

事例から学ぶ費用対効果評価の留意事項

日本製薬工業協会

データサイエンス部会

2022 年度継続タスクフォース 7

「医療経済評価に関連する統計的手法の整理と紹介」

Ver 1.0

2023 年 3 月

目次

1. はじめに	4
2. 費用対効果評価の概要及びフレームワーク	5
3. 事例から学ぶ留意事項	10
3.1. Step 1：分析の目的，比較対照の決定	11
留意事項①：分析対象集団は，分析前協議で十分議論しておく	11
留意事項②：比較対照技術は，分析前協議で十分議論しておく	12
3.2. Step 2：分析の立場と費用の範囲の決定	13
留意事項③：費用に含める診療・治療行為は，実臨床を想定して設定する	13
3.3. Step 3：効果指標と分析の決定	14
留意事項④：追加的有用性がない場合，費用最小化分析となる	15
留意事項⑤：追加的有用性を適切に判断可能なロジックを構築する	16
3.4. Step 4：モデルの構築とパラメータの推定	17
3.4.1. モデルの構築	17
留意事項⑥：分析期間は十分に長い期間を設定する	18
留意事項⑦：各健康状態の起こりうる推移を精査する	19
留意事項⑧-1：代替エンドポイントと真のエンドポイントの関係性を明確にする	21
留意事項⑧-2：代替エンドポイントと真のエンドポイントの関係性を明確にする	22
留意事項⑨：臨床試験の患者集団と実臨床の患者集団に違いを考慮する	24
留意事項⑩：疫学研究を用いる場合，未測定交絡因子の存在を考慮する	25
留意事項⑪：臨床上許容できない有害事象の発現を考慮する	26
留意事項⑫：治療期間を適切に設定する	27
留意事項⑬：治療のスイッチを適切に設定する	28
3.4.2. パラメータの推定	29
留意事項⑭：医療情報データベースを活用する場合，適切にデータベースを選定し，費用や健康状態の定義を適切に設定する	31
留意事項⑮：健康状態と QOL 値の大小関係の整合性を確認する	31

留意事項⑯：QOL 値で捕捉できない症状の改善があるかもしれない	33
留意事項⑰：健康関連 QOL から QOL 値へのマッピングは適切に行う	34
留意事項⑱：健康状態の QOL 値を推定する際、データソースの一貫性を保つ.....	34
留意事項⑲：異なる尺度で推定した QOL 値を用いない.....	35
留意事項⑳：有害事象発現による QOL 値の減少を適切に設定する	36
留意事項㉑：生存時間のモデリング（外挿も含む）は慎重に行う	37
留意事項㉒：間接比較を行う場合、試験間の効果修飾因子の分布の違いを考慮する	37
留意事項㉓：特に海外データの QOL 値を用いる場合、国民標準値との整合性を確認する	
39	
3.5. Step 5：費用と効果の比較：ICER などの算出	40
3.6. Step 6：分析結果の解釈・報告.....	41
留意事項㉔：分析（ex. 確率的感度分析）方法は適切に報告する	41
留意事項㉕：不確実性が結果に与える影響を評価し、解釈する	42
留意事項㉖：評価の妥当性を説明する	44
4. おわりに	46

1. はじめに

日本では 2016 年に費用対効果評価の試行的導入、2019 年に本格導入が行われ、費用対効果評価の経験が徐々に蓄積され始めている。この過程で試行的導入における分析で顕在化した課題については「分析前協議」の対応事項を再検討する等の対応がなされた[1]。また、2023 年 1 月 8 日時点で評価対象品目に指定された 9 品目の評価結果が保健医療経済評価研究センター（Center for Outcomes Research and Economic Evaluation for Health, 以下 C2H）から公開されている[2]ⁱ。一方で、製薬企業の実務担当者ⁱⁱの観点から、これらの費用対効果評価の蓄積されてきた経験が十分に共有・整理されてはいないと考える。

海外に目を向けると、英国では政府（保健省）が評価対象技術を指定し、医療技術評価機関である NICE（National Institute for Health and Care Excellence）が医薬品、医療技術などについて経済評価を行い、NHS（National Health Service）に対して給付範囲に加えるべきかどうかの勧告（recommendation）を行っている。ここでの議論は、一部黒塗りはあるものの、一般公開されており、その中で企業分析に対して日本の公的分析班に相当する ERG（Evidence Review Group）がどのような指摘をしているか、これらの議論を踏まえて Appraisal Committee はどのような意思決定を下したかが確認できる。日本においても前述の通り、C2H が本格導入後の評価対象品目で評価が終了した品目に対して、評価結果を公表し、企業分析と公的分析班で評価時に議論となった点を確認できるようになっている。NICE では費用対効果評価を保険給付の可否判断に用いており、日本では価格調整の判断に用いているため、費用対効果評価の活用方法は異なるものの、NICE や日本の実際の費用対効果評価の先行事例から私たちが学ぶべきことは多いと考える。

費用対効果評価のためのモデル構築に取り組む実務担当者は様々な実務的な問題に遭遇していると考えられるが、これらの実務的な問題を踏まえた留意事項が整理された報告書等は私達の知る限り、日本にはまだないと考える。そこで、本報告書は、費用対効果分析モデル構築に取り組む実務担当者を対象に、NICE や日本で行われた費用対効果評価で議論となった点等を紹介し、可能な範囲で課題解決のためのヒントを提供することを目的に作成した。なお、議論となった点に関しては網羅的ではないものの、それぞれのトピックとして典型的な問題を取り上げるように留意している。2 章では費用対効果評価の概要とフレームワークを説明し、3 章では事例を通して費用対効果評価で議論となった点を留意事項として紹介する。

ⁱ 中央社会保険医療協議会費用対効果評価専門組織における総合的評価の議事録も公開が始まった。

参考 URL：https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431_00302.html

ⁱⁱ 主に HEOR 担当者、データサイエンス担当者。

2. 費用対効果評価の概要及びフレームワーク

医療技術の価値に見合った価格を考えるためには、評価対象技術と比較対照技術について「効果」と「関連する費用」の両面から相対的に評価することが必要となる。この医療技術の効率性を評価する手法の一つとして、費用対効果評価がある。費用対効果評価は評価対象技術にかかる「費用」のみに着目するのではなく、それを使用することによって生み出される「効果」が「費用」の増加に見合うかどうかを比較対照技術と比較することで医療技術の効率性を評価する方法である（図 2-1）。「中央社会保険医療協議会における費用対効果評価の分析ガイドライン第3版」（以下、国内ガイドライン）では、中央社会保険医療協議会（以下、中医協）において評価対象に選定された医薬技術の費用対効果評価を実施する際に用いるべき分析方法が提示されている[3]。また、実際に製薬企業が評価結果を提出する際は「医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引」（以下、分析データの提出様式）に沿って、分析結果を記載することとなる[4]。

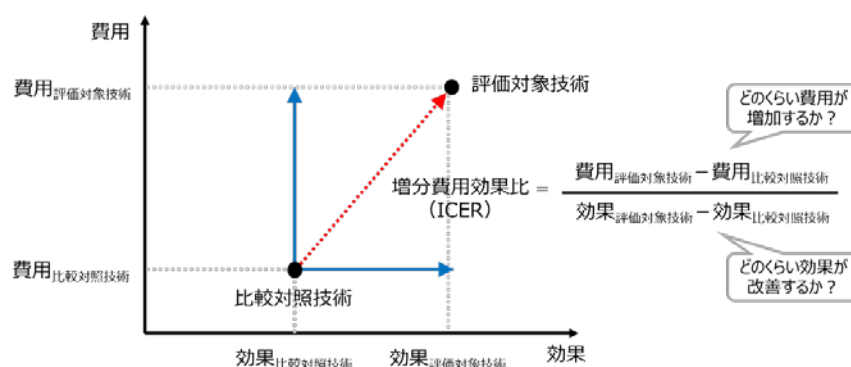


図 2-1 効果と費用の相対的な評価

次章で費用対効果評価の「事例から学ぶ留意事項」を紹介するが、本章ではその際に最低限必要と考える知識を6つの実施手順（フレームワーク）に沿って紹介する（図 2-2）[5]。費用対効果評価に関する詳細は成書等を参照されたい[6][7][8]。

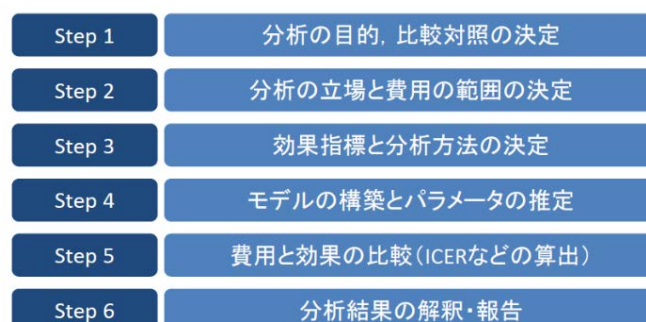


図 2-2 費用対効果評価の6つの実施手順

Step 1：分析の目的，比較対照の決定

費用対効果評価では，まず分析の目的を明確にし，分析対象集団と比較対照技術を決める必要がある．国内ガイドラインでは分析対象集団を「品目の指定時点において，評価対象技術の適応症となる患者」，比較対照技術を「評価対象技術が分析対象集団への治療として導入された時点で，臨床的に幅広く使用されており，評価対象技術によって代替されると想定されるもののうち，治療効果がより高いものを一つ選定」としている．費用対効果評価では，分析対象集団に対して，費用と効果の関係について評価対象技術と比較対照技術との相対的な評価を下す必要があるため，分析対象集団，比較対照技術に何を設定するかは重要な論点となる．

Step 2：分析の立場と費用の範囲の決定

どの意思決定者の立場で分析を行うか，誰にとっての費用を考えるかによって，分析に含める費用の範囲が決まる（表 2-1）が，国内ガイドラインでは分析の立場を原則として「公的医療の立場」としている．そして，「公的医療の立場」では費用の範囲は「公的医療保険制度の範囲」となる．もし公的介護費に与える影響が大きい場合には国内ガイドラインで「「公的医療・介護の立場」の分析を行ってもよい」としているが，「公的医療の立場」を行った上での追加的な分析として扱われるⁱⁱⁱ．

表 2-1 費用対効果評価の目的と分析の立場の関係

分析の立場	費用の範囲
公的医療の立場	公的医療保険制度における医療費のみ（自己負担分を含む）．
公的医療・介護の立場	公的医療保険制度における医療費＋公的介護費（自己負担分を含む）．
限定された社会の立場	公的医療の立場や公的医療・介護の立場に生産性損失を加えたもの．
患者の立場	患者・家族が負担する医療費，OTC，介護費用など．

Step 3：効果指標と分析方法の決定

効果指標は QOL 値^{iv}で重みづけた生存時間である質調整生存年(quality-adjusted life year；QALY)が最もよく用いられる．国内ガイドラインでも原則として QALY を用いることとしている．

分析方法は効果指標によって表 2-2 のように決まるが，国内ガイドラインでは評価対象技術に追加的有用性が示されていると判断できる場合は費用効用分析（cost-utility analysis；CUA）を行い，一方でアウトカムは比較対照技術と同等と考えられ，したがって追加的有用性を有する

ⁱⁱⁱ 言い換えると，日本において，価格調整の検討は「公的医療の立場」の基本分析の結果に基づくため，「公的医療・介護の立場」の結果は価格調整には用いられない．

^{iv} QOL 値は死亡を 0，完全な健康を 1 とした間隔尺度であり，効用値とも表現される．中医協における費用対効果評価の制度化の議論の中で「QOL 値」という用語で統一されてきたため，本報告書では QOL 値という表現を用いることとした．

と判断できない場合は費用最小化分析（cost-minimization analysis：CMA）を行うと定められている。

表 2-2 分析方法と効果指標

分析方法	効果指標
費用最小化分析（CMA）	アウトカムを同等とにおいて費用のみを比較
費用効果分析 （cost-effectiveness analysis：CEA）	QALY以外（生存年，イベント発生割合 等）
費用効用分析（CUA）	QALY
費用便益分析 （cost-benefit analysis：CBA） ^v	金銭換算された便益

Step 4：モデルの構築とパラメータの推定

分析期間を設定し，疾患の予後と治療の流れである治療のスイッチ，治療期間などをモデル化することで疾患推移モデルを構築する．代表的な疾患推移モデルとして決定樹モデル，マルコフモデル等がある（図 2-3）．一つの疾患領域でも様々なモデルのバリエーションがあるため，ケースごとに臨床的に妥当で最適と考えられる疾患推移モデルを構築する必要がある[9]．そして，構築した疾患推移モデルに対して，各健康状態における費用及び効果，健康状態間の推移確率といったモデルパラメータを推定する．

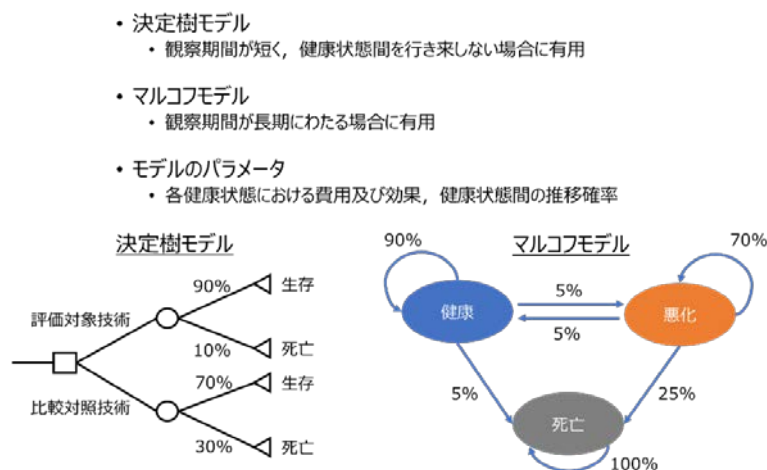


図 2-3 代表的な疾患推移モデルとモデルパラメータ

モデル中の数値は仮想の値である。

^v バイアスなく正確に金銭換算した便益を推定することが難しく，費用便益分析の使用は限定的である。

参考資料：European network for Health Technology Assessment のガイドライン，

https://www.eunetha.eu/wp-content/uploads/2018/03/Methods_for_health_economic_evaluations.pdf

モデルパラメータを推定する際、効果（QOL 値を含む）の検討には国内や海外の疫学・臨床試験データ、費用の検討には国内の治療環境での診療報酬等の医療費関連データを用いる。国内ガイドラインでは、最終的な費用対効果評価の指標となる ICER（incremental cost-effectiveness ratio；増分費用効果比）等を算出するために必要な有効性・安全性・QOL 値等のデータは「研究の質やエビデンスレベルが高く、かつ日本における現実の臨床成績を反映しているものを優先的に使用する」としており、都合のよい研究だけを参照した分析を避け、透明性の高い公正な評価を行うためにシステマティックレビューに基づく先行研究の抽出と選択が必要となる。また、評価対象技術と比較対照技術を直接比較したエビデンスがない場合、間接比較が行われる。

