure 3-9 Forest plot : 有効性のさまざまな効果修飾因子検討の事例 (ダビガトラン)



The hazard ratio and interaction p-values were calculated from Cox regression model with all three treatment groups and each specified subgroup variable in the model.

引用元: CTD 5.3.5.1-4, 試験 1160.26, U-3249-01, Table 15.2.3.2:2

図 2.5.4.6.2: 1 人口統計学的特性および他の基準値の特性別のワルファリン投与群と比較した Bibr 1048 MS 150 mg 1 日 2 回投与群の脳卒中/全身性塞栓症のハザード比および 95%信頼区間: ランダム化集団

Source: ダビガトラン申請資料概要 (2011)

2) Funnel plot

3.3.1.2 項の地域間での一貫性評価で示す funnel plot (Figure 3-15) は、検討する部分集団に対し縦軸にその推定値、横軸にその参加者数をプロットしたものである。本プロットは、内因性又は外因性の民族的要因の検討でも使用可能であり、各要因の水準間で治療効果の規則性や外れ値等を検討するのに有用である。

(2) 効果修飾因子の組み合わせでサブグループ解析を行うアプローチ

複数の効果修飾因子の複合的な効果も考慮しながら説明を行おうとするアプローチである。Figure 3-10 は、対象疾患の重症度と BMI が効果修飾因子の候補である場合の検討を示す。たとえば、対象疾患の重症度と BMI が効果修飾因子である場合、重症度を 3 段階 (軽度, 中等度, 重度), BMI を 2 段階 (25 未満, 25 以上) に分けたとすると 2 つの効果修飾因子の組み合わせによるサブグループは 6 通りとなる。この 6 通りのサブグループ間で治療効果がどのように異なるかをみることで、重症度と BMI の複合的な治療効果を説明しようとするアプローチである。評価項目が連続変数の場合は、Response surface model も有用である。懸念事項としては、サブグループ数が多くなればなるほど、参加者数も少なく十分な情報が得られないサブグループが増え、解釈が困難になる可能性がある。

Figure 3-10 効果修飾因子の組み合わせでサブグループ解析を行う事例（仮想事例）

✓ 分割表でまとめる

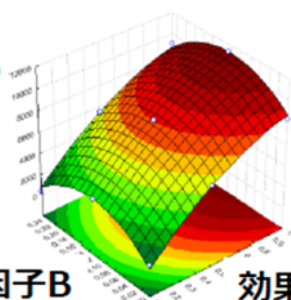
		BMI	
		25未満	25以上
重症度	軽度	平均±95%CI	平均±95%CI
	中等度	平均±95%CI	平均±95%CI
	重度	平均±95%CI	平均±95%CI

フォレスト・プロット
や3Dの棒グラフ
で表せるかも
しれない

✓ モデルを使って解析

効果修飾因子が連続量ならResponse surface modelで提示

興味のある
評価項目



効果修飾因子B

効果修飾因子A

Source: 第7回 ICH E17 ワークショップ (2022)

(3) 統計モデルを用いたアプローチ

統計モデルには複数の効果修飾因子を同時に組み入れることができるため、それぞれの効果修飾因子の影響、これらの複合的な影響、すなわち効果修飾因子間の交互作用を検討することもできる。上記のようなサブグループ解析では、それぞれの効果修飾因子についてどこかに閾値を設定して段階（順序カテゴリー）に分けて考える必要があるが、統計モデルでは効果修飾因子を順序カテゴリーとしても、連続量としても扱うことができ、どちらの場合でも交互作用の検討が可能である。

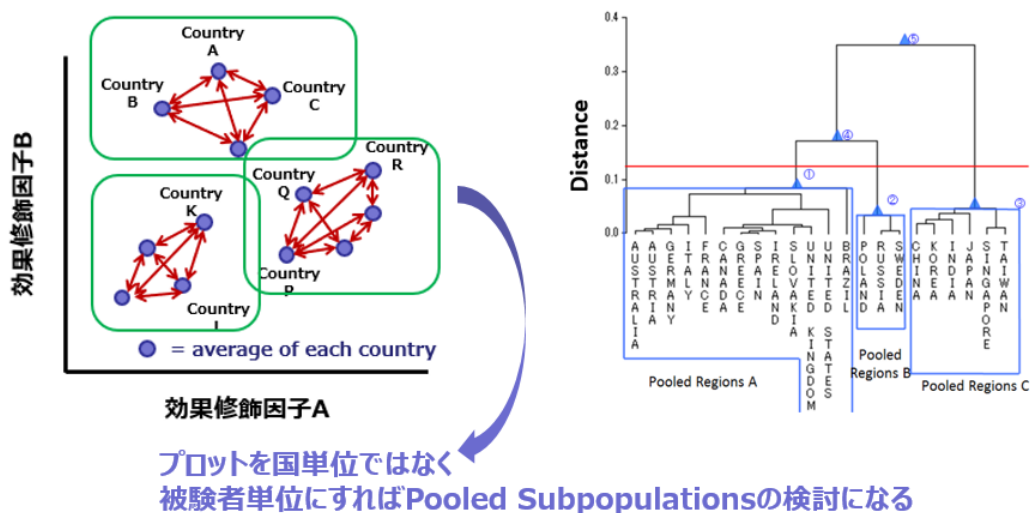
1) 交互作用の検討

統計モデルに、投与群、効果修飾の候補因子、投与群とその因子の交互作用を要因として含めて解析し、交互作用項の検定結果から効果修飾因子を探索するアプローチである。交互作用の検定結果が有意な場合に、その因子が効果修飾因子である可能性が示唆されるが、一般的に交互作用の検定は検出力が低く、また検定の有意水準を定める明確な基準がないため、その結果の解釈には注意を要する。検定結果のみから判断するのではなく、臨床的妥当性も考慮し、検討すべきである。

2) 効果修飾因子のクラスタリングで併合解析を行うアプローチ

複数の効果修飾の候補因子間の類似度に基づき、参加者を複数のグループに分け、各グループの治療効果と特徴（候補因子の分布）から、効果修飾因子を探索するアプローチである。グループ化の際に、国又は地域単位、あるいは参加者単位の2通りがある。国又は地域単位でグループ化した場合は、併合された国又は地域間での治療効果の違いを検討でき、参加者単位でグループ化した場合は、併合された属性別集団間での治療効果の違いを検討できる。ただし、類似度を定める明確な基準がなく、検討したい因子が多いとグループ化された各グループの特徴を説明できない可能性がある。

Figure 3-11 効果修飾因子のクラスタリングで併合解析する事例（仮想事例）

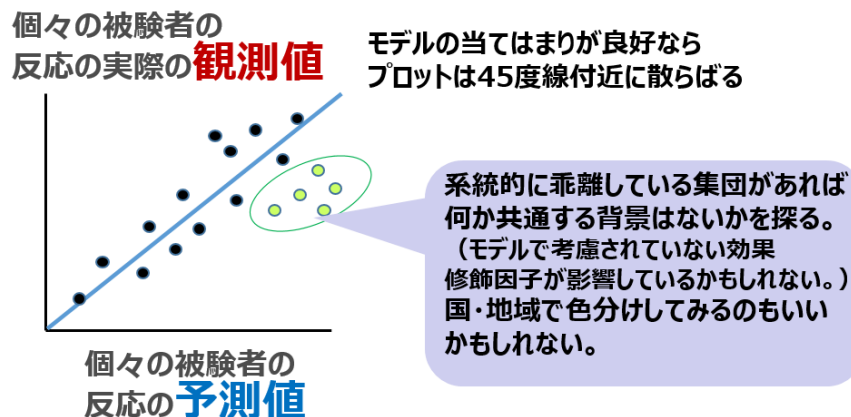


Source: 第7回 ICH E17 ワークショップ (2022)

3) 残差プロット

構築したモデルの診断のために用いられる散布図であり、何かしらの規則性、分散の偏り、外れ値の検出等を視覚的に捉えることができる。Figure 3-12 の例は、Y 軸に観測値、X 軸に予測値とした散布図であり、一般的な残差プロット（Y 軸に残差、X 軸に予測値）の図とは異なる Q-Q plot であるが、本質は変わらず表現の違いだけである。系統的に乖離している場合、効果修飾因子が影響している可能性があるが、それらがどの集団に属しているかを特定することは、この図から判断できない。

Figure 3-12 Q-Q plot の事例（仮想事例）



Source: 第7回 ICH E17 ワークショップ (2022)

4) Galbraith plot

Galbraith Plot は、基準化した推定値を精度 [標準誤差 (SE) の逆数] に対してプロットする手法である。最初に各推定値をその標準誤差で除し標準化推定値又は z 統計を計算し、 $1/SE$ (横軸) に対する各 z 統計量 (縦軸) をプロットした散布図である。外れ値を同定するために有用な手法であり、もし、同質であれば、データ点は原点を通る回帰直線 ± 2 標準誤差の内部に分布する (5.1.1 項 Figure example 1-2 参照)。

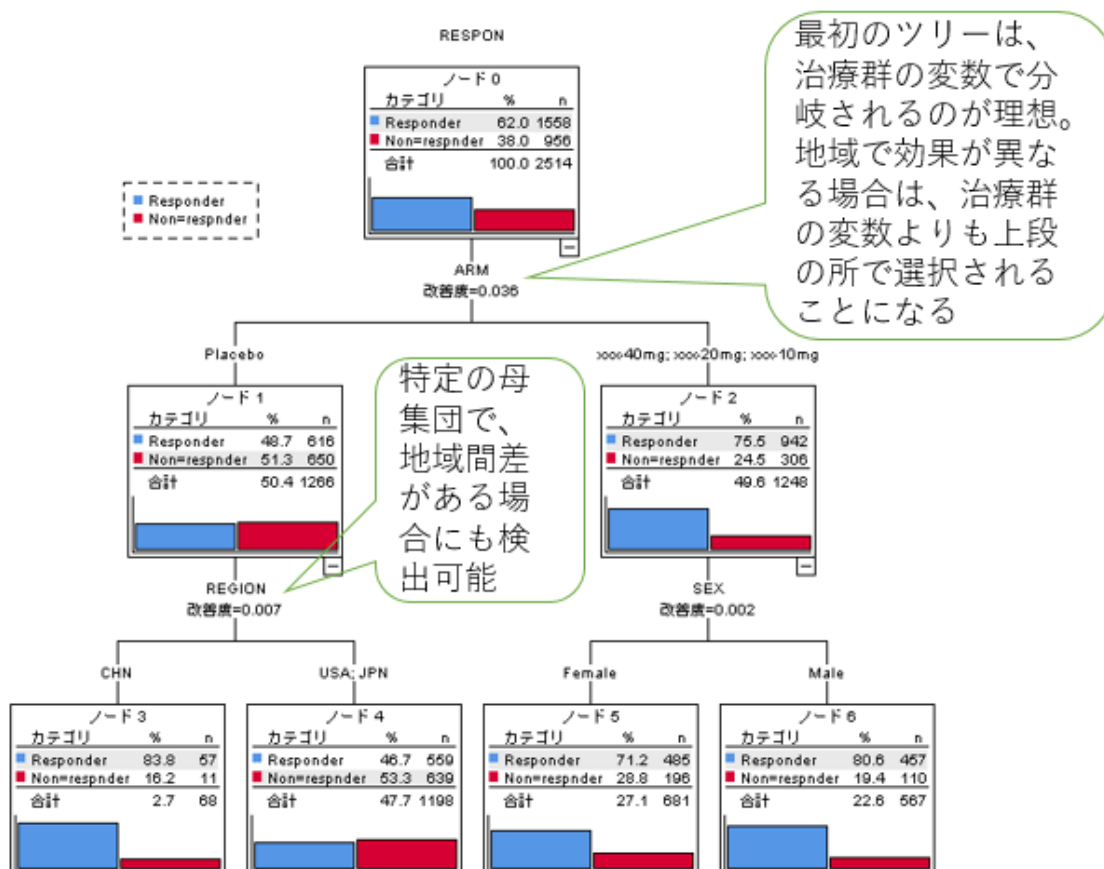
5) CART 分析

CART は決定木分析のアルゴリズムで、分岐が必ず 2 つになることを特徴としている。決定木分析は、判別、予測を目的とする一般的なデータマイニング手法 (機械学習) の一つであり、自動交互作用検出法から発展した手法であることから、交互作用の検出に適している。意思決定に必要な因子と判断基準を樹形図の形で表現する方法であり、その分析結果は目的変数に影響の強い因子から順に自動的に選択され、階層的に並ぶため、それぞれの因子間の相互 (階層) 関係を理解しやすいという特徴がある。利用に際しては、以下のメリットもある。

- データの分布を仮定せず非線形のデータや説明変数の多い (次元の高い) データに適用可能
- 欠測値を含めたさまざまな尺度 (質的変数か量的変数か等) が混在しているデータにも適用可能
- 結果の可視化

ただし、p 値等の統計的判断基準がなく、どの階層までを採択するかは恣意的な判断となってしまうため、留意する必要がある。

Figure 3-13 CART 分析の事例（仮想事例）



3.3.1.2 地域間での一貫性の評価

地域間での一貫性を評価する方法も、ICH E17 Training Material Module 6に記載があるとおり、記述的な要約、グラフ表示、モデルに基づく推定、治療と地域間の交互作用の検討等がある。次項以降に地域間での一貫性を説明するためのアプローチの例を示す。2.1項で説明のとおり、一貫性評価とは、地域間で結果が一致していることを指すのではなく、地域間での結果の差異の有無を多角的、体系的に評価し、差異がある場合はその要因を特定し、ベネフィット・リスクへの影響を検討することを意味する。したがって、地域間で差異がみられた場合には、その差異を効果修飾因子の違い又はその分布の違いによって説明できないかを検討する。

(1) One factor at a time アプローチ

地域別の検討でも、要因別と同様に、単純で典型的な方法として、地域別に集団に分け、治療効果の違いを検討するサブグループ解析がある。

1) Forest plot

地域間での一貫性を評価するグラフィカルアプローチの1つとしても forest plot が有用である。

事例：フィネレノン

Figure 3-14 は、フィネレノンの MRCT での、国別の腎不全発現のプラセボに対するハザード比 (95%信頼区間) の forest plot を示す。ミネラルコルチコイド受容体拮抗薬であるフィネレノンは、2型糖尿病を合併する慢性腎臓病患者を対象とする MRCT を 2 試験実施しており、申請者は、いずれの試験も国により腎不全発現にバラツキがみられ、試験内及び試験間でハザード比が 1 を上回った国に地理的要因や民族的要因の明らかな類似性がみられなかったことを説明している。

Figure 3-14 Forest plot : 地域間の一貫性評価の事例 (フィネレノン)

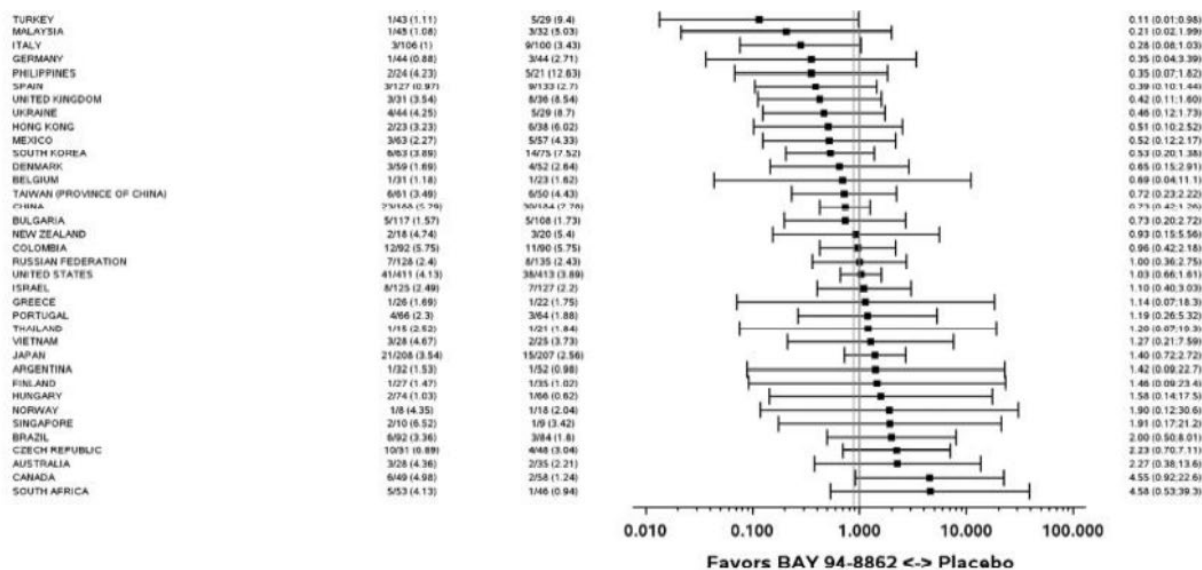


図 5 試験 16244 における「腎不全」の発現状況の国別フォレストプロット (FAS、投与群を因子とした Cox 比例ハザードモデル)

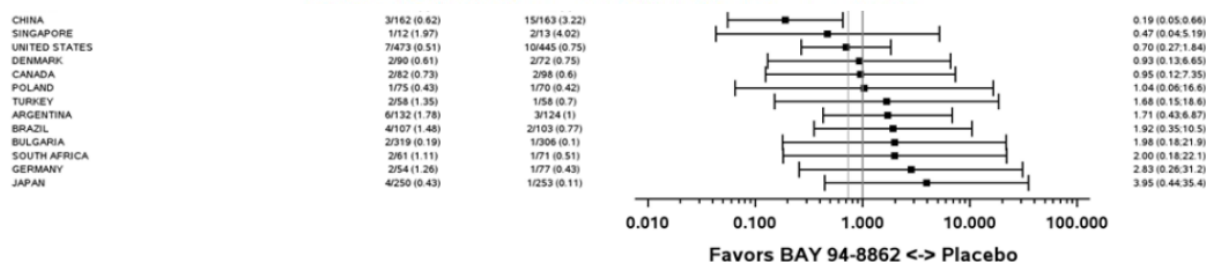


図 6 試験 17530 における「腎不全」の発現状況の国別フォレストプロット (FAS、投与群を因子とした Cox 比例ハザードモデル)

Source: フィネレノン審査報告書 (2022)

2) Funnel plot

Funnel plot は、検討する部分集団に対し縦軸にその推定値、横軸にその参加者数でプロットしたものである。

事例：ダビガトラン

Figure 3-15 は、ダビガトランの第 III 相 MRCT (RE-LY 試験) の結果に基づく、国又は地域別の脳卒中及び全身性塞栓症の発現のワルファリンに対するハザード比の funnel plot を示す。縦軸をハザード比、横軸をイベント数とし、点線で 95%信頼区間を示す。いずれの国の複合イベント発現もバラツキの範囲内であり、また、ハザード比が 1 を上回った国に地理的要因がみられていないことが示されている。

