

日本で承認された患者数の特に少ない希少疾病用医薬品の 臨床データパッケージ等に関する研究

研究分担者 荒戸照世（北海道大学大学院医学研究科・教授）
研究協力者 金子真之（北里大学大学院薬学研究科・特任助教）
研究協力者 前田浩次郎（北海道大学大学院医学研究科・大学院生）

研究要旨

患者数が特に少ない希少疾病用医薬品（以下、ウルトラオーファンドラッグ）の開発促進に資するよう、少数の被験者でも合理的に有効性・安全性を評価するための留意事項・参考とすべき事項等をまとめたガイダンス文書を作成するために、これまで我が国で承認されたウルトラオーファンドラッグの臨床データパッケージを審査報告書を用いて調査し、エビデンスの構築方法、利用データソース、試験デザイン等について分析した。その結果、（ ）エビデンスの構築にあたり必ずしも無作為化比較試験の実施が必須ではないこと、（ ）用量反応試験において同じ患者を用いた用量漸増試験を実施する等の工夫をすること、（ ）継続投与試験／長期投与試験の実施、（ ）国際共同試験の実施や海外臨床試験データの活用、（ ）日本人を対象とした臨床研究や使用実態調査結果を参考資料として活用、（ ）エンドポイント、対照群及び症例数設定における工夫、（ ）適切な製造販売後調査等によるエビデンスの補強、（ ）未承認問題検討会議における開発要請を行うこと等が、ウルトラオーファンドラッグの開発促進に繋がることとなった。

A．研究目的

希少疾病用医薬品の中でも患者数が特に少ない疾病（本研究では国内患者数が1,000人未満程度を想定）に対する医薬品（以下、ウルトラオーファンドラッグと略）については、治験の困難さや市場規模の小ささなどの障害が特に大きく、開発が進まない状況にある。ウルトラオーファンドラッグの開発にあたり、少数の被験者でも合理的に有効性・安全性を評価するために留意すべき事項等をまとめたガイダンス文書を作成することにより、その開発の推進を図りたいと考えている。そのために、これまで我が国で承認されたウルトラオーファンドラッグの臨床データパッケージ

を調査し、エビデンスの構築方法、データソース、試験デザイン等について分析するとともに、その結果に基づき、ガイダンスに記載すべき留意事項・参考とすべき事項をまとめることを目的とした。

B．研究方法

2001年4月から2014年3月までに承認された希少疾病用医薬品の審査報告書を医薬品医療機器総合機構（PMDA）のホームページより入手し、これを基に臨床試験内容を調査した。審査報告書による情報が不十分な場合には、PMDAのホームページで公開されている申請資料概要を参照した。また、「薬事・

食品衛生審議会における事前評価について」（平成23年4月23日付薬食審査発0428第6号）に基づき申請がなされた品目については、「医療上の必要性の高い未承認薬・適応外薬検討会議 公知申請への該当性に係る報告書」の内容も併せて検討した。

調査内容として、臨床データパッケージの資料構成（国内外の評価資料及び参考資料）、実施された臨床試験数、臨床試験の内容（評価項目、デザイン及び症例数）、無作為化比較試験、用量反応試験、継続投与試験/長期投与試験の実施状況を調べた。加えて、薬効分類、新有効成分か否か、海外承認の有無、製造販売後の対応内容も調べた。

資料の取り扱いには以下のとおりとした。

- 健康成人及び他の効能を対象とした臨床試験成績や調査結果等が評価資料として提出されている品目もあったが、これらは評価資料として取り扱わなかった。一方、腫瘍用薬において固形癌を対象として実施している第Ⅰ相試験等は評価資料に含めた。
- 日本を含む国際共同治験は国内臨床試験として取り扱った。
- 特に断りがないかぎり、無作為化比較試験、用量反応試験、継続投与試験/長期投与試験は評価資料を対象として集計した。
- 投与回数等の検討を行った試験も用量反応試験として取り扱った。個人内で用量調節が可能な試験は任意漸増試験と見なした。
- 用量反応試験において、任意漸増試験のみ、あるいは強制的漸増試験のみが実施されている品目を、それぞれ任意漸増試験あるいは強制的漸増試験として扱った。

症例数の設定根拠に関する情報（有意水準、検出力）は、添付資料概要より入手した。ま

た、「未承認薬使用問題検討会議」あるいは「医療上必要性の高い未承認薬・適応外薬検討会議」（以下、両者を合わせて未承認薬使用問題検討会議と略）における開発要請については、厚生労働省のHPに掲載されている検討品目リスト^{1) 2)}を用いて確認した。

なお、これらの作業は2つのグループにより別々に行い、結果を相互に確認した。

C. 研究結果

1. 日本における希少疾病用医薬品の承認状況

2001年4月～2014年3月までに、我が国で承認された希少疾病用医薬品は136品目であり、そのうち、ワクチン8品目、抗HIV薬等平成10年11月12日医薬審第1015号に基づく事前評価対象品目16品目を除く112品目のうち、ウルトラオーファンドラッグが36品目、ウルトラオーファンドラッグ以外のオーファンドラッグ（患者数が1,000以上の希少疾病用医薬品。以下、ウルトラ以外のオーファンドラッグと略）が76品目であった。ウルトラ以外のオーファンドラッグについては、76品目中、審査報告書が入手可能な74品目を本研究の検討対象とした。

ウルトラオーファンドラッグ36品目の内訳は、その他の代謝性医薬品が13品目（うち他に分類されない代謝性医薬品7品目、酵素製剤6品目）、腫瘍用薬が8品目、生物学的製剤が7品目（うち血液製剤類4品目、その他の生物学的製剤3品目）が6品目、その他が8品目であった（表1、図1）。ウルトラ以外のオーファンドラッグ74品目の内訳は、腫瘍用薬が25品目、その他の代謝性医薬品が10品目、循環器官用薬が9品目、中枢神経系用薬が6品目、感覚器官用薬が5品目、生物学的製剤が5品目、ホルモン剤が4品目、末梢神経系用薬が3品目、消化器官用薬が2品目、その他が5品目であった（図1）。