Step 5：費用と効果の比較：ICER などの算出

Step 5 では、評価対象技術が比較対照技術に比べて追加的有用性があると判断した場合に費用と効果を比較するために主に ICER を算出する。QALY を効果の評価指標とした場合、ICER は比較対照技術から 1QALY を追加で得るために必要となる追加費用となり、この値に応じて、費用対効果が良いかどうかを判断する。日本ではこの ICER に応じて価格調整が行われる[10]。

$$\text{ICER} = \frac{\text{費用}_{\text{評価対象技術}} - \text{費用}_{\text{比較対照技術}}}{\text{効果}_{\text{評価対象技術}} - \text{効果}_{\text{比較対照技術}}}$$

Step 6：分析結果の解釈・報告

モデル構造・モデルパラメータの妥当性、分析対象集団に対する分析に用いたデータソースの代表性など、分析手順の中で生じる様々な不確実性を考慮した感度分析やシナリオ分析を行う。パラメータの不確実性に対する代表的な感度分析の手法として、1次元感度分析、確率的感度分析などある。国内ガイドラインでは確率的感度分析を行った際には「使用したパラメータの分布^{vi}についても明らかにするとともに、費用効果平面上の散布図と費用効果受容曲線を提示する」としている。

前述のフレームワークを踏まえると、適切に費用対効果評価を行う際には主に以下の項目に留意する必要があると考える（図 2-4）。次章以降はこのフレームワークに沿って、網羅的ではないが NICE や日本で行われた費用対効果評価で議論となった点等を留意事項として紹介すると共に、その課題解決方法について考えていく。

^{vi} 確率的感度分析はモデルのパラメータに確率分布を当てはめる等により、増分費用や増分効果、ICER の分布を得るための手法である。「使用したパラメータの分布」とはモデルのパラメータに想定した確率分布のことである。

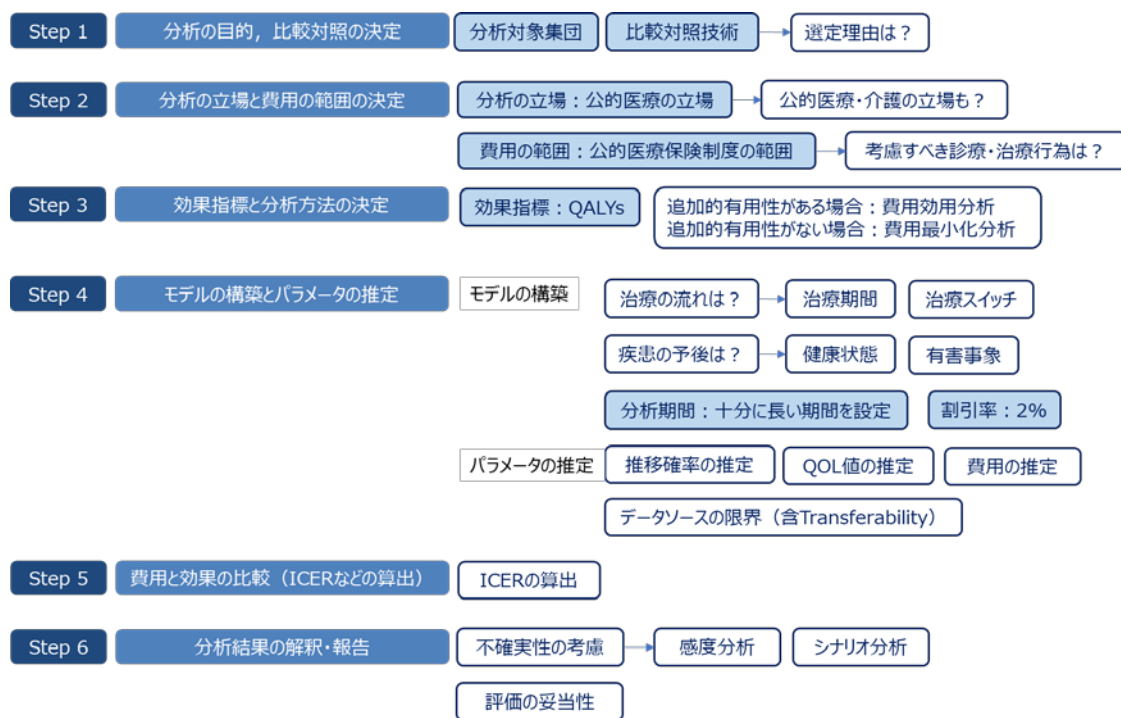


図 2-4 費用対効果評価実施時の主な留意事項

水色部分は、分析データの提出様式の「分析条件の設定の要約」の記載項目である。

各ステップに留意事項があるだろう。

3. 事例から学ぶ留意事項

本章では、前述のフレームワークに沿って、NICE 及び日本で行われた費用対効果評価で議論となった点等を以下の留意事項として紹介し、可能な範囲で解決方法（太字）について考える。

フレームワーク	留意事項
Step 1：分析の目的、比較対照の決定	留意事項①：分析対象集団は、分析前協議で十分議論しておく
	留意事項②：比較対照技術は、分析前協議で十分議論しておく
Step 2：分析の立場と費用の範囲の決定	留意事項③：費用に含める診療・治療行為は、実臨床を想定して設定する
Step 3：効果指標と分析の決定	留意事項④：追加的有用性がない場合、費用最小化分析となる
	留意事項⑤：追加的有用性を適切に判断可能なロジックを構築する
Step 4：モデルの構築とパラメータの推定	留意事項⑥：分析期間は十分に長い期間を設定する
	留意事項⑦：各健康状態の起こりうる推移を精査する
	留意事項⑧：代替エンドポイントと真のエンドポイントの関係性を明確にする
	留意事項⑨：臨床試験の患者集団と実臨床の患者集団に違いを考慮する
	留意事項⑩：疫学研究を用いる場合、未測定交絡因子の存在を考慮する
	留意事項⑪：臨床上許容できない有害事象の発現を考慮する
	留意事項⑫：治療期間を適切に設定する
	留意事項⑬：治療のスイッチを適切に設定する
	留意事項⑭：医療情報データベースを活用する場合、適切にデータベースを選定し、費用や健康状態の定義を適切に設定する
	留意事項⑮：健康状態とQOL値の大小関係の整合性を確認する
	留意事項⑯：QOL値で捕捉できない症状の改善があるかもしれない
	留意事項⑰：健康関連QOLからQOL値へのマッピングは適切に行う
	留意事項⑱：健康状態のQOL値を推定する際、データソースの一貫性を保つ
	留意事項⑲：異なる尺度で推定したQOL値を用いない
	留意事項⑳：有害事象発現によるQOL値の減少を適切に設定する
	留意事項㉑：生存時間のモデリング（外挿も含む）は慎重に行う
	留意事項㉒：間接比較を行う場合、試験間の効果修飾因子の分布の違いを考慮する
	留意事項㉓：特に海外データのQOL値を用いる場合、国民標準値との整合性を確認する
Step 6：分析結果の解釈・報告	留意事項㉔：分析（ex. 確率的感度分析）方法は適切に報告する
	留意事項㉕：不確実性が結果に与える影響を評価し、解釈する
	留意事項㉖：評価の妥当性を説明する

3.1. Step 1：分析の目的，比較対照の決定

前述の通り，分析対象集団と比較対照技術に何を設定するかで ICER の値は大きく異なる．そのため，国内ガイドラインの記載に基づき，分析対象集団と比較対照技術を適切に決める必要がある．日本の試行的導入においても慢性 C 型肝炎，C 型代償性肝硬変を対象とした費用対効果評価において，分析対象集団，比較対照技術に関して企業分析と再分析で設定が異なったことが報告されている．そのため，これらの事例を踏まえて分析の枠組み等について分析前協議を充実させることとなった[1][11]．本項では分析対象集団と比較対照技術が議論となった事例をそれぞれ紹介する．

分析対象集団の設定が論点となった事例として，日本で分析対象集団をサブグループごとに分けて評価を行うか否かで企業分析と専門組織で考え方に乖離が生じた事例がある．

留意事項①：分析対象集団は，分析前協議で十分議論しておく		
事例：慢性閉塞性肺疾患に対するテリルジーの費用対効果評価（C2H 1901）[12][13]		
専門組織での主な検討事項として，企業分析と専門組織の間で分析対象集団の設定に以下の乖離があった ^{vii} ．最終的に12のサブグループに分けて評価が行われた．		
	分析対象集団	設定根拠
企業分析	サブグループに分けず， <u>1つの集団として評価</u>	臨床試験は1つの集団として評価しており，事後的にサブグループごとに評価することでサンプルサイズが極端に小さいサブグループが発生し，分析結果の不確実性が高くなる．
専門組織	<u>12のサブグループに分けて評価</u>	臨床実態等に基づいて分析対象集団を設定することが重要で，極端にサンプルサイズが小さくなった場合の取扱いについては，分析後に専門組織で検討すればよい．

分析対象集団をサブグループに分けることでサブグループあたりの被験者数が少なくなり，結果の不確実性が高くなる．そのため，分析対象集団をサブグループに分けることは追加的有用性の有無を検討する際に影響を与えるだろう．また，トリンテリックスの事例（C2H1905）[14]では，効能・効果が「うつ病・うつ状態」であり，分析対象集団は重症度に応じて「軽度のうつ

^{vii} 企業は分析の枠組み決定の段階であっても分析実行の際の弊害（サブグループによっては分析結果が不確実になる）を勘案してサブグループに関する検討をすべきという立場をとった一方，専門組織は分析の枠組み決定の段階では分析をまだ実行していないので，まずは分析対象集団を設定すること，その結果分析実行の際に弊害が発生した場合は分析後のアプレイザルにより勘案が可能である，という立場をとったと考えられる．

病・うつ状態の患者」と「中等度以上のうつ病・うつ状態の患者」の二つが設定された。しかしながら、「軽度のうつ病・うつ状態の患者」を対象としたランダム化比較試験が存在しなかったため、当該集団における費用対効果は分析不能となり、当該集団の臨床上の意義や患者割合、追加情報の確保の見込み等を踏まえ、当該集団を除外した上で品目全体の評価とする旨の対応が行われた^{viii}[15]。分析対象集団が複数となった場合には、実務担当者は分析のための工数が増えるため分析期間への影響、具体的には遅延も考慮する必要があるだろう^{ix}[15]。70歳未満と70歳以上で別々に分析対象集団を定義したキムリアの事例(C2H1902)[16]では、追加的有用性の評価のために利用した臨床試験のデザイン上、「元々想定されていないサブ集団を構築したことによって追加的有用性を示すことが困難になった場合には、切り分け前の全体集団、この場合であれば年齢の区切りのない全体集団の結果を参照することを容認する旨のコメントを得ている」といった協議が行われていた。以上より、**分析対象集団に関して、臨床的観点からのサブグループの設定と結果の不確実性に対する対処方法は分析前協議でよく議論しておく必要がある。**

続いて、比較対照技術の投与方法の設定が論点となった事例(NICE TA164)[17]である。NICEの評価において企業分析と Appraisal committee（以下、評価委員会）^xで比較対照技術の投与方法の考え方が異なり、費用対効果評価の結果に大きな影響が生じた。

留意事項②：比較対照技術は、分析前協議で十分議論しておく		
事例：痛風の高尿酸血症に対するフェブキシソスタットの費用対効果評価（TA164）[17] 比較対照技術の投与方法に関して、企業分析と評価委員会の間で以下の乖離があった。		
	比較対照技術	設定根拠
企業分析	アロプリノール 300mg/日： <u>固定用量</u>	臨床の意見と市場調査のデータからアロプリノールが英国で最も一般的に使用されており、尿酸降下治療を受けている患者の最大99%が治療を受けていた。
評価委員会	アロプリノール <u>漸増投与：</u> <u>900mg/日まで</u>	多くの GPs が300mg/dayでアロプリノールを処方しているものの、これはベストプラクティスではない。ベストプラクティスの比較対照技術は、最大900mg/dayまでの漸増投与のアロプリノールと結論づけた。

^{viii} 「中等度以上のうつ病・うつ状態の患者」の分析対象集団では、新規抗うつ薬の間で効果が同等であるとして費用最小化分析を実施された。

^{ix} 現状の運用においては、分析結果の提出の遅れには当局は厳しい態度で臨んでおり、提出が期限に間に合わなかった場合、企業はその理由について説明を行い、当局の納得を得る必要がある。

^x 英国では日本の専門組織に当たる組織として NICE における Appraisal committee がある。Appraisal committee は、企業分析や ERG による分析を踏まえて当該品目の費用対効果評価を取りまとめ、その結果に基づき意思決定に係るガイダンスを出している。

評価委員会は、企業分析が算出した固定用量のアロプリノールに対するフェブキソスタットのICERは高用量のアロプリノールの有効性を比較対照技術に考慮していなかったため、かなり過小評価されていたとした。

国内ガイドラインでは比較対照技術は「評価対象技術が分析対象集団への治療として導入された時点で、臨床的に幅広く使用されており、評価対象技術によって代替されると想定されるもののうち、治療効果がより高いものを一つ選定する(4.1)」,「4.1」において、一意に決めることが難しい場合は、無作為化比較試験等における比較対照技術、価格算定上の類似技術、費用対効果の程度等も考慮して最も妥当と考えられる比較対照技術を両者の協議により選定する。」とされている。そのため、治療実態を踏まえつつ、治療効果がより高い治療方法を考慮する必要があるだろう。なお、国内のテリルギーの事例(C2H1901) [13] では専門組織が「ガイドラインの趣旨に鑑み、最も安価で費用対効果のよいものを比較対照技術とすることが妥当」と判断し、比較対照技術が選定された[18]。分析データの提出様式では「比較対照を選定した理由」の明記が求められており、適切に比較対照技術を選定することは重要なステップとなる。

以上のような分析対象集団及び比較対照技術に関する留意事項に対して、第1回分析前協議票[19]に「分析対象集団」、「比較対照技術名」、「比較対照技術を選定した理由」の記載が肝となるだろう。比較対照技術に関しては想定される分析対象集団ごとに治療実態(“臨床的に幅広く使用されているもの”, “評価対象技術によって代替されると想定されるもの”, “治療効果がより高いもの”)に関する適切なエビデンスを開発段階から計画的に収集しておく必要がある。具体的には、診療ガイドライン、既存論文の調査はもちろん、必要に応じて間接比較によるエビデンス構築、医療情報データベースを活用した治療実態調査等が有用だと考える。なお、この段階で治療実態や疾病負担を調査することはメディカルニーズの根拠情報にもなるため、これらの情報はCTD(Common Technical Document)を作成する上でも重要な情報となるだろう。

3.2. Step 2: 分析の立場と費用の範囲の決定

分析の立場は国内ガイドラインに基づいて、基本的に「公的医療の立場」、費用の範囲は「公的医療保険制度の範囲」を設定することになる。その上でどのような診療・治療行為の費用が発生するかがポイントとなる。例えば、NICE TA605 [20]では、注射剤である評価対象技術の注射部位を特定するための超音波ガイダンスの使用頻度について議論された。

留意事項③：費用に含める診療・治療行為は、実臨床を想定して設定する

事例：慢性流涎症に対するゼオミン(A型ボツリヌス毒素製剤)の費用対効果評価(TA605) [20]

ゼオミン投与時に正しい注射部位への注射を誘導するために超音波画像が使用されることがある。超音波ガイダンスを用いない場合には解剖学的ランドマークが使用される。注射部位の特定方法に関して、企業分析と評価委員会の間で以下の乖離があった。