Figure 3-15 Funnel plot : 有効性の地域間の一貫性評価の事例 (ダビガトラン)

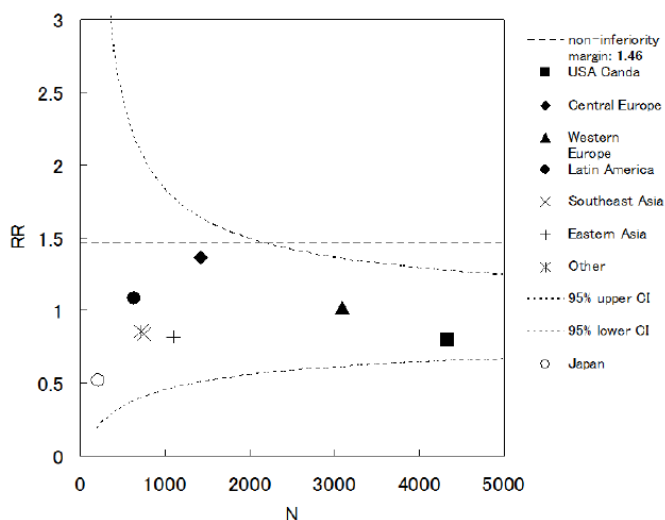


図 3-3 脳卒中及び全身性塞栓症の地域別のハザード比 (110 mg 群 vs ワルファリン群) ¹⁰

Source: 国際共同治験での民族的要因の検討方法 3-layer approach の実践 (2018 年日本製薬工業協会)

(2) モデルに基づく推定

モデルに基づく国／地域別の治療効果の推定方法は、既に「国際共同治験での民族的要因の検討方法 3-layer approach の実践」(2018 年日本製薬工業協会) で、以下の 3 つの方法が紹介されている。

1) 縮小推定量の利用

Quan らが地域間の一貫性を評価するために提案した経験縮小推定や James-Stein 型縮小推定を用いて地域ごとの治療効果を推定する方法である。

2) 地域間の類似性を考慮した樹形アプローチ

Quan らの縮小推定量では地域間の類似性が考慮されないという点を是正するため、Guo らが提案したのは、PhRMA の multi-regional clinical trial group が推奨する human development index (HDI) に基づき、国家間の HDI による類似性を考慮したベイジアン の枠組みで解析する方法である。

3) 標準化法を応用した地域別の治療効果の推定

治療効果に影響する民族的要因が特定できている場合は、地域ごとの治療効果の推定に層別解析の理論を応用した標準化法を適用することができる。標準化法を利用すると比較的簡便な方法で全体のデータを利用した地域ごとの推定値を得ることができる。

これら 3 つの方法には Table 3-1 に示す長所と短所があり、各手法を理解した上で選択する必要がある。これらの方法の詳細は、「国際共同試験での民族的要因の検討方法 3-layer approach の実践」(2018 年日本製薬工業協会) を参照すること。

Table 3-1 モデルに基づく国/地域別の治療効果の各手法の長所と短所

	長所	短所
縮小推定量	全体の推定値に向かって縮小されるため一貫性が示しやすい	民族的要因の類似性や異質性が反映されない(例数とばらつきのみ依存)
樹形アプローチ	地域間の総合的な類似性が反映される	計算が煩雑 反映される類似性は一般的なものであり、当該薬に固有の要因の影響を直接反映しているわけではない
標準化法	計算方法が分かりやすく、結果も解釈しやすい	影響要因が特定され、データとして収集されている必要がある。 多くの影響要因を考慮できない

(3) 治療と地域間の交互作用の検討

3.3.1.1 項のさまざまな視点による効果修飾因子検討の forest plot の事例 (Figure 3-9) と同様に、国又は地域別のサブグループ解析は交互作用の有無の検討を目的として実施するため、交互作用の p 値を示すことがある。しかしながら、前述のとおり、交互作用の p 値はあくまで参考値として取り扱うこと。治療効果に地域間差が生じる原因には、GCP や試験実施計画書からの逸脱等試験実施上の問題も想定されるため、その可能性を否定した上で、交互作用の p 値が小さい場合には、地域間の治療効果の違いを、他の効果修飾因子で説明できないか検討する必要がある。

3.3.2 安全性評価

ICH M4E(r2)の記載のとおり、内因性又は外因性の民族的要因に基づいて治療及び管理を患者個々に考慮するために、安全性に関しても、Layer 2 で効果修飾因子の検討を主軸とした評価をす

ることが重要である。安全性評価対象には、有害事象、臨床検査値、バイタルサイン、心電図、その他所見等があり、多岐にわたり、関連する項目のパターンや関連性を評価することが重要である。比較的発現頻度が高い有害事象に関しては、以下に示す手法も有用である。しかしながら、重篤な有害事象を含め、発現頻度が低い有害事象に関しては、集団としての評価は困難であり、個々の参加者での発現パターンや背景要因等に基づき、参加者間に共通の要因の有無を評価した上で、リスク因子を特定することが必要となる。また、安全性評価は臨床試験で得られた結果のみならず、非臨床試験、薬物相互作用試験、母集団薬物動態データ、利用可能な場合には市販後情報や類薬の情報も含め、包括的、論理的に示す必要がある。

ICH M4E(r2)	<p>2.7.4.5 特別な患者集団及び状況下における安全性</p> <p>2.7.4.5.1 内因性要因</p> <p>本項では、ICHESにて「内因性民族的要因」と定義される人口統計学的要因や他の要因に基づいて治療及び管理を患者個々に考慮するための安全性データを要約すること。内因性要因には、年齢、性別、身長、体重、除脂肪体重、遺伝子多型、身体組成、他の疾患及び臓器機能不全が含まれる。申請適応症が小児でも考えられる場合は、小児における安全性も分析すること。これらの要因が安全性に与える影響については他項に記載されることになるが、腎疾患又は肝疾患患者等における安全性は、薬物動態、その他の情報とともに本項で要約すること。もし、高血圧、心疾患、糖尿病等の特定の合併症のある患者の例数が十分であれば、合併症が治験薬の安全性にどのような影響を与えるのか分析すること。患者部分集団における解析を行う場合は、作成した有害事象の表又は記述を相互に参照すること。</p> <p>2.7.4.5.2 外因性要因</p> <p>本項では、ICHESにて「外因性民族的要因」と定義される人口統計学的要因や他の要因に基づいて治療及び管理を患者個々に考慮するための安全性データを要約すること。これらの要因は、患者をとりまく環境に関する要因である。例としては、医療環境、他剤の使用状況（第2.7.4.5.3項、薬物相互作用を参照）、喫煙、飲酒、食事の習慣等が挙げられる。たとえば、代謝プロフィール、試験結果、市販後調査又は類薬情報等により飲酒と治験薬との相互作用が示唆される場合、それらの内容を本項に示すこと。</p>
-------------	--

3.3.2.1 さまざまな視点による効果修飾因子の検討

(1) One factor at a time アプローチ

1) 頻度分布表

安全性評価でも、有効性評価と同様に、単純で典型的な方法として、内因性又は外因性民族的要因の何らかの要因によって集団を部分集団に分け、有害事象や他の安全性所見（臨床検査値、バイタルサイン、心電図等）の発現頻度の違いを検討するサブグループ解析がある。

事例：フィネレノン

フィネレノンの審査報告書で、PMDAは、本薬の作用機序から血圧低下のリスクが想定されることから、当該注意喚起の必要性について説明を求めたところ、申請者は、以下に示すベースライン時の収縮期血圧別の血圧低下関連事象の発現状況も含めて、説明している。

Table 3-2 頻度分布表：安全性のさまざまな視点による効果修飾因子検討の事例（フィネレノン）

表 56 全体集団におけるベースライン時の収縮期血圧値別の血圧低下関連事象の発現状況
(試験 16244：安全性解析対象集団)

	収縮期血圧値 100 mmHg 未満		収縮期血圧値 100 以上 130 mmHg 未 満		収縮期血圧値 130 以上 160 mmHg 未 満		収縮期血圧値 160 mmHg 以上	
	プラセボ群 (16 例)	本薬群 (9 例)	プラセボ群 (761 例)	本薬群 (779 例)	プラセボ群 (1916 例)	本薬群 (1897 例)	プラセボ群 (138 例)	本薬群 (141 例)
すべての 血圧低下関連事象 ^a	12.5 (2)	11.1 (1)	6.0 (46)	8.9 (69)	3.0 (57)	3.9 (74)	2.2 (3)	3.5 (5)
投与中止に至った事象	0 (0)	0 (0)	0 (0)	0.1 (1)	0 (0)	0 (0)	0 (0)	0 (0)
重篤な事象	0 (0)	0 (0)	0.3 (2)	0.5 (4)	0.2 (4)	0.4 (7)	0 (0)	0 (0)
入院に至った事象	0 (0)	0 (0)	0.3 (2)	0.4 (3)	0.2 (4)	0.4 (7)	0 (0)	0 (0)
死亡に至った事象	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

発現割合% (発現例数)

a : MedDRA PT 「血圧低下」「低血圧」「起立性低血圧」

表 57 全体集団におけるベースライン時の収縮期血圧値別の血圧低下関連事象の発現状況
(試験 17530：安全性解析対象集団)

	収縮期血圧値 100 mmHg 未満		収縮期血圧値 100 以上 130 mmHg 未 満		収縮期血圧値 130 以上 160 mmHg 未 満		収縮期血圧値 160 mmHg 以上	
	プラセボ群 (26 例)	本薬群 (14 例)	プラセボ群 (1168 例)	本薬群 (1173 例)	プラセボ群 (2351 例)	本薬群 (2389 例)	プラセボ群 (113 例)	本薬群 (107 例)
すべての 血圧低下関連事象 ^a	34.6 (9)	28.6 (4)	4.5 (53)	7.5 (88)	2.0 (48)	3.8 (91)	4.4 (5)	5.6 (6)
投与中止に至った事象	0 (0)	7.1 (1)	0 (0)	0 (0)	0.1 未満 (1)	0.1 未満 (1)	0 (0)	0 (0)
重篤な事象	0 (0)	0 (0)	0.2 (2)	0 (0)	0 (0)	0.1 (3)	1.8 (2)	0 (0)
入院に至った事象	0 (0)	0 (0)	0.2 (2)	0 (0)	0 (0)	0.1 (3)	1.8 (2)	0 (0)
死亡に至った事象	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

発現割合% (発現例数)

a : MedDRA PT 「血圧低下」「低血圧」「起立性低血圧」

Source: フィネレノン審査報告書 (2022)

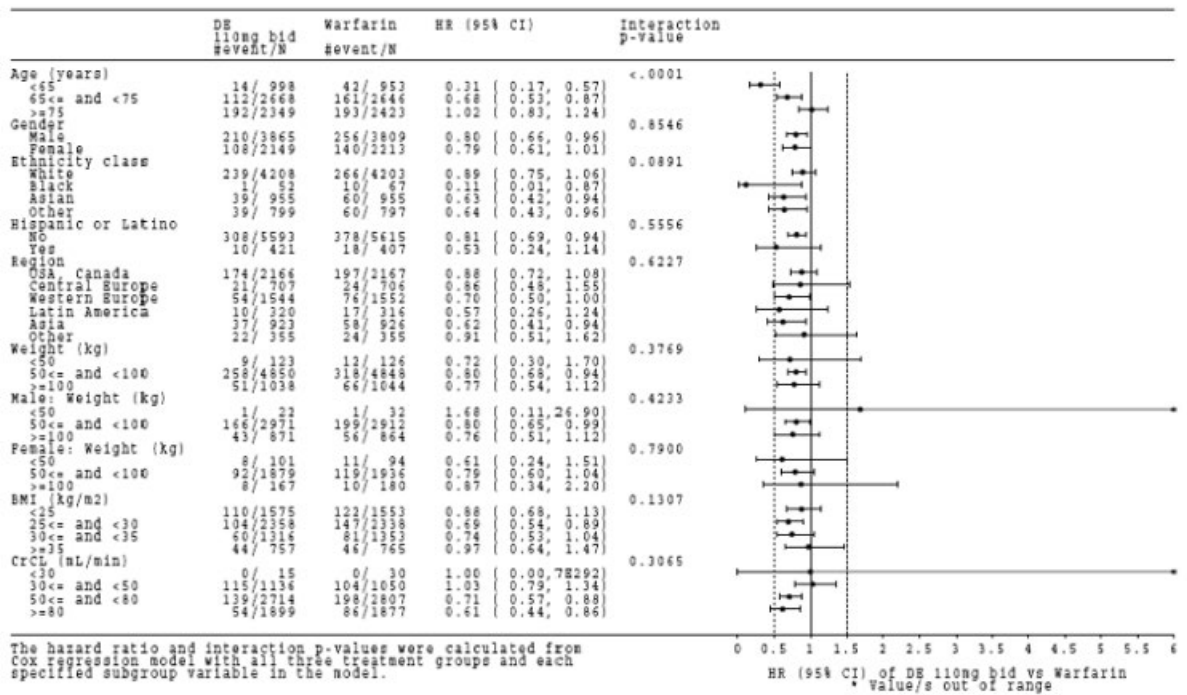
3) Forest plot

有効性評価と同様に、安全性評価でもグラフィカルアプローチとして forest plot が有用な場合がある。

事例：ダビガトラン

Figure 3-16 はダビガトランの第 III 相 MRCT (RE-LY 試験) の大出血発現のワルファリンに対するハザード比の forest plot を示す。Forest plot は、特定の有害事象発現と内因性又は外因性民族的要因の関連を評価するのに有用な方法である。ただし、前述とおり、交互作用の p 値はあくまで参考値として取り扱い、p 値が小さい場合には、その要因が効果修飾因子の可能性を考慮し、更なる検討が必要となる。

Figure 3-16 Forest plot : 安全性のさまざまな視点による効果修飾因子検討の事例 (ダビガトラン)



引用元 : CTD2.7.4, 付録 1, Figure 2.1.1.1.1.26

図 2.5.5.1.7: 1 試験 1160.26 における人口統計学的特性別の大出血のハザード比および 95%信頼区間の比較 (BIBR 1048 MS 110 mg 1 日 2 回投与群 vs ワルファリン投与群) : ランダム化集団

Source: ダビガトラン申請資料概要 (2011)

3) 寄与危険度比 (ARR)

1 つのサブグループが他のサブグループと異なっているかどうかを特定する 1 つの方法として、サブグループ間の差の ARR を算出する方法がある (くすりの安全性を科学する 2012)。ARR は発現頻度の高い事象や、発現頻度が高く薬剤に関連する可能性がある他の所見 (臨床検査値、バイタルサイン等) には有用であるが、発現頻度が低い事象に関しては、このような解析を実施することはできない。

ARR = (被験薬 X の有害事象発現割合 - 対照薬 X の有害事象発現割合) / (被験薬 Y の有害事象発現割合 - 対照薬 Y の有害事象発現割合)

X, Y はサブグループ (e.g. X = 65 歳以上, Y = 65 歳未満)

(2) 統計モデルを用いたアプローチ

安全性評価に関しても、特定の有害事象発現に影響を与える要因を探索するために、統計モデルを導入して変数選択法で影響の強い要因を探索することが有用な場合もある。

事例：エンホルツマブ ベドチン（遺伝子組換え）

がん化学療法後に増悪した根治切除不能な尿路上皮癌に対する治療薬として承認されたエンホルツマブ ベドチンの審査報告書では、PMDA が皮膚障害の発現機序及びリスク因子について申請者の見解を求めたところ、申請者は、一般的なサブグループ解析に加え、リスク因子の候補として、22 項目を設定し、多変量ロジスティック回帰モデルによる変数選択等に基づき、各項目と当該皮膚障害との関連を検討し、リスク因子として、非白人人種、ヘモグロビン値高値、腫瘍径が小さいことを説明している（オッズ比 [95%信頼区間] ①非白人人種 1.947 [1.182, 3.210] , ②ヘモグロビン値が 1 g/dl 上昇 1.024 [1.009, 1.039] , ③腫瘍径が 1 mm 増大 0.991 [0.984, 0.997])
Source: エンホルツマブ ベドチン（遺伝子組換え）審査報告書（2021）

統計モデルには複数の効果修飾因子を同時に組み入れることができるため、それぞれの効果修飾因子の影響、これらの複合的な影響、すなわち効果修飾因子間の交互作用を検討することもできることから、安全性評価でも使用されることがある。

事例：ダビガトラン

ダビガトランの審査報告書で、PMDA は、申請者が出血の危険性が高い患者として、年齢 75 歳以上、P 糖蛋白阻害剤との併用、消化管出血の既往を挙げた根拠を求めた。1 例として、申請者が説明した 75 歳以上の根拠として、以下の 2 点を説明している。

- 1160.26 試験で出血の発現頻度を年齢 75 歳以上及び未満で層別解析した結果、本薬投与群、ワルファリン投与群のいずれにおいても、大出血、すべての出血の年間イベント率は年齢 75 歳以上の患者で高かった。
- Cox 回帰モデル（年齢、ベースライン時のクレアチニンクリアランス、性別、アスピリン併用の有無を共変量とし、投与群と各共変量の交互作用項を入れたモデル）を用いて解析した結果、ワルファリンに対する本薬のハザード比は加齢とともに上昇する傾向が示され、本薬 110 mg 群では 85 歳を超えたところで、本薬 150 mg 群では 75 歳を超えたところでハザード比が 1 を超えた。

Figure 3-17 統計モデルを用いたアプローチ：安全性のさまざまな視点による効果修飾因子検討の事例（ダビガトラン）

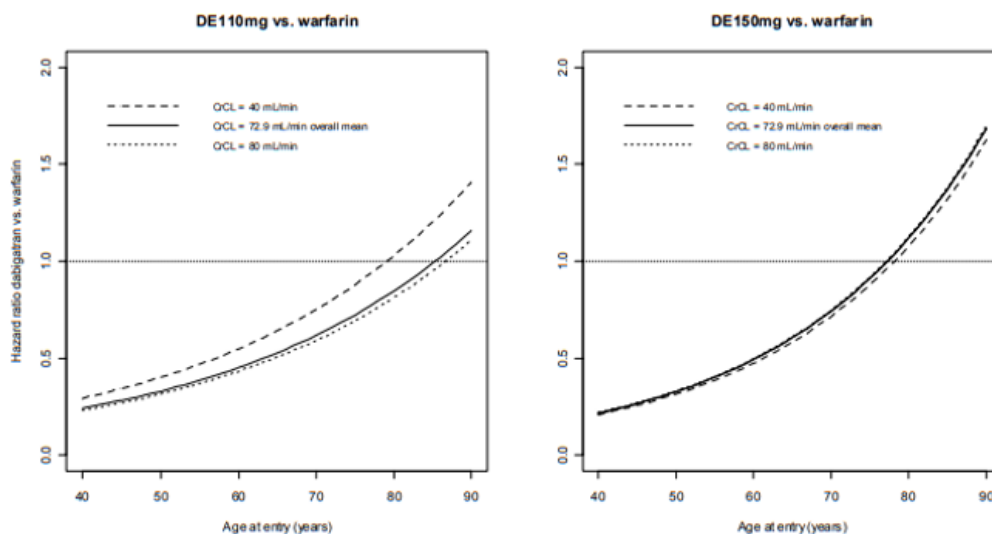


図 2.7.4.2.2.1.1.2: 8 試験 1160.26 における特定のクレアチニンクリアランス値での年齢に対する大出血のハザード比（ランダム化集団）
引用元：CTD 2.7.4, 付録 1, Table 2.1.1.1.7.3