ウルトラオーファンドラッグ 36 品目中、新有効成分含有医薬品として開発された品目が 25 品目、既に海外で承認を取得している品目が 30 品目であった(表 1)。また、ウルトラ以外のオーファンドラッグ 74 品目中、新有効成分含有医薬品として開発された品目が 43 品目、既に海外で承認を取得している品目が 60 品目であった。

2. 臨床試験の概略

(1) 臨床データパッケージの構成

ウルトラオーファンドラッグ 36 品目のうち、臨床データパッケージに国内臨床試験成績が評価資料として添付されている品目が 28 品目(77.8%)、海外臨床試験成績が評価資料として添付されている品目が 24 品目(66.7%)であった(図 2A)。国内臨床試験成績が添付されていない 8 品目中 6 品目で海外臨床試験成績が添付されており、その内訳は、酵素製剤が 3 品目、腫瘍用薬 1 品目、他に分類されない代謝性医薬品が 1 品目、その他の呼吸器官用薬が 1 品目であった。臨床試験成績が添付されていない品目は、医学薬学上公知であると判断され承認されたイマチニブメシル酸塩(販売名:グリベック錠、効能・効果:FIP1L1-PDGFR 陽性の好酸球増多症、慢性好酸球性白血病)とサリドマイド(販売名:サレドカプセル、効能・効果:らい性結節性紅斑)であった(表 2)。また、36 品目中 21 品目(58.3%)で国内で実施された臨床研究や各種の調査結果などが参考資料として添付されていた(図 2A)。

一方、ウルトラ以外のオーファンドラッグでは、74 品目のうち、臨床データパッケージに国内臨床試験成績が評価資料として添付されている品目が 71 品目(95.9%)、海外臨床試験成績が評価資料として添付されている品目が 24 品目(32.4%)であった(図 2D)。国内臨床試験成績が添付されていない 3 品目すべてにおいて海外臨床試験成績が添付され

ており、評価資料として臨床試験成績が添付されていない品目はなかった。また、74 品目中 20 品目(27.0%)で国内で実施された臨床研究や各種の調査結果などが参考資料として添付されていた(図 2D)。海外臨床試験成績及び国内参考資料が添付されている品目の割合はウルトラオーファンドラッグよりも少なかった。

ウルトラオーファンドラッグを日本における患者数が 1,000 人未満と想定し検討を行ったが、患者数が更に少ない場合(500 人未満、100 人未満)には(図 2B、図 2C)、国内臨床試験成績が添付されている割合が減少し(500 人未満:20/27 品目(74.1%)、100 人未満:7/12 品目(58.3%))、国内参考資料が添付されている割合が増加する傾向があり(500 人未満:15/27 品目(55.6%)、100 人未満:9/12 品目(75.0%))、100 人未満では顕著であった。

ウルトラオーファンドラッグでは、評価資料として海外のみで実施された国際共同治験が添付されている品目や参考資料として他効能における日本を含む国際共同治験成績が添付されている品目があったが、日本を含む国際共同治験成績が添付されている品目はなかった。一方、ウルトラ以外のオーファンドラッグでは、74 品目中 7 品目(9.5%)で日本を含む国際共同治験が実施され、7 品目中 6 品目が腫瘍用薬であった。

(2) 添付された臨床試験(評価資料)数

ウルトラオーファンドラッグの評価資料として提出された国内臨床試験の数は、1 試験が最も多く(36 品目中 16 品目)、次いで 2 試験が 9 品目であり、0 試験(実施されていない)が 8 品目であった(図 3)。一方、評価資料として提出された海外臨床試験の数は、0 試験が最も多かったが(12 品目)、中には 5 試験以上実施されている品目もあった。

ウルトラ以外のオーファンドラッグの評価資料として提出された国内臨床試験の数も、1試験が最も多く(74品目中37品目)、2試験が21品目であり、0試験(実施されていない)が3品目であった(図3)。一方、評価資料として提出された海外臨床試験の数は、0試験が最も多かったが(33品目)、中には5試験以上実施されている品目もあった。

このようにウルトラオーファンドラッグとウルトラ以外のオーファンドラッグの添付された臨床試験数の傾向に違いはなかった。

(3) 治験にエントリーされた症例数

ウルトラオーファンの国内臨床試験(評価資料)における全有効性評価症例数は0~60例未満であった。一方、海外臨床試験(評価資料)における全投与症例数は0~10,000例超と幅が広がった。ドルナーゼ アルファ(遺伝子組換え)(販売名:プルモザイム吸入液)では、国内資料としては使用実績調査が参考資料として添付されたのみであったが、海外では希少疾病用医薬品に指定されているにもかかわらず、臨床試験の全症例数は1,500例におよんでいた。また、塩酸メフロキン(販売名:メファキン錠)のように海外で大規模な試験が実施され、全投与症例数が10,000例を超える品目もあった。

2. 実施された臨床試験内容

(1) 無作為化比較試験

用量反応試験を含む無作為化比較試験の実施状況を調査したところ、ウルトラオーファンドラッグ36品目中、国内で無作為化比較試験が実施された品目は、国内のみで開発された乾燥スルホ化人免疫グロブリン(販売名:献血ベニロン- 静脈注用)のチャグ・ストラウス症候群及びアレルギー性肉芽腫性の効能追加の1品目のみであった。一方、海外で実施された無作為化比較試験成績が添付されている品目は18品目あり、無作為化比

較試験が実施されていない品目が17品目あった(図4)。無作為化比較試験が2試験以上添付されている品目は8/36品目(22.2%)であり、1試験のみの添付が11/36品目(30.6%)であった(図5)。

ウルトラ以外のオーファンドラッグでは74品目中、国内外で無作為化比較試験が実施された品目が10品目、国内無作為化比較試験のみが15品目、海外無作為化比較試験のみが20品目であり、無作為化比較試験が添付されていない品目が29品目あった(図4)。無作為化比較試験が2試験以上添付されている品目は32/74品目(43.2%)であり、1試験のみの添付が13/74品目(17.6%)であった(図5)。

このようにウルトラオーファンドラッグではウルトラ以外のオーファンドラッグに比し、国内無作為化比較試験の実施割合が極めて低かった。また、無作為化比較試験が2試験以上実施されている品目が少なく、実施されていない品目が多かった。