	注射部位特定	設定根拠
企業分析	<u>超音波画像を使用</u>	SIAXI試験では56%が超音波ガイダンスを用いており、この頻度をNHSの臨床診療と同様であると想定した。
評価委員会	<u>超音波画像の使用</u> <u>頻度は低い</u>	臨床専門医に確認したところ、超音波ガイダンスは施設によって使用頻度が異なるものの、英国のNHSの臨床診療ではめったに使用しない。

評価委員会は、実臨床での超音波ガイダンスの使用頻度はSIAXI試験よりも低いと考えた。そのため、企業分析は超音波ガイダンスのコストを過大評価した可能性があり、結果的にICERの推定値が小さくなるだろうと考えた。

本事例では、臨床試験という特殊な環境における追加的な超音波画像取得を費用に含めたことで、実臨床における分析を求める評価委員会との解釈に差異を生んだと言える。**分析に含める診療・治療行為は、厳密な管理下で介入を行う臨床試験だけでなく、治療実態調査や臨床医の意見を踏まえて、実臨床に即して設定する必要がある。**また、介護者等の費用を分析で考慮したい場合、国内ガイドラインでも「公的医療・介護の立場」の分析を行ってもよい」とされている。そのため、2020年10月よりレセプト情報・特定検診等情報データベース・介護データベースの連結解析が開始された点を考慮すると、今後、C2Hと「医療介護レセプト連結データベース」を活用した治療実態調査も想定されるだろう^{xi}。費用を具体的にどのように推定するかはStep 4で触れる。

3.3. Step 3：効果指標と分析の決定

国内ガイドラインでは評価対象技術が追加的有用性を有すると判断できる場合には、効果指標はQALYを用いることを原則とし、QALYを用いた費用効用分析を行い、ICERを算出する^{xii}。一方で、評価対象技術が追加的有用性を有すると判断できない場合には比較対照技術と費

^{xi} NICEでは費用の範囲を「公的な医療費及び福祉費（National Health Serviceの医療費及びPersonal Social Servicesの福祉費）」としており、介護費も費用に基本的に含めることが可能である。一方で、日本では「公的医療・介護の立場」は追加的な位置づけであり、価格調整の検討に分析結果が考慮されない点に留意する必要がある。

^{xii} ただし、評価対象技術が比較対照技術に比べて効果が同等以上(増分効果の大きさが非負)で、かつ費用が安い場合はICERを算出せずに優位(dominant)であるとする。

用のみを比較する費用最小化分析を行う。NICE でも実際に費用最小化分析が行われた事例（NICE TA497）[21]があるため、以下に紹介する。なお、国内でも同様に費用最小化分析が行われた事例が存在する[2]。

留意事項④：追加的有用性がない場合、費用最小化分析となる

事例：非放射線学的軸性脊椎関節炎に対するゴリムマブの費用対効果評価（TA497）[21]

企業分析のネットワークメタアナリシスの結果を踏まえて、評価委員会はゴリムマブの臨床的有效性と有害事象のプロファイルが比較対照技術（アダリムマブ、エタネルセプト、セルトリズマブペゴル）と同等と判断した。

企業分析では、ゴリムマブの費用は治療の1年目以降のアダリムマブの費用と同じ 9,156ポンドであり、エタネルセプトの費用（9,295ポンド）よりも低いことが示された。セルトリズマブペゴルは最初の10バイアルを無料で提供するPatient Access Schemeのため、1年目の費用（5,720ポンド）はゴリムマブの費用よりも低くなるが、その後の費用（9,295ポンド）はゴリムマブよりも高かった。また、ゴリムマブと比較対照技術で治療開始、管理、モニタリングの費用に違いはなかったため、取得費用のみの比較が行われた。

以上の結果を踏まえ、評価委員会はゴリムマブを推奨した。

	費用（£）1年目	2年目以降
ゴリムマブ	9,156	9,156
アダリムマブ	9,156	9,156
エタネルセプト	9,295	9,295
セルトリズマブペゴル	5,720	9,295

前述の通り、追加的有用性の有無に基づいて費用効用分析と費用最小化分析のいずれを行うかが決定される。追加的有用性の評価方法の詳細は国内ガイドラインを参照されたいが、ランダム化比較試験（RCT）のシステマティックレビューだけでなく、協議次第では非 RCT（観察研究等）のシステマティックレビュー、既存の観察研究やレジストリデータなどを再解析した結果、間接比較、単群試験を用いることも可能となっている。そして、分析データの提出様式では、当然のことだが追加的有用性があると判断する場合、その理由を明記することとなっている（表 3-1）。また、分析データの提出様式では判断の根拠となったデータとして「RCT のメタアナリシス」、「単一の RCT」、「前向きと比較観察研究」、「RCT の間接比較」、「単群試験の比較」、「臨床データなし」が選択肢として用意されているため、実務上はこれらのうちいずれかを選択し、追加的有用性ありと判断した理由を説明していくことになる。

表 3-1 追加的有用性の有無に関する評価のまとめ方

対象集団	
介入	
比較対照	
アウトカム	
追加的有用性の有無	☐ 追加的有用性あり ☐ 「追加的有用性なし」あるいは「ありとは判断できない」
判断の根拠となったデータ	☐ RCTのメタアナリシス ☐ 単一のRCT ☐ 前向き比較観察研究 ☐ RCTの間接比較 ☐ 単群試験の比較 ☐ 臨床データなし
追加的有用性の有無を判断した理由	

【出典 医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引[4]】

日本の本格導入において、追加的有用性の判断が企業分析と公的分析で異なった事例が存在する[12][13][22]。また、企業分析と公的分析で同じ追加的有用性の判断結果であっても、判断理由が異なるケースも存在する[23]。

留意事項⑤：追加的有用性を適切に判断可能なロジックを構築する		
事例：発作性夜間ヘモグロビン尿症に対するユルトミリスの費用対効果評価（C2H 1903）[22]		
追加的有用性の評価結果が企業分析と公的分析で異なった。		
	追加的有用性	評価内容
企業分析	<u>あり</u>	「遊離C5関連ブレークスルー溶血（BTH；Breakthrough hemolysis）発生率の減少」「輸血回避」「投与間隔の延長によるQOLの向上」において、ラブリズマブがエクリズマブに対して追加的有用性を有することを主張した。
専門組織	<u>なし</u>	遊離C5関連 BTH：301試験と302試験の結果によると、遊離C5関連を含むBTHの発生率において、ラブリズマブはエクリズマブと比べて統計学的な有意差を示していない。遊離C5関連BTHは、実診療ではほとんど使用されないため、追加的有用性の評価指標としての適切性が明らかではない。ラブリズマブの投与にあたっては患者の体重に応じて用量調整が行われる一方で、エクリズマブによる治療では、患者の体重によらず固定用量が投与される。したがって、高体重の患者では投与量が不足することにより、エクリズマブ群において遊離C5関連BTHが多く発生した可能性が否定できない。

		輸血回避：301試験と302試験の結果によると、輸血の発生率において、ラブリズマブはエクリズマブと比べて統計学的な有意差を示していない。 投与間隔の延長によるQOLの向上：離散選択実験に基づく調査結果は、患者ではなく、一般住民の選好を調査したものであり、追加的有用性を評価する際のアウトカム指標として適切ではない。 上記の理由からラブリズマブはエクリズマブに対して追加的有用性がないと評価した。
--	--	--

追加的有用性の判断が異なった場合、その後に行われる費用対効果評価は費用効用分析ではなく費用最小化分析となる（又はその逆もありえる）ため、追加的有用性の評価結果は企業分析に大きな影響を与えることになるだろう。RCTのメタアナリシスで臨床的及び統計的に有意な差が示されたり、優越性を検証した単一のRCTがあれば特に問題は生じないかもしれないが、有効性に関して非劣性を検証した臨床試験で承認取得した場合、代替エンドポイントで検証試験を行った場合、統計的検出力が確保されていない評価項目で追加的有用性を説明したい場合、分析対象集団でサブグループ解析が必要となった場合などは追加的有用性を適切に判断可能なロジックを構築し、可能な限り分析前協議で相談しておくことが望まれる。以上を踏まえて、追加的有用性の検討には十分な時間とリソースを確保しておく必要があると考える。

3.4. Step 4：モデルの構築とパラメータの推定

3.4.1. モデルの構築

疾患推移モデルは、疾患の予後と治療の流れを適切に表現する必要がある。具体的には、いくつかの健康状態を定義し、健康状態間の推移を考え、さらに治療を割り当てることになる（図3-1）。疾患の予後と治療の流れに関しては、企業分析と評価委員会で考え方に乖離があった事例がNICE TAでは散見される。本項ではモデル構築の過程で分析期間、健康状態の可逆性、健康状態の設定方法、安全性の考慮方法、医療技術の治療期間・治療内容に関して議論があった事例を紹介する。

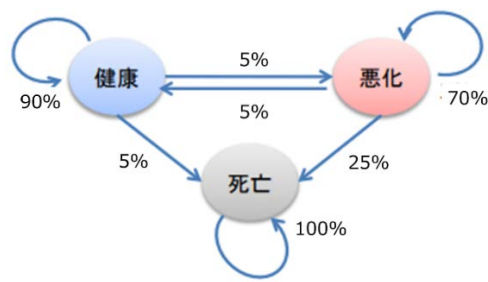


図 3-1 疾患推移モデル：マルコフモデルの場合
モデル中の数値は仮想の値である。

● 分析期間について

分析期間はモデル構築時の重要な要素である。国内ガイドラインでは「評価対象技術の費用や効果におよぼす影響を評価するのに十分に長い分析期間を用いる」としており、基本的には長期間にわたって分析を行うことになるが、NICE TA164 [17]では分析期間の設定が議論となった。

留意事項⑥：分析期間は十分に長い期間を設定する

事例：痛風の高尿酸血症に対するフェブキソスタットの費用対効果評価（TA164）[17]
分析期間の設定に関して、企業分析と評価委員会の間で以下の乖離が生じた。

	分析期間	設定根拠
企業分析	<u>2年間</u>	1年間を評価した臨床試験データと外挿による更なる1年間に基づいて設定した。なお、感度分析では5年間を解析した。
評価委員会	<u>生涯</u>	生涯にわたる治療を必要とする慢性疾患では生涯にわたる期間が適切であると考えた。

短期間での効果のみが問題となる医療技術もあるが、一般的には分析期間は十分に長い期間を設定することになる。分析期間を長く設定することにより、臨床試験の評価期間を超えた期間における治療効果等を推定（外挿）する必要があるが生じる。そのため、長期間における評価は不確実性を生じることにより留意が必要である^{xiii}。分析期間以外に割引率もモデル構築時の重要な要素だが、国内ガイドラインでは「年率 2%」、感度分析では「年率 0%から 4%の範囲で変化させる」と明記されていることから、割引率が論点となることはないだろう^{xiv}。参考までに NICE では

^{xiii} 外挿による不確実性が生じたとしても分析期間は十分に長い期間を設定する必要がある。そのため、外挿による不確実性は感度分析で影響を評価することになる。なお、分析データの提出様式では分析期間が生涯より短い場合は設定した分析期間で十分とした理由を記載することになっている。

^{xiv} 割引率の設定方法として、民間資本の生産性を反映した市場利子率を割引率に用いることが望ましく、日本では市場利子率の近似として国債等の利子率を用いる考え方がある。

Technology Appraisals の methods guide で割引率は 3.5%と設定している [24].

なお、ここまで検討してきた内容を整理することで分析データの提出様式で定義されている「分析条件の設定の要約」の項目（表 3-2）が整理できたこととなる。

表 3-2 「分析条件の設定の要約」の項目

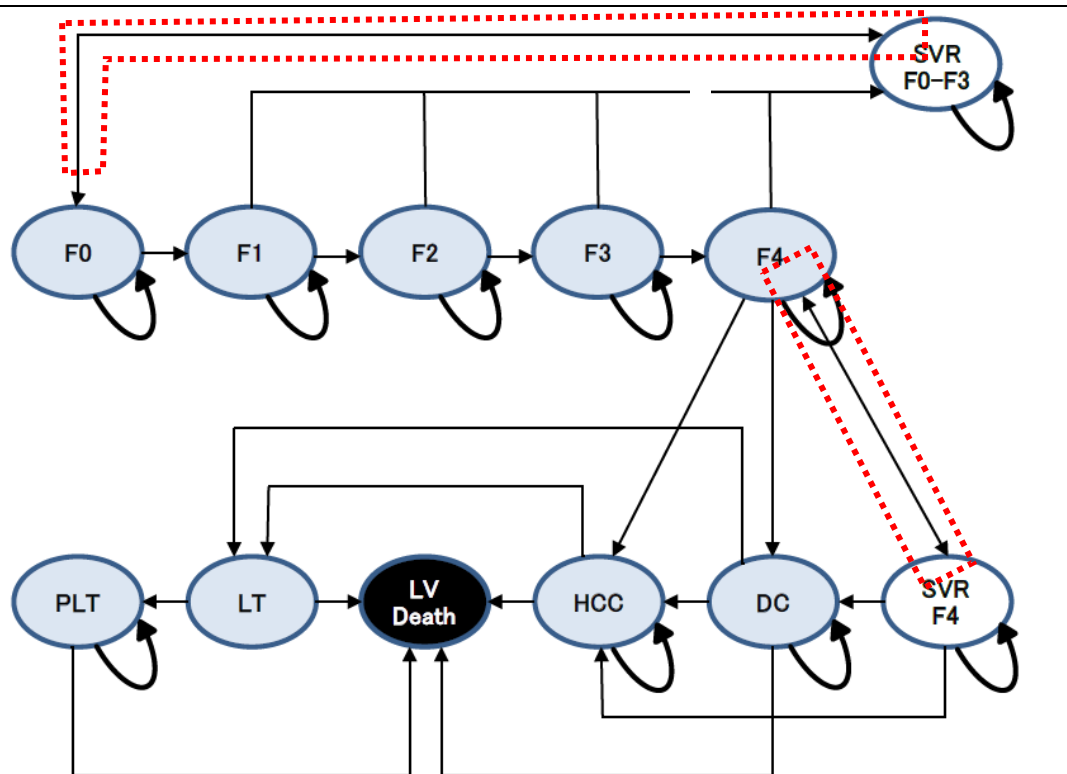
		(例)
分析対象とする集団		疾患 A をもち、標準療法(治療 B)が無効であった患者
比較対照		無治療
比較対照を選定した理由		
分析の立場と費用の範囲		公的医療の立場 公的医療費のみ
効果指標		QALY、生存年
分析期間		生涯
割引率		費用・効果ともに年率 2%

【出典 医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引 [4]】

- 健康状態の可逆性について

続いて、健康状態の可逆性（疾患の予後）が議論となった事例（NICE TA413）[25]である。

留意事項⑦：各健康状態の起こりうる推移を精査する
事例：C型慢性肝炎に対するエルバスビル/グラゾプレビル[®]の費用対効果評価（TA413）[25] 企業分析では以下の疾患推移モデルを構築した。これまでに行われてきたいくつかのC型肝炎の評価とは異なり、SVR（持続陰性化）を達成した後の再感染を可能にした点（以下の図で赤点線部分）は評価委員会から適切なアプローチと評価されたが、再感染した患者がさらに健康状態F0（線維化なし）に戻ることができる点に評価委員会は懸念を示した。これはC型肝炎によって引き起こされた肝障害が可逆的であることを想定しており、評価委員会はこの仮定は適切ではないとした。



