

## 経済評価の種類

- 5.2.11 リファレンスケースでは、費用効果（特に費用効用）分析が経済評価法として推奨される。これは選択肢間の費用の差が健康効果の変化という点で妥当であるかどうか確認しようとするものである。健康効果はQALYで表すべきである。
- 5.2.12 費用効果分析は、費用便益分析と比較して、より広範に使用および公表されており、当機構がNHS/PSSの固定予算から最大の健康効果を得ることを重視していることから、費用効果分析に重点を置くことは妥当である。その広範な使用を考えると、QALYは、死亡率およびHRQL効果の両方を反映する、健康効果の最適な包括的評価項目とみなされる。代替評価項目の存在（例：健康当量年）が認められているが、これらの方法を用いている経済評価は少なく、その強みおよび弱みが完全に確立されていない。QALYにおける基礎的な仮定事項（例：constant proportional trade-offおよび健康状態間の加法独立性）が特定の症例で不適切と考えられる場合は、この効果に対するエビデンスを作成すべきであり、代替評価項目を用いた分析を追加の非リファレンスケース分析として示すことができる。

## 時間地平

- 5.2.13 臨床的効果および費用対効果を推定する時間地平は、比較する技術間の費用またはアウトカムの重要な差を全て反映するのに十分な長さであるべきである。
- 5.2.14 技術の多くは、患者の生涯にわたって費用およびアウトカムに影響する。慢性疾患の治療の場合は特にそうである。そのような例では、臨床的効果および費用対効果の時間地平は生涯とすることが適切である。代替技術間の生存効果の差の影響を定量化するためには、生涯にわたる時間地平に、死亡の項目も必要である。また、生涯にわたる時間地平には、外挿によるモデリングも必要なことが多い。治療の影響が臨床試験の追跡期間を越えるために不明な場合は、将来の治療効果について異なる仮定を反映した複数の代替シナリオを比較する分析を提示すべきである（モデリングについては、セクション5.7を参照）。この仮定には、さらなる便益はないという制限をした仮定と、より楽観的な仮定を含めるべきである。治療に予測される影響よりも短い期間に時間地平に限った分析は、通常、費用および便益の最良推定にはならないと考えられる。
- 5.2.15 生涯よりも短い時間地平は、選択肢間で死亡率への影響に差がない場合、並びに費用およびHRQLの差が比較的短い期間に関連する場合（例：急性感染症の場合）に妥当と判断される。時間地平および臨床試験の期間を超えたデータの外挿に関する不確実性の検討は、評価の重要な要素の1つである。

### 5.3 アウトカムに関するエビデンスの総合

- 5.3.1 臨床的効果の分析の目的は、比較中の技術間の平均的な臨床的効果をバイアスなく推定することである。臨床的効果の分析は、品質が最良の関連する全ての試験から得たデータに基づくべきである。また、典型的な患者の範囲、正常な臨床的状況、臨床的に意味のあるアウトカム、関連する比較対象による比較、相対的・絶対的効果の評価項目並びに不確実性に関する適切な評価項目を考察する必要がある。

#### システマティックレビュー

- 5.3.2 明確に記述されている全てのデータ源を使って、全ての健康効果を特定、定量化する必要がある。リファレンスケースの場合、明確に立てられた疑問に関連するデータの信頼性が高く妥当な要約を得るための、系統的な位置付け、組み入れ、評価、およびエビデンス統合と定義されるシステマティックレビューから、アウトカムに関するエビデンスを得るべきである<sup>1</sup>。
- 5.3.3 診断技術の評価は、その他の医療技術について本ガイドで推奨されるように、システマティックレビューの一般原則に従うべきである。ただし、例えば診断検査の感度および特異度に関する研究のメタアナリシスの特性は、治療介入効果のレビューとは異なるとみなされる。これは現在進展している方法論的研究の一領域である。

#### 関連する研究

- 5.3.4 Head-to-head RCTでは、相対的な治療効果に関する最も有効なエビデンスが得られる。ただし、このようなエビデンスが常に入手できるとは限らず、ベースラインの健康効果を定量化するために十分ではないことがある。したがって、RCTデータを補完するために、非無作為化試験のデータが必要になる場合がある。アセスメントで使用する試験のデザインにより生じる可能性のあるバイアスを調べて報告すべきである。