ウルトラオーファンドラッグを日本における患者数が1,000人未満と想定し検討を行ってきたが、患者数が500人未満と想定した場合、無作為化比較試験が2試験以上添付されている割合は22.2%(6/27品目)で、1,000人未満と同じであったが、100人未満と想定した場合には2/12品目(16.7%)で少なかった(図5)。100人未満で無作為化比較試験が実施されている品目は、「ガルスルファーゼ(遺伝子組換え)(販売名:ナグラザイム点滴静注液)」と「ドルナーゼ アルファ(遺伝子組換え)(販売名:プルモザイム吸入液)」で、試験数はそれぞれ2試験と4試験であった。ガルスルファーゼ(遺伝子組換え)の2試験は、第1相試験とプラセボ対照の第2相試験であり、第2相試験の総投与症例数は39例であった。

(2) 用量反応試験

ウルトラオーファンドラッグ 36 品目中、国内外で用量反応試験が実施された品目は 11 品目、国内用量反応試験のみが 7 品目、外用量反応試験のみが 11 品目であり、用量反応試験が実施されていない品目が 7 品目あった(図 6)。ウルトラ以外のオーファンドラッグでは、74 品目中、国内外で用量反応試験が実施された品目が 26 品目、国内のみが 28 品目、海外のみが 11 品目であり、用量反応試験が実施されていない品目が 9 品目であった(図 6)。ウルトラ以外のオーファンドラッグに比しウルトラオーファンドラッグの方が、国内用量反応試験が実施された品目の割合がやや低く、用量反応試験が実施されていない品目の割合がやや高かった。

ウルトラオーファンドラッグのうち海外で承認を取得していない 6 品目中 5 品目で国内用量反応試験が実施されていた。試験が実施されていない 1 品目は効能追加であった。腫瘍用薬では 8 品目中 5 品目で、国内外で用量反応試験が実施されており、残りは国内開発品であるタミバロテン(販売名: アムノレイク錠)(国内用量反応試験実施)、海外用量反応試験のみが添付されていたネララビン(販売名: アラノン G 静注用)、公知申請で承認されたイマチニブメシル酸塩(販売名: グリベック錠)であった(表 3)。一方、酵素製剤 6 品目では、海外の用量反応試験結果を踏まえ開発がなされ、国内では用量設定試験は実施されていなかった(表 4)。

国内用量反応試験成績が添付されていたウルトラオーファンドラッグ 18 品目 9 品目で任意漸増試験(同一対象)が、4 品目で強制的漸増試験(同一対象)が実施されていた(図 7)。海外用量反応試験成績が添付されていたウルトラオーファンドラッグでは、22 品目 8 品目で任意漸増試験が、1 品目で強制的漸増試験が実施されていた。ウルトラ以外のオーファンドラッグでは、国内用量反応試験が添付されていた 54 品目中 25 品目、海外用量

反応試験が添付されていた 37 品目中 10 品目で同一対象における漸増試験が実施されていた。海外よりも国内で同一症例を対象とした漸増試験が実施されている割合が多く、患者数が少なくなるほど(1,000 人未満 500 人未満 100 人未満と想定した場合ほど)同一対象漸増試験の占める割合が多くなる傾向が認められた(図 7)。

国内で用量反応試験を実施していない場合の用量設定の根拠として、海外用量反応試験成績、オープン試験の成績、臨床研究や使用実態調査、文献報告に加え、健康成人で実施した PK/PD 試験成績や非臨床試験結果を用いている事例が認められた。効能追加の開発では、既承認効能における推奨用量が用いられていた。

(3) 継続投与試験 / 長期投与試験

ウルトラオーファンドラッグ 36 品目中 14 品目で継続投与試験あるいは長期投与試験成績が添付されていた。そのうち国内外で継続投与試験 / 長期投与試験が実施されていた品目が 3 品目、国内のみが 3 品目、海外のみが 8 品目であった。

国内継続試験 / 長期投与試験が添付されていた 6 品目中全品目の試験で、既に臨床試験に参加した被験者が再度エントリーされており、6 品目中 3 品目がその他の生物学的製剤(モノクローナル抗体)であった。また、海外継続投与試験 / 長期投与試験が添付されていた 11 品目 6 品目がすべての酵素製剤(6 品目)であり、2 品目がその他の生物学的製剤(モノクローナル抗体)であった。11 品目中 9 品目の試験で被験者が再度エントリーされており、この 9 品目に酵素製剤 5 品目が含まれていた。被験者が再度エントリーしていなかった品目はアルグルコシダーゼ アルファ(遺伝子組換え)(販売名: マイオザイム点滴静注用)とドルナーゼ アルファ(遺伝子組換え)(販売名: プルモザイム吸入液)

であった。アルグルコシダーゼ アルファ(遺伝子組換え)では、他社で開発された rhGAA 製剤の臨床試験に参加した患者を対象とした継続投与試験及び臨床試験の組み入れ基準を満たさない重度の障害を有し他に治療法がない患者に対する継続提供試験が実施されていた。

(4) 参考資料

国内外のウルトラオーファンドラッグの参考資料として、レトロスペクティブ調査(適応外使用症例報告収集)、使用実績調査、臨床研究、先進医療・継続提供試験・継続調査、compassionate use の成績、他効能の試験、公表文献、ガイドライン等が用いられていた。

アルグルコシダーゼ アルファ(遺伝子組換え)(販売名：マイオザイム点滴静注用)では、海外で実施された自然経過歴に関する疫学研究が参考資料として添付され、この調査結果が臨床試験の既存対照群として用いられていた。腫瘍用薬では、他の癌腫における臨床試験成績が参考資料として添付され、安全性の評価に用いられている品目が散見された。インフリキシマブ(遺伝子組換え)(販売名：レミケード点滴静注用、効能・効果：ベーチェット病)では、当該効能における無作為化比較試験は実施されていなかったが、他の効能における無作為化比較試験が参考資料として添付されていた。