Source: [ダビガトラン申請資料概要（2011）](#)

3.3.2.2 地域間での一貫性の評価

(1) One factor at a time アプローチ

地域別の検討でも、要因別と同様に、単純で典型的な方法として、地域別に集団に分け、治療効果の違いを検討するサブグループ解析がある。また、特定の有害事象発現に対し、グラフィカルアプローチも有用である。

事例：ダビガトラン

ダビガトランの申請資料（CTD 2.7.4）では、試験実施地域別の大出血のハザード比と参加者数、それに応じた 95%信頼区間を以下の頻度分布表（Table 3-3）と funnel plot（Figure 3-18）で示している。

Table 3-3 頻度分布表：安全性の地域間での一貫性の評価（ダビガトラン）

表 2.7.4.2.3.2.1: 3 試験実施地域別の大出血の年間イベント率，ハザード比と信頼区間

地域または国	Subjects with event / Number of randomized subjects (Yearly event rate %)			Hazard ratio (95% CI)	
	DE 110 bid	DE 150 bid	Warfarin	DE 110 bid vs Warfarin	DE 150 bid vs Warfarin
Total	318 / 6015 (2.67)	375 / 6076 (3.11)	396 / 6022 (3.36)	0.79 (0.68, 0.92)	0.93 (0.81, 1.07)
North America	174 / 2166 (3.92)	208 / 2200 (4.61)	197 / 2167 (4.45)	0.88 (0.72, 1.08)	1.04 (0.86, 1.27)
Latin America	10 / 320 (1.82)	15 / 320 (2.73)	17 / 316 (3.18)	0.57 (0.26, 1.24)	0.86 (0.43, 1.72)
Central Europe	21 / 707 (1.53)	23 / 706 (1.69)	24 / 706 (1.77)	0.86 (0.48, 1.55)	0.96 (0.54, 1.69)
Western Europe	54 / 1544 (1.74)	65 / 1555 (2.07)	76 / 1552 (2.46)	0.70 (0.50, 1.00)	0.84 (0.61, 1.17)
Southeast Asia	11 / 378 (1.60)	15 / 381 (2.09)	24 / 375 (3.56)	0.44 (0.22, 0.91)	0.59 (0.31, 1.13)
Eastern Asia	26 / 545 (2.43)	21 / 552 (1.95)	34 / 551 (3.22)	0.74 (0.45, 1.24)	0.60 (0.35, 1.03)
Other	22 / 355 (3.30)	28 / 362 (4.13)	24 / 355 (3.64)	0.91 (0.51, 1.62)	1.15 (0.67, 1.98)
Japan	8 / 107 (5.53)	5 / 111 (3.33)	5 / 108 (3.31)	1.68 (0.55, 5.15)	1.02 (0.29, 3.51)

引用元：CTD 2.7.4, 付録 1, Table 2.1.1.1.1.3, Table 2.1.1.1.1.14, Table 2.1.1.1.1.24, Figure 2.1.1.1.1.26, Figure 2.1.1.1.1.27; 付録 3, Table 15.3.5.3: 3, Table 15.3.2.1: 9; 付録 4, Table 15.3.2.1: 2, Table 15.3.2.1: 9; CTD 5.3.5.1-4, 試験 1160.26, U-3249-01, Table 15.3.5.3: 3, Table 15.3.5.3: 6

Source: [ダビガトラン申請資料概要 \(2011\)](#)

Figure 3-18 Funnel plot：安全性の地域間の一貫性評価の事例（ダビガトラン）

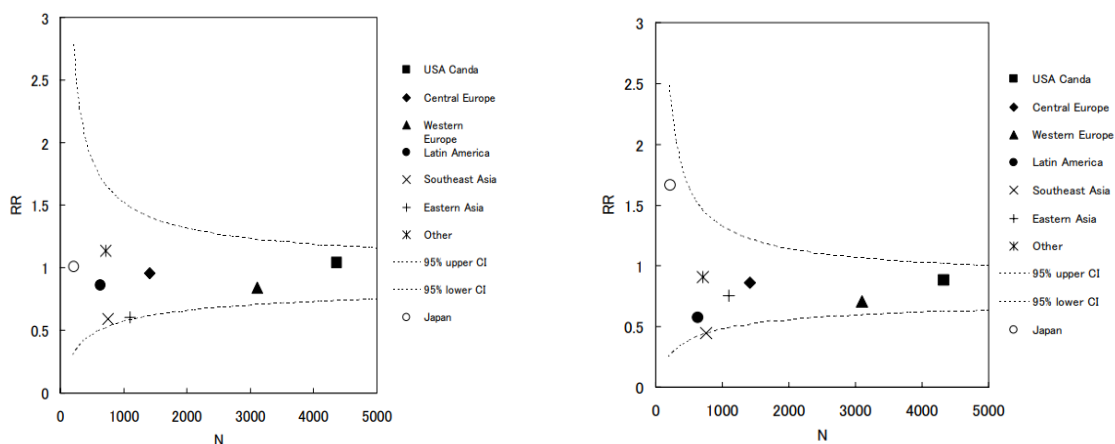


図 2.7.4.2.3.2.1: 3 試験実施地域別の大出血のハザード比と実施例数 (DE 150 bid vs Warfarin)

図 2.7.4.2.3.2.1: 2 試験実施地域別の大出血のハザード比と実施例数 (DE 110 bid vs Warfarin)

RR：相対リスク，N：実施例数
注：95%信頼区間は，全患者データから算出した 95%信頼区間（対数値）をそれぞれの実施地域における実施例数の平方根で補正したもの。
引用元：CTD 2.7.4, 付録 1, Table 2.1.1.1.1.3, Table 2.1.1.1.1.24; 付録 3, Table 15.3.5.3: 3; 付録 4, Table 15.3.2.1: 2; CTD 5.3.5.1-4, 試験 1160.26, U-3249-01, Table 15.3.5.3: 3

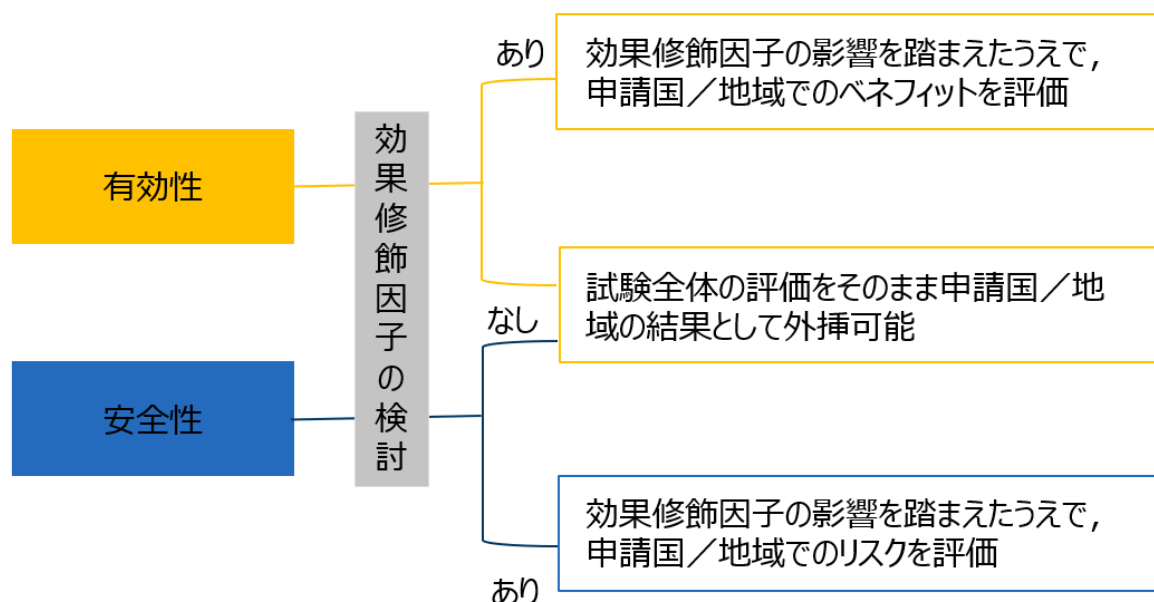
Source: [ダビガトラン申請資料概要 \(2011\)](#)

(2) 治療と地域間の交互作用の検討

3.3.1.2 項のさまざまな視点による効果修飾因子検討の forest plot の事例 (Figure 3-14) と同様に，国又は地域別のサブグループ解析は交互作用の有無の検討を目的として実施するため，交互作用の p 値を示すことがある。しかしながら，前述のとおり，交互作用の p 値はあくまで参考値として取り扱い，p 値が小さい場合には，地域間の治療効果の違いを，他の効果修飾因子で説明できないかを検討する必要がある。

3.4 Layer 3 : 申請国／地域でのベネフィット・リスクの評価

Layer 3 は、Layer 1, Layer 2 の検討結果に基づいて、有効性と安全性の両面から被験薬の有用性を評価し、最終的に申請国／地域でのベネフィット・リスクバランスを評価する階層である。Layer 2 までの評価は、MRCT 全体としてどのような知見が得られたか、その知見は MRCT の外部情報（過去の試験から得られた知見、MRCT と同時期に得られた MRCT 外の知見）と整合するかを総合的に評価するものであり、MRCT に実際に参加した国のみならず、広く世界で共有できる評価である。これらの評価を基礎として、それぞれの国や地域において、どのようなベネフィット・リスクが想定できるかを説明するのが Layer 3 である。Layer 2 の評価で、有効性、安全性ともに臨床的に意義のある効果修飾因子又はその候補が特定されなかった場合には、試験全体の評価をそのまま申請国／地域の結果として外挿可能と判断する。有効性又は安全性で効果修飾因子又はその候補が特定された場合には、その効果修飾因子の影響を踏まえた上で、申請国／地域でのベネフィット・リスクを評価する必要がある。



次項以降で、効果修飾因子が特定された場合に、申請国／地域でのベネフィット・リスクを評価する際に留意すべき事項を示す。

3.4.1 効果修飾因子が特定された場合の申請国／地域でのベネフィット・リスク評価

効果修飾因子が特定された場合には、申請国／地域での効果修飾因子の分布に基づき、被験薬の有効性及び安全性を推定する。効果修飾因子の分布の情報源は臨床試験データに加え、文献や申請国の統計データ等の外部情報も参照した上で、申請国／地域での患者集団を効果修飾因子で特徴づける。効果修飾因子が特定された事例として、チカグレロルの PLATO 試験に基づき効果

修飾因子で申請国／地域で投与対象となる患者集団を限定した事例を以下に、詳細は 5.1 項に示す。

さらに、実施可能な場合には、効果修飾因子を用いた予測モデルに基づいて、申請国／地域の患者集団での治療効果を推定し、そのベネフィット・リスクを評価することも有用である。

事例：チカグレロル

ICH E17 Training Material Module 6に記載されている PLATO 試験では、アスピリンの維持用量が高い集団でチカグレロルの治療効果が低下することが明らかとなった。これにより、各地域の添付文書でアスピリンの維持用量に関する注意が記載された。

- FDA：維持用量 100 mg を超えるアスピリンの併用は、チカグレロルの効果を低下させるため、避けるべき
- EMA：Special Warning 「チカグレロルと維持用量 300 mg 以上のアスピリンの同時投与は推奨されない」
- NMPA：アスピリン維持用量を 75～100 mg とする注意の記載
- PMDA：用法・用量の注意「アスピリン（維持用量として 81～100 mg/日）と併用すること」

4 CTD

4.1 CTD で検討すべき内容

本章では、ICH E17 の理念に基づいた CTD を作成するための基本方針を提案する。ICH E17 の理念に基づいた CTD を作成するためには、従来と異なる情報の追加及び基本的な考え方を考える必要があると考えた。

1 点目は、治療効果に影響を及ぼす重要な効果修飾因子についての検討を充実させる必要がある。具体的には、3 章に記載したとおり、医薬品の治療効果の予測及び理解に役立てるため、申請国／地域だけではなく、試験全体で効果修飾因子になり得る要因がないか探索した結果を示す必要がある。

2 点目は、申請国／地域での承認の可否を判断する情報が必要であるため、試験全体の評価を踏まえた申請国／地域でのベネフィット・リスク評価を示す必要がある。しかし、これは MRCT の全体集団と申請国／地域の集団との結果を比較した一貫性の検討が必要ということの意味するものではなく、申請国／地域への一般化可能性や、申請国／地域の集団で他国／地域と異なる反応を示さないこと、また、効果修飾因子が見つかった場合でも適切に管理すればベネフィット・リスクは良好であることが示せればよいと考える。

MRCT の結果に関する一貫性の検討については、2.1 項で述べたとおり、単一の評価（上述したような、議論対象の MRCT の主要／副次評価項目について、全体集団の結果と同じ傾向が申請国／地域の集団の結果でも示されているか否か等）のみに基づき結論付けられるものではなく、多角的・体系的に評価することが重要である。このため、ICH E17 の理念に従い、MRCT の試験全体の結果について、5 つの視点に基づき考察することが必要であると考えた。

上記の要素を追加した CTD を作成するために、本書では 3 章で紹介した 3-layer approach を使用することを推奨する。3-layer approach は、これらの要素を盛り込み、系統立ててわかりやすく CTD を作成するために有用であると考えた。

このような基本方針に基づき、CTD に記載すべき内容及びその記載箇所を以下に提案する。

4.2 CTD の記載内容及び記載箇所

4.1 項に記載したとおり、本書では、ICH E17 の理念に基づいた CTD を作成するために、3-layer approach に基づいて結果及び考察を示すことを推奨する。3-layer approach の概略は、3 章を参照する。

また、MRCT の試験全体の結果、並びに効果修飾因子の検討及び地域間での一貫性の評価について考察する場合は、ICH E17 に示された 5 つの視点に基づき検討することを推奨する。ICH E17 の概要については、2 章を参照する。

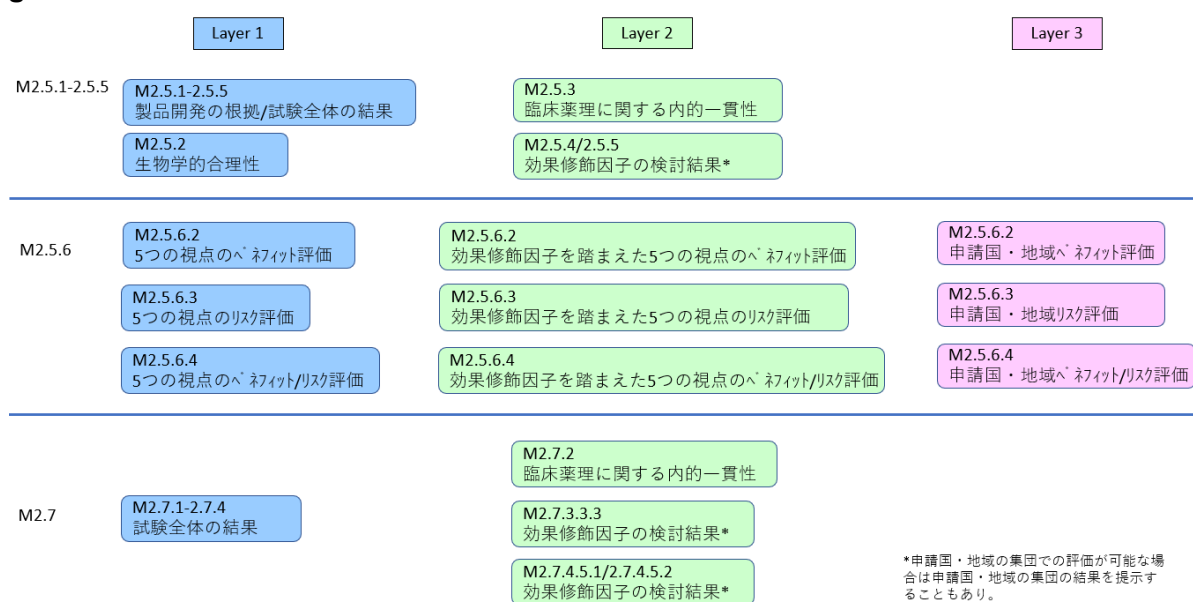
ICH E17 の理念に基づいた CTD を作成するための基本的な構成要素及び CTD 内の推奨される記載箇所を Layer ごとに Figure 4-1 に示す。

ICH E17 の理念に基づいた CTD の全体像を把握しやすくするため、本項及び Figure 4-1 では、特に対応する必要があると思われる要素にフォーカスして示しており、従来から一般的に記載されている試験全体の結果及び特に変更が生じない要素については、簡略化して示した。

まず、Layer 1 では、MRCT の試験全体の結果と、それに基づく 5 つの視点の考察及びベネフィット・リスク評価を記載する。次に、Layer 2 では、試験全体の結果に対する影響要因の検討及び地域間での一貫性の評価として、効果修飾因子に関する検討結果と、それに基づく 5 つの視点の考察及びベネフィット・リスク評価を記載する。さらに、Layer 3 では、Layer 2 までの試験全体の結果でのベネフィット・リスク評価を踏まえた、申請国／地域でのベネフィット・リスク評価を記載する。

各 Layer に記載すべき内容の概要を 4.2.1 項から 4.2.3 項に記載する。

Figure 4-1 ICH E17 の理念に基づいた CTD を作成するための基本的な構成要素



本書で提案する基本方針に基づいた CTD を作成するために、特に対応が必要である M2.5.6 については、詳細な説明を 4.3 項に示す。また、本項の基本的な考え方に従い作成した CTD の事例を 5 章に示す。

4.2.1 試験全体の結果の評価 (Layer 1)

