

## 論文の内容の要旨

論文題目：日本及び米国の承認医薬品の薬効評価に影響を与える  
治験デザイン要素に関する探索的研究

氏名：福永 悟史

### 【背景・目的】

現在の薬効評価で一般的に実施されているランダム化比較試験の結果の最も本質的な要素である「被験薬群と対照薬群の薬効の差」は、第一に被験薬と対照薬の薬効及びその差に依存する。しかし、同じ被験薬・対照薬の組み合わせでも、体格や疾患の重症度等の被験者集団や製薬企業が設定する治験デザインまたは製薬企業の治験経験などの治験条件によって観察される薬効の差が変化する可能性がある。先行研究の多くは、効果の差を定義づける被験薬と対照薬の組み合わせ、及び治験毎の属性を制御した上で、異なる疾患領域の多数の薬剤や治験の組み合わせに共通して存在する関係を探査するというアプローチは採っておらず、また「実薬（対照）と被験薬」の組み合わせを「プラセボ（対照）と被験薬」と同じ枠組みで考慮した分析研究は少ない。日本の治験を対象にした実証研究も存在しない。臨床試験の計画・実施においてきわめて重要な役割を果たす製薬企業属性等との関係、特に国内臨床試験開始時までに蓄積された国内外の関連する先行治験の情報量（臨床開発経験）の影響に着目した分析は行われていない。

本研究では被験薬群と対照薬群の効果の差の大きさの弁別度である効果量に着目し、治験、被験薬と対照薬の組合せ、疾患領域の違いを統計学的に制御した上で、複数の疾患で共通して効果量と関係する治験デザイン及び臨床開発経験の要素を探査し、先行研究では未検討の重要な諸問題（上述）を論じる上での定量的なエビデンスを得ることを目的とした。そして、分析結果を踏まえて当局の施策や製薬企業の開発戦略のあり方が効果量に与える影響を考察した。

### 【方法】

国内治験の検討では、1970年から2011年に日本で承認された医薬品90品目の承認申請に用いられた第Ⅱ相またはⅢ相無作為化二重盲検平行群間比較試験145試験を研究対象とした。国内治

験及び海外治験の比較では、1997年から2011年に日本または米国で承認された医薬品を対象にした。効果量と治験条件の関連性は混合効果モデルによって解析した。効果量を目的変数とし、治験条件として、治験デザイン（例：群の数）、薬効評価の難しさ（例：評価指標の種類）、開発相、トレンド（承認年）、企業の同種同効薬の国内開発経験／被験薬と同一薬の海外臨床開発経験の有無、身体的要因（年齢、女性割合）を説明変数とし、治験条件と効果量との関係を解析した。疾患・被験薬・治験を階層型のランダム効果とし、さらに被験薬と対照薬の組み合わせをクロスランダム効果として混合効果モデルに加えることで、それらが被験薬及び対照薬の効果の差の弁別度を示す効果量に与える影響を制御した。

## 【結果】

### 予備的検討：身体的要因（薬物動態）と承認用量の国内外差

「患者の身体的要因」のばらつきは、本研究の分析ではモデルにおける制御因子として扱われ、必ずしも直接の分析の対象としていない。医薬品開発・承認審査の種々の判断に「患者の身体的要因」がどの程度反映されているのかを予備的に確認した。薬効評価に直接的に関係する「患者の身体的要因」の一つである薬物動態の人種差が、承認医薬品の有効性及び安全性の評価さらに用法・用量設定の決定的要因であるのかを、2001年から2008年までの国内承認品目を対象に、「患者の身体的要因」の重要な例である日本人と白人の薬物動態の違いと、日米における承認用量の違いの関心の検討により間接的に確認した。患者集団の薬物動態の日本人と白人の差と、医薬品開発・承認審査の重要な帰結（判断）である国内承認用法・用量の設定の間には、容易に観察可能な関係性は見出せなかった。国内承認用法・用量の設定を含む薬効評価の結果及びそれに基づく判断は、薬物動態のみで単純に決まるのではなく、それ以外の要因も重要な役割を果たすと考えられた。この予備的結果を踏まえ、「患者の身体的要因」を分析モデルにおいて制御因子として取り扱い、「患者の身体的要因以外の要因」である治験条件等が薬効評価で果たす役割を次に検討することとした。