3. pivotal 臨床試験のデザイン

(1) 評価項目

同じ効能・効果における患者数の違いと評価項目の関係を明らかにするために、ウルトラオーファンドラッグ、ウルトラ以外のオーファンドラッグともに品目数の多かった腫瘍用薬における pivotal 試験の主要評価項目を比較した。ウルトラオーファンドラッグで無作為化比較試験が実施されていたのは8品目中2品目であった(表3)。1品目は「ダサ

チニブ水和物(販売名：スプリセル錠、効能・効果：イマチニブ抵抗性の慢性骨髄性白血病、フィラデルフィア染色体陽性急性リンパ性白血病)」であり、第 相試験の主要評価項目は細胞遺伝学的寛解(CyR)で、症例数は150例であった。もう1品目は「スニチニプリンゴ酸塩(販売名：スーテントカプセル、効能・効果：膵神経内分泌腫瘍)」であり、第 相試験の主要評価項目はPFSで、症例数は171例であった。他の品目では主に奏功率を主要評価項目とした非盲検非対照試験が実施されており、症例数は6~39例であった。

ウルトラ以外のオーファンドラッグである腫瘍用薬25品目中が無作為化比較試験が実施されていたのは8品目であり(表3)、主要評価項目が奏功率等である場合の症例数は110~190例程度/試験であったが、PFS、TTP、OSの場合には、症例数は160~920例程度/試験であり大規模なものもあった。また、他の品目で非盲検非対照試験が実施され、主要評価項目は主に奏功率で症例数は10~90例であった。

次にウルトラオーファンドラッグにおいて品目数の多かった酵素製剤では、6品目すべてにおいて無作為化比較試験が実施されており(表4)、アルグルコシダーゼ アルファ(遺伝子組換え)(販売名：マイオザイム点滴静注用、効能・効果：糖原病)では真のエンドポイントである生存率を主要評価項目とした2用量比較試験が実施され、総投与症例数は18例であった(設定根拠は記載されていない)。これを除く5品目でプラセボ対照比較試験が実施され、アガルシダーゼ ベータ(遺伝子組換え)(効能・効果：ファブリー病)では、代替エンドポイントとして「腎のGL-3蓄積に関する形態学的評価」がなされ、症例数は29例×2群であった。代替エンドポイントを用いた理由として、腎不全はファブリー病の進行及び死亡の主な原因であるが、腎機能が正常な患者集団において腎機能

の維持をエンドポイントとする試験を実施するには多数の患者を対象とした長期間の臨床試験が必要であり、患者数が極めて少ない本疾患においてこうした臨床試験を実施することは困難と考えられたことやファブリー病における糸球体硬化症の主要原因が GL-3 蓄積による血管障害であること等が記されていた。ムコ多糖症治療薬 3 品目のうち、イデフルスルファーゼ（遺伝子組換え）（販売名：エラプレース点滴静注用）とガルスルファーゼ（遺伝子組換え）（販売名：ナグラザイム点滴静注用）においても、臨床試験で用いられた有効性評価項目と真のエンドポイントと考えられる生命予後との関係について説明が求められていた。イデフルスルファーゼ（遺伝子組換え）では、患者数が限られていること及び疾患の進行が遅いことから真のエンドポイントである生存期間を主要評価項目とした臨床試験の実施は困難であり、歩行距離の短縮は、患者における肺、筋骨格系及び心血管系の終末器官不全に関連しており、これらすべてが患者の生存期間に影響を与えらると思われると回答している。

(2) 対照群の設定

先に述べたとおり、ウルトラオーファンドラッグ 36 品目中、無作為化比較試験が添付されている品目は 19/36 品目（52.8%）であり、これらでは言うまでもなく対照群が設定されていた（図 5）。一方、評価資料が単群試験のみから構成されている品目（群間比較用量反応試験等が実施されている品目は含まない）も 11 品目あった。7/11 品目が海外で承認を取得している品目であり、5/11 品目が効能追加であった（重複あり。いずれも医学薬学上公知であると判断され承認されたイマチニブメシル酸塩とサリドマイドが含まれる）。

なお、シナカルセト塩酸塩（販売名：レグパラ錠）のように、評価資料である海外第

相試験と国内第 相試験はいずれも非盲検長期投与試験であったが、参考資料のうち海外第 相試験 1 試験でプラセボ群（参照群）を設定した無作為化二重盲検比較試験が実施されているものの、統計解析手法に基づいた目標被験者数設定がおこなわれていない（本薬品 8 例、プラセボ群 4 例）品目もあった。

(3) 症例数の設定

ウルトラオーファンドラッグであるクラドリピン（販売名：ロイスタチン注）の国内第 相試験、ヘミン（販売名：ノーモサング点滴静注）の国内臨床試験、シナカルセト塩酸塩（販売名：レグパラ錠）の国内第 相試験では、患者数を踏まえ、実施可能性の観点から症例数が設定されていた。クラドリピンの目標症例数は 10 例、ヘミン（3 試験）の目標症例数は 2～5 例、シナカルセト塩酸塩の目標症例数は 2 例以上（総投与症例数 7 例）であった。

一方、海外で実施されたプラセボ対照無作為化比較試験では、申請資料概要に症例数設定根拠として有意水準、検出力が記載されている 9 試験すべてにおいて、有意水準を 0.05（両側）、検出力を 80～90% に設定し、症例数が算出されていた。症例数は効果に依存するため、カナキヌマブ（遺伝子組換え）（販売名：イラリス皮下注用）のように検出力が 90% であっても症例数が各群 10 例と算出される場合もあれば、バシリキシマブ（遺伝子組換え）（販売名：シムレクト）のように検出力を 80% と設定しても目標症例数が 1 群 170 例になるものもあった。また、アガルシダーゼ ベータ（遺伝子組換え）（販売名：ファブラザイム点滴静注用）の海外第 相試験では、症例数を 1 群 28 例と想定し、有意水準を 0.05（両側）としたときの検出力の範囲が記載されていた。

なお、ウルトラ以外のオーファンドラッグで、症例集積が困難なため有意水準を両側

0.1 に設定した品目（ルフィナミド（販売名：イノベロン錠、効能・効果：Lennox-Gastaut 症候群）の国内第 Ⅰ 相試験）があったが、総投与症例数は 59 例（プラセボ群 30 例、本薬群 29 例）であった。