本項では、Layer 1 として示すべき、MRCT の試験全体の結果の評価について説明する。

Layer 1 では、製品開発の根拠及び MRCT の試験全体の結果を示す。これらは、従来から ICH M4E(r2)に従って CTD に示されてきた内容であり、本書で特に追加で対応を推奨する事項はない。ICH M4E(r2)に記載のとおり、Layer 1 の構成要素として、試験全体の結果を M2.5.1 から M2.5.6 及び M2.7.1 から M2.7.4 に記載する。

ただし、ICH E17の理念に基づいたCTDとするために、本書では、前述のICH E17の5つの視点を用いて、MRCTの試験全体の結果を多角的・体系的に考察し、ベネフィット・リスクを評価することを推奨する。試験全体の結果について、考察すべきICH E17の5つの視点の詳細を3.2項、4.3.2.1項及び4.3.3.1項に示す。

試験全体の結果についてのICH E17の5つの視点に基づく考察は、M2.5.6.2、M2.5.6.3及びM2.5.6.4に記載することを推奨する。また、5つの視点の考察に用いた根拠情報は、M2.5.1からM2.5.6及びM2.7.1からM2.7.4の該当箇所に示しておく必要がある。

4.2.2 結果に対する効果修飾因子の検討及び地域間での一貫性の評価（Layer 2）

本項では、Layer 2として示すべき、試験全体の結果に対する効果修飾因子の検討及び地域間での一貫性の評価について説明する。

Layer 2は、ICH E17の理念に基づいたCTDを作成するために最も重要な要素である。3.3項に記載したとおり、Layer 2の目的は、効果修飾因子又はその候補を探索し、国／地域又は集団間で差異が認められた場合は、効果修飾因子又は分布の違いによって、それらを説明することである。したがって、Layer 2で示すべき主な内容は、効果修飾因子の探索又は地域間での一貫性を評価するために実施したサブグループ解析及びその他の検討結果と、それに基づくICH E17の5つの視点からの効果修飾因子の考察、並びにこれらの検討結果を踏まえたベネフィット・リスク評価となる。

Layer 2までの情報は、世界各国／地域で共有できるため、申請国／地域以外又は申請国／地域とは直接関係がないと思われる集団で違いが認められた場合でも、その原因を考察することが重要である。また、Layer 2での効果修飾因子の検討をどの程度実施すればよいかについては、薬剤／申請ごとに検討が必要であるが、充実させればさせるほど良いということではなく、Layer 3を検討するために必要なものを特定して実施することが望ましい。本書では、米国、EU、日本及び中国での申請を念頭に記載する。

効果修飾因子の探索及び地域間での一貫性を評価するための解析方法の事例は、3.3項に示した。これらのサブグループ解析並びにその他の効果修飾因子又は地域間の一貫性の検討結果は、M2.5.4及びM2.5.5、並びにM2.7.3.3.3、M2.7.4.5.1及びM2.7.4.5.2に記載することを推奨する。

実施した効果修飾因子又は地域間の一貫性の検討結果は、ICH E17の5つの視点に基づいて、多角的・体系的に考察することを推奨する。効果修飾因子の検討結果について考察すべきICH E17の5つの視点を3.3項、4.3.2.2項及び4.3.3.2項に示す。

治療効果に影響を与える内因性・外因性要因（効果修飾因子）及び国／地域間の差に関する考察の方向性は、Table 4-1を参考とすることを推奨する。

効果修飾因子の検討結果及びそれに基づく5つの視点の考察で、試験結果に重大な影響を及ぼす可能性のある効果修飾因子が特定されず、地域間差も認められなかった場合は、現在までの検討結果では、試験全体の結果を申請国／地域に適用することを否定するエビデンスは得られな

ったとの結論になり、MRCT の試験全体の結果をそのまま申請国／地域に適用することが可能となる。

試験結果に重大な影響を及ぼす可能性のある効果修飾因子が特定された場合は、要因の分布に関する地域間差の有無にかかわらず、その影響を考慮して各国／地域での被験薬の有効性又は安全性を評価する必要がある。効果修飾因子について、考察すべきポイントの詳細を3.3項、4.3.2.2項及び4.3.3.2項に示す。

Table 4-1 治療効果に影響を与える内因性・外因性要因（効果修飾因子）及び国／地域間の差に関する考察の方向性

		内因性・外因性要因（効果修飾因子）	
		なし	あり
国／地域間の差	なし	【ケース A】 全体の結果は頑健であり、全体的な結論をすべての国／地域に適用できる。	【ケース B】 生じた影響は世界的に共通であり、効果修飾因子に関する同じ検討アプローチをすべての国／地域に適用できる。
	あり	【ケース C】 国／地域間と内因性・外因性要因の観点の交絡因子を特定するために、今まで検討されていない因子（国／地域以外）をさらに検討する必要がある。 ①さらに検討した結果、新たな内因性・外因性要因（効果修飾因子）が特定された場合は、ケース B と同様に対応する。 ②さらに検討しても、全体の結果に影響を及ぼす効果修飾因子が特定されなかった場合は、他の要因（たとえば、統計学的不確実性又は治験実施状況等）を考察する。影響が認められた国／地域でのベネフィット・リスクを Layer 3 で評価する際も本件を考慮する。	【ケース D】 国／地域間と内因性・外因性要因の観点の交絡因子が存在する可能性があるため、国／地域間で認められた影響が他の内因性・外因性要因の観点から説明できないか検討する。 ①国／地域間で認められた影響が他の内因性・外因性要因の観点から説明できる場合は、ケース B と同様に対応する。 ②国／地域間で認められた影響が他の内因性・外因性要因の観点から説明できない場合は、ケース C と同様に、他の要因（たとえば、統計学的不確実性又は治験実施状況等）を考察する。影響が認められた国／地域でのベネフィット・リスクを Layer 3 で評価する際も本件を考慮する。

Source: Yoshida et al. 2015 を改変

サブグループ解析等の結果に基づく 5 つの視点からの効果修飾因子の考察と、それを踏まえたベネフィット・リスク評価の結果は、M2.5.6.2、M2.5.6.3 及び M2.5.6.4 に記載することを推奨する。

また、MRCT で、申請国／地域のみで評価可能な数の参加者を組み入れられた場合は、Layer 2 の一部として、申請国／地域の集団での結果を示すことが有用な場合もある。ただし、この場合も、全体集団と申請国／地域の集団との結果の一貫性を検討するのではなく、ICH E17 の基本的な理念に則り、試験全体としての一貫性（すなわち、5 つの視点に基づく結果の評価）を検討することが重要であるため、申請国／地域の集団での結果は、Layer 2 のサブグループ解析の一部として示すことを推奨する。

申請国／地域の集団の結果を示す場合は、M2.5.4 及び M2.5.5、M2.7.3.3.3 及び M2.7.4.5 に記載することを推奨する。

4.2.3 各地域でのベネフィット・リスク評価（Layer 3）

本項では、Layer 3 として示すべき、申請国／地域でのベネフィット・リスク評価について説明する。本書では、米国、EU、日本及び中国での承認申請に用いる CTD の作成を想定して提案する。

Layer 3 の目的は、Layer 1 及び Layer 2 での検討結果を基に、各国／地域のベネフィット・リスクを考察又は推定し、治療効果を推定したい国／地域又は集団の特徴を、効果修飾因子を用いて表現することである。したがって、Layer 3 で示すべき主な内容は、Layer 1 及び Layer 2 の試験全体でのベネフィット・リスク評価（エビデンスの強さ、限界、不確実性等を含む）を踏まえた申請国／地域でのベネフィット・リスク評価となる。具体的には、Layer 2 での効果修飾因子の評価結果を鑑みた上で、申請国／地域とその他の国／地域での内因性又は外因性の患者背景要因の違いや不均衡について考察し、内因性又は外因性の患者背景要因の違いによる有効性及び安全性への影響が Layer 1 及び Layer 2 の結果から得られているようであれば、それらの要因の地域的な違い又は不均衡による有効性及び安全性への影響を評価する。申請国／地域でのベネフィット・リスク評価について、特に考察すべきポイントの詳細を 3.4 項及び 4.3.2.3 項に示す。

情報のソースは、開発プログラムで収集されたデータのみならず、公表文献や各国又は世界の統計情報、リアルワールドデータ等から得られる場合もある。

申請時までの情報に基づく検討は、患者背景要因が主であり、申請国／地域の医療環境等の環境要因の検討は限定的であることが多いと思われることから、環境要因に関する本格的な検討は市販後に実施されることになると想定される。

Layer 3 では、あくまで Layer 1 及び Layer 2 の検討結果に基づき、申請国／地域でのベネフィット・リスクを評価することを推奨する。Layer 3 には、基本的に申請国／地域でのベネフィット・リスク評価を記載するが、他国／地域の評価が申請国／地域の評価の参考になる場合もあるかもしれない。MRCT で申請国／地域でのサブグループ解析を実施した場合であっても、本書では、上述したとおり、Layer 3 ではなく Layer 2（サブグループ解析の一部）として取り扱う。

申請国／地域でのベネフィット・リスク評価は、M2.5.6.2、M2.5.6.3 及び M2.5.6.4 に記載することを推奨する。

4.3 M2.5.6（ベネフィットとリスクに関する結論）の各項及びそのソース文書で検討・記載すべき事項

前項までに説明した、5 つの視点、3-layer approach、効果修飾因子の検討結果は、最終的にすべて 2.5.6 ベネフィットとリスクに関する結論項に集約される。そのため、2.5.6 の項別に、ICH

M4E(r2)ガイドラインで示されている記載事項を踏まえ、各項及びその引用元で留意すべき事項を次項以降で説明する。

4.3.1 M2.5.6.1 治療の背景

本項では、申請医薬品の治療の背景を簡潔に記載すること。M2.5.6.2 及び M2.5.6.3 の Layer 2 及び Layer 3 の生物学的合理性を検討するにあたり重要な、関連する部分集団の差異が知られている場合には、本項で説明すること。

ICH M4E(r2)	本項では、申請医薬品の治療の背景を簡潔に記載すること。治療の背景では、治療の対象となる疾患又は症状、治療の対象となる集団、現行の治療のベネフィット及びリスクについて記載する。疾患又は症状を理解する上での重大な限界、並びに現行の治療のベネフィット及びリスクの重大な不確実性についても記載する。関連する部分集団における差異が知られている場合には、それらについて記載すること。申請医薬品に関するベネフィット及びリスクの情報を本項に記載すべきではなく、当該情報はそれぞれ 2.5.6.2 及び 2.5.6.3 項に記載すること。
-------------	--

4.3.1.1 M2.5.6.1.1 疾患又は症状

本項では、対象となる患者集団に最も関連する、又は重大な影響を及ぼす疾患若しくは症状の特徴を記載すること。地域や関連する部分集団の差異が知られている場合には、本項で説明すること。

ICH M4E(r2)	本項では、対象となる患者集団に最も関連する、又は重大な影響を及ぼす疾患若しくは症状の特徴（たとえば、発病率、罹病期間、罹患率、死亡率、健康に関連する生活の質等）を記載すること。申請医薬品の目的とする適応症に関連する疾患の側面に焦点をあてること。必要な場合には、疾患の社会的影響又は公衆衛生への影響（たとえば、感染症の不十分な予防及び管理が及ぼす影響）についても記載すること。
-------------	---

4.3.1.2 M2.5.6.1.2 現行の治療

本項では、対象となる患者集団における主要な治療法とともに、有効性、安全性、忍容性、利便性又は患者による選好の観点から、新たな治療の医学的必要性について記載すること。薬理学的分類が同じ治療薬で、関連する部分集団の差異が知られている場合には、本項で簡潔に説明すること。また、現在用いられている治療が地域間で異なる場合は、ここで簡潔に説明すること。

ICH M4(r2)	本項では、対象となる患者集団における主要な治療法（すなわち臨床ガイドラインにおいて推奨されている治療、最もよく用いられている治療等）とともに、有効性、安全性、忍容性、利便性又は患者による選好の観点から、新たな治療の医学的必要性について記載すること。薬理学的分類の異なる治療薬が存在する疾患領域では、薬効分類によりまとめて簡潔に記載してもよい。対象となる患者集団に対し用いられる他の治療についても、これが日常診療又は臨床ガイドラインとして確立している場合には記載してもよい。当該治療には、医学的処置、手術、適応外使用されている医薬品、他の非薬物的介入（たとえば、食事療法、理学療法等）が含まれる。現在用いられている治療に関する地域間の主要な違いについても記載することができる。対象となる患者集団に対して現在利用可能な治療法がない場合には、その旨を記載すること。
------------	---

4.3.2 M2.5.6.2 ベネフィット

本項では、申請医薬品のベネフィット・リスク評価において議論される主要なベネフィットに関するデータの事実に基づき要約する。ICH E17 の理念に基づいた CTD とするために、本書では、3-layer approach, 5つの視点で Clinical Data Package に含めた試験の結果を体系的にまとめ、被験薬のベネフィットを評価することを推奨する。

次項以降に、本項での考慮すべき事項を、3-layer approach, 5つの視点で纏めて提示する。

ICH M4(r2)	<p>本項では、申請医薬品のベネフィット・リスク評価において議論される主要なベネフィットに関するデータの事実に基づく要約を記載すること。ベネフィットとは申請医薬品の好ましい効果である。ベネフィットは試験における評価項目の組み合わせで表される場合もある（たとえば、喘息の増悪、入院の頻度、喘息に関連した死亡数等により喘息コントロールの改善のベネフィットを表す）。</p> <p>代替エンドポイントがベネフィット評価の根拠になっている場合、代替エンドポイントによる臨床的ベネフィットの予測可能性とその根拠について説明すること。また、ベネフィットには、患者のコンプライアンスの改善につながる利便性（たとえば、より便利な用法・用量、又は投与経路）、又は患者以外の人に及ぼすベネフィット（たとえば、ワクチンによる集団免疫のベネフィット）等、申請医薬品の重要な特性が含まれる場合もある</p> <p>申請医薬品の主要なベネフィットを特定するにあたり、以下の特性について検討すること。</p> <ul style="list-style-type: none"> ● ベネフィットの臨床的重要性（たとえば、生存期間の延長、治癒、疾患修飾、症状軽減、患者のコンプライアンスの改善、機能又は生活の質の改善、疾患進行の抑制、感染症の予防、診断） ● 対照と比較した試験対象集団における効果の絶対的な差の大きさ；場合によっては、対照との相対的な差を示すことも有用である [たとえば、反応率が申請医薬品の群で20%、対照群で8%の場合、絶対的な差は12%（すなわち、20%-8%）、相対的な差は2.5（すなわち、20%/8%）] <p>各主要なベネフィットについて記載するにあたり、上記の点に加えて以下の点についても考慮すること。</p> <ul style="list-style-type: none"> ● 主要なベネフィットの経時的な変化（たとえば、効果の発現時期、効果の持続性） ● 年齢、性、人種、臓器機能、疾患重症度、遺伝子多型等の関連する部分集団を考慮した主要なベネフィットの差異 <p>本項では、各主要なベネフィットに関連するエビデンスの強さ、限界及び不確実性の分析、並びにこれらの情報の解釈についても記載すること。必要に応じて以下の点を考慮すること。</p> <ul style="list-style-type: none"> ● 試験デザインの考察（たとえば、実薬に対する優越性又は非劣性、プラセボに対する優越性、盲検性、比較対照なし） ● データ収集の完全性及び追跡期間 ● 臨床試験の数、試験間での結果の一貫性 ● 曝露量（たとえば、薬物血中濃度）とベネフィットの関係 ● 臨床試験結果の日常診療への一般化可能性（たとえば、試験対象集団と目的とする適応の患者集団との臨床的に重要な差異） <p>対象となる患者集団におけるベネフィットを予測する代替エンドポイントの信頼性</p>
------------	---

4.3.2.1 Layer 1

Layer 1 では、Clinical Data Package に含めた試験全体の結果に基づき、被験薬の主要なベネフィットを説明する。Layer 1 では、ICH E17 の5つの視点に基づき、各主要なベネフィットに関連するエビデンスの強さ、限界及び不確実性の分析、並びにこれらの情報の解釈を記載する。この5つの視点は、ICH M4E(r2)で求められている内容とも一致する。

- 生物学的合理性
 - 得られた結果は、臨床的、薬理的又は機序的な考察に基づき合理的に説明できるか
 - 曝露量（たとえば、薬物血中濃度）とベネフィットの関係
- 内的一貫性
 - 生物学的又は医学的に関連のある複数の評価項目の結果で、同じ結論を支持する結果が得られているか
 - 主要なベネフィットの経時的な変化（たとえば、効果の発現時期、効果の持続性）
- 外的一貫性
 - 臨床試験の数、試験間での結果の一貫性
 - 類薬/類似した試験集団での結果との一貫性

注：斜体は ICH E17 ガイドラインも考慮した補足説明を示す。

- 統計学的な不確実性
 - 試験デザインの考察（たとえば、実薬に対する優越性又は非劣性、プラセボに対する優越性、盲検性、比較対照なし）
 - データ収集の完全性及び追跡期間
 - 対象となる患者集団におけるベネフィットを予測する代替エンドポイントの信頼性
 - 得られた所見がどの程度確からしいか
 - 推定に無視できないほどのバイアス（GCPの遵守等も考慮）は混入していないか
 - 推定の精度はどの程度高いか
- 臨床的意義
 - ベネフィットの臨床的重要性（たとえば、生存期間の延長、治癒、疾患修飾、症状軽減、患者のコンプライアンスの改善、機能又は生活の質の改善、疾患進行の抑制、感染症の予防、診断）
 - 臨床試験結果の日常診療への一般化可能性（たとえば、試験対象集団と目的とする適応の患者集団との臨床的に重要な差異）
 - 各主要なベネフィットに関連するエビデンスの強さ、限界及び不確実性の分析の結果が、臨床的な判断や治療方針の決定の根拠を与える重要な知見か

4.3.2.2 Layer 2

Layer 2 では、ICH M4E(r2)の以下について、エビデンスの強さ、限界を適切に評価するために、地域、民族的要因に関して、5つの視点を考慮した上で、効果修飾因子を主軸とした評価の結果を考察することを推奨する。

- 年齢、性、人種、臓器機能、疾患重症度、遺伝子多型等の関連する部分集団を考慮した主要なベネフィットの差異

さまざまな視点による効果修飾因子及び地域間での一貫性を評価するための解析方法の事例を3.3.1.2項に示す。これら多角的に検討した結果の詳細は、M2.5.4、並びにM2.7.3.3.3に記載し、本項では結果に基づく考察を、効果修飾因子を主軸に5つの視点に基づき簡潔に説明することを推奨する。