F0：線維化なし，F1：門脈域の線維性拡大，F2：線維性架橋形成，F3：小葉構造のひずみを伴った線維性架橋形成，F4：代償性肝硬変，DC：非代償性肝硬変，HCC：肝細胞癌，LT：肝移植（一年目），PLT：肝移植後，LV Death：肝臓関連死，SVR：持続陰性化，逆向きの矢印は再感染を示す（SVR, F0-F3 からの再感染は F0 へ，SVR, F4 からは F4 へ移行）

【出典：費用対効果評価の実践[9]】

ERGの疾患推移モデルを確認できなかったため，企業分析の疾患推移モデルを引用し，作成。

一方，ERGはSVRを達成した後に再感染した患者はSVR前の線維化の健康状態に戻ること
を想定したモデルを構築した。臨床専門家はERGの想定が合理的で，臨床診療をよりよく反
映しているとし，評価委員会は再感染後の想定はERGのモデルがより合理的であると結論し
た。

この事例から学ぶべきことは**各健康状態の推移は実際に起こりうるかを臨床的に精査する必要**
があるということである。

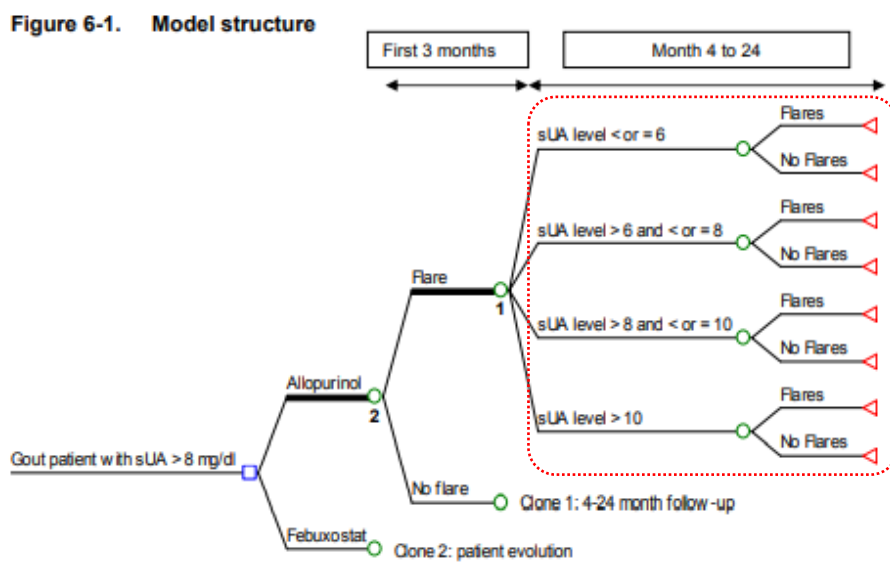
- 代替エンドポイントを利用した健康状態の設定方法について

疾患推移モデルは健康状態を結ぶことで疾患の予後を表現していくが，NICE TA164 [17]で
は健康状態として代替エンドポイントを設定することの妥当性が議論となった。具体的には，真
のエンドポイントと代替エンドポイントの関連を評価する統計解析の方法に関してERGから疑
義が挙がり，代替エンドポイントを用いた健康状態の設定方法が議論となった。

留意事項⑧-1：代替エンドポイントと真のエンドポイントの関係性を明確にする

事例：痛風の高尿酸血症に対するフェブキシostatの費用対効果評価（TA164）[17]

企業分析では、血清尿酸値は痛風関節炎の発現及び痛風結節のサイズ縮小の代替エンドポイントの位置づけであり、疾患推移モデルとして、血清尿酸値を用いた以下の決定樹モデルを採用した。



【出典：NICE TA164 [17]】

sUA：血清尿酸値，Flare：痛風関節炎

ERGは、血清尿酸値と痛風関節炎の発現の関係に大きな不確実性があると指摘した。具体的には、企業分析で行われた痛風関節炎の発現に対する血清尿酸値のモデル解析について以下の点を指摘した。

- 4ヶ月後から24ヶ月後（図中の「Month 4 to 24」）において、血清尿酸値の値に応じて痛風関節炎の発現割合が異なることを想定してモデルが構築されていた（図中の赤枠点線）。これは企業が企業分析の多変量解析の結果、痛風関節炎の発現に血清尿酸値は関連すると想定したことに基づくが、他の重要な共変量を考慮した異なる多変量解析の結果では痛風関節炎の発現と血清尿酸値の間に統計的有意差が認められなかった。しかしながら、企業分析ではこの多変量解析を用いていなかった。また、backward stepwise analysisによる変数選択を行った多変量解析を提示する必要がある^{xv}。
- 企業分析のモデル解析では、血清尿酸値を4つのカテゴリー [6mg/dL以下， 6mg/dL超

^{xv} NICE の資料上では「The ERG argued that a proper multivariate analysis, in which a backward stepwise analysis is carried out, should be presented.」となっていたが、変数選択したモデルによる調整した結果を用いるべき一例という理解でいいだろう。

8mg/dL以下, 8mg/dL超10mg/dL以下, 10mg/dL超] に分けて, 痛風関節炎の発現と血清尿酸値の4つのカテゴリーの間に線形性を仮定していたが, 線形性を仮定する必要はない^{xvi}

ERGの指摘を踏まえ, 評価委員会でも代替エンドポイントである血清尿酸値と真のエンドポイントである痛風関節炎の発現の関連性について議論が行われた。その結果, 血清尿酸値と臨床上のベネフィットである痛風関節炎の発現抑制, 腎機能障害の発現抑制, 痛風結節のサイズ及び数の減少との間の関係はいくつかの不確実性が残っていることに留意が必要だが, 血清尿酸値が約6mg/dLを超えると痛風関節炎の発現の可能性が高いと仮定することは合理的であると結論づけた^{xvii}。

上記の事例は, 代替エンドポイントである血清尿酸値を主要評価項目に設定した臨床試験がデータソースとなっていることが主な原因だが, 対面助言, 承認申請時でも説明が求められるように代替エンドポイントの妥当性を明確に示す必要があるだろう。なお, 同じ疾患領域の異なる薬剤の費用対効果評価 (NICE TA506 [26]) でも代替エンドポイントである血清尿酸値と死亡の関連について議論が生じていた。

留意事項⑧-2: 代替エンドポイントと真のエンドポイントの関係性を明確にする

事例: 痛風の高尿酸血症に対するレシヌラドの費用対効果評価 (TA506) [26]

代替エンドポイントである血清尿酸値と死亡の関連に関して, 企業分析と評価委員会の間で以下の乖離が生じた。

	血清尿酸値と死亡の関連	設定根拠
企業分析	<u>あり</u>	6つの観察研究のメタアナリシスの結果を提示し, 高尿酸血症又は痛風患者に対して, 尿酸降下治療と死亡リスクの低下が関連していることを示した。そして, 尿酸降下治療により血清尿酸値を低下させた場合の死亡リスクの低下をモデル化するためにStack(2013) [27]の研究を使用し, レシヌラドによる治療は高い血清尿酸値に起因する死亡リスクを34% 減少させると想定した。

^{xvi} つまり, モデルに強い仮定を置くべきではないと ERG はコメントした。

^{xvii} 検証試験では代替エンドポイントを主要評価項目として承認を取得している場合でも, 費用対効果評価の段階で代替エンドポイントの妥当性が再度議論されるケースもある。

		※臨床試験の主要評価項目は血清尿酸値の目標値達成割合であり、死亡リスクはメタアナリシスの検討結果である。
評価委員会	<u>なし</u>	メタアナリシスは疫学研究に基づいており、レシヌラド以外の尿酸降下治療でも血清尿酸値の低下と平均余命の関係を検証したランダム化比較試験が提示されていないとし、尿酸降下治療が痛風患者の余命を伸ばすことを証明するために観察研究を使用することに懸念を示した。評価委員会は既知及び未知の交絡因子により観察研究の結果は交絡するリスクがあり、特にStackら（2013）[27]では潜在的な交絡因子の中でも腎機能低下が考慮されていなかったことに留意した。腎機能低下は血清尿酸値と高度に関連があり、腎機能低下は死亡リスクを高めることが知られている。

費用対効果評価に必要な効果は真のエンドポイントに対する効果であり、一方で承認取得のために必要な効果は代替エンドポイントに対する効果が許容されるケースがある。これらの事例は各評価に求められる効果の違いから生じたエビデンスのギャップによるものといえるだろう。同様のエビデンスギャップは様々な NICE TA で言及されており、抗がん剤領域における全生存時間、糖尿病領域における心血管系のリスクに関する事例も併せて以下に紹介する。

- NICE TA214 [28]：転移性乳がん患者に対するペバシズマブとタキサン併用の評価
ERG は臨床的有用性の評価は主として E2100 試験 1 試験に基づいていることに限界があると考えた。具体的には、E2100 試験が非盲検であること、疾患進行後の治療に関するデータ収集が実施されていないことを挙げた。また、E2100 試験では群間で全生存期間に統計的有意差が認められなかった。この理由が治療のクロスオーバーによるものかどうかは、治療のクロスオーバーに関するデータが収集されていなかったため、判断できないとした。
- ☆ がん領域において、癌腫や治療ラインによっては承認取得前の臨床試験で全生存期間ではなく奏効率や無増悪生存期間に関して有効性の検証を行い、承認申請するケースが増えている[29][30]。
- NICE TA315 [31]：2 型糖尿病に対するカナグリフロジンの併用療法の評価
評価委員会はカナグリフロジンの短期的な有害事象は管理可能だが、特に心血管系のイベントについては、長期的な結果に関する更なるデータが必要であると結論づけた。
- ☆ 承認取得前の臨床試験で心血管系の有害事象の詳細なデータを取得することは困難だろう。FDA は 2 型糖尿病治療薬の心血管系合併症の発症リスクの評価に関し

て、既往歴のある患者での臨床試験を行うことを提案している[32].

また、日本で臨床試験と実臨床の対象集団の違いに関して議論となった事例がある。以下の事例（C2H 1902 [16]）は臨床試験の被験者集団と実際の患者集団で平均年齢が異なる点が議論となった。

留意事項⑨：臨床試験の患者集団と実臨床の患者集団に違いを考慮する		
事例：びまん性大細胞型B細胞リンパ腫に対するキムリアの費用対効果評価（C2H 1902）[16] 分析対象集団の平均年齢に関して、企業分析と公的分析の間に以下の乖離があった。		
	平均年齢	設定根拠
企業分析	非公開	臨床試験の被験者背景に基づいて設定した。
公的分析	57歳	臨床試験の被験者集団の年齢は実際の患者集団に比べて低いとし、レセプト情報・特定健診等情報データベース（National Database of Health Insurance Claims and Specific Health Checkups of Japan；NDB）により推計された患者年齢が当該分析対象集団における平均年齢としてより適切であるとした。
公的分析は上記の設定根拠に基づき、治療開始年齢を57歳として再分析を実施した。		

日本で費用対効果評価が行われたコララン（C2H 1906[33]）でも同様に患者集団の平均年齢と投与量に関して、公的分析は臨床試験の結果と実臨床で想定される値に乖離があると指摘し、NDBを用いた解析結果を提示した。

このような臨床試験と実臨床のギャップに関する議論は、承認申請時にも規制当局から照会事項を受ける場合があり、臨床試験の限界といえるだろう。つまり、**費用対効果評価で臨床試験の結果を用いる際には以下の点に留意する必要があるだろう。**

◇ 臨床試験の結果を費用対効果評価で用いる際には不確実性や限界がある

- ex. 臨床試験で代替エンドポイントを主要評価項目に設定して検証試験を実施した場合、真のエンドポイントについて効果は検証されず、真のエンドポイントに関するエビデンスには不確実性がある
- ex. 費用対効果評価の比較対照群が臨床試験の比較対照群に設定されないケースや単群試験で臨床試験が行われる場合もある
- ex. 統計的有意差が議論できるケースは適切に症例数設計を行った評価項目の結果に対してのみであり、多くの評価項目は検証目的で設定されていないため本来統計的有意性に基づく議論はできない

☆ 臨床試験の結果は実臨床における有用性（effectiveness^{xviii}）を示しきれていない

- ex. 臨床試験は一般的には日常臨床とは異なる実験的な条件下で、かつ参加意欲のある患者に対して行われることが多いため、結果の外的妥当性（一般化可能性、代表性）が不十分となる懸念がある
- ex. 臨床試験の結果は臨床効果を過大評価する可能性がある；臨床試験と実臨床ではアドヒアランスが異なることがあり、臨床試験は実臨床に比べてアドヒアランスが良いことが多い

☆ 臨床試験から得られる安全性情報は限られている

- ex. 5 toos の問題^{xix}があり、臨床試験における安全性評価には限界がある

上記の臨床試験の限界は費用対効果評価の結果に大きな影響を与えることとなり、NICE の非推奨につながる臨床データの欠陥として報告されている[34]。続いて、疫学データを用いて構築したエビデンスが実際に費用対効果評価で検討された NICE TA279[35]を紹介する。

留意事項⑩：疫学研究を用いる場合、未測定交絡因子の存在を考慮する

事例：骨粗鬆症脊椎圧迫骨折に対する経皮的椎体形成術及び経皮的後弯矯正術の費用対効果評価（TA279）[35]

臨床試験では死亡に関して統計的有意差が認められなかったが、観察研究の結果も踏まえて死亡に対する効果を検討した。

	死亡に対する効果
企業分析	米国のMedicareデータを用いて大規模な観察研究を行い、年齢、性別、人種、チャールソン併存疾患指数、及びその他の併存疾患で調整した結果、最適な疼痛管理治療に対して経皮的椎体形成術ではハザード比が0.76（95%信頼区間：0.75, 0.77）、経皮的後弯矯正術では0.56（95%信頼区間：0.55, 0.57）であった。
評価委員会	企業分析では複数の併存疾患を調整したが、交絡の可能性が残っている。ただし、 <u>さらなる交絡を考慮することで効果は小さくなる可能性があるが、ベネフィットの大きさを考えると効果がなくなる可能性は低いと考え</u>

^{xviii} Efficacy は理想的な条件下における治療効果であり、effectiveness は現実的な条件下（実臨床）における有用性である。費用対効果評価では「effectiveness」を用いる。

^{xix} too few：症例数が少ない、too narrow：腎機能・肝機能障害、妊婦などの特殊な患者は除外されている、too median-aged：高齢者や小児は除外されている、too simple：投与方法が単純で併用薬などが使われていない、too brief：投与期間が短く、長期投与の結果が不明である（Rogers（1987）による定義）

	た。また、生物学的妥当性の観点からも議論し、椎体の高さ及び脊椎の湾曲を改善することで肺機能、消化及び可動性が改善され、その結果、死亡に対するベネフィットが得られると考えた。
評価委員会は経皮的椎体形成術と経皮的後弯矯正術の両方が最適な疼痛管理治療と比較して生存期間を延ばすことと仮定することは合理的であるが、死亡に対する効果の正確な大きさは不明であると結論した。	