(4) 比較のために工夫された試験デザイン

ウルトラオーファンドラッグの臨床試験において特記すべき試験デザインとして、乾燥スルホ化人免疫グロブリンで、国内で 23 例を対象にプラセボ対照投与期無作為化二重盲検前後比較試験が実施されていたことが挙げられる。

また、酵素製剤のうち唯一、プラセボ対照試験が実施されていなかったアルグルコシダーゼ アルファ（遺伝子組換え）（販売名：マイオザイム点滴静注射用）では、前述のように海外乳児型 GSD 患者 168 例を対象とした自然経過歴に関する疫学研究が行われ、この調査結果が海外第 Ⅰ 相用量設定試験において既存対照群（168 例中 61 例を抽出）として用いられていた。

ヘミン（販売名：ノーモサング点滴静注）では、参考資料であるが、予防投与における効果を検討するために個人内比較試験が実施されていた。

4. 有効性の評価

ウルトラオーファンドラッグでは、無作為化比較試験で統計学的有意差が認められなかったにもかかわらず承認された品目があった。

アガルシダーゼ アルファ（遺伝子組換え）（販売名：リプレガル点滴静注用）では、米国で実施された無作為化プラセボ対照二重盲検試験において主要評価項目（BPI 項目 3 のスコアの低下）に統計学的な有意差は認められず、英国で実施された無作為化プラセボ対照二重盲検試験においても主要評価項目（心臓中 CTH 量）に統計学的な有意差は認められなかったが、承認を取得した。米国試験の

副次評価項目のうち 4 項目（抗てんかん薬等の服用を中止することができた期間、性状系球体の割合、メサングウム肥厚系球体の割合、血漿中 CTH 濃度の変化量）で有意差が認められ、英国試験の副次評価項目 3 項目のうち 2 項目（左室心筋重要の変化量、血漿中 CTH 濃度の変化）でも有意差が認められていた。また、本邦における第 Ⅰ 相オープン試験の主要評価項目 3 項目のうち 2 項目（血漿中 CTH 濃度、尿沈渣中 CTH 量）で有意な変化がみとめられていた。機構における審査では、本邦及び米国試験における被験者ごとの血漿中、尿沈渣中及び腎組織中 CTH 濃度の変化と腎病理組織学的評価との関係から、全体として有効性が示唆されていると判断されていた。

ラロニダーゼ（遺伝子組換え）（販売名：アルドラザイム点滴静注液）では、海外第 Ⅰ 相試験（無作為化二重盲検比較試験）において主要評価項目が 2 項目設定されており、努力肺活量の予測正常値に対する割合に統計学的有意差は認められたが、6 分間歩行距離には有意差は認められなかった。機構における審査では「現時点で疾患に対する確立した評価法がないことから、種々の臨床症状の指標となり得る海外第 Ⅰ 相試験で設定された主要評価項目 2 項目、副次評価項目 4 項目及び三次評価項目 12 項目を以て有効性を評価せざるを得ない」と判断されていた。

5. 製造販売後における対応

ウルトラオーファンドラッグ、ウルトラ以外のオーファンドラッグいずれにおいても、検討したすべての品目で再審査期間が付与されていた。また、ウルトラオーファンドラッグ 36 品目中 30 品目（83.3%）、ウルトラ以外のオーファンドラッグ 74 品目中 53 品目（71.6%）で製造販売後に投与全症例を対象とした調査が要求されていた（図 8）。

加えて、ウルトラオーファンドラッグ 2 品目、ウルトラ以外のオーファンドラッグ 7 品

目で製造販売後臨床試験が要求されていた。ウルトラオーファンドラッグであるバシリキシマブ（遺伝子組換え）（販売名：シムレクト注射用）では、日本人小児患者に対する投与経験がないことから、日本における小児の用量・用量を検討することを目的とした試験が求められた。アガルシダーゼ ベータ（遺伝子組換え）（販売名：ファブラザイム点滴静注用）では、心ファブリー病は臨床試験が実施された古典型ファブリー病と病態に差異があることから、心ファブリー病に対する有効性及び安全性の明確化を目的とした製造販売後臨床試験が求められていた。

6. 未承認薬使用問題検討会議

今回検討したウルトラオーファンドラッグ 36 品目中 16 品目（44.4%）、ウルトラ以外のオーファンドラッグ 74 品目中 10 品目（13.5%）で未承認薬使用問題検討会議において開発要請がなされていた（図 9）。特に 2011 年以降に承認されたウルトラオーファンドラッグでは、14 品目中 11 品目（78.6%）で未承認薬使用問題検討会において開発が要請されていた。

国内で臨床試験を実施することなく承認されたウルトラオーファンドラッグ 8 品目中 6 品目で未承認薬使用問題検討会議において開発要請がなされていた。また、酵素製剤 6 品目中、未承認薬問題検討会で検討されなかった 3 品目（いずれも 2006 年以前の承認）では国内臨床試験が実施されていたが、未承認薬問題検討会議で開発要請がなされた 3 品目（いずれも 2007 年以降の承認）では国内臨床試験を実施することなく承認されていた。

D. 考察

エビデンスの構築：

一般的に、医薬品の承認時点でその有効性・安全性に関する高いエビデンスを示す手段の一つは、2 本以上の無作為化比較試験に

おいて適切な対照群に対する統計学的有為差を示すことである。PMDA が発出した「新医薬品承認審査実務に関わる審査員のための留意事項」（平成 20 年 4 月 17 日）³⁾にも、「結果の信頼性を担保するため、原則として「2 本以上の無作為化比較試験」において、有効性が検証されていることが望ましい」ことが記載されているが、一方で「特に、希少疾病医薬品又は既存治療が確立していない重篤な疾患領域の医薬品については、その臨床的意義等を踏まえ、個別に判断すべき事項が多いと考えられることから、必ずしも本文書における各適用のみによって最終判断を行うものではないが、本文書の趣旨を十分に理解し、適切なデータのもとに科学的な評価を行う必要がある」という記載もある。ウルトラオーファンドラッグは希少疾病用医薬品の中でも患者数が特に少ない疾病（本検討では 1,000 人以下と想定）を対象としていることから、こうした考え方を理解し、適切な臨床開発を行う必要がある。事実、今回検討したウルトラオーファンドラッグ 36 品目のうち評価資料として無作為化比較試験が 2 試験以上添付されていた品目は、検討した 36 品目中 8 品目でしかなく、無作為化比較試験が 1 試験であるものが 11 品目、なしが 17 品目であった。このように患者数が極めて少ない場合には、許容可能なデータがあれば必ずしも 2 本以上の無作為化比較試験の実施が必要ではないことが伺える。