### 研究1：効果量に関連する治験条件（国内治験：1970年から2011年）

第2相・第3相試験の両方を含むすべての治験の効果量を対象にした分析1においては、「優越性」、「サンプルサイズ」、「用量調節方法（可変用量）」、「主要評価指標」、「被験薬と同一作用機序の対照薬」、「主観的指標（全般改善

表1. 効果量と関連する治験条件（国内治験）

説明変数		Phase 2/Phase 3		Phase 3のみ		
		分析1 (N = 182)	分析1' (N = 182)	分析2 (N = 149)	分析2' (N = 149)	
		Estimate	Estimate	Estimate	Estimate	
治験 デザイン	対照薬の種類と 治験の統計学的 目的 (統計学的基準なし) <sup>†</sup>	優越性	0.192**	0.156*	0.208***	0.143*
		非劣性:実薬対照	-0.002	-0.027	0.009	-0.063
		用量設定	0.072	0.058	-	-
		群の数	0.063	0.047	0.045	0.033
		サンプルサイズ(人)	-0.001***	-0.001**	-0.001***	-0.001***
		投薬期間(週)	0.005	0.003	0.004	0.001
		用量調節方法(可変用量)	-0.096*	-0.098*	-0.119**	-0.138***
		主要評価指標	0.088**	0.078*	0.062	0.043
		被験薬と同一作用機序の対照薬	-0.134**	-0.150***	-0.169***	-0.198***
		評価指標の種類	主観的(評価スコア)	-0.150**	-0.132**	-0.113*
評価の 難しさ		(客観的評価指標 <sup>†</sup> )	-0.122**	-0.101*	-0.096*	-0.063
	1治験実施施設あたりの被験者数		0.008	0.007	0.012**	0.009
開発相	Phase 2		-0.064	-0.058	-	-
トレンド	薬事承認取得年		0.007**	0.004	0.005	-0.001
開発 経験	製薬企業の同一薬の海外開発経験		-	0.075*	-	0.082**
	製薬企業の同効薬の国内開発経験		-	0.079	-	0.178**
身体的 要因	平均年齢		0.005	0.006	0.012***	0.015***
	女性被験者の割合		-0.409*	-0.437**	-0.795***	-0.726***
定数			-1.979***	-1.988***	0.012	0.076

<sup>†</sup> 参照カテゴリー, \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

度)、「女性被験者の割合」は効果量と関係していることがわかった。第3相の治験のみを対象にした分析2の結果は分析1と同様の傾向であった。分析1の説明変数に企業の開発経験を示す変数を加えた分析1'においては「同種同効薬の国内開発経験」が、開発相を第3相に限定した分析2'では「同種同効薬の国内開発経験」及び「同一薬の開発経験」が効果量と有意に正の関連があった。

## 研究2：効果量に関連する治験条件（国内治験と海外治験の比較：1997年から2011年）

国内外の治験デザインはプラセボ対照試験の割合、群の数、サンプルサイズ、固定または可変用量、1施設あたりの被験者数の点で差異があることがわかった。混合効果モデルによる多変量回帰分析の結果(表2)、分析3(国内治験)及び分析4(海外治験)の両方で「優越性」は効果量と有意に正の関係があ

表2. 効果量と関連する治験条件(国内及び海外治験)

説明変数	日本		海外
	分析3 (N = 94)	分析3' (N = 94)	分析4 (N = 92)
統計学的目的	Estimate	Estimate	Estimate
優越性(プラセボ対照)	0.217**	0.965**	0.353***
用量反応試験	0.029	0.053	-0.047
群の数	0.104*	0.142**	-0.082**
治験デザイン			
サンプルサイズ(人)	-0.001**	-0.001*	0.000
投薬期間(週)	-0.005	-0.005	0.028***
用量調節方法(可変用量)	-0.072	0.025	-0.081
主要評価指標	0.111***	0.114***	-0.004
評価の難しさ			
評価指標の種類			
主観的(評価スコア)	-0.101	-0.112	0.116
(客観的評価指標†) 主観的(全般改善度)	-0.101	-0.115	0.149
1治験実施施設あたりの被験者数	-0.012	0.004	0.000
開発相			
Phase 2	-0.117	-0.102	0.099
トレンド			
薬事承認取得年	0.023**	0.020*	-0.030***
身体的要因			
平均年齢	0.031***	0.027***	-0.009
女性被験者の割合	0.014	-0.224	-0.001
交互作用			
優越性(プラセボ対照)*群の数	-	-0.235**	-
定数	-2.156***	-2.173***	0.553*