近年、一般的な患者集団における医療技術の有用性や安全性、使用実態や経済効果を検討するデータソースとしてリアルワールドデータが注目されているが、**疫学研究を行う際には疫学研究特有のバイアス、データの質、限界等に留意してエビデンスを構築する必要がある**[36][37]。NICE でも疫学データからのエビデンス構築という観点で TSD17^{xx} [38]が公開されており、ランダム化が行われていない観察研究から因果効果を推定する際の解析手法（傾向スコアを用いた手法、操作変数法及び回帰不連続デザインなど）が解説されている。製薬協成果物でもその概要をまとめた資料を公開しているため、疫学研究の結果を用いる際には参照されたい[39]。

● 安全性の考慮方法について

続いて、安全性をどのように疾患推移モデルで考慮するかである。そもそも各健康状態は有効性に関する疾患の予後を示したものが多だろう。そのため、安全性をどのようにモデルに組み込むかは重要なポイントとなる。NICE TA605 [20]は企業分析で有害事象の発現を考慮しなかったことの妥当性が議論された事例である。

留意事項⑪：臨床上許容できない有害事象の発現を考慮する		
事例：慢性流涎症に対するゼオミン（A型ボツリヌス毒素製剤）の費用対効果評価（TA605）[20] 有害事象の発現に関して、企業分析と評価委員会の間で以下の議論があった。		
	有害事象の発現	設定根拠
企業分析	<u>考慮せず</u>	有害事象の発現を比較するために利用できるデータが十分ではなかったため、評価対象技術及び比較対照技術のいずれにも有害事象の発現を想定しなかった。

^{xx} NICE は医療経済評価に関連する技術的なガイド（TSD：Technical Support Document）を作成しており、これまでに 21 のガイドが公開されている（2023 年 1 月 8 日時点）。

評価委員会	本来考慮すべきだが、企業分析の見解を受け入れた	比較対照技術に含まれる抗コリン作用薬の重篤な副作用の負荷を考慮すると、有害事象の発現をモデルに含めていない企業分析は保守的な結果になっていると考えた。企業分析では経済モデルに有害事象を含めるべきだったと考えるが、企業のモデルの構造は意思決定上受け入れられると評価委員会は結論づけた。
-------	-------------------------	---

上記の事例は、比較対照技術に比べて評価対象技術の安全性リスクの懸念が少なかったため、評価委員会が有害事象の発現を考慮しないことを許容したが、**基本的には臨床上許容できない有害事象の発現は考慮する必要があるだろう**。国内ガイドラインでも「評価対象技術や比較対照技術の費用のみでなく、有害事象や将来の関連する合併症等の費用も含めて推計する」とされており、分析データの提出様式では「対象となる医薬品・医療機器の性質」において「主な有害事象」を記載する必要がある。また、**評価対象技術と比較対照技術の間で主な有害事象の発現頻度が異なるのであれば、費用だけでなく QOL 値への反映も忘れずに考慮する必要がある**。具体的には、有害事象の発現頻度と発現時の QOL 値の減少をモデルに反映させる必要がある。有害事象発現による QOL 値の減少を考慮した事例については後述する。

- 医療技術の治療期間・治療内容について

疾患推移モデルを考える際には、評価対象技術及び比較対照技術の治療の流れとして、治療期間や評価対象技術の治療終了後の治療内容も設定する必要がある。NICE TA423 [40]は治療期間の設定が論点となった事例である。

留意事項⑫：治療期間を適切に設定する		
事例：2つ以上の化学療法後の局所進行性乳がん/転移性乳がん患者に対するエリブリンの費用対効果評価（TA423）[40]		
治療期間に関して、企業分析とERGの間で以下の乖離があった。		
	治療期間	設定根拠
企業分析	最大6ヶ月	患者が治療を受けることができる期間の上限を最大6ヶ月と固定した。企業はこの根拠として、既存文献と臨床試験の結果を用いて説明した。
ERG	最大期間を設定すべきではない	治療費用を過小評価する可能性があるため、治療の最大期間に上限を設定することは妥当ではないとした。
評価委員会は臨床専門家から一部の患者は化学療法にsensitiveな病態を有しており、そのよ		

うな場合は6ヶ月を超えてエリブリンを継続する可能性があり、これらの患者はその後の治療ラインでもよく奏効する可能性があることを確認した。一方、その他の患者はおそらく化学療法にnegativeな病態であり、急速に疾患が進行するかもしれず、このような患者には6ヶ月以降の治療を行わない可能性があることを確認した。

以上を踏まえ、評価委員会は6ヶ月を超えて治療を継続するというERGの考え方に同意したが、6ヶ月後も治療を継続する可能性のある患者の割合及びその後の治療期間については大きな不確実性があると考えた。そして、この不確実性を解決することは不可能とし、企業分析の仮定は楽観的かもしれないが、一方で、ERGの仮定は治療費用を最大に評価する最も悲観的なシナリオであるとした。

この事例では評価委員会は治療期間の設定に関して不確実性があると結論づけたが、**治療期間の設定に不確実性がある場合はシナリオ分析で検討する必要があるだろう。**

日本の試行的導入においても、オプジーボ、キイトルーダの費用対効果評価において、企業分析と再分析で治療期間の設定が異なり、両分析で異なる結果が得られたと報告されている[1]。このように治療期間をどのように設定するかが議論となった事例は多く存在し、**治療期間の設定根拠を説明できるように準備しておく必要があるだろう。**

上記の NICE TA423 [40]ではさらに治療スイッチの設定についても議論があった。

留意事項⑬：治療のスイッチを適切に設定する

事例：2つ以上の化学療法後の局所進行性乳がん/転移性乳がん患者に対するエリブリンの費用対効果評価（TA423）[40]

治療のスイッチに関して、企業分析とERGにおいて異なる仮定を設定した。

	治療のスイッチの仮定
企業分析	疾患進行時に全患者がエリブリンから治療を切り替えると仮定し、 <u>PFS^{xxi}データを用いて治療スイッチのタイミングを考慮した。</u>
ERG	PFSデータではなく <u>治療中止までの時間のデータを用いて</u> 、疾患進行後もエリブリンの治療は可能と仮定し、さらに中止した患者の60%は二次治療に切り替えると想定した。

費用対効果評価の実施時に上市して間もない評価対象技術の治療実態を予測することは限界があるものの、臨床的知見を踏まえて尤もらしい仮定を設定したり、必要に応じてシナリオ分析

^{xxi} PFS とは progression free survival（無増悪生存期間）であり、疾患進行又は死亡までの時間である。

で不確実性を検討しておくことが重要だろう。

以上のように、疾患推移モデルを構築するには、疾患の予後及び治療の流れが適切に表現できるように留意する必要がある。同様の分析対象集団における費用対効果評価の先行研究や NICE のガイダンスが存在する場合には、論点となった部分も含めて参考にしながらモデルを構築するといいたいだろう。一方、NICE では特に議論にならなかったが、日本における公的分析、総合的評価の過程で日本独自に疾患推移モデルについて新たな指摘が挙がる可能性もあるだろう。今後、このような事例については積極的な情報共有が望まれるところである。なお、既存研究によるモデルがない疾患領域で新規に疾患推移モデルを構築する際は NICE TSD13 [41] が参考になり、製薬協成果物でも概要をまとめた資料を公開している[42]。

3.4.2. パラメータの推定

疾患推移モデルに対して、各健康状態における費用及び効果（QOL 値）、健康状態間の推移確率といったモデルパラメータを推定する必要があるが、ここでも様々な留意事項が存在する。

● 費用の推定について

国内ガイドラインでは費用の推定について様々なルールが定められているが、これは費用対効果評価の試行的導入において多くの品目で費用の算出において議論が交わされ、その結果がガイドラインに反映された模様である[1]。まず最も重要な点は費用に関しては国内データを用いる必要があるという点である。国内ガイドラインでも「各健康状態の費用の推計において、(中略) 実臨床を反映した国内におけるレセプトのデータベースを用いることを推奨する」としている。レセプトデータベースを用いて費用を推計する際は、以下の各健康状態の定義、外れ値処理などについて留意する必要があるだろう。レセプトデータベースのハンドリング時の留意点は別途資料を参照されたい[36][43]。

<各健康状態の定義>

☆ 健康状態は ICD-10 コードのみで定義可能か？

- レセプト病名が想定される場合は処方薬剤、診療行為などを組み合わせて健康状態を定義するといった工夫が必要である
- 健康状態の定義はバリデーションが取れた定義を採用もしくは臨床専門医の確認に基づき設定する

☆ 「疑い病名」を含めるか否か？

- 検査を行うために疑い病名を付ける可能性がある場合などは、処方薬剤なども組み合わせて判断するなどの工夫が必要である

☆ 入院を考慮する必要がある場合、どのように入院を定義するか？

- ex. 入退院日をどのように特定するか？

☆ DPC レセプトに記録されている入退院日は必ずしも実際の入退院日と一致する

とは限らないため、例えば、「DPC レセプト」と「医科入院レセプト」をセットにして「入院料」の算定を用いて、入院日を特定するなどの工夫が必要である

- ex. 分析対象疾患による入院をどのように特定するか？
 - ✧ DPC レセプトでは傷病名区分が「医療資源を最も投入した傷病名」,「主傷病名」,「入院の契機となった傷病名」などの 7 つに分類されているため、どの傷病名区分のときに集計対象とするか？
 - ✧ 医科入院レセプトでは傷病名区分コードがないため、傷病名、医薬品情報、診療行為などとの組み合わせを利用してレセプトを同定する工夫が必要だろう
- ✧ 医療費に含める範囲はどこまでか？
 - 対象となるレセプトの総点数か？
 - ✧ 対象となるレセプトのうち、薬剤処方日/検査実施日に算定された点数（非関連医療費を除く）の合算のみを用いるか？対象疾患、対象となる医療技術によってケースバイケースの検討が必要である。
 - もしくは、関連する医療を特定した積み上げ式で費用を算定するか？
 - ✧ 費用を積み上げて算定する場合は「単価」と「数量」を調査して、分析データの提出様式の記載のように整理して費用を推定する（表 3-3）

表 3-3 費用パラメータの推定（費用を積み上げで算定する場合）

【治療 A の費用】

項目名	単価	数量	小計
(例) D007 血液化学検査 5. LDL-コレステロール	180	1	180
(例) D026 検体検査判断料 3 生化学的検査(I)判断料	1,440	1	1,440
	総計		9,690

【出典 医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引[4]】

<外れ値処理>

- ✧ 例えば、年間の医療費に対して外れ値を除外するのか？月ごとの医療費に対して外れ値を除外するのか？
 - レセプトの総点数で検討した場合、両者間で推定される費用は大きく異なるケースがある。例えば、手術を伴う疾患領域の場合には月ごとの医療費に対して外れ値を検討すると手術実施月のデータを除外してしまう可能性がある。特に手術費用が他の治療費に比べて高額だった場合に問題となる。

また、上記以外にも国内ガイドラインでは「既存の疾病費用分析やレセプトデータを用いた分析などの場合、単価は医療資源が消費された時点ではなく、分析実施時点にそろえたものを用い

る。その際に、診療報酬改定率を乗じる等により調整してもよい」としている。全ての診療報酬内容を1つずつ現在価値へ補正することは困難なため、ネット改定率を用いて診療報酬の現在価値へ補正することが一案であり、ネット改定率の詳細は福田ら（2019）の論文を参照されたい[44]。そして、データベース研究では常に考慮すべきことだが、どの医療情報データベースを用いるかで患者網羅性、患者追跡性が変わるため、対象疾患に合わせてデータベースを選定しなければならない。必要に応じて、NDBを活用することも一案であり、C2Hとの分析前協議時に希望を申し込むことができる。いずれにしても費用の推定に関しては、企業がデータ抽出に関する定義書を作成し、分析前協議で認識を擦り合わせておくことは有用だろう。また、実務担当者自ら医療情報データベースに触れて、費用の推計方法を決めていくことが理想である。

留意事項⑭：医療情報データベースを活用する場合、適切にデータベースを選定し、費用や健康状態の定義を適切に設定する

● QOL 値の推定

続いて、QOL 値の推定に関する留意事項を紹介する。NICE TA における QOL 値は EQ-5D が主に使われていたが、疾患推移モデルの各健康状態に対応した QOL 値がすべて利用可能であることは少ない。そのため、しばしば QOL 値は先行研究から類似の集団の数値を引用したり、何らかの仮定をおいて回帰分析などのマッピングを用いた推定を行うことになる。しかしながら、企業が選択した QOL 値のうち最終的に採用されたのは 58%（87/150）であったという報告がある[45]。以下にいくつかの事例を紹介する。

QOL 値は生活の質を表すものであり、0（死亡）から 1（完全な健康状態）までの値を基本的にとるが、各健康状態とその QOL 値が合理的な関係になっている必要がある。NICE TA505 [46]では stable disease と post progression の健康状態に対する QOL 値の大小関係について、評価委員会から指摘があり、企業は再解析を行った^{xxii}。

留意事項⑮：健康状態とQOL値の大小関係の整合性を確認する	
事例：再発又は難治性の多発性骨髄腫に対するレナリドミド及びデキサメタゾンを含むイキサゾミブの費用対効果評価（TA505）[46]	
健康状態とQOL値の大小関係に関して、企業分析とERGで以下の乖離があった。	
	QOL値の推定
企業分析	QOL値を推定するためにTOURMALINEMM1試験の患者レベルのEQ-5Dデータを用いて、経時測定データの構造を考慮した混合効果モデルに

^{xxii} 英国では評価委員会が結果の解釈のために製薬企業や ERG に追加分析等を指示する場合がある。

	よるモデリングを行った。奏効状態（最良総合効果），入院，Grade 3/4のTRAE（treatment-related adverse event），新しい原発性悪性腫瘍，患者が死亡前3ヶ月以内であるか否かを説明変数に含めた。なお，これらの説明変数は臨床医の意見に基づいて選択した。その結果，各奏効状態のQOL値を以下のように推定した。 <table><tr><th>健康状態</th><th>QOL値</th></tr><tr><td>very good partial response</td><td>0.712</td></tr><tr><td>partial response</td><td>0.674</td></tr><tr><td>stable disease</td><td>0.653</td></tr><tr><td>post progression</td><td>0.654</td></tr></table>	健康状態	QOL値	very good partial response	0.712	partial response	0.674	stable disease	0.653	post progression	0.654
健康状態	QOL値										
very good partial response	0.712										
partial response	0.674										
stable disease	0.653										
post progression	0.654										
ERG	企業分析について，以下の観点から合理的でないと考えた。 ● <u>post progressionのQOL値がstable diseaseよりも高くなっている</u> ● PFSのQOL値は，最良総合効果に基づいているため，QOL値が実際に測定された時点の奏効状態ではなく，過大評価されている可能性がある ● 年齢と前治療で調整されていないため，バイアスがある										

その後，ERGの意見を踏まえ，企業分析として臨床医の意見も考慮して年齢，人種，性別を説明変数に追加し，最良総合効果の代わりに EQ-5Dを評価した時点の奏効状態（腫瘍縮小効果）を用いて再解析を行い，解析結果を更新した。その結果，stable diseaseからpost regressionに0.028ほど減少した。評価委員会は更新した結果は最初に企業が提出した解析よりも尤もらしいと判断した。

	最初の解析	更新後
very good partial response	0.712	0.689
partial response	0.674	0.690
stable disease	0.653	0.678
post progression	0.654	0.650

続いて EQ-5D-3L の検出感度の問題である。EQ-5D-3L は「移動の程度」、「身の回りの管理」、「ふだんの活動」、「痛み/不快感」、「不安/ふさぎ込み」の5つのドメインに対して、「できない」、「いくらか問題がある」、「問題がない」の3つの水準で回答する。いずれかのドメインで1つでも水準が改善すれば EQ-5D のスコアは1（完全な健康状態）に近づくが、なかなか改善が見

いだせない疾患領域も存在し, EQ-5D-3L を用いた QOL 値では症状の改善を捉えきれないケースがある (NICE TA605[20]).