4.3.2.3 Layer 3

Layer 3 は、Layer 1、Layer 2 の検討結果に基づいて、被験薬の主要なベネフィットを評価し、最終的に申請国／地域でのベネフィットを評価する。Layer 2 の評価で、有効性の効果修飾因子又はその候補が特定されなかった場合には、試験全体の評価をそのまま申請国／地域の結果として外挿可能と判断する。有効性の効果修飾因子又はその候補が特定された場合には、その効果修飾因子の影響を踏まえた上で、申請国／地域での主要なベネフィットを評価することを推奨する。

Layer 2 までの結果で、効果修飾因子又はその候補が特定された場合又は地域間で差が認められた場合に、Layer 3 で考慮すべきポイントを以下に記載する。

- 効果修飾因子又はその候補が特定された場合（Table 4-1 のケース B 及び D）：
申請国／地域で、内因性又は外因性の患者背景因子に違いや不均衡があるか。また、特定された効果修飾因子又はその候補に関する患者背景因子に違いや不均衡が認められた場合、その違いや不均衡が申請国／地域でのベネフィットに影響を与えるか。
- 地域間の一貫性の検討で申請国／地域を含む集団に差が認められた場合（Table 4-1 のケース C 及び D）：
申請国／地域を含む集団で治療効果に差が認められたが、それに関する効果修飾因子が特定されなかった場合、又はその違いが他の内因性又は外因性の民族的要因の観点から説明できない場合、その差が申請国／地域のベネフィットにどのような影響を与えるか（他にどのような原因が考えられるか）。
なお、申請国／地域を含む集団で治療効果に差が認められたが、その差が効果修飾因子を用いて説明できる（効果修飾因子として特定できる）場合、又は他の内因性又は外因性の民族的要因の観点から説明できる場合は、上記の「効果修飾因子又はその候補が特定された場合」を参照する。

4.3.3 M2.5.6.3 リスク

本項では、申請医薬品のベネフィット・リスク評価において議論される主要なリスクに関するデータの事実に基づき要約する。ICH E17 の理念に基づいた CTD とするために、本書では、3-layer approach, 5つの視点で Clinical Data Package に含めた試験の結果を体系的にまとめ、被験薬の主要なリスクを評価し、各主要なリスクをモニタリング、最小化又は管理することが可能かを説明することを推奨する。

次項以降に、本項での考慮すべき事項を、3-layer approach, 5つの視点で纏め、その引用元での検討事項を次項以降に提示する。

ICH M4(r2)	<p>本項では、申請医薬品のベネフィット・リスク評価において議論される主要なリスクに関するデータの事実に基づく要約を記載すること。リスクには申請医薬品に関連する有害事象及び他の好ましくない効果が含まれる。考慮されるリスクとしては、薬物相互作用、非臨床データにおいて特定されたリスク、患者以外の人に対するリスク（たとえば、胎児、申請医薬品を取り扱う人及び投与する人）、薬理学的分類又は製品の現時点での知見に基づくリスクについても含まれる。製品の誤用、乱用又は転用の可能性についても考慮すること。本項で記載する主要なリスクには、申請資料の他の項（たとえば、リスク管理計画、添付文書）で記載しているすべての重要なリスクが含まれないこともある</p> <p>申請医薬品の主要なリスクを特定するにあたり、以下のリスクの特性について検討すること。</p> <ul style="list-style-type: none"> ● 重篤性及び/又は重症度 ● 発現率 ● 可逆性 ● 忍容性 <p>各主要なリスクについて記載するにあたり、上記の点に加えて以下の点についても考慮すること。</p>
------------	--

<ul style="list-style-type: none"> ● 発現率は、一般的に、対照との比較及び対象となる患者集団における自然発現率に照らした絶対的な差として示される。場合によっては、対照との相対的な差を示すことも有用である。発現率が申請医薬品の群で8%、対照群で5%の場合、絶対的な差は3%（すなわち、8%-5%）、相対的な差は1.6（すなわち、8%/5%）である。 ● リスクをモニタリング、最小化又は管理できるか否か ● 年齢、性、人種、体重、臓器機能、疾患重症度、合併症、併用療法又は遺伝子多型等の関連する部分集団を考慮した主要なリスクの差異 ● 試験対象集団における有害事象の経時的な変化（すなわち、発現又は消失時期や、発現率が投与初期に高くその後低下するか、比較的一定であるか、又は曝露の累積とともに上昇するか等） <p>本項では、各主要なリスクに関連するエビデンスの強さ、限界及び不確実性の分析並びにこれらの情報の解釈についても記載すること。必要に応じて以下の点を考慮すること。</p> <ul style="list-style-type: none"> ● 試験デザインの考察（たとえば、実薬との比較、プラセボとの比較、盲検性、比較対照なし） ● リスク評価の妥当性（たとえば、患者数、試験数及び試験デザイン、曝露期間、モニタリングの頻度） ● 医薬品開発において特定された安全性の課題に対応するための検討（たとえば、非臨床試験の知見に対応するために計画された眼科検査） ● データ収集の完全性及び追跡期間 ● 申請用法・用量が投与された関連する部分集団の患者数 ● 非臨床情報又は同じ薬効分類の薬剤に共通する作用の情報を含む、既に知られている有害事象の発現機序 ● リスクに影響する患者特性に関する情報の完全性（たとえば、喫煙歴、併用療法） ● 試験間の結果の一貫性 ● 曝露量（たとえば、薬物血中濃度）とリスクの関係 ● 臨床試験結果の日常診療への一般化可能性（たとえば、試験対象集団と目的とする適応の患者集団との臨床的に重要な差異） <p>各主要なリスクを管理するための方策について、その方策によりそのリスクを適切に管理できることを保証する合理的な理由も含めて記載すること。リスク管理計画の詳細な記載の重複は必要ない。すべての主要なリスクを特定・記載した後に、全体的なリスク管理の方策を記載することで十分な場合もある。</p>
--

4.3.3.1 Layer 1

Layer 1 では、Clinical Data Package に含めた試験全体の結果に基づき、被験薬の主要なリスクを説明する。Layer 1 では、ICH E17 の 5 つの視点に基づき、各主要なリスクに関連するエビデンスの強さ、限界及び不確実性の分析、並びにこれらの情報の解釈を記載する。この 5 つの視点は、ICH M4E(r2)で求められている内容とも一致する。

- 生物学的妥当性
 - 得られた結果は、臨床的、薬理学的又は機序的な考察に基づき合理的に説明できるか
 - 非臨床情報又は同じ薬効分類の薬剤に共通する作用の情報を含む、既に知られている有害事象の発現機序
 - 曝露量（たとえば、薬物血中濃度）とリスクの関係
- 内的一貫性
 - 関連する評価項目（たとえば、有害事象と臨床検査値）で一致した結果が得られているか
 - 有害事象の経時的な変化、可逆性
- 外的一貫性
 - 試験間での結果の一貫性

- 類薬／類似した試験集団での結果との一貫性
- 被験薬の他の適応症での結果との一貫性
- 医薬品開発において特定された安全性の課題に対応するための検討（たとえば、非臨床試験の知見に対応するために計画された眼科検査）
- 統計学的な不確実性
 - 試験デザインの考察（たとえば、実薬との比較、プラセボとの比較、盲検性、比較対照なし）
 - リスク評価の妥当性（たとえば、患者数、試験数及び試験デザイン、曝露期間、モニタリングの頻度）
 - データ収集の完全性及び追跡期間
 - 申請用法・用量が投与された関連する部分集団の患者数
 - リスクに影響する患者特性に関する情報の完全性（たとえば、喫煙歴、併用療法）
- 臨床的意義
 - 臨床試験結果の日常診療への一般化可能性（たとえば、試験対象集団と目的とする適応の患者集団との臨床的に重要な差異）
 - 各主要なリスクに関連するエビデンスの強さ、限界及び不確実性の分析の結果が、臨床的な判断や治療方針の決定の根拠を与える重要な知見か

4.3.3.2 Layer 2

Layer 2 では、ICH M4E(r2)の以下について、エビデンスの強さ、限界を適切に評価するために、地域、民族的要因に関して、5つの視点を考慮した上で、効果修飾因子を主軸とした評価の結果を考察することを推奨する。

- 年齢、性、人種、体重、臓器機能、疾患重症度、合併症、併用療法又は遺伝子多型等の関連する部分集団を考慮した主要なリスクの差異

さまざまな視点による効果修飾因子及び地域間での一貫性を評価するための解析方法の事例を3.3.2項に示す。これら多角的に検討した結果の詳細は、M2.5.5、並びにM2.7.4.5.1及びM2.7.4.5.2に記載し、本項では結果に基づく考察を、効果修飾因子を主軸に5つの視点に基づき簡潔に説明することを推奨する。

4.3.3.3 Layer 3

Layer 3 は、Layer 1、Layer 2 の検討結果に基づいて、被験薬の主要なリスクを評価し、最終的に申請国／地域でのリスクを評価する。Layer 2 の評価で、安全性の効果修飾因子又はその候補が特定されなかった場合には、試験全体の評価をそのまま申請国／地域の結果として外挿可能と判断する。安全性の効果修飾因子又はその候補が特定された場合には、その効果修飾因子の影響を踏まえた上で、申請国／地域での主要なリスクを評価することを推奨する。

Layer 2 までの結果で、効果修飾因子若しくはその候補が特定された場合又は地域間で差が認められた場合に、Layer 3 で考慮すべきポイントは [4.3.2.3 項](#) を参照する。

4.3.4 M2.5.6.4 ベネフィット・リスク評価

本項では、[4.3.2 項](#) 及び [4.3.3 項](#) で考察したベネフィット・リスク評価について、Layer ごとのバランスを考慮し、結論を簡潔に記載する。また、各 Layer の評価を個別に述べるだけでなく、Layer 1 から Layer 3 までの評価結果を通して、全体的に一貫した流れのある構成とすることを推奨する。これ以外については、ICH M4E(r2)に従って記載する。

5 Example

本項では、集団又は地域間での有効性及び安全性の多角的かつ体系的な一貫性評価に基づいたベネフィット・リスクの説明の参考として、CTDのM2.5又はM2.7.3での有効性及びベネフィットのLayer 2及びLayer 3の検討及び考察の記載例を示す。記載例は、ICH E17 Training Material Module 6に提示されているチカグレロル（5.1項）及び仮想事例2つ（5.2項及び5.3項）を題材とし、それぞれ背景情報、CTD記載例、及び5つの視点でのまとめを示す。なお、これら3事例の背景情報及び試験結果は、審査関連情報及び公表資料〔Carroll and Fleming 2013, 第7回 ICH E17 ワークショップ（2022）〕を参考としたものの、CTD記載のイメージ共有を容易にするために単純化した架空設定であり、仮想事例では対象疾患に関連する正確性、適切性、及び現実性は考慮していない。

5.1 Example 1：チカグレロル

5.1.1 背景情報

対象疾患	急性冠症候群
品目	チカグレロル
申請国	米国
申請パッケージ中のMRCT	1試験 国際共同第III相試験 PLATO 試験（Pivotal試験）
Pivotal試験のデザイン	
デザイン概略	ランダム化、二重盲検、クロピドグレル対照、国際共同第III相試験 参加国：米国他、計43ヵ国
投与群	チカグレロル群（以下、本薬群）、クロピドグレル群（以下、対照群） ランダム化比1：1、層別因子：なし いずれの群も基礎治療としてアスピリン維持用量75～100 mg/日を併用（ステント留置例では留置6ヵ月後まで325 mg/日を投与可）
投与期間	6～12ヵ月
主要評価項目	心血管死、心筋梗塞、及び脳卒中の複合イベントのいずれかの初回発現までの期間
副次評価項目	主要評価項目の各コンポーネントの初回発現までの期間等 安全性 薬物動態（PK）
参加者数	18624名（各群約9300名）、うち米国人1413名（各群約700名）
Pivotal試験全体での有効性の結果（Layer 1）	<ul style="list-style-type: none"> 主要評価項目の複合イベントの発現率は対照群と比べて本薬群で有意に低かった（ハザード比0.84）。 心血管死及び心筋梗塞それぞれの発現率は対照群と比べて本薬群で有意に低く、脳卒中は群間で有意な差はなかった。
その他	<ul style="list-style-type: none"> 本薬（チカグレロル）は対照薬クロピドグレルの類薬である。 早期臨床試験や公表文献から、特定の集団で本薬の有効性が異なると予想される臨床的・薬理学的特徴は知られていない。 推奨されるアスピリン維持用量範囲は、試験計画時の米国の診療指針（75～325 mg/日）と米国外の診療指針（75～100 mg/日）で異なる。 実医療では、急性冠症候群の治療や臨床経過に地域間の違いはないと考えられる。

5.1.1 CTD 記載例

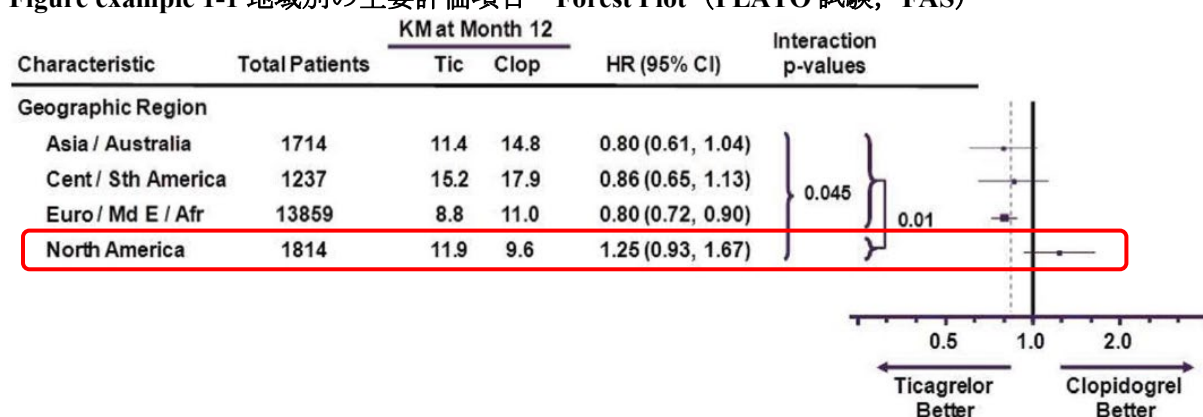
5.1.1.1 M2.5.4 又は M2.7.3.3.3 記載例

M2.X.X.X 特別な患者集団での有効性

M2.X.X.X.1 主要評価項目のサブグループ解析

主要評価項目のサブグループ解析を 31 種類の背景因子（内因性、外因性、地理的地域）について事前に計画し、実施した。その結果、地理的地域以外の背景因子では、本薬の有効性がサブグループ間で明らかに異なる傾向は認められなかった（Appendix-Figure X-X_未掲載）。地理的地域別では、North America（米国、カナダ）を除く 3 地域ではいずれもハザード比の点推定値が 1 を下回ったのに対して North America では 1 を上回り、North America と North America 外で有効性に違いが認められた（Figure example 1-1）。

Figure example 1-1 地域別の主要評価項目—Forest Plot（PLATO 試験, FAS）



Source: Carroll and Fleming 2013

HR = hazard ratio; KM = Kaplan-Meier.

M2.X.X.X.2 事後的追加解析による参加国の主要評価項目への影響の検討

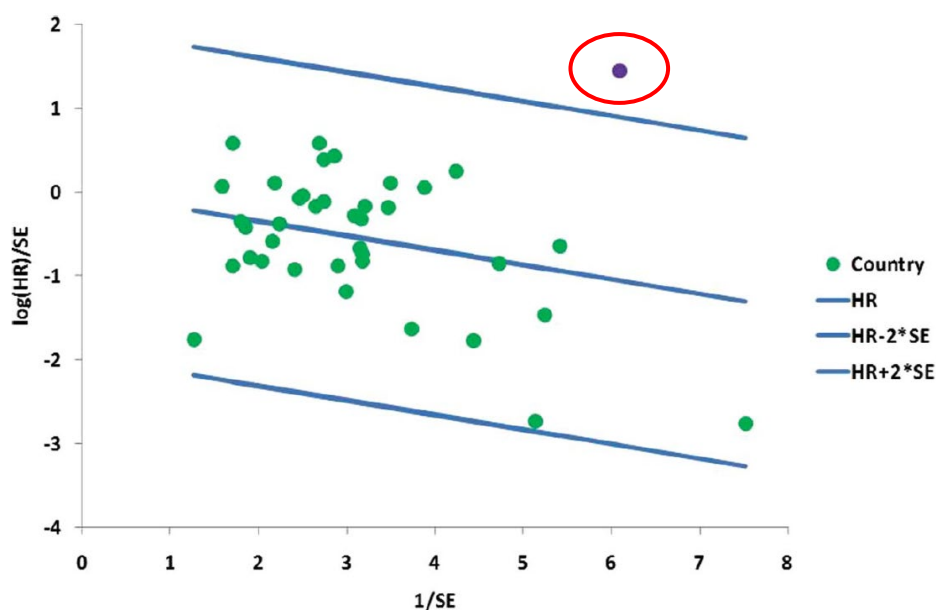
地理的地域間で違いが認められた要因を探索するために、以下の事後的追加解析を主要評価項目について実施した。なお、試験運用に関する事項（試験手順、治験薬輸送、施設管理状況、脱落率、治験実施計画書からの逸脱率等）を精査した結果、これらは地理的地域間の有効性の違いの原因ではないと考えられた。

(1) 参加国別の検討

各国の主要評価項目の結果をプロットして分布を視覚的に評価した。その結果、Galbraith Plot 及び Q-Q plot のいずれでも米国（ハザード比 1.27）のみが分布から乖離し、米国の結果の異質性が認められた（Figure example 1-2, Figure example 1-3）。Funnel plot では米国のみが疑似的 95%信頼区間から乖離したことから（Figure example 1-4）、米国の結果には何らかの系統的な要因が存在することが示唆された。本試験で米国は組入れ参加者数及びイベント数が 2 番目に多く（全体 1878 イベントのうち米国 151 イベント）、投与群と米国の交互作用は有意であった（ $p=0.0095$ ）。また、主要評価項目のハザード比が全体集団で 0.84 のときに米国で 1.27 以上となる確率は 0.006

未満であった (Appendix-Table X-X_未掲載)。しかしながら、米国のみで検証できるだけの検出力はないため、米国での結果が偶然の結果である可能性も否定できない。