ウルトラオーファンドラッグを日本における患者数が 1,000 人未満と想定し検討を行ってきたが、その中でも患者数が更に少ない場合には、無作為化比較試験が 2 試験以上添付されている割合が減少する傾向があった。ウルトラオーファンドラッグであっても一定の症例数が確保できる場合においては、複数の無作為化比較試験（用量反応試験や海外試験を含む）を実施することも可能であるが、症例数が更に少ない場合には、無作為化比較

試験 1 本あるいは少数例の単群試験のみで有効性を評価せざるを得ない場合も想定される。

用量反応試験：

用量反応試験は適切な用量を探索するために重要な試験であり、並行群間比較用量 反応試験、クロスオーバー用量 反応試験、強制的漸増試験、任意漸増試験等、種々のデザインが知られている。今回検討したウルトラオーファンドラッグでも 36 品目中 29 品目で用量反応試験が実施されており、ウルトラオーファンドラッグであっても可能な限り用量反応試験を実施することが望ましいことは言うまでもない。ただし、症例数が極めて少なく複数の用量群をおいた並行群間比較試験や群漸増試験の実施が困難な疾患も想定される。今回検討したウルトラオーファンドラッグのうち国内用量反応試験が添付されていた 18 品目中 13 品目、海外用量反応試験が添付されていた 22 品目中 9 品目で強制的漸増試験や任意漸増試験が実施されていた。また、患者数が少なくなるほど、同一対象漸増試験の占める割合が多くなる傾向が認められた。ウルトラオーファンドラッグでの用量設定の方法として、強制的漸増試験により用量設定を行うことや任意漸増試験により各個人において適切な用量を探索することが有用と考えられた。後者については用量設定期を設けた試験も散見されている。しかしながら、こうした方法が適用可能な疾患に限られること、持ち越し効果の影響があることなどに留意する必要がある。

用量反応試験を実施しない場合には、薬理作用、健康成人で実施した PK/PD 試験成績、治験での投与経験、海外用量反応試験成績、臨床研究や使用実態調査、文献報告等も用いて、用法・用量設定の妥当性を説明する必要がある。効能追加の開発では、既存効能における推奨用量も十分参考にすべきである。

用量設定の根拠を説明する際に海外用量反応試験成績を参考にすることは可能であるが、今回検討したウルトラオーファンの腫瘍用薬では（国内開発品であるタミバロテン、国内臨床試験が実施されていないネララビン、公知申請で承認されたイマチニブメシル酸塩を除く）すべての品目で海外用量試験成績に加え国内用量反応試験が実施されていた。このような品目もあることから、用量反応試験を実施しない場合には、参照とした資料の適切性を含め用いる用量の妥当性を十分検討する必要がある。

継続投与試験 / 長期投与試験：

今回検討したウルトラオーファンドラッグ 36 品目中 14 品目で継続投与試験あるいは長期投与試験が実施され、ほとんどの品目で被験者が再エントリーされていた。患者数が特に少ない希少疾病に対する酵素補充療法や長期投与が想定されるモノクローナル抗体製剤等において、継続投与試験は医薬品を長期間投与した際の安全性及び有効性を確認する等の薬効評価の観点のみならず、患者（被験者）が治療を継続できるメリットがあることからも有用と考える。

国際共同治験や海外試験成績の利用：

国内患者数が少なく国内で十分なエビデンスレベルの高い臨床試験が実施出来ないことが想定される場合の有効な開発戦略として、国際共同試験の実施により症例数を増やし、適切な試験を計画立案することが考えられる。今回検討したウルトラオーファンドラッグでは日本を含む国際共同治験は実施されていなかったが、今後、臨床開発戦略の一つとして、積極的に実施を検討すべきと考える。

加えて、海外試験成績を利用して開発することが考えられる。今回の検討対象の中では 36 品目中 24 品目で海外臨床試験成績が評価資料として添付されており、海外で承認を取

得している 30 品目では 23 品目で海外臨床試験成績が添付されていた。また、日本における臨床試験成績が評価資料として添付されていなかった 8 品目のうち、医学薬学上公知であると判断され承認された品目を除いた 6 品目すべてにおいて海外臨床試験成績が添付されており、海外データが有効性・安全性の担保に利用されていることがわかる。特に、ブルモザイム（遺伝子組換え）や塩酸メフロキンのように、日本では希少疾病であるが、海外では比較的症例数の多い臨床試験が実施されている品目では、海外臨床試験成績の利用が有用であると言える。また、国内では無作為化比較試験が実施された品目は 1 品目のみであったが、18 品目で海外で実施された無作為化比較試験成績が評価資料として添付されていた。このように、より高いエビデンスを確立するために海外臨床試験成績を利用することも有効な手段であると考えられる。

臨床研究や使用実態調査の利用：

国内患者数が少なく国内で十分な治験が実施出来ないことが想定される場合の有効な開発戦略の一つとして、治験成績に加えて日本における臨床研究や使用実態調査の結果を参考資料として用いることが挙げられる。今回検討したウルトラオーファンドラッグ 36 品目中 21 品目で国内外の臨床研究や各種の調査結果などが参考資料として添付されていたが、ウルトラオーファンドラッグの中でも更に患者の数が少なくなるにつれ、国内参考資料が添付されている割合が多くなっていた。臨床研究や使用実態調査はエビデンスレベルや資料の信頼性が必ずしも高いとは言えないが、日本の医療実態下での使用成績を利用することができるという利点がある。また、アルグルコシダーゼ アルファ（遺伝子組換え）で、海外の自然経過歴に関する疫学研究が参考資料として添付され、臨床試験の既存対照群として用いられていた事例があるように海