留意事項⑯：QOL値で捕捉できない症状の改善があるかもしれない

事例：慢性流涎症に対するゼオミン（A型ボツリヌス毒素製剤）の費用対効果評価（TA605）[20]

EQ-5D-3Lの限界に関して、企業分析と評価委員会の間で以下の議論があった。

	EQ-5D-3Lによる評価
企業分析	SIAXI試験ではEQ-5D-3Lを用いてQOL値を測定した。しかしながら、企業はこの質問表は唾液分泌過多*の症状の測定に対して感度が低いと考えた。多くの患者が複数のドメインに対して、2（いくらか問題がある）のスコアで回答したため、QOL値の改善を示すためには1（問題がない）と回答する必要があった。しかしながら、唾液分泌過多を有する患者の原疾患がパーキンソン病、外傷性脳損傷などであったことから、<u>全てのドメインが原疾患の重症度の影響を受け、「問題がない」と回答することは難しく、QOL値の改善は困難だと考えた。</u> *：慢性流涎症の治療は唾液分泌過多の症状を改善することが目的である
評価委員会	SIAXI試験がQOL値推定のための最も適切なデータソースであったが、<u>EQ-5D-3Lデータと導出されたQOL値が健康関連QOLのベネフィットを完全に捉えることは難しいとした。</u>さらに、EQ-5D-3Lは社会的孤立を含む唾液分泌過多の心理的影響を捉えていない可能性がある。また、介護者の仕事量や潜在的なさらなる社会的孤立を含む、介護者のQOLへの影響は捉えられていないと考えた。 以上より、SIAXI試験からのEQ-5D-3Lの結果は治療による効用の増加を過小評価する可能性があり、意思決定においてこれを考慮するとした。

日本では EQ-5D-3L ではなく EQ-5D-5L を用いて QOL 値を測定することで、NICE が推奨する EQ-5D-3L より QOL 値の変化をより鋭敏に捉えることが出来る^{xxiii}（天井効果の改善）[47].

国内ガイドラインでは「QALY を算出する際の QOL 値は、選好に基づく尺度 (preference-based measure: PBM) により測定したものをを用いることを原則とする」としているが、疾患特異

^{xxiii} EQ-5D-3L は 3 つの水準で各ドメインを評価するが、5L は 5 つの水準で各ドメインを評価する。水準数が増えることでより小さな QOL 値の変化も捉えることが可能となる。

的 QOL 尺度から EQ-5D への回帰分析によるマッピング手法の使用を検討するケースもあるだろう [48]. マッピングに関しては分析対象集団と異なる対象集団に対して構築されたマッピングアルゴリズムを採用したことで議論が生じた事例 (NICE TA423) [40]を紹介する.

留意事項⑰：健康関連QOLからQOL値へのマッピングは適切に行う

事例：2つ以上の化学療法後の局所進行性乳がん/転移性乳がん患者に対するエリブリンの費用対効果評価 (TA423) [40]

マッピングの適用に関して、企業分析とERGの間で以下の議論があった。

	QOL値の算出方法
企業分析	臨床試験の健康関連QOLからQOL値を推定するためにCrott and Briggsのマッピングアルゴリズムを使用した。このアルゴリズムは局所進行性だが転移性ではない乳がん患者でかつベースラインの健康状態は良好な患者のデータから開発されていた。 <u>分析対象集団に比べて健康状態が良好な患者集団について開発されたマッピングアルゴリズムを用いていた。</u> これにより企業分析の進行後のQOL値が進行前に比べてわずかに減少しただけであった。
ERG	ERGは企業分析の方法は妥当ではないと考え、より適切であると考えられるLloyd et al.の研究[49]からのQOL値を使用した。これにより進行前と進行後の間でQOL値が約20%で低下した。

評価委員会は臨床専門家の意見を踏まえ、企業分析のわずかなQOL値の減少はエリブリンの臨床試験から得られたものだが、QOL値の変化を過小評価している可能性があると考えた。また、Lloyd et al.の研究からの推定値が、企業分析のマッピングから得られた推定値よりも正確かどうか、確信を持って判断することができなかった。進行後の健康状態の最も妥当なQOL値は企業分析の推定値とERGの推定値の間にある可能性が高いと結論した。

各健康状態の QOL 値は一つのデータソースから推定できれば問題ないが、実際には一つのデータソースから推定しきれないケースが多いだろう。そのような場合、各健康状態における QOL 値もシステマティックレビューにより、網羅的に検討する必要がある。NICE TA550 [50]の企業分析では複数のデータソースを用いて QOL 値を推定した際に ERG から疑義事項が挙がった。

留意事項⑱：健康状態のQOL値を推定する際、データソースの一貫性を保つ

事例：甲状腺髄様がんに対するバンデタニブの費用対効果評価 (TA550) [50]

QOL値の推定方法に関して、企業分析とERGの間で以下の議論があった。

害事象の影響を考慮したが、有害事象発現による QOL 値の減少の考え方について指摘があった。

留意事項⑳：有害事象発現によるQOL値の減少を適切に設定する

事例：未治療の限局性前立腺癌に対するパデリボルフィンの費用対効果評価（TA546）[52]

Shimizu et al. により限局性前立腺がんの日本人男性のコホートのデータから腸に問題を有する男性患者のEQ-5Dは腸に問題のない男性に比べて有意に低いことが報告されていた（0.84 vs 0.94）。この文献を用いて、Hummel et al. は前立腺がんの強度変調放射線治療のモデルにおいて腸機能障害を経験している患者のQOL値を推定した。具体的には、Shimizu et al. の文献より腸に問題のない男性に対する腸に問題のある男性の相対値 0.8936 ($=0.84/0.94$) を算出し、70歳の腸機能障害のある男性患者のQOL値を0.727 ($=0.813$ （英国の70歳の健康男性のEQ-5D） $\times 0.8936$ ；相対値を乗算）と報告した。この算出方法は NICE TSD12[51]の経済モデルでのQOL値の適用に関する乗法アプローチと一致している。

一方、企業分析では乗法アプローチを用いずにベースライン値を0.88と推定し、0.88と0.72の差（絶対値としての減少）である0.16を腸機能障害のQOL値の減少分と推定した。この計算方法について、ERGは元のShimizu et al. により報告されている絶対減少の0.10 ($=0.94-0.84$) よりも腸機能障害の影響を過大評価していると指摘した。そのため、評価委員会はERGが使用した0.10の減少を支持した。

以上のように QOL 値に関しても様々な留意事項があるが、QOL 値の測定と利用に関しては NICE TSD8 から NICE TSD12 の概要をまとめた製薬協成果物が参考になるだろう[45]。なお、分析データの提出様式では QOL 値の詳細として、以下に示すような表で健康状態ごとに測定国、測定した集団の詳細、使用した尺度、測定者数、引用文献などを示すように求めている。

変数名	測定国	測定した集団の詳細	使用した尺度	測定者数	reference
(例)健康状態 A の QOL 値	日本	疾患 A を持つ集団で、 治療等には制限なし。 平均年齢 60 歳。	EQ-5D-5L	n=60	[12]

【出典 医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引[4]】

● 健康状態間の推移確率の推定について

モデルパラメータの推定に関する興味深い日本の事例として、健康状態間の推移確率に関して統計的有意差が示されないと評価対象技術と比較対照技術の間に差をつけた推定値を設定すべきではないという事例（C2H1903）があった[22]。統計的有意性は症例数設計が適切に行われていない限り、臨床的に意味のない違いを見出していたり、主要なエンドポイントでないため症例数が十分になく、臨床的に重要な違いが認められていても有意差が認められないケースもある。

り、解釈が困難となる。この場合、点推定値の臨床的考察、及び信頼区間の幅の大きさにより不確実性の大きさを示し、結果に与える影響を考察するという対応も今後必要だろう^{xxiv}。健康状態間の推移確率の推定^{xxv}についても多岐に渡る議論があるが、その方法論や事例の詳細は NICE TSD に一つ一つ詳細に紹介されていることから、ここでは事例紹介や各論には触れず、必要に応じて、それぞれの NICE TSD 又はその概要をまとめた製薬協成果物を適宜参照いただきたい。

◇ Evidence Synthesis TSD series (TSD1~7) [53][54][55][56][57][58][59]

- ネットワークメタアナリシスの方法論及びモデルの妥当性評価（異質性、不一致性の検討）を解説

◇ Survival analysis TSD (TSD 14) [60][61]

- 生存時間曲線のモデリング（外挿も含む）とモデルの適合性の評価方法を解説

◇ 日本でも生存時間曲線の外挿方法が議論となった事例（C2H1902[16]）がある

留意事項②①：生存時間のモデリング（外挿も含む）は慎重に行う

◇ Treatment switching TSD (TSD16) [61][62]

- 実臨床、臨床試験のいずれでも発生し得る治療スイッチング^{xxvi}が臨床成績に与えるバイアス及びその調整法を解説

◇ Population-adjusted indirect comparisons (MAIC and STC) (TSD18) [63][64]

- 間接比較やネットワークメタアナリシスに用いる複数の試験のうち、一部の試験では個別患者データを利用できることを利用して、試験間の効果修飾因子の分布の違いを調整して治療効果を推定する手法（母集団調整法）を解説

◇ 日本でも企業分析と公的分析で異なる間接比較の方法を採用した事例がある（C2H2003 [65]）

留意事項②②：間接比較を行う場合、試験間の効果修飾因子の分布の違いを考慮する

^{xxiv} 一方で、価格調整は基本分析により求められた ICER の値が用いられている。そのため、信頼区間を用いて不確実性を評価した感度分析の結果を考慮した上で価格調整が行われた事例は現時点で報告されていない（2023 年 1 月 8 日時点）。

^{xxv} ネットワークメタアナリシスを用いた場合、評価対象技術と比較対照技術の差（比）が推定される。さらに比較対照技術のベースラインリスクもネットワークメタアナリシスの推定対象となることがある。

^{xxvi} スwitching とは、抗がん剤のランダム化比較試験で対照群に割り付けられた患者が、PFS 後などに治療を対照治療から新治療に切り替えることである。スイッチングが存在する場合、ITT 解析で推定された対照群の効果に新治療の効果が混在し、バイアスが含まれ群間差が過小評価される。費用対効果評価を行うにあたり、対照治療に対する新治療のベネフィットをより正しく示すためスイッチングの調整が必要になる。

● 海外データの活用について

Step 4 の最後に海外データを日本の費用対効果評価で活用する際の留意点について触れたい。費用対効果評価は多くのデータソースを寄せ集めて、様々なモデルパラメータを推定する必要があるが、全てのモデルパラメータが国内のデータソースから推定可能なケースは非常に稀である。海外データを用いて推定する際には下記の点に注意する必要があるだろう。

- ✧ 日本と海外では治療実態及び医療環境（保険制度）が異なる可能性がある
- ✧ 日本と海外では医療費が異なる可能性がある
- ✧ 日本と海外では QOL 値が異なる可能性がある
- ✧ 日本と海外ではベースラインリスク（対照群の自然な発症リスク）、治療効果が異なる可能性がある

国際薬剤経済アウトカムリサーチ学会 (ISPOR) のタスクフォースの報告書では、ローカルデータの使用を推奨しているパラメータとしてベースラインリスク、QOL 値、費用の単価、医療資源消費量を挙げ、比較的一般化可能性が高いパラメータとして治療効果を挙げている[66]。国内ガイドラインも表 3-4 のように同様の方針となっている。海外データを評価に用いる際はその妥当性をよく検討し、必要に応じて、感度分析やシナリオ分析でその不確実性を評価するとい

表 3-4 データソースに関する国内ガイドラインの記載

パラメータ	データソースに関する国内ガイドラインの記載（括弧内の数値は項番号）
有効性・安全性・QOL 値等のデータ	「原則として、研究の質やエビデンスレベルが高く、かつ日本における現実の臨床成績を反映しているものを優先的に使用する（9.1）」、 「国内外でデータに明確な異質性が存在する際には、国内データを優先して使用する（9.2）」 ✓ 国内外で異質性がないと考えられる場合、海外を含めた全データを利用することが可能
QOL 値	「国内での調査結果を優先して使用することを推奨する（8.4）」、「国内における研究がないあるいは不十分で、海外で質の高い研究がなされている場合は、海外で測定されたものを使用してもよい（8.4.1）」 ✓ 基本は国内データだが、国内データが無ければ海外データの利用もやむなし
費用	「各健康状態の費用の推計においては、日本における平均的な使用量や標準的な診療過程等が反映されている必要がある（10.3）」、「海外データを用いる際には、医療資源消費量について、国内外における医療技術の使用実態等の違いに配慮する必要がある。単価は国内のもの

	のを反映させなければならない (10.12) 」 ✓ 単価は国内データ必須，消費量は基本国内データを利用
--	---

データソースの選択に関して，QOL 値の推定で海外データを用いた際に議論となった日本の事例（C2H 1902 [16]）を紹介する．

留意事項②③：特に海外データのQOL値を用いる場合，国民標準値との整合性を確認する

事例：びまん性大細胞型B細胞リンパ腫に対するキムリアの費用対効果評価（C2H 1902） [16]

QOL値の推定方法に関して，企業分析と公的分析において以下の議論があった。

	QOL値の推定方法										
企業分析	JULIET試験（日本人を含む国際共同治験）ではEQ-5Dを直接測定していなかったため，Chen et al.の論文に基づいてPFSの健康状態のQOL値を0.83とした。										
公的分析	企業分析はPFSの健康状態を0.83としたが，この値は日本人の70歳以上のEQ-5D-5Lの国民標準値（男性：0.866，女性：0.828）とほぼ同様の値であり，70歳以降におけるQOL値が過剰推計されている可能性がある．そのため，びまん性大細胞型B細胞リンパ腫患者を対象としてCAR-T 療法の費用効果分析を実施したLin JK et al. の論文を用いて，より実態を反映しているQOL値，すなわちPFSの健康状態として 0.70 を 70歳以降から用いた。 <table><tr><td></td><td>企業分析</td><td>公的分析</td></tr><tr><td>70歳未満の集団</td><td>全分析期間： 0.83</td><td>70歳未満までの分析期間：0.83 <u>70歳以降の分析期間：0.70</u></td></tr><tr><td>70歳以上の集団</td><td>全分析期間： 0.83</td><td><u>全分析期間：0.70</u></td></tr></table> ※日本人を対象としたデータが無かったため，公的分析でも海外データを用いているが，日本の国民標準値を考慮して企業分析とは異なる海外のデータソースが引用された。			企業分析	公的分析	70歳未満の集団	全分析期間： 0.83	70歳未満までの分析期間：0.83 <u>70歳以降の分析期間：0.70</u>	70歳以上の集団	全分析期間： 0.83	<u>全分析期間：0.70</u>
	企業分析	公的分析									
70歳未満の集団	全分析期間： 0.83	70歳未満までの分析期間：0.83 <u>70歳以降の分析期間：0.70</u>									
70歳以上の集団	全分析期間： 0.83	<u>全分析期間：0.70</u>									