#### 試験の選択およびデータの抽出

- 5.3.5 システマティックレビューは、バイアスの可能性を最小にするために、あらかじめ決められたプロトコールに従って実施すべきである。プロトコールには、レビューの特性を規定し、これによってバイアスのリスクを軽減し、再現性のあるレビューになるよう確保する。
- 5.3.6 研究戦略が作成されており、文献検索が実施されると、候補となる試験のリストが作成される。いずれの試験も、レビューの組み入れ基準に適合するかどうかを判断するために評価されなければならない。不適格な試験の記録を除外の根拠とともに保管し、文献検索および試験選択のプロセスの頑健性を評価できるようにする。研究戦略によって収集された全ての記録を複数のレビューアが評価した場合、決定プロセスの妥当性が高くなる。レビューア間の見解の相違を解消する手順を報告すべきである。

---

<sup>1</sup> 当機構のアセスメントグループは、NHSレビュー及び普及センター（Centre for Reviews and Dissemination）発行のシステマティックレビュー実施に関する一般的なガイドラインに従っている（Undertaking Systematic Reviews of Research on Effectiveness. York: NHS Centre for Reviews and Dissemination）

### 批判的吟味

- 5.3.7 個別の試験結果の妥当性は、その全体的デザインおよび実行、並びに決定問題への関連度に応じて変わる。したがって、組み入れ基準に適合する試験それぞれについて、批判的吟味を行うべきである。可能な場合は、未発表の試験および一部発表されている試験の妥当性の評価には、発表済み試験を評価するための基準を用いるべきである。

### 治療効果の修飾因子

- 5.3.8 多くの因子が、システマティックレビューから得られる相対的治療効果の推定値全体に影響を与える可能性がある。各試験間の相違は偶然生じることがある。その他、結果に相違を生じさせるよくある原因としては、患者特性（年齢、性別、疾患の重症度、アウトカムの選択および評価項目など）、治療環境、追加の日常治療、並びに、試験実施年（医療技術は発展していくため）の違いなどがある。対象エリア全体のレビューまたは当該臨床分野の専門家との討論のいずれかによって、これらの治療効果の潜在的な修飾因子をデータ分析前に特定する必要がある。

### メタアナリシス

- 5.3.9 比較可能なアウトカムを評価項目に用いている、関連性があり妥当なデータが十分ある場合は、メタアナリシスによるアウトカムデータの統合が適切である。
- 5.3.10 フォレストプロットは、個々の試験結果を示すのに有用なツールである。分析に含める各試験について、データの特性と考えられる限界（つまり、対象集団、介入、環境、標本サイズおよびエビデンスの妥当性）を完全に報告するべきである。
- 5.3.11 統計的統合（プール）では、不均一性（つまり、偶然で説明されるばらつきに加えて、あらゆるばらつき）の評価を行うべきである。結果の統計的な不均一性は、ある程度、変量効果モデル（固定効果モデルの対義語）を用いて説明することができる。ただし、不均一性の程度および理由をできるだけ完全に解明するべきである。既知の臨床的不均一性（例：患者特性または介入の投与の投与量または頻度）は、サブグループ解析およびメタ回帰分析を慎重に使うことで管理することが可能である。特定の試験の関連性に疑いがある場合は、当該試験を除外する影響を確認する感度分析を実施すべきである。メタアナリシスにおいて、特定イベントのリスクが試験のコントロール群間で著しく異なる場合は、相対的治療効果の評価項目が異なるベースラインのリスク間で一定かどうかを評価するべきである。

このことは、経済決定モデルで相対的治療効果という評価項目を用いるべきであり、当該モデルのベースラインの割合がメタアナリシスの試験の比較対象イベント率と大きく異なる場合に、特に重要である。

- 5.3.12 関連する技術のグループが、認められている「区分 (class)」の一部として正式に指定されているかどうかにかかわらず、似ているが必ずしも同一ではない効果を有する場合がある。当機構が1回の評価内で多数の関連する技術の評価している場合、当該評価についての最終スコープで別途指定されない限り、通常は区分の効果および個々の効果の両方に基づく分析を実施すべきである。