外にしかない研究結果等を利用することも有用であろう。

なお、こうした国内外資料の事例として、臨床研究、先進医療、継続提供、適応外使用症例報告や公表文献、場合によってガイドライン等が考えられる。

評価項目：

ウルトラオーファンドラッグの臨床試験においても主要評価項目として真のエンドポイントを用いることが望ましいが、今回検討した酵素製剤に見られたように真のエンドポイントである生存期間ではなく代替エンドポイントを用いることにより、治験の実施そのものが可能になったり、治験期間の短縮や少ない症例数での治験が可能となったりする場合もある。このように代替エンドポイントによる評価が可能となる場合も考えられるが、代替エンドポイントと真のエンドポイントとの関係を明確に説明する必要がある。

腫瘍用薬では「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」（薬食審査発第 1101001 号、平成 17 年 11 月 1 日）⁴⁾において、「患者数が多い癌腫では、延命効果を中心に評価する第 相試験の成績を承認申請時に提出することを必須」とされており、希少癌では必ずしも延命効果が要求されるわけではないことが示唆されている。今回、検討したウルトラオーファンドラッグの腫瘍用薬では延命効果を確認したものは 1 品目に過ぎず、奏功率で有効性が認められれば承認される可能性も考えられる。

また、副次評価項目として、可能な限り多くのエンドポイントを評価し、それらの一貫性や関連性を評価することは、薬効を総合的に評価する上で有用である。アルグルコシダーゼ アルファ（遺伝子組換え）やラロニダーゼ（遺伝子組換え）のように主要評価項目が統計学的有意差を示さなくとも副次評価項

目の動きを総合的に判断して承認に繋がるケースも考えられる。

対照群の設定：

一般に対象疾病の病因や性質（自然経過を含む）、医薬品の作用機序、臨床試験で用いられる評価項目や得られる効果の大きさ等も勘案して、同時比較対照群の設置の必要性を判断する。作用機序が明確で期待される効果が推定可能な場合や疾患の自然経過が明確な場合等では単群の試験により有効性を評価することも可能であろう。

一方、症例数が極めて少なくても、シナカルセット塩酸塩の海外第 Ⅰ 相試験（参考資料）に見られるように参照群（統計的な検出力を担保しないものの比較のために置かれる対照群（プラセボあるいは既存薬））をおいた無作為化比較試験の実施を検討した方が有効性評価がしやすい場合もある。

更に、外部対照を用いるといった開発戦略も考えられる。前述のアルグルコシダーゼ アルファ（遺伝子組換え）では、自然経過歴に関する疫学研究が行われ、この調査結果が海外第 Ⅰ / Ⅱ 相用量設定試験において既存対照群として用いられていた。こうした場合、外部対照の患者群と臨床試験の対象に違いがないことや調査時期による影響がないことを担保することが重要と考えられる。アルグルコシダーゼ アルファ（遺伝子組換え）では、対照群として疫学研究 168 例中 61 例を抽出していたが、このようにマッチングなどの手法を積極的に導入することも、外部対照と臨床試験の対象との違いを少なくし、エビデンスを高めるための手段と言えよう。調査時期による影響を除くためには、可能であれば同時期での対照を用いることが望ましいが、既存対照を用いる場合には、その妥当性を説明する必要があるであろう。患者登録（レジストリー）は患者のリクルートを円滑にするの

みならず、適切な外部対照を入手する手段として有用であることが考えられる。

臨床試験デザイン：

対照との比較を行う場合には、通常、並行群間比較試験の実施が考えられる。並行群間比較試験に比べ少ない症例数で比較が可能なクロスオーバー試験を実施することも考えられるが、適応できるケースが限定されること、持ち越し効果や時期効果が存在し結果にバイアスが生じる可能性があることに留意する必要がある。持ち越し効果等の影響も含めて評価を行うために、例えば、乾燥スルホ化人免疫グロブリンで用いられた対照の投与期間を置いた投与期無作為化二重盲験前後比較試験を実施することも考えられる。この試験デザインは、PMDA が「希少疾病で主観的な評価項目を設定せざるを得ないような疾患領域において、盲検下でプラセボとの比較が可能な試験デザインを検討することは、今後の医薬品開発において有用と考えている」ことが審査報告書に記載されている。こうしたデザインが適用できる疾患は限定されているが、症例数が少なくてもエビデンスレベルの高いデータを得るために工夫された方法の一つと考えられる。

一方、疾患の自然経過が明確な場合には、閾値を設定した非盲検非対照試験により有効性を評価することも可能であろう。

また、ヘミンの参考資料にあるように予防投与における効果を検討するような場合には、個人内比較試験を実施することにより有効性を評価することも考えられる。このような場合には、発症頻度を踏まえ、観察期間と投与期間を適切に設定する必要があることは言うまでもない。

なお、欧州の Guideline on Clinical Trials in Small Populations⁵⁾では、臨床試験デザインの段階で考慮すべき事項としてランダム化の方法、Response-adaptive methods、

Sequential designs、*n*-of-1 trials を挙げている。adaptive methods や *n*-of-1 trials により承認された品目はまだ認められておらず、実施にあたっては規制当局との議論が必要であろう。

症例数：

一般に、有意水準は 5%に、検出力は 80% ~ 90%に設定されることが多いが、ウルトラオーファンドラッグの無作為化比較試験を実施する場合、症例数が少なく十分な検出力を確保することができないことも想定される。今回検討した品目の無作為化比較試験では、有意水準を 0.05(両側)、検出力を 80~90%に設定し、症例数が算出されていたが、ウルトラオーファンドラッグの中でも患者数が特に少ない場合等においては、検出力を下げることを考慮することも必要になってくるであろう。

ウルトラ以外のオーファンドラッグであるルフィナミドでは症例集積が困難であることを理由に有意水準が両側 0.1 に設定されていた。不十分な検出力の下で機械的に有意水準 5%の評価を行うよりも、との関連性を示したうえで調整されたに基づく評価を行うことは有益との考えもあるようだが⁶⁾、こうした考えに基づき症例数設定を行う場合には、事前に規制当局と十分議論する必要があるであろう。