✓ パラメータの推定に関するまとめ

パラメータの推定で重要なことは，各パラメータに対して利活用可能なデータソースの中から最適なデータソースを選択し，最適な方法で推定することである．そして，その妥当性を明確に説明することである．また，評価結果にインパクトを与えるパラメータとそうでないパラメー

タを意識した上で、リソースの割り方を考えることも重要であり、最終的に残った不確実性は感度分析やシナリオ分析により意思決定をサポートするための情報を提供する必要がある。

3.5. Step 5：費用と効果の比較：ICER などの算出

費用と効果の比較に関して、基本的には国内ガイドラインに沿って「追加的有用性が示されていると判断される場合には、各群の期待費用と期待効果から増分費用効果比」を算出することとなる。

2011 年に報告された少々古い論文ではあるが、NICE で評価された品目において、企業分析と ERG が算出した ICER の値には差があることが報告されている[67]。企業分析の基本分析の ICER は、13 ケースで 1QALY あたり 30,000 ポンド以上、30 ケースで 30,000 ポンド未満だった。一方、ERG の基本分析の ICER は 25 ケースで 30,000 ポンド以上、15 ケースで 30,000 ポンド未満と報告された。同じ品目に対する企業分析と ERG の ICER の差を取った報告ではないため、両者の ICER の分布に比較可能性はないものの、企業分析の方が ERG よりも ICER が低く報告されている傾向が示唆されている。

また、日本では費用対効果評価を価格調整のために用いているが、NICE では NHS での使用の可否判断に用いている。NHS での使用の可否判断に際して、NICE は 1QALY あたり 20,000 ～30,000 ポンド、致死性疾患、終末期における治療については 50,000 ポンドという閾値が設定されているが [24]、2015 年に報告された論文では、調査対象となった 114 例のうち、41 例が閾値を上回り、そのうち 54% (22/41) が推奨又は制限付推奨となっていた[68]。つまり、NICE では、ICER は医療技術評価の意思決定プロセスにおいて最も影響力のある項目ではあるが、ICER 以外の項目（治療効果が改善したという強固な臨床エビデンス、アンメットニーズ、代替治療の有無、社会的なベネフィット、オフファン薬等）も考慮されていることがわかる。このような状況を踏まえ、この論文では患者アクセスを確保するためには閾値を超えた ICER だったとしても、透明性、頑健性があり、具体的なエビデンスを用いて患者アクセスを確保する説明の必要性について言及している。一方、日本では総合的評価において配慮が必要とされた品目は価格調整に用いる ICER の基準値が調整される（図 3-2）が、ICER 以外の項目（例えば、倫理的・社会的な価値などの多面的な価値）をさらに考慮して価格調整が行われた事例は現時点では報告されていない^{xxvii}。

^{xxvii} ICER 以外の項目の考慮に関しては、英国と日本で費用対効果評価の活用方法に違いがある点も勘案する必要があるだろう。

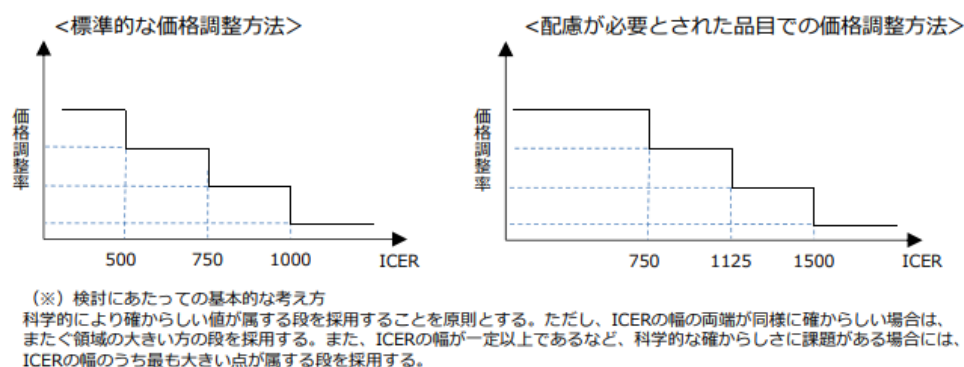


図 3-2 日本の ICER に基づく価格調整（加算部分）の方法[69]

3.6. Step 6：分析結果の解釈・報告

分析結果を解釈する上で、不確実性を有するパラメータ（推定値のばらつきの大きなパラメータ、実際のデータではなく仮定に基づき設定したパラメータ、諸外国のデータで国内のデータと異質性を有する可能性があるパラメータ等）や割引率については感度分析を行う必要がある。また、モデル構築時に仮定を置いていた場合にはその仮定を変えたシナリオ分析を行う必要もあるだろう[4]。一次元感度分析や確率的感度分析の方法論の詳細は成書を参照されたいが、確率的感度分析で報告方法が曖昧として NICE から指摘を受けた事例（NICE TA315）[31]を紹介する。

留意事項②④：分析（ex. 確率的感度分析）方法は適切に報告する

事例：2型糖尿病に対するカナグリフロジン併用治療の費用対効果評価（TA315）[31]

企業分析で確率的感度分析を実施したが、ERGは企業分析の提出資料について確率分布のパラメータが十分に記載されておらず、モデルに透明性がないため、意思決定に関して不確実性が十分に調査されたか不明とした。

パラメータ	確率分布	確率分布のパラメータ		確率分布のパラメータ	
		タイプ	範囲	タイプ	範囲
ベースライン時の年齢	正規分布	平均	$[0, \infty)$	SD	$[0, \infty]$
2型糖尿病の罹患期間	一様分布	最小値	$[0, \text{最大値}]$	最大値	$[\text{最小値}, \infty]$
低血糖発現率	ポアソン分布	平均	$[0, \infty)$	—	
...					

Table 66 Parameter distributions for PSA in ECHO-T2DM (reproduced from MS Table 46 p271)から抜粋して作成

※上記のような表が報告されていたが、各モデルパラメータに対してどのような確率分布を設定したかは把握できるが、具体的な確率分布を規定するパラメータ（例：正規分布の平均、標準偏差）の値が記載されていない。確率分布のパラメータの取り得る値の範囲のみしか記載

されておらず、具体的な値が記載されていなかった。

分析データの提出様式では、分析に用いたパラメータを下表のように整理するように求めている（表 3-5）。確率的感度分析を行う際には、上述の事例のような不明確な記載を避けて、設定根拠と共に確率分布の詳細について記載する必要がある。また、確率分布の設定に困った場合は統計担当者に意見を求めるとよいだろう。

表 3-5 分析に用いたパラメータの記載方法

変数名	値	(該当する場合) 95% CI	分布[該当する場合]	設定根拠
(例) 年齢	57 歳	-	-	
(例) 全生存期間	25 ヶ月	-	ワイブル (scale, shape)	
(例) 評価対象技術の有効性(ハザード比)	0.65	0.54-0.76	対数正規 (μ, σ^2)	
(例) 健康状態 A の QOL 値	0.71	0.62-0.77	正規 (μ, σ^2)	
(例) 評価対象技術の一月あたり費用	100,000 円	-	対数正規 (μ, σ^2)	

【出典 医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引 [4]】

不確実性について NICE は 3 つの種類に分けて整理している（表 3-6）[24]。また、ISPOR の報告書[70]ではパラメータ推定における不確実性を表 3-7 のように整理しており、回帰モデルと結びつけることでわかりやすく説明している。それぞれの不確実性に対して、適切な対応方法を選択して結果に与える影響を解釈する必要がある。

留意事項②⑤：不確実性が結果に与える影響を評価し、解釈する

表 3-6 NICE の methods guide [24]における不確実性の分類と対応

不確実性の種類	説明	対応方法
structural uncertainty (構造的な不確実性)	モデル構築に必要な仮定及び判断に関する不確実性。 モデルの設計上の特徴（ex. 想定される標準的な治療経路）、エビデンスの妥当性に関する判断、健康状態の分け方、パラメータの分布に関する仮定、パラメータの統計的な推定方法が含まれる。	これらの構造的な仮定は明確に記載し、それを裏付ける根拠を示すべきである。 構造的な不確実性の影響はシナリオ分析により検討する。

不確実性の種類	説明	対応方法
choice of data sources (データソースの選択)	費用, QOL値, 相対的な有効性の推定値とその期間などのキーとなるパラメータ推定に用いるデータソースの選択.	感度分析によりデータソースの選択が結果に与える影響を検討する. 例えば, 基本分析で用いたデータソースとは別のデータソースを用いて評価を行う.
parameter precision (パラメータの精度)	パラメータの精度に関する不確実性. モデルパラメータ, 例えば効果, QOL値, 資源消費の平均値に関する不確実性.	確率的感度分析を用いて不確実性を評価することが望ましい. 個々のパラメータ間に相関があるかは慎重に検討し, 必要に応じて確率的感度分析に反映すべきである.

表 3-7 ISPOR の報告書に[70]におけるモデルパラメータの不確実性の分類

分類	内容	回帰モデルで考えた際の解釈
stochastic uncertainty (確率論的な不確実性)	同一患者内のランダムな変動	エラー項
parameter uncertainty (パラメータの不確実性)	興味のあるパラメータの推定の不確実性	推定値の標準誤差
heterogeneity (異質性)	患者の特徴に起因する患者間の変動	回帰係数 (患者の特徴によって変動する程度)
structural uncertainty ^{xxviii} (構造的な不確実性)	モデルの本質的な構造に関する不確実性	回帰モデルの形 (ex. 線形, 対数線形)

また, 構築したモデルの妥当性を確認することは重要である. 分析データの提出様式でも「モデルや分析結果の外的妥当性について記載する. 特に, 推計された費用や効果について既存の臨床研究や統計, 長期的な疫学調査との整合性等について記載する」としている. 外的妥当性以外にも ISPOR のタスクフォースでは表面的妥当性, 検証又は内的妥当性, クロスバリデーション, 予測妥当性を挙げており (表 3-8), **費用対効果評価を行った際にはどの手法を用いて評価の妥当性を確認したかを明記することは重要である**[71]. 外的妥当性と予測妥当性による確認が最も強力とされている.

^{xxviii} 表 37 の structural uncertainty とは異なる点に留意する.

表 3-8 妥当性の種類

種類	内容
face validity (表面的妥当性)	臨床の専門家によるモデルの構造, データソース, 分析のフレームワーク (分析対象集団, 介入, 転帰, 仮定, 分析期間) 及び結果の評価であり, 主観的である.
verification or internal validity (検証又は内的妥当性)	プログラムコードの確認 ^{xxix} .
cross validity (交差妥当性)	同様の問題を分析した他のモデルの結果との比較.
external validity (外的妥当性)	モデルの結果と実際の結果を比較.
predictive validity (予測妥当性)	現在実施中の前向きにイベントを観察している研究とモデルの予測結果を比較 ^{xxx} .

NICE の事例でも妥当性に関する記載があり, 例えば, NICE TA423 [40]では以下のように妥当性を説明している.

留意事項②⑥: 評価の妥当性を説明する
事例: 2つ以上の化学療法後の局所進行性乳がん/転移性乳がん患者に対するエリブリンの費用対効果評価 (TA423) [40] ✓ 基本分析である5年間の分析期間の生存データは, PSM (Partitioned survival model) を採用することで試験結果をそのまま用いた ➤ 評価に用いた試験のEMBRACE 試験は 95% の死亡を追跡できていたため, 試験期間を超えた部分である外挿部分の不確実性は小さかった ✓ 評価を行うにあたり, 各種資料を参考にした ➤ 10年, 20年の分析期間のシナリオ分析では, 既存論文 (Tremblay et al) に従い, 外挿を行った ➤ 費用は主に NICE advanced breast cancer guidelinesと最近のNHS reference costsを参照した ➤ 有害事象のコストはHRG (Healthcare resource group) /DRG (Diagnosis-related

^{xxix} 計算が正しく行われており, モデルの仕様と整合しているかを確認することである. 主に二つのステップがあり, 個々の式の検証, その式を正確にプログラムコードとして実装しているかを確認する.

^{xxx} 予測妥当性の限界は, この妥当性の確認結果は必ず将来に出てくることであり, 早急な意思決定には役立たないことである. しかしながら, 過去に研究が開始され, 費用対効果評価の評価時点ではその結果がまだわかっていないが, 近い将来に結果が発表されるような場合には有用とされている.

groups) のアプローチに基づいて算定した

- ✓ モデルで使用したQOL値と有害事象発現によるQOL値の減少は、出来るだけ保守的に設定した
- ✓ 重要な有害事象を考慮するために、2%以上の発現率があるgrade 3以上の有害事象を評価に含めた

Step 6 では感度分析やシナリオ分析を行い、評価結果の不確実性が結果の解釈に与える影響を明らかにし、さらに、これまで実施してきた費用対効果評価の透明性（モデルがどのように構築されているかを確認できるか）と妥当性（モデルが現実をどの程度再現できているか）を報告することが非常に重要である[71]。今後、これらの具体的なノウハウがさらに蓄積・整理されていくことで、より質の高い費用対効果評価の推進につながると考える。

4. おわりに

本報告書では、NICE 及び日本における費用対効果評価の事例を通して、実務担当者が費用対効果評価を行う際に留意すべき事項を紹介した。網羅的な事例調査を踏まえた論点整理までは至っていないが、先行事例から学ぶことは非常に多いと考える。以下に主な留意事項をまとめた。

- ✓ 分析対象集団，比較対照技術は分析協議前に十分に議論しておく
- ✓ 追加的有用性は適切に判断できるようにロジックを構築する
- ✓ 疾患推移モデルは疾患の予後と治療の流れを適切に表現する
- ✓ 各モデルパラメータは最適なデータソースを選択し，適切な方法で推定する
- ✓ 様々な不確実性が存在するため，不確実性が結果に与える影響を評価する
- ✓ 評価内容に関する透明性を確保し，評価の妥当性を報告する

医療資源の効率的配分の意思決定に資する情報を提供するために，費用対効果評価は利用可能な情報を最大限活用しながら，合理的と思われる方法を適宜選択していく必要がある。そのため，費用対効果評価では正しい唯一の値が定まるものでもなく，企業分析と公的分析で異なる見解が示された事例も散見される。これは利用可能な情報が必ずしも費用対効果評価を見据えて取得されたエビデンスでなかったり，そもそも必要な情報が評価時に存在しないことで，企業分析と公的分析で不確実性に対するアプローチが乖離したことが一因であろう。実務担当者にとって重要なことは，透明性を確保しながら評価の経緯や不確実性を明示し，様々な限界に向き合いながら最善で妥当と考える方法を選択し，その理由を報告書に示すことだと考える。本報告書では留意事項に対する対策案も可能な範囲で述べてきたつもりだが，個々のケースに応じて必要な対策を考えないといけない。また，蓄積されてきた事例に基づいて産官学で方法論などについて議論を深めることは有用であり，現在，本格導入後の経験を踏まえて費用対効果評価の制度の課題を整理し，議論されている模様である[15][72][73]。