#### 間接治療比較および混合治療比較

- 5.3.13 入手可能であれば、リファレンスケース分析では、*head-to-head RCT*のデータを提示すべきである。*head-to-head RCT*が存在し、直接比較では入手できない情報を追加することを検討する場合、混合治療比較分析によるエビデンスを提示することが可能である。この混合治療比較は、十分に記述して、リファレンスケース分析の追加として提示しなくてはならない（「混合治療比較」には介入を直接的に比較した試験と間接的に比較した試験が含まれる）。1つの*RCT*内で比較されていない複数の技術の評価している場合は、一連の対比較の*head-to-head RCT*のデータを提示すべきである。また、直接比較では入手できない情報を追加することを検討する場合、混合治療比較の枠組みを用いた複合分析の提示を検討すべきである。*head-to-head RCT*のデータが得られない場合、間接治療比較を用いるべきである（「間接比較」とは試験のネットワークで得たデータの統合である）。また、混合治療比較でも間接治療比較でも、標準的メタアナリシスの適正実施の原則に従うべきである。
- 5.3.14 当機構には、*head-to-head RCT*によるデータへの選好がある。入手可能な場合は、リファレンスケース分析で*head-to-head RCT*のデータを提示すべきである。
- 5.3.15 「間接治療比較」とは、関心のある技術を*head-to-head*で比較していない試験からのデータの統合であり、その他の介入と比較した複数の試験によるネットワークからデータを取り出して関心のある技術同士を間接的に比較する手法である。当該技術を他の介入と比較する試験のネットワークのデータにより間接的に比較した「混合治療比較」とは、関心対象の介入を直接比較した試験および間接的に比較した試験を含む分析である。間接治療比較や混合治療比較を実施する場合は、システムティックレビューおよびメタアナリシスの適正実施の原則に慎重に従うべきである。含めた治療比較の選択の根拠など、*RCT*の同定および選択の根拠を説明すべきである。統合の方法を明確に記述することが求められる。個別の試験の方法および結果を文書化すべきである。特定の試験の関連度に疑問がある場合は、これらの試験が除外される感度分析を提示すべきである。対比較の結果間の不均一性および技術に関する直接および間接エビデンス間の不一致を報告すべきである。

- 5.3.16 head-to-head RCTのデータが理想よりも少ない場合がある（例：標本サイズが小さいまたは外的妥当性に懸念がある場合）。このような場合は、混合治療比較によってエビデンスを追加することが検討される。また、リファレンスケース分析とは別に混合治療比較およびその組み入れの根拠を提示すべきである。これに関しても、適正実施の原則を適用する。
- 5.3.17 複数の技術を評価している場合、各技術を直接比較したRCTのデータを一連の対比較で提示すべきである。混合治療比較の枠組みを用いた追加分析の提示を検討することができる。このような場合、評価委員会は、統合方法および試験の組み入れまたは除外の適切性を特に参照して、両方の分析の結果を検討する。
- 5.3.18 技術（および/または比較対象）のhead-to-head RCTのデータが入手できない場合がある。そのような場合、間接治療比較分析を検討すべきである。
- 5.3.19 間接治療比較または混合治療比較の枠組みを使ってエビデンスを結合する場合、試験の無作為化を維持しなければならない。データが観察的データとして扱われ、バイアスの可能性および不確実性の増大を調整するための適切な手段が講じられない限り、異なる無作為化試験の単一の投与群の結果を比較することは許容されない。
- 5.3.20 間接治療比較または混合治療比較の枠組みを使った解析には、それらがエビデンスのネットワーク化に関連する場合、評価の範囲で定義されていない比較対象介入（プラセボなど）が含まれる場合がある。比較対象介入の組み入れおよび除外の根拠は明確に報告すべきである。これに関しても、適正実施の原則を適用する。
- 5.3.21 混合治療比較によるエビデンスは様々な方法で提示可能である。エビデンスのネットワークは表形式で提示可能である。また、直接治療比較および間接治療比較が明確に特定され、各比較の試験数が示されていれば、エビデンスを図示することも可能である。
- 5.3.22 head-to-head試験のメタアナリシス、あるいは混合比較または間接比較に含めるのに、関連する妥当なデータが十分入手できない場合、当該分析を質的概要に限定して、各試験を批判的に吟味し、その結果を提示しなければならない可能性がある。このような場合、評価委員会は分析結果のレビュー時に特に注意する。