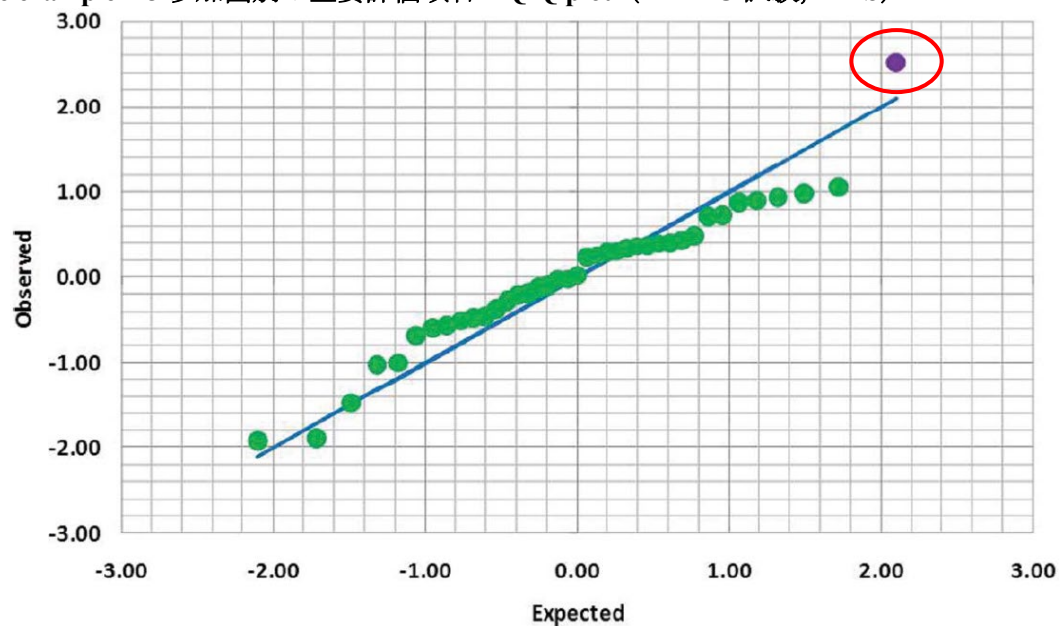
Figure example 1-2 参加国別の主要評価項目—Galbraith Plot (PLATO 試験, FAS)



Source: Carroll and Fleming 2013

SE = standard error. 赤丸が米国。

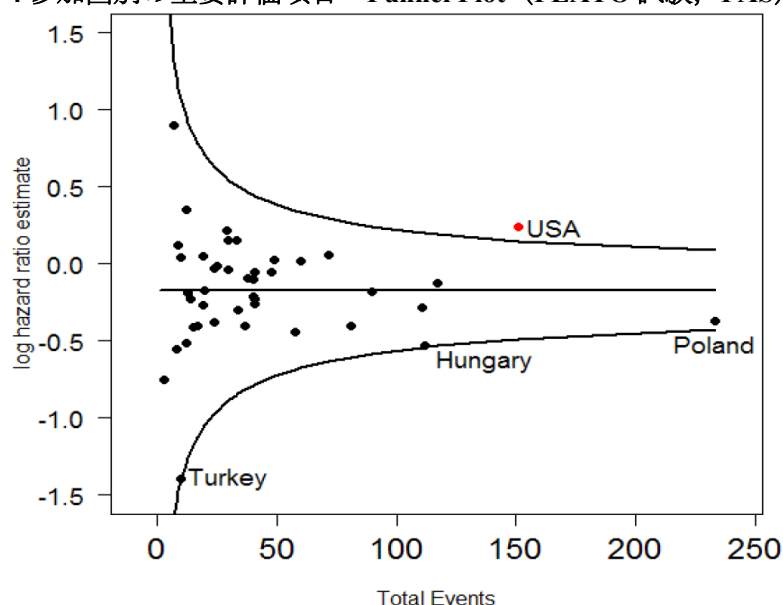
Figure example 1-3 参加国別の主要評価項目—Q-Q plot (PLATO 試験, FAS)



Source: Carroll and Fleming 2013

横軸：正規分布，縦軸：観測値。赤丸が米国。

Figure example 1-4 参加国別の主要評価項目—Funnel Plot (PLATO 試験, FAS)



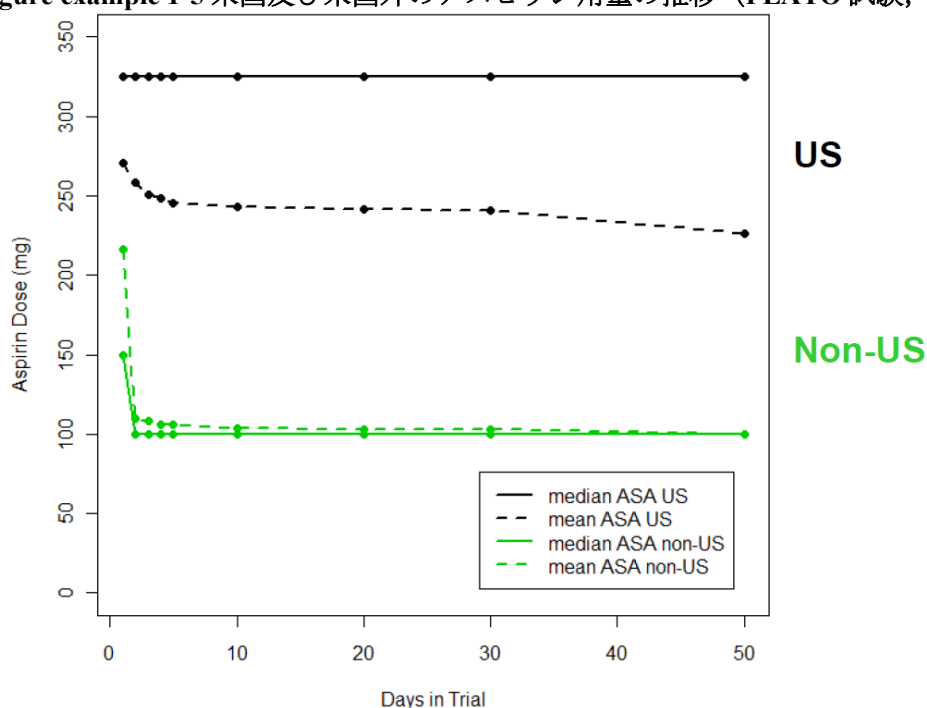
Source: チカグレロル 米国審査関連情報 (2010)

(2) 米国と米国外で本薬の有効性の結果が異なった要因の探索

米国と米国外で本薬の有効性の結果が異なった要因を探索するため、ベースライン及びベースライン後の特性に地域間で不均衡がないか検討した。投与群、地域（米国、米国外）、及び投与群と地域の交互作用を共変量とする Cox 回帰分析を実施した後、投与群と各特性の交互作用をモデルに加えて解析した。これらの結果、投与群との交互作用が明らかで、かつ米国と米国外で不均衡である特性は、基礎治療として併用するアスピリンの用量であった（投与群とアスピリン用量の交互作用： $p=0.003$ ）。アスピリン用量の平均値及び中央値は、試験期間を通じて米国外に比べて米国で高く（Figure example 1-5）、100 mg 超であった参加者の割合は米国で 57%、米国外で 8%、300 mg 以上は米国で 54%、米国外で 2%であった。米国内外でのアスピリン用量別の参加者の割合は、群間で同程度であった（Appendix-Table X-X_未掲載）。

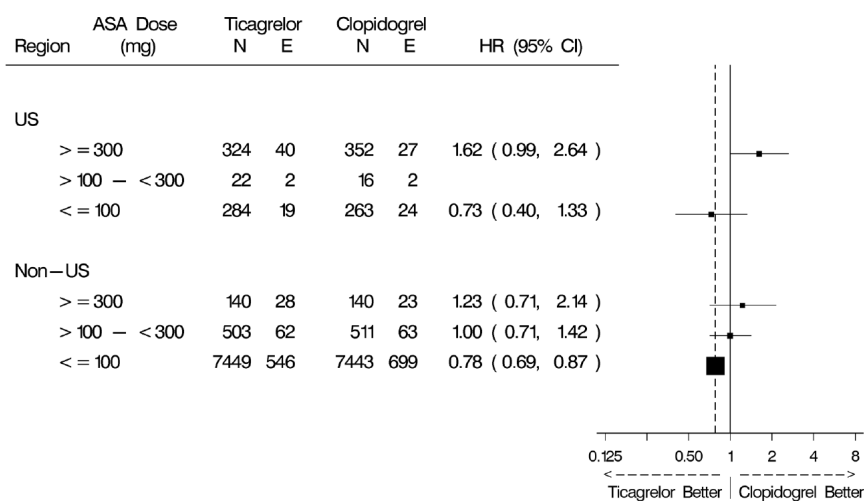
また、米国と米国外それぞれについてアスピリン用量別に主要評価項目のサブグループ解析を実施した結果、米国内外で類似していた。米国では中用量の参加者が少なくハザード比は推定不能であったが、低用量と比べて高用量でハザード比が高く、米国外では用量が高いほどハザード比が高かった。また、米国内外ともに、100 mg 以下の集団のみでハザード比の点推定値が 1 を下回った（Figure example 1-6）。したがって、基礎治療のアスピリンの用量は本薬の効果修飾因子であり、米国のみで本薬の有効性が示されなかった要因は、米国で高用量のアスピリンを併用した参加者の割合が高かったためであると考えられた。ただし、これらは事後的追加解析であり多重性を調整していないことや、モデルの感度が高い（アスピリンが高用量であった米国外の参加者数やアスピリンデータの欠測の取扱いに影響される）という限界がある。

Figure example 1-5 米国及び米国外のアスピリン用量の推移 (PLATO 試験, FAS)



Source: チカグレロル 米国審査関連情報 (2010)

Figure example 1-6 米国及び米国外のアスピリン用量別の主要評価項目—Forest Plot (PLATO 試験, FAS)



Source: チカグレロル 米国審査関連情報 (2010)

E = events; N = number of patients.

アスピリン高用量の併用により有効性が低下する可能性は、作用機序からも理論的に推測される。アスピリンは低用量ではCOX-1を阻害して血小板凝集を抑制するが、高用量では内皮細胞のCOX-2も阻害し、結果的に血管抵抗性が高まることが知られている。また、本薬はP2Y₁₂受容体をほぼ完全に遮断し、COX-1阻害を介した血小板凝集も抑制すると考えられることから、アスピリンの抗血小板作用は限定的となる (Warner et al. 2011)。

5.1.1.2 M2.5.6.2 記載例

M2.5.6.2 ベネフィット

<全体の試験結果に基づくベネフィット主張の後に記載>

早期臨床試験成績や PLATO 試験の PPK 解析から、急性冠症候群患者での本薬の PK に臨床的意味のある影響を与える内因性・外因性民族的要因は認められなかった。また、PLATO 試験の主要評価項目のサブグループ解析の結果では、地理的地域を除き、参加者背景因子によって本薬の有効性が明らかに異なる傾向は示唆されなかった。

PLATO 試験の地理的地域別の主要評価項目は、Asia/Australia, South America/Central America, 及び Europe/the Middle East/Africa ではハザード比の点推定値が 1 を下回ったのに対し、North America (米国及びカナダ) では 1 を上回った。そこで、複数の事後的追加解析により参加国の有効性への影響を検討した結果、米国の結果の異質性が認められ、また、米国の結果には何らかの系統的な要因が存在することが示唆された。この要因を検討するため、投与群との交互作用が明らかで、かつ米国と米国外で不均衡である特性を探索した結果、基礎治療であるアスピリンの用量が該当した。

米国と米国外でアスピリン用量 (1 日用量 100 mg 以下, 100 mg 超 300 mg 未満, 300 mg 以上) 別に主要評価項目のサブグループ解析を実施した結果、米国内外の結果は類似しており、アスピリン用量が高いほどハザード比が高い傾向が示唆された。また、米国内外ともに 100 mg 以下のみでハザード比の点推定値が 1 を下回り、100 mg 超では本薬による明らかなベネフィットが示されなかったが、100 mg 以下では米国内外で一貫したベネフィットが示唆された。したがって、基礎治療のアスピリンの用量は本薬の効果修飾因子であり、米国のみで本薬の有効性が示されなかった要因は、米国で高用量のアスピリンを併用した参加者の割合が高かったためであると考えられた。これらは事後的追加解析であり多重性を調整していないことや、モデルの感度が高いという限界があるものの、低用量 (75~100 mg/日) のアスピリン併用下では、各国で一貫して本薬のベネフィットは対照薬を上回ることが期待できると考える。

なお、陳旧性心筋梗塞患者を対象とした North America を含む国際共同第 III 相 PEGASUS 試験で、低用量 (75~100 mg/日) のアスピリン併用下での本薬の有効性に地域差は認められなかったことから、地域間での本薬の有効性の一貫性は支持される (Bonaca et al. 2015)。

米国での本薬の有効性は PLATO 試験では米国外の地域と比較して劣る結果であったが、上記のとおり併用するアスピリンを低用量 (75~100 mg/日) とすることにより、米国人患者での本薬の対照薬を上回るベネフィットが期待できると考える。

5.1.2 5つの視点のまとめ

生物学的合理性	高用量アスピリンの本薬有効性への影響は作用機序からも推測される。 本薬 PK に臨床的に影響を与える内因性民族的要因はない。
内的一貫性	アスピリン用量別の本薬有効性が米国内外で類似していた。
外的一貫性	類似疾患の試験でも低用量アスピリン併用時は本薬有効性が地域間で一貫して認められた。
統計学的な不確実性	偶然の結果である可能性は否定できないものの、米国結果の異質性が複数の評価方法で支持された。 アスピリン用量の交互作用が有意であったが、事後的追加解析・多重性未調整・高感度モデルであるという限界がある。
臨床的意義	併用するアスピリンが低用量の場合のみ、心血管イベント発現リスクが対照薬を下回る。

5.2 Example 2 : 仮想事例 ABC123

5.2.1 背景情報

対象疾患	乳癌
品目	ABC123
申請国	日本
申請パッケージ中の MRCT	1 試験 国際共同第 III 相試験 301 試験 (Pivotal 試験)
Pivotal 試験のデザイン	
デザイン概略	ランダム化, 二重盲検, DEF456 対照, 国際共同第 III 相試験 参加国: 6 カ国 (中国, 米国, ドイツ, 日本, フランス, イタリア)
投与群	ABC123 群 (以下, 本薬群), DEF456 群 (以下, 対照群) ランダム化比 1:1, 層別因子: 病変部位 (内臓転移あり, 内臓転移なし), 手術補助療法歴 (ホルモン療法実施歴あり, ホルモン療法実施歴なし)
投与期間	固定せず (進行又はその他の中止基準に該当するまで) *投与中止後は死亡又はその他の中止基準に該当するまで生存追跡調査を実施
主要評価項目	無増悪生存期間 (PFS)
副次評価項目	全生存期間 (OS), 奏効率 (ORR), 奏効持続期間 (DOR) 等 安全性 PK
参加者数	600 名 (各群 300 名), うち日本人 60 名 (各群 30 名)
Pivotal 試験全体での有効性の結果 (Layer 1)	<ul style="list-style-type: none"> OS は対照群と比べて本薬群で有意に延長した。 PFS, ORR, 及び DOR はいずれも対照群と比べて本薬群で良好であった。
その他	<ul style="list-style-type: none"> 本薬 (ABC123) は対照薬 DEF456 の類薬である。 乳癌の診断や治療といった医療環境に地域間の違いはないと考えられる。 早期臨床試験や公表文献から, 特定の集団で本薬の有効性が異なると予想される臨床的・薬理学的特徴は知られていない。

5.2.2 CTD 記載例

5.2.2.1 M2.5.4 又は M2.7.3.3.3 記載例

M2.X.X.X 特別な患者集団での有効性

M2.X.X.X.1 PFS 及び OS のサブグループ解析

人種別に加え、乳癌の予後因子とされる年齢、病変部位、ホルモン療法実施歴別に PFS 及び OS を検討した結果、サブグループ間で本薬の有効性が異なる傾向は認められなかった。参加者数が 36 名と限られた、アジア人又は白人以外の人種は 95%信頼区間が広いものの、いずれのサブグループでも PFS 及び OS ともに対照群に対するハザード比の点推定値は 1 を下回り、全体の結果と一貫していた (Figure example 2-1, Figure example 2-2)。

Figure example 2-1 参加者背景因子別の PFS (301 試験, ITT 集団)

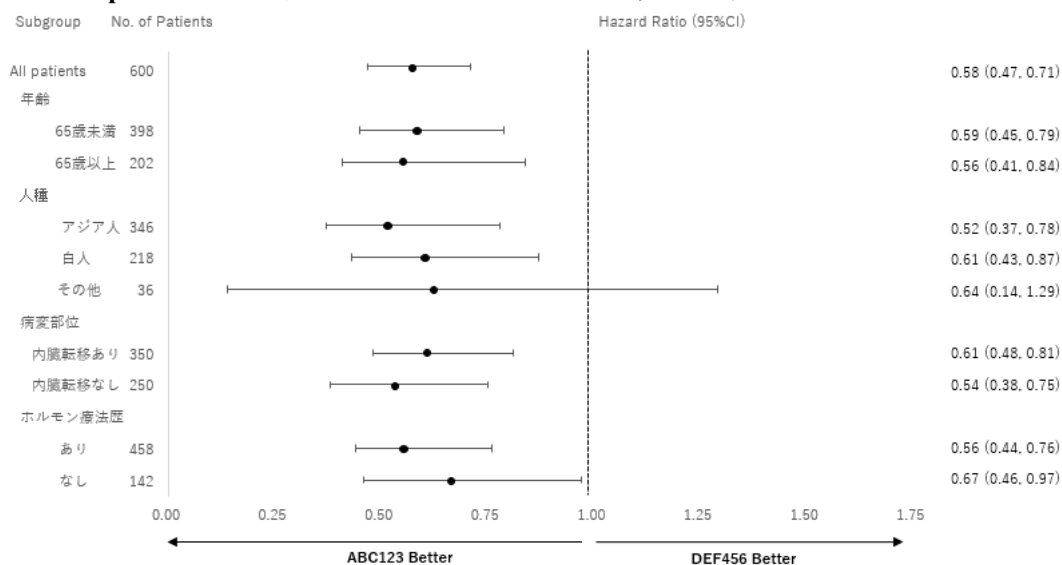
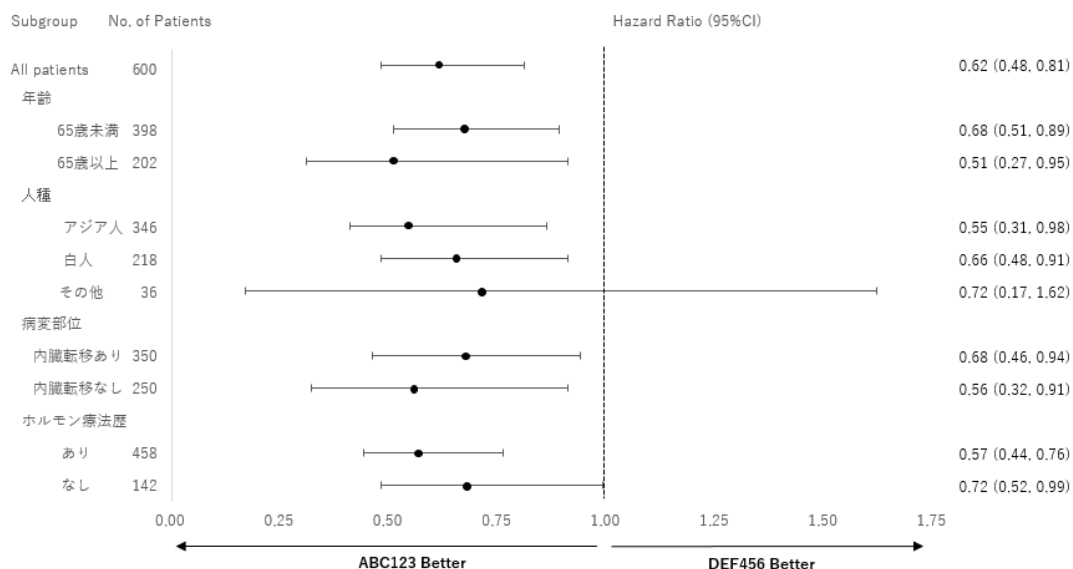