製薬企業の実務担当者は「評価対象品目に選定されるから費用対効果評価を行わなければ…」というマインドセットではなく，自分達が心血を注いで研究・開発してきた医療技術の価値を費用対効果評価でどのように示すかを研究開発段階から検討し，不確実性に向き合いながら評価の肝となる必要なエビデンスを計画的に取得し，評価の準備を進めることでより良い適切な評価が行えるだろう。本報告書が費用対効果評価を行う際の一助になれば幸いである。

参考文献

1. 中央社会保険医療協議会. 試行的導入における課題と対応について. 費薬材-3 (H30.11.21) . 2018. <https://www.mhlw.go.jp/content/12404000/000405906.pdf>.
2. 国立保健医療科学院ウェブサイト. 費用対効果評価の制度化以後(2019.4~)に選定された品目. <https://c2h.niph.go.jp/results/item.html>.
3. 国立保健医療科学院保健医療経済評価研究センター. 中央社会保険医療協議会における費用対効果評価の分析ガイドライン . 2022. https://c2h.niph.go.jp/tools/guideline/guideline_ja.pdf.
4. 国立保健医療科学院. 医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引. 2019. https://c2h.niph.go.jp/tools/system/Reporting_Format_Jap.pdf.
5. 日本製薬工業協会医薬品評価委員会. 医薬品の価値の科学的な評価ーデータサイエンス担当者のための費用対効果評価の現状と手法解説ー . 2016. https://www.jpma.or.jp/information/evaluation/results/allotment/2014ds_tf3.html.
6. Drummond M, Sculpher, M. J., Claxton, K., Stoddart, G. L., Torrance, G.W., (監訳) 久繁哲徳, 橋本英樹. 保健医療の経済評価第 4 版. 篠原出版新社. 2017.
7. 鎌江伊三夫. 医療技術評価ワークブック. じほう. 2016.
8. 五十嵐中, 佐條麻里. 「薬剤経済」わかりません!! . 東京図書. 2014.
9. 日本製薬工業協会医薬品評価委員会. 費用対効果評価の実践ーNICE 技術評価ガイダンスの事例から分析モデルのバリエーションを学ぼうー . 2018. <https://www.jpma.or.jp/information/evaluation/results/allotment/cost-effectiveness.html>.
10. 中央社会保険医療協議会. 費用対効果評価制度の当面の運用について. 総-9 (R3.2.10) . 2021. <https://www.mhlw.go.jp/content/12404000/000736552.pdf>.
11. 国立保健医療科学院ウェブサイト. 医薬品・医療機器の費用対効果評価における分析前協議について . 2020. https://c2h.niph.go.jp/tools/system/preliminary_consultation_20200714.pdf.
12. 中央社会保険医療協議会. 医薬品等の費用対効果評価案について. 総-4 (R3.3.24) . 2021. <https://www.mhlw.go.jp/content/12404000/000757518.pdf>.
13. 国立保健医療科学院ウェブサイト. [C2H1901] フルチカゾン/ウメクリジニウム/ビラントロール (テリルジー). 2021. <https://c2h.niph.go.jp/results/C2H1901.html>.
14. 国立保健医療科学院ウェブサイト. [C2H1905] ボルチオキセチン(トリンテリックス). 2021. <https://c2h.niph.go.jp/results/C2H1905.html>.
15. 中央社会保険医療協議会. 令和 4 年度費用対効果評価に関する検討について (費ー 1 参考 1) 費用対効果評価制度の見直しに関する検討 (その 1) . 2021. <https://www.mhlw.go.jp/content/12404000/000852904.pdf>.

16. 国立保健医療科学院ウェブサイト. [C2H1902] チサゲンレクルユーセル (キムリア). 2021. <https://c2h.niph.go.jp/results/C2H1902.html>.
17. NICE. TA164. Febuxostat for the management of hyperuricaemia in people with gout. 2008. <https://www.nice.org.uk/guidance/ta164>.
18. 中央社会保険医療協議会. 費用対効果評価専門組織からの報告について (総-4) 医薬品等の費用対効果評価案について. 2021. <https://www.mhlw.go.jp/content/12404000/000757518.pdf>.
19. 国立保健医療科学院ウェブサイト. 医薬品, 医療機器及び再生医療等製品の費用対効果評価における分析前協議について. 2022. https://c2h.niph.go.jp/tools/system/preliminary_consultation.pdf.
20. NICE. TA605. Xeomin (botulinum neurotoxin type A) for treating chronic sialorrhoea. 2019. <https://www.nice.org.uk/guidance/ta605>.
21. NICE. TA497. Golimumab for treating non-radiographic axial spondyloarthritis. 2018. <https://www.nice.org.uk/guidance/ta497>.
22. 国立保健医療科学院ウェブサイト. [C2H1903] ラブリズマブ (ユルトミリス). 2021. <https://c2h.niph.go.jp/results/C2H1903.html>.
23. 国立保健医療科学院ウェブサイト. [C2H2001] ポサコナゾール(ノクサフィル). 2021. <https://c2h.niph.go.jp/results/C2H2001.html>.
24. NICE. Guide to the methods of technology appraisal 2013. <https://www.nice.org.uk/process/pmg9/chapter/foreword>.
25. NICE. TA413. Elbasvir–grazoprevir for treating chronic hepatitis C. 2016. <https://www.nice.org.uk/guidance/ta413>.
26. NICE. TA506. Lesinurad for treating chronic hyperuricaemia in people with gout. 2018. <https://www.nice.org.uk/guidance/ta506>.
27. Stack AG, Hanley A, Casserly LF, Cronin CJ, Abdalla AA, Kiernan TJ, et al. Independent and conjoint associations of gout and hyperuricaemia with total and cardiovascular mortality. Qjm. 2013;106(7):647-58.
28. NICE. TA214. Bevacizumab in combination with a taxane for the first-line treatment of metastatic breast cancer. 2011. <https://www.nice.org.uk/guidance/ta214>.
29. (ICER) IfCaER. ICER Publishes White Paper Evaluating Policy Options to Strengthen FDA's Accelerated Approval Pathway for Prescription Drugs. 2021. <https://icer.org/news-insights/press-releases/icer-publishes-white-paper-evaluating-policy-options-to-strengthen-fdas-accelerated-approval-pathway-for-prescription-drugs/>. Accessed 6/21, 2022.
30. Anna Kaltenboeck AM, Steven D. Pearson. White Paper: Strengthening the Accelerated Approval Pathway: An Analysis of Potential Policy Reforms and their Impact on

- Uncertainty, Access, Innovation, and Costs. 2021. Institute for Clinical and Economic Review (ICER) https://icer.org/wp-content/uploads/2021/04/Strengthening-the-Accelerated-Approval-Pathway_-ICER-White-Paper_-April-2021.pdf.
31. NICE. TA315. Canagliflozin in combination therapy for treating type 2 diabetes. 2014. <https://www.nice.org.uk/guidance/ta315>.
 32. U.S. Department of Health and Human Services FaDA, Center for Drug Evaluation and Research (CDER). Guidance for Industry. Diabetes Mellitus – Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. 2008. <https://www.fda.gov/media/71297/download>.
 33. 国立保健医療科学院ウェブサイト. [C2H1906] イバブラジン(コララン). 2021. <https://c2h.niph.go.jp/results/C2H1906.html>.
 34. Hendrich J, Griffiths EA. Exploring the Flaws in Clinical Data that Lead to Rejection of Nice Submissions. Value Health. 2014;17(7):A329-30.
 35. NICE. TA279. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures. 2013. <https://www.nice.org.uk/guidance/ta279>.
 36. 佐藤俊哉, 山口拓洋, 石黒智恵子(編). これからの薬剤疫学—リアルワールドデータからエビデンスを創る—: 朝倉書店; 2021.
 37. 日本製薬工業協会医薬品評価委員会. データベース研究入門. 2015. <https://www.jpma.or.jp/information/evaluation/results/allotment/rwd.html>.
 38. Rita Faria MHA, Andrea Manca, Allan J Wailoo. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. 2015. <http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD17-DSU-Observational-data-FINAL.pdf>.
 39. 日本製薬工業協会医薬品評価委員会. 非ランダム化比較試験データを用いて治療効果を推定するための統計的手法. 2020. <https://www.jpma.or.jp/information/evaluation/results/allotment/non-randomization.html>.
 40. NICE. TA423. Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens. 2016. <https://www.nice.org.uk/guidance/ta423>.
 41. Kaltenthaler E, Tappenden, P., Paisley, S., Squires, H. NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. . 2011. <http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD-13-model-parameters.pdf>.
 42. 日本製薬工業協会医薬品評価委員会. NICE TSD の紹介・モデル作成—モデル作成のた

- め の エ ビ デ ン ス 収 集 と レ ビ ュ ー ー . 2021.
https://www.jpma.or.jp/information/evaluation/results/allotment/nice_tsd_model.html.
43. 奥村泰之, 佐方信夫, 清水沙友里, 松居宏樹. 研究部レポート ナショナルデータベースの学術利用促進に向けて : レセプトの落とし穴. Monthly IHEP. 2017(268):16-25.
 44. 福田治久, 佐藤大介, 福田敬. レセプトデータを用いた医療費分析における診療報酬改定の補正方法. 保健医療科学. 2019;68(2):147-57.
 45. 日本製薬工業協会医薬品評価委員会. NICE TSD の紹介・QOL 値ーQOL 値の測定と費用 効 果 分 析 へ の QOL 値 の 利 用 ー . 2021.
https://www.jpma.or.jp/information/evaluation/results/allotment/nice_tsd.html.
 46. NICE. TA505. Ixazomib with lenalidomide and dexamethasone for treating relapsed or refractory multiple myeloma. 2018. <https://www.nice.org.uk/guidance/ta505>.
 47. 池田俊也, 白岩健, 五十嵐中. 日本語版 EQ-5D-5L におけるスコアリング法の開発. 保健医療科学 = Journal of the National Institute of Public Health. 2015;64(1):47-55.
 48. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to Estimate Health-State Utility from Non-Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task Force Report. Value Health. 2017;20(1):18-27.
 49. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. Br J Cancer. 2006;95(6):683-90.
 50. NICE. TA550. Vandetanib for treating medullary thyroid cancer. 2018.
<https://www.nice.org.uk/guidance/ta550>.
 51. Ara R, Wailoo A. NICE Decision Support Unit Technical Support Documents 12: The Use of Health State Utility Values in Decision Models. 2011. National Institute for Health and Care Excellence (NICE).
 52. NICE. TA546. Padeliporfin for untreated localised prostate cancer. 2018.
<https://www.nice.org.uk/guidance/ta546>.
 53. 日本製薬工業協会医薬品評価委員会. NICE TSD Evidence Synthesis シリーズの紹介 : ネットワークメタアナリシスを中心として . 2020.
https://www.jpma.or.jp/information/evaluation/results/allotment/lofurc000000c6n6-att/nice_tsd_evidence_synthesis.pdf.
 54. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE Decision Support Unit Technical Support Documents 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2014. <http://nicedsu.org.uk/wp-content/uploads/2017/05/TSD2-General-meta-analysis-corrected-2Sep2016v2.pdf>.
 55. Sofia Dias AJS, Nicky J Welton, AE Ades. NICE DSU Technical Support Document 3: Heterogeneity: Subgroups, meta-regression, bias and bias-adjustment. 2011.

- <http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD3-Heterogeneity.final-report.08.05.12.pdf>.
56. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. NICE Decision Support Unit Technical Support Documents 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials. 2014. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD4-Inconsistency.final_.15April2014.pdf.
 57. Sofia Dias NJW, Alex J Sutton, AE Ades. NICE DSU Technical Support Document 5: Evidence synthesis in baseline natural history model. 2011. <http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD5-Baseline.final-report.08.05.12.pdf>.
 58. Sofia Dias AJS, Nicky J Welton, AE Ades. NICE DSU Technical Support Document 6: Embedding evidence synthesis in probabilistic cost-effectiveness analysis: software choices. 2011. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD6-Software.final_.08.05.12.pdf.
 59. AE Ades DMC, Stefanie Reken, Nicky J Welton, Alex J Sutton, Sofia Dias. NICE DSU Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: A reviewer's checklist. 2012. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD7-reviewer-checklist.final_.08.05.12.pdf.
 60. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials-extrapolation with patient level data. 2011. <http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf>.
 61. 日本製薬工業協会医薬品評価委員会. 費用対効果評価における生存時間解析の諸問題. 2020. https://www.jpma.or.jp/information/evaluation/results/allotment/c_e_a_problems.html.
 62. Nicholas R Latimer KRA. NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching. 2014. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD16_Treatment_Switching.pdf.
 63. Phillippo DM, Ades, A.E., Dias, S., Palmer, S., Abrams, K.R., Welton, N.J. . NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE. . 2016. <http://nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf>.
 64. 日本製薬工業協会医薬品評価委員会. 間接比較において母集団調整法は有用か？－Matching Adjusted Indirect Comparison 及び Simulated Treatment Comparison－. 2022. https://www.jpma.or.jp/information/evaluation/results/allotment/indirect_comparison.html.
 65. 国立保健医療科学院ウェブサイト. [C2H2003] トラスツズマブ デルクステカン (エン

- ハーツ) 2021. <https://c2h.niph.go.jp/results/C2H2003.html>.
66. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health*. 2009;12(4):409-18.
 67. Kaltenthaler E, Boland A, Carroll C, Dickson R, Fitzgerald P, Papaioannou D. Evidence Review Group approaches to the critical appraisal of manufacturer submissions for the NICE STA process: a mapping study and thematic analysis. *Health Technol Assess*. 2011;15(22):1-82, iii-iv.
 68. Griffiths EA, Hendrich JK, Stoddart SD, Walsh SC. Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. *Clinicoecon Outcomes Res*. 2015;7:463-76.
 69. 中央社会保険医療協議会. 費用対効果評価について 骨子(案)概要. 費薬材一参考(H31.3.27). 2019. <https://www.mhlw.go.jp/content/12404000/000493967.pdf>.
 70. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. *Value Health*. 2012;15(6):835-42.
 71. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making*. 2012;32(5):733-43.
 72. 中央社会保険医療協議会. 令和4年度費用対効果評価に関する検討について(費一1) 費用対効果評価制度の見直しに関する検討(その2). 2021. <https://www.mhlw.go.jp/content/12404000/000852903.pdf>.
 73. 中村洋. 日本版費用対効果評価制度の課題と今後の改革の方向性: 今後の改革の方向性. 2021. <https://www.ihep.jp/publications/monthly-ihep/?y=2021>.

2022 年度データサイエンス部会 KT7

作成担当者

東 美恵	エーザイ株式会社
荒西 利彦	日本イーライリリー株式会社
奥山 ことば	MSD 株式会社*
中島 章博	帝人ファーマ株式会社*

レビュー担当者

渥美 淳	東レ株式会社
河田 祐一	中外製薬株式会社
直井 一郎	住友ファーマ株式会社
野島 俊秋	興和株式会社
町田 光陽	塩野義製薬株式会社
吉田 瑞樹	ファイザーR&D 合同会社*

*TF リーダー

監事

酒井 弘憲	エーザイ株式会社
-------	----------