## 5.4 健康効果の測定および価値評価

- 5.4.1 費用効果分析では、健康効果の価値を適切な時間地平のQALYで表記すべきである。リファレンスケースでは、HRQLの測定値の変化は患者から直接報告されるべきであり、患者HRQLの測定値の変化の価値（つまり効用）は選択法を使った、一般の人々の選好をもとに測定されるべきである。EQ-5Dは成人のHRQLに推奨される評価項目である。EQ-5Dの効用値を導き出す方法を完全に記述すべきである。EQ-5Dデータが利用できない、またはその疾患や治療効果に適切ではない場合は、価値評価の方法を完全に記述し、EQ-5Dで用いた方法と同等とすべきである。疾患固有の、選好による評価項目を使って収集したデータは非リファレンスケース分析で示すことができる。公表文献の推定効用値の使用については、それらが系統的に特定、選択されていることを示すエビデンスによる裏付けがなければならない。
- 5.4.2 QALYは、生存期間をその期間のHRQL評価で重み付けした評価項目である。HRQLの「重み付け」には通常、HRQL自体の変化およびその評価を示す2つの要素がある。治療の結果であるHRQLの変化に関する情報は患者から直接（および介護者の健康に対する治療の影響も重要な場合は、介護者から直接）報告されるべきである。患者によって報告されるHRQLの変化に対する価値付けは、イギリス人集団の代表標本で選択するという方法を使って導かれた一般の人々の選好に基づくべきである。
- 5.4.3 患者からHRQLの変化について直接情報を得ることができない場合は、（医療従事者からではなく）患者の介護者からデータを得るべきである。患者（または介護者）で測定したHRQLの変化の価値付けは、イギリス人集団の代表標本の一般の人々の選好の価値評価に基づくべきである。
- 5.4.4 HRQLへの技術の効果を定量化するためには、EQ-5D（標準化され妥当性が証明されている一般的ツール）が望ましい。分類システムが異なれば、効用値が異なる。したがって、異なるシステムを用いた結果は常に比較できるとは限らない。当機構の業務の特性である比較性と、評価全体にわたる一貫性の必要を考えると、HRQLの測定および評価には単一の分類システム、すなわちEQ-5Dが望ましい。
- 5.4.5 EQ-5Dは広く使われているHRQLの評価項目であり、多くの異なる患者集団で妥当性が確認されている。EQ-5Dは、移動の程度、身の回りの管理、普段の活動、痛み/不快感、および不安/ふさぎ込みという、健康に関する5つの領域からなる。本システムは、標準化された記述システムを用いて、患者がHRQLを表現できるように設計されている。EQ-5Dの分類システムでは、選択ベースの価値の評価方法（時間得失 [time trade-off] 法）による大規模なイギリス人集団の研究で明らかになった一連の選好値が利用可能である。この一連の値を患者の自己報告によるHRQLの記述に適用して、健康関連の効用値を導くことができる。

- 5.4.6 EQ-5Dツールを使ったデータが常に利用できるとは限らない。EQ-5Dデータが利用できない場合は、関連する臨床試験に含まれる他のHRQL測定項目によるEQ-5D効用データをマッピング（「cross-walking」ともいう）する方法をEQ-5D効用データ推定に使える。これは、適切なマッピング機能が明らかにされており、その妥当性が証明されれば、可能である。マッピングは経験的データに基づいて行い、当該マッピング機能の統計的特性を明確に記述するべきである。
- 5.4.7 EQ-5D効用データが利用できない場合は、関連する臨床試験に含まれる標準化され妥当性が確認されたHRQL評価項目に基づく、健康状態の記述の直接評価を提出することができる。このような場合は、その記述の評価では、イギリス人集団の代表標本において時間得失法（time trade-off）法を使い、上限を「full health（完全な健康）」とし、EQ-5Dの評価に使う方法と方法論的に一致するようにすべきである。
- 5.4.8 疾患に固有の、選好に基づいた評価項目を使って関連する臨床試験から直接収集されたデータは、別の経済分析で提示すべきである。
- 5.4.9 EQ-5Dは、全ての場合に適切な健康関連効用の評価項目というわけではない。EQ-5Dが不適切と考えられる場合は、EQ-5Dの特性が特定の集団に適さない理由について経験的エビデンスを示すべきである。これらの特性には、EQ-5Dの内容妥当性、構造妥当性、反応性および信頼性が含まれる。代替評価項目が望ましい場合は、分析を提出する際に、その理由を示し、使用するツールの特性に関する経験的データをその根拠として示すべきである。また、ツールの選択が獲得QALYの評価にどの程度影響を及ぼすかについて評価委員会の理解をサポートするエビデンスを示すべきである。EQ-5D以外のHRQL評価項目に基づく健康状態の記述の直接評価を使う場合、当該評価方法はEQ-5Dに使う方法と同等でなければならない（セクション5.4.5を参照）。
- 5.4.10 現行版のEQ-5Dは小児に使うために設計されていない。必要な場合は、標準化され妥当性が証明された別の選好ベースのHRQL評価項目、例えば小児専用設計されたHealth Utility Index 2（HUI 2）などを用いることを検討すべきである。
- 5.4.11 健康関連効用値を文献から入手する場合、系統的で、透明性が高いデータ特定方法を用いるべきである。特定のデータセットを選択する根拠を明確に説明すべきである。リファレンスケースの基準に適合しない健康関連効用データには、そのデータを得るのに用いた方法の詳細な説明、およびそれらの方法が値に及ぼす影響の検討を記載すべきである。複数の妥当な健康関連効用データが利用可能な場合は、感度分析を実施すべきである。