† 参照カテゴリー, \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

った。一方、「群の数」は国内治験では群の数が多いと効果量が大きく、全試験がプラセボ対照試験の海外治験では小さかったが、分析3(国内治験)に「優越性試験(プラセボ対照)×群の数」の交互作用項を加えたところ係数が負で有意になり、国内でもプラセボ対照試験は群の数が多いと効果量は小さいことが示唆された(分析3')。国内治験のみで「サンプルサイズ」及び「平均年齢」が大きいと効果量は小さく(分析3)、海外治験のみで「投薬期間(週)」が長いと効果量は大きかった(分析4)。国内治験では「薬事承認取得年」が近年ほど効果量が大きく(分析3)、海外では近年ほど小さくなった(分析4)。

### 【考察】

研究1の分析結果は、同種同効薬の国内開発経験及び同一薬の海外開発経験と効果量の正の関係は、過去の開発において治験条件が設定された経緯及び治験結果と計画時の想定との異同や、治験実施医療機関及び治験担当医師に関する情報及び経験から得たノウハウを、製薬企業が将来的に開発する際に参照できる形で残すことで新医薬品の有効性の検証に実質的かつ観察可能な影響を与えうることを示唆している。治験デザインと効果量に関する探索的分析の結果は、効果量に直接的に影響を与える要素を検討する足掛かりとして今後の治験の企画・運用に利用できる。また、分析結果に基づき、製薬企業が新たな治験において設定した治験条件を基にどの程度の効果量が得られるかを推定することにも役立つと考えられる。

研究2での分析結果から、国内及び海外治験ともに、海外先行研究で示唆されていたようにプ

ラセボ対照試験における群の数は、実薬投与確率の変化により被験者及び評価者の効果への期待を変えることで、プラセボ効果を介して効果量に影響したと考えられた。一方、国内治験でのみ認められた薬事承認取得年と効果量の正の関係は、海外での開発先行を前提に「キャッチアップ」型で実施されている日本の開発状況の反映であると考えられた。また、海外治験では基礎的な治療方法等の進歩によりプラセボ群でもかつてより大きな改善が認められた可能性、近年ほど既存治療でも効果が認められない治験薬の効きにくい被験者集団が組み入れられやすくなった可能性が考えられた。同様に、サンプルサイズと効果量の関係が国内治験でのみ強く見られたことは、新薬の開発が欧米先行・主導の下で実施されており、多くの国内治験が先行する海外治験成績を参考に計画されたという状況を反映するものと考えられた。

本研究の意義は、第一に、先行研究において調整されていなかった「治験薬の属性」（対象疾患、被験薬及び対照薬）及び「治験毎の属性」（各治験）と関係する固有の効果をとともにモデル的に制御した上で、複数の疾患領域の治験を同時に分析し、複数疾患で共通して効果量に影響する治験条件を探索することに成功したことである。第二の意義は、先行研究で積極的に着目されることのなかった、製薬企業の国内及び海外における開発経験、すなわち国内治験開始時まで蓄積された情報量に着目して効果量との関係を解析した点であり、この点において本研究の新規性は高いと考える。分析の結果、治験プロトコルの特徴として直接に表現されない要素、すなわち開発企業の経験・ノウハウに基づく開発方針が、効果量と何らかの形で関係していることを初めて見出した。第三の意義は、国内治験及び海外治験それぞれについて効果量と関係する要素を同じモデルを適用して解析し、比較を行ったことであり、企業のグローバル開発戦略（特に欧米での開発が日本での開発に先行すること）と当局の申請データ受け容れ・承認に際しての要求等の歴史的状況が効果量の大きさに反映している可能性が示された。国際共同開発・国際同時開発の推進に伴って臨床試験実施の帰結（潜在的な成功確率）が変化する可能性が高いことを本研究は示している。

本研究は、新薬承認に係る法規制や臨床開発ガイドラインに示される承認条件・治験実施条件をどのように設定すれば、薬効評価の実体に歪みが生じない形で、パブリックヘルスの観点から望ましい帰結が得られるかを考えるための材料を与えているものと思料する。治験のグローバル化および国際ハーモナイゼーションが、各国固有の環境・条件のもとで治験結果にどのような方向・大きさの影響を及ぼすか、そして、治験結果が承認審査に活用され、医療における新薬の使い方が規定されることにより各国の、そして世界各国のパブリックヘルスはどのように変化するかを推察しつつ、具体的な政策を立案していく必要がある。