また、クラドリピンの国内第 相試験、ヘミンの国内臨床試験、シナカルセト塩酸塩の国内第 相試験のように、患者数を踏まえ、実施可能性の観点から症例数を設定し、その中で有効性および安全性を検討せざるを得ない場合も想定される。

腫瘍用薬では第 相試験のプロトコールデザインの一つとして「第 2.5 相試験」と言われる無作為化比較試験がある⁷⁾。この試験の目的は対照群との比較であるが、検証試験で

はないため、代替エンドポイントや通常用いられる 5%有意水準よりも大きい有意水準が許容される。このような考え方をウルトラオーファンドラッグの開発に用いることも考えられるのではないだろうか。

製造販売後における情報収集：

治験で得られる情報には限界があることから、製造販売後調査等により安全性・有効性情報を収集することは情報量を増やす上で有用である。特に、ウルトラオーファンドラッグに関しては、治験で得られる情報が希少疾病用医薬品の中でもより少ないことが想定されることから、製造販売後調査の位置づけはより高いと考えられる。今回、検討したウルトラオーファンドラッグでは 80%以上の品目で、ウルトラ以外のオーファンドラッグでも 70%以上の品目は製造販売後に投与全症例を対象とした調査/全症調査がなされていた。製造販売後に投与全症例を対象とした調査を行うことに加え、例えば、治験で用量反応の検討が十分できなかった場合に推奨用量を再確認することや治験の選択基準以外の対象における有効性・安全性を確認するための製造販売後臨床試験を実施することも想定される。加えて、製造販売後、これら以外にエビデンスレベルを高める方法として、製造販売後調査中に無治療群等の対照も含めて調査することも考えられる。ウルトラ以外のオーファンドラッグであるが、抗ヒト胸腺細胞ウサギ免疫グロブリン(販売名：サイモグリブリン点滴静注用、効能・効果：骨髄幹細胞移植の前治療では、「治験計画時に適切な対象を設定すべきであったにもかかわらず、やむなく事後的に外部対照と比較していたことから、製造販売後調査の実施にあたり、対照となりうるような外部データを用い、可能限り多くの背景因子を考慮した評価を行うべきである」ことが述べられている。こうした方法は製造販売後にエビデンスを高めるために有

用であるが、そのためには外部の利用できるデータがあることが望ましい。患者レジストリー等の構築等を考えていくことは希少疾病用医薬品の開発のみならず市販後にエビデンスを高めるためにも有用と考える。

未承認薬使用問題検討会の利用：

制度に関しては、未承認薬使用問題検討会議で開発要請がなされることも、希少疾病用医薬品あるいは患者数が特に少ない疾病に対する医薬品等の開発の推進に役立つ手段と考えられる。事実、今回検討したウルトラオーファンドラッグ 36 品目中 13 品目で、未承認薬使用問題検討会議で開発要請がなされ、特に 2011 年以降に承認された品目では 11 品目中 9 品目で開発要請がなされていた。また、国内臨床試験を実施することなく承認されている 8 品目のうち 6 品目で、未承認薬使用問題検討会議で開発要請がなされており、未承認薬使用問題検討会議を利用することは、開発を促進するための手段の一つであると考えられる。

E．結論

2001 年 4 月から 2013 年 3 月までに承認されたウルトラオーファンドラッグの臨床データパッケージを調査し、ガイダンスに記載すべき内容として、少数の被験者でも合理的に有効性・安全性を評価するために留意すべき事項をまとめた（下記）

エビデンスの構築

- (1) 無作為化比較試験
- (2) 用量反応試験
- (3) 継続投与試験の実施

利用データソース：

- (1) 国際共同治験や海外試験成績の利用
- (2) 臨床研究や使用実態調査の利用

Pivotal 試験の計画

- (1) 評価項目
- (2) 対照群の設定

(3) 試験デザイン

(4) 症例数の設定

製造販売後における情報収集

なお、制度面では、未承認薬使用問題検討会議を利用することも開発を促進するために有効な手段と考えられた。

F．健康危険情報 なし

G．研究発表

【学会発表】

- 1) 荒戸照世，金子真之，前田浩次郎，成川衛：日本で承認された患者数が特に少ない希少疾病用医薬品の臨床データパッケージについて．第4回レギュラトリーサイエンス学会学術大会，東京，平成26年9月5日
- 2) 前田浩次郎，金子真之，成川衛，荒戸照世：患者数が特に少ない希少疾病用医薬品の臨床評価のポイントについて．第35回日本臨床薬理学会学術総会，松山，平成26年12月6日

H．知的財産権の出願・登録状況 なし

(参考資料)

- 1) 第 21 回未承認薬使用問題検討会議 資料 5 「未承認薬使用問題検討会議」での検討結果等について
<http://www.mhlw.go.jp/shingi/2009/06/dl/s0618-10f.pdf>
- 2) 医療上必要性の高い未承認薬・適用外薬検討会議での検討結果を受けて開発企業の募集又は開発要請を行った医薬品のリスト
http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iyakuhin/kaihatsuyousei/index.html
- 3) 「新医薬品承認審査実務に関わる審査員のための留意事項」(平成 20 年 4 月 17 日)

<http://www.pmda.go.jp/topics/file/h200417koho.pdf>

- 4) 「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」(薬食審査発第 1101001 号、平成 17 年 11 月 1 日)
http://www.japal.org/contents/19910204_9.pdf
- 5) Guideline on Clinical Trials in Small Populations (London, 27 July 2006, Doc. Ref. CHMP/EWP/83561/2005)
http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC50003615.pdf
- 6) Small Clinical Trials による薬効評価の考え方
日本製薬工業会 医薬品評価部委員会 データサイエンス部会 タスクフォース 2 2013 年 5 月
http://www.jpma.or.jp/information/evaluation/allotment/pdf/trials_01.pdf
- 7) Rubinstein L, Crowley J, Ivy P, et al. Randomized phase II designs. Clin Cancer Res. 2009; 15(6): 1883-90.