## 5.5 資源の使用および費用に関するエビデンス

### NHSおよびPSSの費用

- 5.5.1 リファレンスケースについては、比較する技術間に費用の差異効果が生じる可能性がある場合は、費用をNHSおよびPSSのコントロール下にある資源と関連づけるべきである。これらの資源をNHSおよびPSSに関連する価格を使って評価すべきである。資源の使用および費用のデータが系統的に特定されていることを示すエビデンスを提示すべきである。
- 5.5.2 資源に対して支払われた取得価格が公定価格と異なる場合（例：医薬品や医療機器がNHSの機関に割引価格で販売された）、リファレンスケース分析には公定価格の方を用いるべきである。感度分析でこの価格からの変動の影響を評価すべきである。NHSへの割引価格に基づく分析は、割引価格の透明性が高くNHS全体で一貫して利用可能であり、当該価格が一定期間保証されている場合にのみ検討する。これらの場合、NHS購買供給局(PASA:Purchasing and Supply Agency)またはウェールズ保健供給機関(Welsh Health Supplies)などから助言を受けられたい。評価のレビュー日は、当該合意が保証される時点までに知らされる。
- 5.5.3 公定価格および国の機関によって合意された価格がない場合（特定の診断技術の場合が該当する）は、国内で公的に入手可能であれば、製造業者から提出される価格を用いることができる。
- 5.5.4 リファレンスケースの視点では、NHS/PSSに関連した財務費用は原価計算の基礎として使用するのが適切だが、これらは任意の資源に関する社会的機会費用を必ずしも完全に反映しない可能性がある。可能な限り、単位原価の推定および特定の資源の価格を評価全体で一貫して用いるべきである。当該費用および価格を特定する際の第1の基準点は、保健省および/またはウェールズ議会政府(Welsh Assembly Government)によって公表されている現行の公定価格である。
- 5.5.5 資源の使用および単位費用データの特定方法は、臨床的効果の特定に関するエビデンスほどには十分には定義されていない。Payment by Results tariffなどのHRGに基づくイギリス国内のデータは有益な情報源であり、それらが適切かつ利用可能な場合は、使用を検討すべきである。HRGによるデータは全ての状況で適切というものではない（例：HRGの定義が広範である、または評価対象の技術に関して平均費用が資源の使用を反映しないと考えられる）。そのような場合、ミクロの原価計算研究など、他のエビデンス情報源がより適切である可能性がある。費用データが文献に由来する場合は、情報源を特定するのに使う方法を定義すべきである。代替りの情報源が複数ある場合は、選択した費用の妥当性を示して、情報源間の不整合を説明すべきである。



該当する場合は、感度分析を使って、別のデータ源の使用が結果に及ぼす影響を評価すべきである。

- 5.5.6 関心疾患に関連する費用および治療の結果として生じる獲得生存年における費用をリファレンスケース分析に含めるべきである。疾患または対象の技術に関連しないとみなされる費用は除外すべきである。
- 5.5.7 技術の導入に新しくインフラを整備する必要がある場合は、当該費用を分析に含めることを検討すべきである。
- 5.5.8 関連する技術のグループを治療の「区分（クラス）」の一部として評価している場合は、通常、各技術の単位原価を用いる分析をリファレンスケースで提示すべきである。例外として、非常に広範な技術および費用を検討する必要がある場合は、少なくとも最大および最小の原価推定値を使った分析を提示すべきである。
- 5.5.9 付加価値税（VAT）は全ての経済評価から除外すべきである。ただし、対象の資源がこの課税の対象である場合は、適切なレート（現在は17.5%）で予算影響の計算に含めるべきである。