Figure example 2-2 参加者背景因子別の OS (301 試験, ITT 集団)

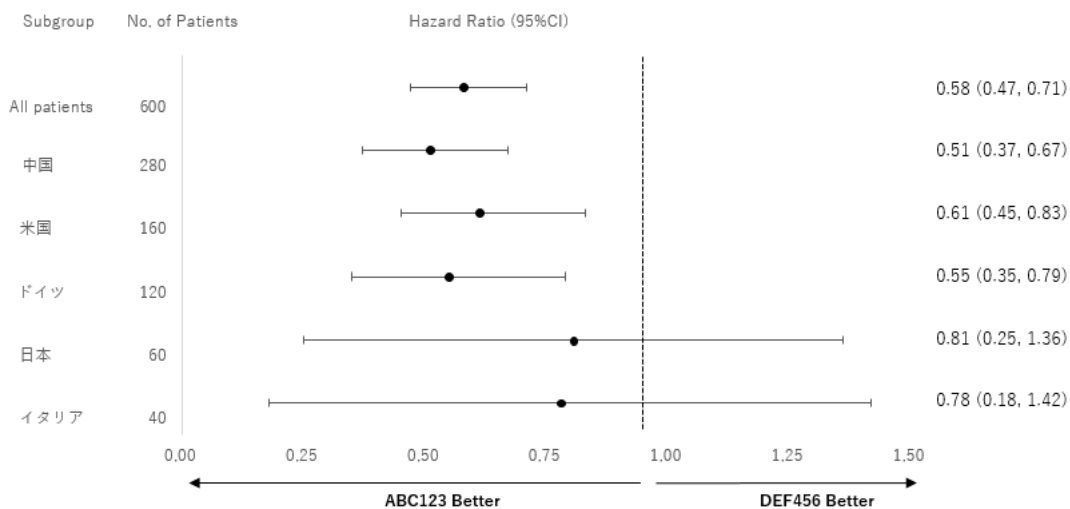


M2.X.X.2 参加国の PFS への影響の検討

参加国間での本薬の有効性の一貫性を評価するため、PFS の参加国別解析、Gail-Simon 検定、J-S shrinkage estimator を用いた解析を行った。

PFS の参加国別解析の結果、参加国間で本薬の有効性が異なる傾向は認められなかった。参加者数が限られた日本及びイタリアは 95%信頼区間が広いものの、すべての国で対照群に対するハザード比の点推定値は 1 を下回り、全体の結果と一貫していた (Figure example 2-3)。

Figure example 2-3 参加国別の PFS—Forest plot (301 試験, ITT 集団)



Gail-Simon 検定を用いて、対照群と比較した本薬群の有効性と参加国での交互作用を検定した結果では、有効性と参加国に質的な交互作用が存在しないことが示唆され ($p = 0.777$)、参加国間での有効性の一貫性は否定されなかった (Appendix-Table X-X_未掲載)。

J-S shrinkage estimator を用い、対照群と比較した本薬群全体の有効性と各国のデータを利用して算出した推定値の加重平均 (縮小推定値) は、いずれの国でも全体に類似した傾向であり、本試験の参加国間での有効性の一貫性は否定されなかった (Appendix-Figure X-X_未掲載)。

5.2.2.2 M2.5.6.2 記載例

M2.5.6.2 ベネフィット

<全体の試験結果に基づくベネフィット主張の後に記載>

国内外早期臨床試験及び国際共同第 III 相試験 301 試験の PPK 解析の結果、PK に臨床的意味のある影響を与える内因性・外因性民族的要因は特定されなかった。また、301 試験で人種に加え、乳癌の予後因子とされる年齢、病変部位 (内臓転移の有無)、及びホルモン療法歴の有無別で PFS 及び OS を検討した結果、一部のサブグループは参加者数が少なく、ばらつきが大きかったものの、サブグループ間で本薬の有効性が異なる傾向は認められなかった。したがって、本薬はこれらの背景因子によらず一貫したベネフィットが期待できる。

さらに、PFS での参加国の有効性への影響の検討では、一部の国は参加者数が少なく、ばらつきが大きかったものの、複数の解析のいずれでも参加国間の本薬の有効性の一貫性は否定されなかった。したがって、全体集団の結果は各参加国に適用可能であり、本薬は各国で一貫したベネフィットが期待できる。

日本人参加者での PFS は 95%信頼区間が広いものの、ハザード比の点推定値は 1 を下回った。また、上述のとおり、参加国間の本薬の有効性の一貫性は否定されなかったことから、日本人患者でも本薬のベネフィットは対照薬を上回ることが期待できる。

なお、類薬である DEF456 の乳癌に対する有効性に国内外差は知られていない。

5.2.3 5つの視点のまとめ

生物学的合理性	類薬である DEF456 でも乳癌に対する有効性に国内外差は知られていない。 早期臨床試験や公表文献から、特定の集団で本薬の有効性が異なると予想される臨床的・薬理学的特徴は知られていない。
内的一貫性	301 試験での参加者背景因子別の検討では、OS 及び PFS のいずれでもサブグループ間で本薬の有効性が異なる傾向は認められなかった。
外的一貫性	類薬である DEF456 でも乳癌に対する有効性に国内外差は知られていない。 早期臨床試験や公表文献から、特定の集団で本薬の有効性が異なると予想される臨床的・薬理学的特徴は知られていない。
統計学的な不確実性	301 試験での参加者背景因子別の PFS の検討では、一部のサブグループは参加者数が少なく、ばらつきが大きかった。 301 試験での参加国別の PFS の検討では、日本を含む一部の国で参加者数が少なく、ばらつきが大きかった。
臨床的意義	乳癌の診断や治療といった医療環境に地域間の違いはないと考えられ、また、301 試験での検討で参加国間の本薬の有効性の一貫性は否定されなかった。

5.3 Example 3 : 仮想事例 GHI789

5.3.1 背景情報

対象疾患	X線基準を満たさない体軸性脊椎関節炎
品目	GHI789
申請国	日本
申請パッケージ中の MRCT	1 試験 国際共同第 III 相試験 301 試験 (Pivotal 試験)
Pivotal 試験のデザイン	
デザイン概略	ランダム化, 二重盲検, プラセボ対照, 国際共同第 III 相試験 参加国: 15 カ国 (米国, ブルガリア, スペイン, デンマーク, ドイツ, ベルギー, イギリス, スイス, 中国, オランダ, フランス, 日本, スウェーデン, ブラジル, カナダ)
投与群	GHI789 群 (以下, 本薬群), プラセボ群 (以下, 対照群) ランダム化比 1 : 1, 層別因子: C-反応性タンパク (CRP) 値又は磁気共鳴映像法 (MRI) 画像所見での炎症徴候有無 [CRP 基準値超かつ MRI で仙腸関節炎あり (CRP+/MRI+), CRP 基準値超かつ MRI で仙腸関節炎なし (CRP+/MRI-), CRP 基準値内かつ MRI で仙腸関節炎あり (CRP-/MRI+)] * *CRP 基準値内かつ MRI で仙腸関節炎なし (CRP-/MRI-) の患者は本試験の対象外であった。
投与期間	52 週
主要評価項目	投与 16 週後の ASAS40 反応率* *複数のスケール評価結果に基づき定義する臨床症状・徴候の改善が認められた参加者の割合
副次評価項目	ASAS20 反応率等の臨床症状・徴候の改善評価, 身体機能の改善評価等 安全性 PK
参加者数	350 名 (各群 175 名), うち日本人 8 名 (本薬群 2 名, 対照群 6 名)
Pivotal 試験全体での有効性の結果 (Layer 1)	<ul style="list-style-type: none"> 投与 16 週後の ASAS40 反応率は対照群と比べて本薬群で有意に高かった。 副次評価項目でも対照群と比べて本薬群の有効性が良好であった。
その他	<ul style="list-style-type: none"> 公表文献で, 体軸性脊椎関節炎は, 男性患者は疾患進行リスクが高く, 薬剤反応性が高いと示唆されている。患者の男女比に人種差や地域差は知られていない。 公表文献で, 体軸性脊椎関節炎は, CRP 高値の患者や MRI 画像所見で炎症徴候が認められる患者は, 疾患進行リスクが高く, 薬剤反応性が高いと示唆されている。炎症徴候が認められる患者割合に人種差や地域差は知られていない。 体軸性脊椎関節炎の診断や治療といった医療環境に地域間の違いはないと考えられる。

5.3.2 CTD 記載例

5.3.2.1 M2.5.4 又は M2.7.3.3.3 記載例

M2.X.X.X 特別な患者集団での有効性

M2.X.X.X.1 参加者背景因子及び参加国別の検討

体軸性脊椎関節炎に対する本薬の有効性に影響を与える可能性がある因子として、性別及び CRP 値又は MRI 画像所見での炎症有無が考えられた。性別では、男性患者の方が疾患進行リスクは高いこと、薬剤反応性は高いことが示唆されている（参考文献）。また、CRP 高値の患者や MRI 画像所見で炎症徴候が認められる患者も、疾患進行リスクは高いこと、薬剤反応性は高いことが示唆されている（参考文献）。なお、これらの患者背景因子に人種差や地域差は知られていない。

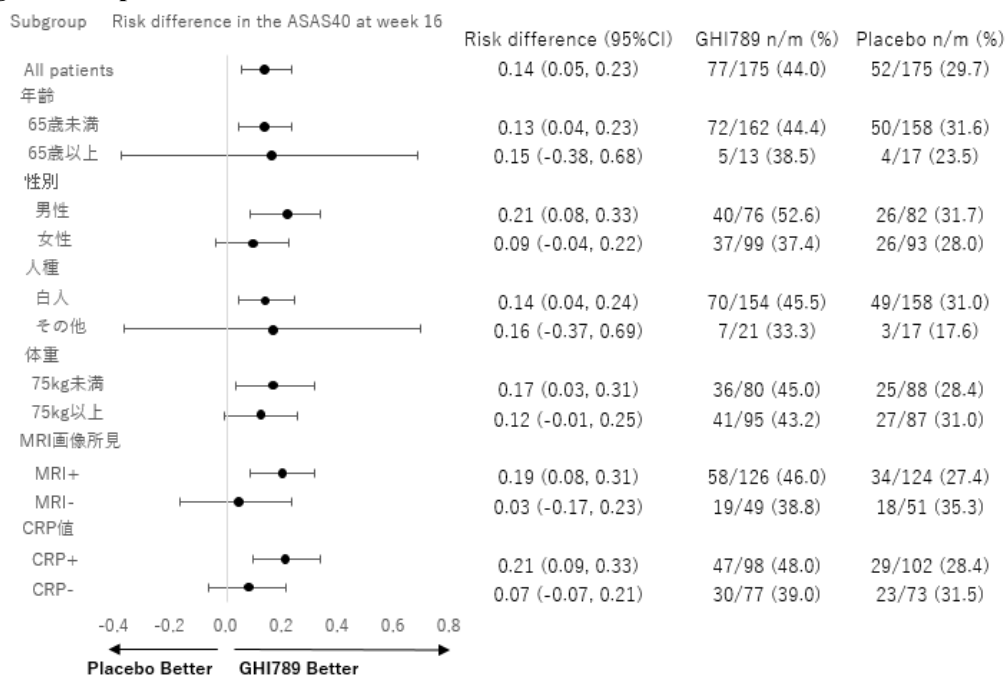
これらの因子を含む参加者背景因子別及び本試験の参加国別に投与 16 週後の ASAS40 反応率を検討した。

• 参加者背景因子別の検討

年齢、性別、人種、体重、CRP 値、MRI 画像所見別では、いずれのサブグループでも投与 16 週後の ASAS40 反応率は対照群と比べて本薬群で高かった。

本薬群の投与 16 週後の ASAS40 反応率の対照群との群間差は、女性と比べて男性、CRP- と比べて CRP+、MRI- と比べて MRI+ で大きく、性別及び CRP 値や MRI 画像所見での炎症徴候有無は本薬の効果修飾因子である可能性が示唆された。体重別では群間差に大きな差はなく、体重の本薬の有効性への影響は示唆されなかった。年齢は 65 歳未満 (91.4%)、人種は白人 (89.1%) がほとんどを占めたため、年齢別や人種別での厳密な評価は困難であるものの、本薬の有効性が明らかに異なる傾向は認められなかった (Figure example 3-1)。

Figure example 3-1 参加者背景因子別の投与 16 週後の ASAS 反応率 (301 試験, ITT 集団)



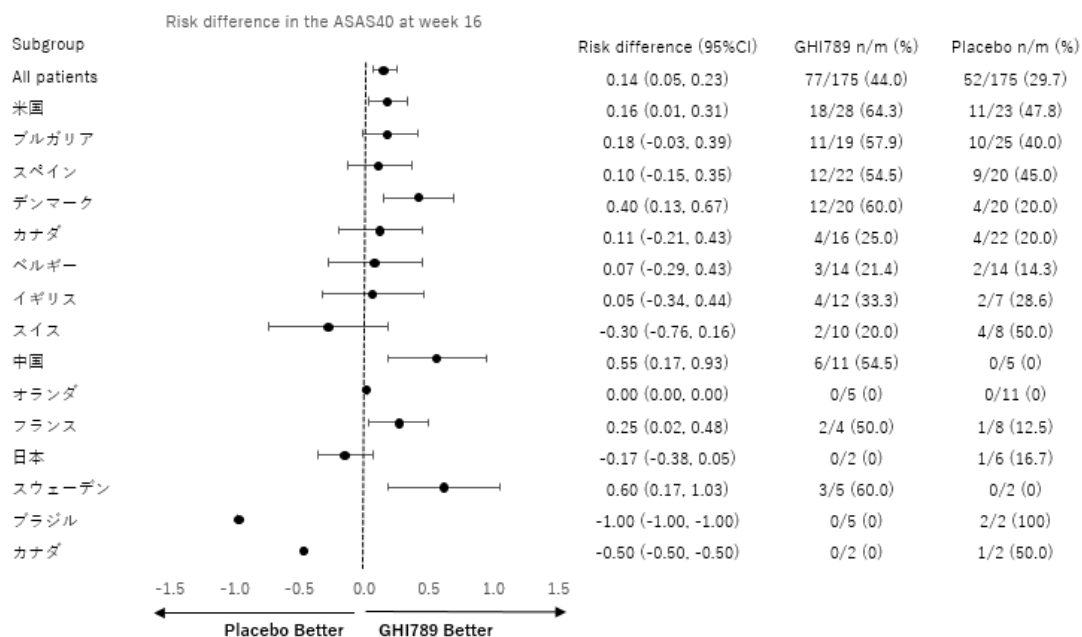
n: number of responders, m: number of patients

● 参加国別の検討

参加国 15 カ国中 9 カ国で参加者数が 20 名未満と限られたことから、参加国間の一貫性評価は困難であった。15 カ国中 5 カ国で本薬群の投与 16 週後の ASAS40 反応率の対照群との群間差が 0 以下であったが、これらはいずれも参加者数が 20 名未満の国であった (Figure example 3-2)。

投与 16 週後の ASAS40 反応率について、参加国間の交互作用の検定で有意差はなく (p = 0.3210) , 参加国間の有効性の一貫性は否定されなかった (Appendix-Table X-X_未掲載)。

Figure example 3-2 参加国別の投与 16 週後の ASAS 反応率 (301 試験, ITT 集団)



n: number of responders, m: number of patients

M2.X.X.X.2 事後的追加解析による効果修飾因子に基づくグループ別検討

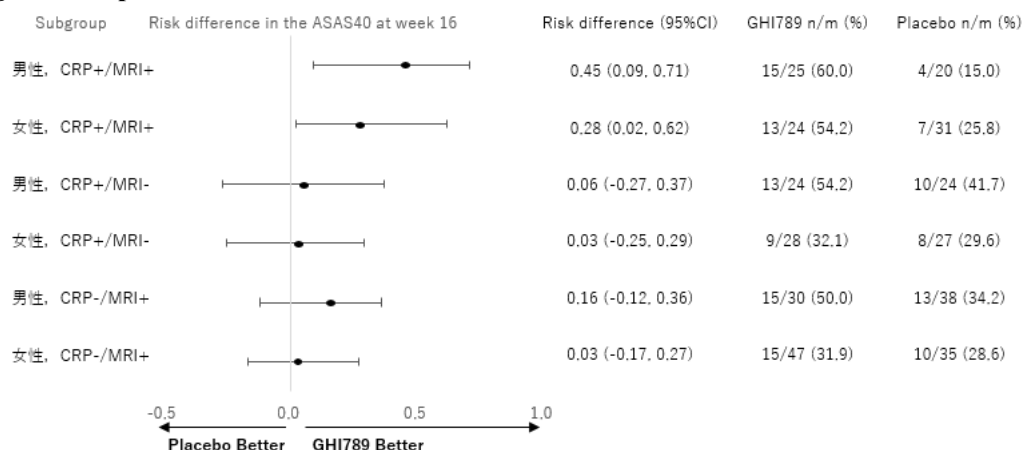
前項記載の検討で、性別及び CRP 値や MRI 画像所見での炎症徴候有無が効果修飾因子の候補に挙げられた。また、複数の参加国で参加者数が限られ、参加国間での有効性の一貫性評価は困難であった。そこで、有効性の一貫性をさらに評価するため、事後的追加解析により、効果修飾因子候補に基づいて類似した参加者又は参加国をグループ化し、投与 16 週後の ASAS40 反応率を検討した。