#### **NHSおよびPSS以外の費用**

- 5.5.10 技術の中には、他の政府機関への費用（または費用削減）にかなりの影響を及ぼすものがある。これらの例外的な状況では、通常、当該テーマを委託する前に、保健省と個別に合意を得ていれば、他の政府機関に対する費用を含めることができる。非リファレンスケースの分析がこれらのより広範な費用を含む場合は、価値評価の方法を明確にすることが求められる。いずれの場合もこれらの費用はNHS/PSS費用とは別に報告すべきである。これらの費用は、増分費用効果比（ICER：QALYが関心アウトカム評価項目である場合）に統合すべきではない。
- 5.5.11 患者が負担する費用については、当該費用がNHSまたはPSSによって償還される場合は含めることができる。償還率が患者間または地域間で異なる場合は、当該費用を全患者で平均すべきである。生産性費用並びにNHSおよびPSSによって償還されない患者負担費用は除外すべきである。

## **5.6 割引**

- 5.6.1 *費用対効果の結果は、分析の時間地平に生じる一連の費用・便益の現在価値を反映すべきである。リファレンスケースでは、費用および効果の双方に年率3.5%の割引率を用いるべきである。結果が使用する割引率に対して高い感度を有する可能性がある場合、費用およびアウトカムに異なる割引率を用いた感度分析および/または割引率を0%～6%の間で変動させる感度分析を検討する。*

5.6.2 現在価値の割引の必要性は経済評価では広く受け入れられているが、その率は区域および経時的に変動する。当機構では、費用および健康効果を同じ率で割り引くことが適切であると考え。費用割引に関するイギリス財務省の推奨に基づく年率3.5%を費用および健康効果の両方に適用すべきである。

## 5.7 モデリング法

5.7.1 利用可能なエビデンスを統合して、当機構の要求について、臨床的效果および費用対効果の推定値を算出するために用いるモデルは、容認されたガイドラインに従うべきである。構造仮定およびデータインプットの完全な文書および妥当性を示すべきである。他に妥当な仮定およびインプットがある場合は、モデルのアウトプットに対するそれらの影響について感度分析を実施すべきである。

5.7.2 セクション5.1.3に示したように、モデリングは、利用可能なエビデンスを統合し、評価委員会の意思決定プロセスに関連する形式で臨床的效果および費用対効果の推定値を算出する重要な枠組みとなる。モデルはほとんどの技術評価で必要である。モデリングが必要になる状況として以下が挙げられる。

- 1つの試験に関連する全てのエビデンスが含まれていない場合
- 試験に参加している患者が、NHSで当該技術を用いると考えられる典型的な患者に適合しない場合
- HRQLおよび生存率への影響ではなく、中間アウトカム評価項目を用いる場合
- 関連する比較対象が使われていないまたは試験に関連するサブグループのエビデンスが含まれていない場合
- 当該技術の費用および便益が試験の追跡期間よりも長期となる場合

5.7.3 質の高いモデルの構成要素の包括的な定義は不可能だとしても、ガイドラインがいくつかある。一般的に、全ての構造仮定について完全に妥当性を確認すべきであり、代替手段の妥当なレビューにおいてデータインプットを明確に文書化し、妥当性を確認すべきである。これは、特に、関連する複数の試験から選択した1つないし2つの臨床試験を用いた際に生じる選択バイアスに類似のバイアスを生む外れ値の選択を避けるために重要である。健康効果の推定値は、システマティックレビューの結果に基づくべきである。モデリングは、多くは、長期の時間地平の費用および健康効果を外挿するために必要とされる。治療効果の外挿に用いる仮定は、臨床的妥当性を有し、その報告では透明性が維持され、妥当性が明確に証明されているべきである。結果について外挿に関する仮定が異なる場合の影響を比較する別のシナリオも検討すべきである。例えば、治療効果の持続期間について、外挿フェーズの治療効果には次のようなシナリオがある。すなわち (i) なし、(ii) 治療フェーズと同一でかつ同じレベルで持続する、または (iii) 長期の間に減少する。