● 参加者グループ別の検討

性別（男性，女性）及び CRP 値又は MRI 画像所見での炎症徴候有無（CRP+/MRI+，CRP+/MRI-，CRP-/MRI+）の組み合わせで、参加者を 6 つにグループ化したとき、全グループで投与 16 週後の ASAS 反応率は対照群と比べて本薬群で高かった。

各グループの群間差からは、性別及び CRP 値又は MRI 画像所見での炎症徴候有無は本薬の効果修飾因子であり、特に炎症徴候が CRP 値又は MRI 画像所見の両方で認められるか、片方のみで認められるかによる影響が大きいことが示唆された。男女ともに CRP+/MRI-又は CRP-/MRI+に比べて CRP+/MRI+で群間差は大きかった。また、CRP+/MRI+，CRP+/MRI-，又は CRP-/MRI+での男女差は小さかったが、いずれでも男性の方が群間差は大きかった。

Figure example 3-3 参加者グループ別の投与 16 週後の ASAS 反応率 (301 試験, ITT 集団)



n: number of responders, m: number of patients

● **参加国グループ別の検討**

各参加国での CRP+/MRI+及び男性の参加者割合, 並びに参加者数バランスを考慮して参加国を4つにグループ化した。参加国グループ別の CRP+/MRI+及び男性の参加者割合を Table example 3-1 に, 投与 16 週後の ASAS 反応率を Figure example 3-4 に示す。

グループ 1 及び 2 は CRP+/MRI+の参加者割合が全体より多く, グループ 1 及び 3 は男性の参加者割合が全体より多かった。グループ間では, CRP+/MRI+及び男性の参加者割合ともに, グループ 1 で最も多く, グループ 4 で最も少なかった。いずれのグループでも CRP+/MRI+及び男性の参加者割合は群間で同程度であった (Appendix-Table X-X_未掲載)。

全参加国グループで, 投与 16 週後の ASAS 反応率は対照群と比べて本薬群で高かった。

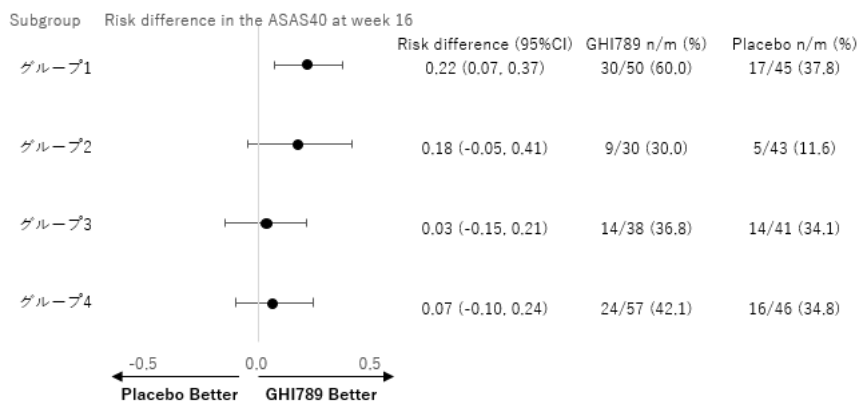
各参加国グループの群間差は同程度であったことから, 全体集団の結果は各参加国に適用可能と考えられた。しかしながら, 群間差が最大及び最小のグループは, CRP+/MRI+及び男性の参加者割合からはグループ 1 及び 4 が想定されるのに対し, 実際にはグループ 1 及び 3 であったことから, 性別, CRP 値, 及び MRI 画像所見以外の効果修飾因子の存在を否定できない。

Table example 3-1 参加国グループ別の CRP+/MRI+及び男性の参加者割合 (301 試験, ITT 集団)

参加国グループ	m	CRP+/MRI+ (%)	男性 (%)
All patients	375	26.7	42.1
グループ 1 カナダ, 米国, デンマーク	95	33.7	51.6
グループ 2 スウェーデン, オランダ, フランス, ドイツ	73	30.1	42.5
グループ 3 ブラジル, ブルガリア, ベルギー	79	26.2	46.6
グループ 4 日本, スイス, イギリス, スペイン, 中国	103	24.1	38.0

m: number of patients

Figure example 3-4 参加国グループ別の投与 16 週後の ASAS 反応率 (301 試験, ITT 集団)



n: number of responders, m: number of patients

5.3.2.2 M2.5.6.2 記載例

M2.5.6.2 ベネフィット

<全体の試験結果に基づくベネフィット主張の後に記載>

国内外早期臨床試験及び国際共同第 III 相試験 301 試験の PPK 解析の結果, PK に臨床的意味のある影響を与える内因性・外因性民族的要因は特定されなかった。

301 試験での参加者の年齢, 性別, 人種, 体重, CRP 値, 及び MRI 画像所見での炎症徴候有無別の投与 16 週後の ASAS40 反応率の検討では, 65 歳以上の高齢者や白人以外の人種が限られたものの, いずれのサブグループでも対照群と比べて本薬群で高く, 本薬はこれらの患者背景因子によらずベネフィットが期待できることが示唆された。

公表文献の報告, 並びに 301 試験での参加者背景因子別及び参加者グループ別の投与 16 週後の ASAS40 反応率の検討から, CRP 値又は MRI 画像所見での炎症徴候有無及び性別を本薬の効果修飾因子と特定した。301 試験で, CRP 値又は MRI 画像所見での炎症徴候有無 (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) 及び性別 (男性, 女性) の組み合わせで参加者を 6 つにグループ化し, 投与 16 週後の ASAS40 反応率を検討した。この結果, 各グループの群間差から, 炎症徴候が CRP 値又は MRI 画像所見の片方のみで認められる患者と比べて両方で認められる患者で, 女性と比べて男性で, 本薬のベネフィットが大きい可能性が示唆された。しかしながら, 全グループで投与 16 週後の ASAS40 反応率は対照群と比べて本薬群で高かったことから, 本薬のベネフィットはこれらの因子の組み合わせによらず期待できる。

301 試験は参加者数が少ない国も多く, 参加国別の有効性の評価は困難であったが, 投与 16 週後の ASAS40 反応率での交互作用の検定結果からは, 参加国間の有効性の一貫性は否定されなかった。また, 男性及び CRP+/MRI+ の参加者割合に基づいて参加国を 4 つにグループ化したとき, 全グループで投与 16 週後の ASAS40 反応率は対照群と比べて本薬群で高かった。CRP 値又は MRI 画像所見での炎症徴候有無及び性別以外の効果修飾因子の存在は否定できないが, 各グループの群間差は同程度であったことから, 全体集団の結果は各参加国に適用可能であると考えられる。したがって, 本薬は各国で一貫したベネフィットが期待できる。

301 試験の日本人参加者数は 8 名（本薬群 2 名，対照群 6 名）と少なく，対照群と比べた本薬群の投与 16 週後の ASAS40 反応率の有効性は示されなかった。しかしながら，本薬の効果修飾因子と特定された CRP 値又は MRI 画像所見での炎症徴候有無及び性別について，患者分布に人種差や地域差は知られておらず，上述の検討で全体集団の結果は各参加国に適用可能であると考えられた。したがって，日本人患者でも全体集団で認められた本薬のベネフィットが期待できる。

5.3.3 5つの視点のまとめ

生物学的合理性	本薬の有効性への影響が示唆された CRP 値又は MRI 画像所見での炎症徴候有無及び性別は，体軸性脊椎関節炎の予後因子であり，薬剤反応への影響が示唆されることが公表文献で報告されている。
内的一貫性	特定された効果修飾因子（CRP 値又は MRI 画像所見での炎症徴候有無及び性別）に基づく，301 試験での患者グループ別の検討では，炎症徴候が CRP 値又は MRI 画像所見の片方のみで認められる患者と比べて両方で認められる患者で，女性と比べて男性で，本薬のベネフィットが大きい可能性が示唆された。一方，特定された効果修飾因子に基づき，301 試験での参加国グループ別の検討では，想定される群間差の大きさ順とは異なった。
外的一貫性	本薬の有効性への影響が示唆された CRP 値又は MRI 画像所見での炎症徴候有無及び性別は，薬剤反応への影響が示唆されることが公表文献で報告されている。
統計学的な不確実性	一部の国では参加者数が少なく，推定値のばらつきが大きかった。
臨床的意義	体軸性脊椎関節炎の診断や治療といった医療環境に国内外で違いはないと考えられ，本薬の効果修飾因子と特定された CRP 値又は MRI 画像所見での炎症徴候有無及び性別について，患者分布に人種差や地域差は知られていない。また，特定された効果修飾因子に基づき，301 試験で参加国を 4 つにグループ化したとき，投与 16 週後の ASAS40 反応率の群間差は，グループ間で同程度であった。

6 まとめ

本書では、ICH E17 の理念及び小宮山らが提案した 3-layer approach (Komiya et al. 2013) に基づき、多角的・体系的に MRCT の結果を評価したうえでの、CTD 作成時の留意事項を示した。

世界中の患者が必要な医薬品に早期にアクセス可能となることを目指し、先進国以外の国を含む MRCT の利用や MRCT を含む国際共同開発が増加している。2018 年に ICH E17 「MRCT の計画及びデザイン的一般原則に関するガイドライン」が実装され、ICH E17 に準拠して MRCT を実施することで、多様な人種、民族的要因を踏まえて、医薬品のベネフィット・リスクを多角的に評価可能であるという利点がある。しかしながら、現状、申請国／地域の結果に注目し、そのうえで他の地域又は試験全体との類似性を評価している Local first の考えに基づく CTD が多い。申請国／地域で類似性の評価が実施可能な参加者数が確保できる場合には、このような評価のアプローチはある程度機能してきた。しかしながら、MRCT 全体の参加者数は試験目的に応じて決められるため、MRCT に参加する国／地域における参加者数は必然的に少なくなるのが予測される。さらに、MRCT に参加する国／地域が多い場合には、各国／地域の参加者数はさらに少数となる可能性があり、少数であることに起因する統計的不確実性が増す。効果修飾因子によるサブグループ解析を行おうとしても統計的不確実性はさらに増す。

本書で強調したことは、MRCT の結果を医薬品の治療効果の予測及び理解に役立てるために、「Local の結果 vs MRCT 全体の結果」に飛びつこうとせず、まずは、効果修飾因子やその候補となりうる因子を探索するという視点で、MRCT 全体のデータを検討すること。さらに、この検討を踏まえて、申請国／地域での患者集団を効果修飾因子で特徴付け、効果修飾因子を使って治療効果を推定することの重要性であった。これらを実践するためには、MRCT の結果を検討する流れを示した、3-layer approach が有用であることも述べた。3-layer approach を CTD に反映させることにより、申請国／地域の承認申請資料や、審査段階における企業と規制当局間の議論に好ましい変化が起きることが期待される。MRCT 全体のデータを用いた Layer 2 までの評価を充実させることは、どのような因子（効果修飾因子、あるいは潜在的な効果修飾因子の表現型としての国／地域なども含む）が治療効果に影響を及ぼしうるかの知見を整理することを意味する。Layer 2 までの検討結果は世界各国で共通利用することができるはずであるが、そのためには Global team が作成する CTD の中で、特定の国／地域への申請を意図しない Layer 2 を充実させることが必要であろう。そうすれば、各申請国／地域は、世界で共有された Layer 2 の知見を基にして、それぞれの申請国／地域の Layer 3 を考察することができる。このことは、各申請国／地域で個別に実施してきた多くの検討を減らすこと、各申請国／地域における CTD の共通部分を増やすこと、審査段階における規制当局との議論がより整理され、どの規制当局に対しても基本的に同じストーリーで説明することも可能になるであろう。また、ある申請国／地域において、他の国／地域で評価した内容を参照したり、それにより議論をブラッシュアップすることも可能になるであろう。「Local の結果 vs MRCT 全体の結果」というまとめは、他の国／地域にとっては直ちに自国の患者集団に適用できる情報にならないかもしれないが、効果修飾因子を軸に議論を進めれば、効果修飾因子が“世界の共通言語”になり、MRCT の結果を各国／地域の患者集団に適用しやす

くなるであろう。医薬品のベネフィット・リスクを世界が共通の視点で行うことは、開発段階のみならず、市販後の育薬の段階になっても、世界で協力しながらエビデンスを積み上げていく、その中で我が国はどのような検討に貢献できたかも論じることができるようになるであろう。

本書では、ICH E17 の 5 つの視点、3-layer approach での解析の事例も多く紹介した。提示の通り、解析手法自体は目新しいものではなく、いずれも従来実施してきた方法が多く含まれている。ICH E17 の 5 つの視点、3-layer approach の framework を中核に据え、利用可能な情報は最大限に利用した上で、申請国／地域のベネフィット・リスクを議論可能な解析手法を計画し、実行することで、MRCT の計画も結果の解釈も、今までとは異なるものになることが期待される。

本書が ICH E17 の理念に基づく MRCT の結果の解釈及び CTD 作成の普及につながる一助になれば幸いである。

7 参考資料

- Baselga J, Cortés J, Kim SB, et al. (2012) Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. *N Engl J Med*; 366(2):109-19.
- Bonaca MP, Bhatt DL, Cohen M, et al. (2015) Long-term use of ticagrelor in patients with prior myocardial infarction. *N Engl J Med*; 372(19):1791-800.
- Brown DM, Emanuelli A, Bandello F, et al. (2022) KESTREL and KITE: 52-Week Results From Two Phase III Pivotal Trials of Brolocizumab for Diabetic Macular Edema. *Am J Ophthalmol*; 238:157-72.
- KJ Carroll and TR Fleming (2013) Statistical Evaluation and Analysis of Regional Interactions: The PLATO Trial Case Study, *Stat Biopharm Res*, 5:2, 91 91-101.
- Klepper MJ and Cobert B (2012) くすりの安全性を科学する 「くすりの安全性を科学する会」 訳サイエンティスト社
- Komiyama O, Hiro S, Isogawa N, et al. (2013) Evaluation of Data for Multi-Regional Trials: A Three-Layer Approach. *Applied clinical trials*, vol 22, Issue 11.
- Swain SM, Kim SB, Cortés J, et al. (2013) Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Oncol*; 14(6):461-71.
- Uyama Y, Shibata T, Nagai N, et al. (2005) Successful bridging strategy based on ICH E5 guideline for drugs approved in Japan. *CLINICAL PHARMACOLOGY & THERAPEUTICS*; 78(2):102-13
- Warner TD, Nylander S, Whatling C. (2011) Anti-platelet therapy: cyclo-oxygenase inhibition and the use of aspirin with particular regard to dual anti-platelet therapy. *Br J Clin Pharmacol*; 72(4):619-33.
- Yoshida K, Awaji N, Takenouchi K, et al. (2015) How Should a Globalized CTD Be Created? An Introduction to the Japanese 3-Layer Approach. *Ther Innov Regul Sci*; 49(1):175-80.
- ICH E5 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の指針 (1998) (<https://www.pmda.go.jp/files/000156571.pdf>, 最終アクセス確認 2023年7月21日)
- ICH E5 「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の指針」に関するQ&Aについて (その2) (2006) (<https://www.pmda.go.jp/files/000156024.pdf>, 最終アクセス確認 2023年7月21日)
- ICH E17 国際共同治験の計画及びデザインに関する一般原則に関するガイドライン (2018) (<https://www.pmda.go.jp/files/000224557.pdf>, 最終アクセス確認 2023年7月21日)
- ICH E17 国際共同治験の計画及びデザインに関する一般原則」に関するトレーニングマテリアル等 (<https://www.pmda.go.jp/int-activities/int-harmony/ich/0022.html>, 最終アクセス確認 2023年7月21日)
- エンホルツマブ ベドチン (遺伝子組換え) 審査報告書 (2021) がん化学療法後に増悪した根治切除不能な尿路上皮癌 (https://www.pmda.go.jp/drugs/2021/P20211011002/800126000_30300AMX00454000_A100_1.pdf, 最終アクセス確認 2023年7月21日)

- 国際共同治験に関する基本的考え方について（薬食審査発第 0928010 号）（2007）
（[http://www.hourei.mhlw.go.jp/cgi-bin/t_document.cgi?MODE=tsuchi&\(japal.org\)](http://www.hourei.mhlw.go.jp/cgi-bin/t_document.cgi?MODE=tsuchi&(japal.org))），最終アクセス確認 2023 年 7 月 21 日）
- 国際共同治験での民族的要因の検討方法 3-layer approach の実践（2018 年日本製薬工業協会）
（https://www.jpma.or.jp/information/evaluation/results/allotment/lofurc000000a0k7-att/3_layer_approach_final.pdf，最終アクセス確認 2023 年 7 月 21 日）
- これからの医薬品評価－国際共同開発での日本人データの意義って何？－（2022 年日本製薬工業協会）（https://www.jpma.or.jp/information/evaluation/results/allotment/gbkspa0000001336-att/DS_202207_e17etc.pdf，最終アクセス確認 2023 年 7 月 21 日）
- 第 2 回 ICH E17 ワークショップ（2022）（<https://www.pmda.go.jp/files/000247710.pdf>，<https://www.pmda.go.jp/files/000247711.pdf>，<https://www.pmda.go.jp/files/000247712.pdf>，<https://www.pmda.go.jp/files/000247713.pdf>，<https://www.pmda.go.jp/files/000247714.pdf>，最終アクセス確認 2023 年 7 月 21 日）
- ダビガトランエテキシラートメタンスルホン酸塩 申請資料概要（2011）非弁膜症性心房細動患者における虚血性脳卒中及び全身性塞栓症の発症抑制
（<https://www.pmda.go.jp/drugs/2011/P201100019/index.html>，最終アクセス確認 2023 年 7 月 21 日）
- チカグレロル米国審査関連情報（2010）（<https://wayback.archive-it.org/7993/20170405212359/https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/UCM221383.pdf>，<https://wayback.archive-it.org/7993/20170405212347/https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/UCM220197.pdf>，最終アクセス確認 2023 年 7 月 21 日）
- フィネレノン審査報告書（2022）2 型糖尿病を合併する慢性腎臓病
（https://www.pmda.go.jp/drugs/2022/P20220317001/630004000_30400AMX00176_A100_1.pdf，最終アクセス確認 2023 年 7 月 21 日）

2022 年度データサイエンス部会タスクフォース 2-2

小宮山靖（ファイザーR&D 合同会社）担当副部長
朝原優子（ノバルティスファーマ株式会社）リーダー
青谷真砂子（アッヴィ合同会社）
河西寧子（ノバルティスファーマ株式会社）
齊藤有希（ヤンセンファーマ株式会社）
佐久間知子（ブリストル・マイヤーズ スクイブ株式会社）
笹川裕次（Meiji Seika ファルマ株式会社）
世古規子（ノバルティスファーマ株式会社）
丸茂雅哉（キッセイ薬品工業株式会社）