- 5.7.4 試験データが、対象集団の特定の健康アウトカムまたはイベントのベースラインリスクを定量する上で十分でないことがある。健康アウトカムのベースラインのリスクおよび比較対象介入を用いた場合の当該疾患の経過を定量することは、経済分析における絶対的健康アウトカムを推定する際に有用なステップとなる。その後、対象集団またはサブグループの健康アウトカムのベースラインリスクに関するデータに無作為化試験の相対的治療効果を適用することができる。これらの推定値のデータ源を特定し、批判的吟味を行う方法を示して、正当化すべきである。
- 5.7.5 技術の使用が診断検査のアウトカムの条件である場合は、検査の精度および関連費用を臨床的効果および費用対効果の評価に組み込むべきである。
- 5.7.6 モデルの作成で使う品質保証の方法を詳述し、モデルの妥当性評価の方法および結果を提示すべきである。また分析結果を集計していない形で提示すべきである。これには、獲得生存年、死亡率（該当する場合は、個別の時点の死亡率）、および当該モデルで予測される臨床イベントの頻度の推定値に関する情報を提示することを含めるべきである。
- 5.8 潜在的バイアスおよび不確実性の評価**
- 5.8.1 モデルへのインプットにおける潜在的な選択バイアスを特定することが重要であり、また、当該モデルでは、技術に関連する決定の不確実性（つまり、各技術の真の費用対効果が意思決定前に確定される場合に別の決定が導かれる確率）を定量することが重要である。
- 5.8.2 モデルを構築する際は仮定を作ることが必要とされている。潜在的バイアスおよびそれに伴う不確実性を「構造的な不確実性」と称することがある。構造的な不確実性の例として、異なる健康状態のカテゴリー化や異なる治療パスの提示などが挙げられる。これらの構造仮定を明確に文書化し、それらを裏付けるエビデンスおよび根拠を示すべきである。代表的な範囲の妥当なシナリオを使った別の分析で費用対効果の推定に対する構造的な不確実性の影響を調査すべきである。
- 5.8.3 潜在的バイアスの2番目のタイプは、費用および効用の違い、相対的効果の推定値、およびそれらの持続期間などの主要なパラメータの値を示すデータ源の選択的使用で生じる。主要パラメータの推定値が異なる影響を（例：別のシナリオにより）感度分析に反映させなければならない。インプットを完全に正当化し、別のインプット値による感度分析で不確実性を調査しなければならない。

- 5.8.4 不確実性の3番目の発生源は、最適な情報源が特定された後のパラメータの精度から生じる（つまり、当該モデルの平均的な健康および費用のインプットに関する不確実性）。パラメータの平均値（の精度）に関連した不確実性を評価するために各パラメータに分布を割り付けるべきである。確率論的感度分析が望ましい。これによってパラメータ関連の不確実性をモデルの結果に同時に反映できるようになる。非線形の決定モデルでは、確率論的方法によって平均費用およびアウトカムの最良推定値が示される。
- 5.8.5 モデルに含める各パラメータについて、平均値、平均値周辺の分布、および情報源並びに裏付けとなるエビデンスの根拠を明確に記述すべきである。確率論的感度分析で選択される分布は、任意ではなく、関心パラメータに関する利用可能なエビデンスを示すように選択し、その選択を正当化すべきである。平均値およびパラメータの関連分布を明らかにするためのデータが不足している場合は、正式な導出法を利用できる。パラメータ値の不確実性を示すために用いることができる妥当な分布が他にある場合は、これらのシナリオによる別の確率論的解析で調査すべきである。
- 5.8.6 個別のパラメータ間の相関度に関するエビデンスを慎重に検討して、確率論的解析に反映すべきである。相関度についての仮定を明確に提示すべきである。
- 5.8.7 適切なモデル構造を実施するために用いる計算法が、時に確率論的感度分析に問題を生じさせることがある。確率論的感度分析の実現可能性を制限するようなモデル構造の利用は明確に規定し、正当化されるべきである。モデルは常に目的に適合し、モデル構造およびインプットパラメータに関連する決定の不確実性を入念に考察できるものでなければならない。「望ましい」モデル構造またはプログラミングプラットフォームの選択によって不確実性の証明が不明確になってはならない。

## 5.9 データおよび結果の提示

### データの提示

- 5.9.1 臨床的効果および費用対効果を推定するためのパラメータは全て表形式で明確に示し、データ源に関する詳細を含むべきである。連続変数については、分析において平均値を提示・使用すべきである。全ての変数については、精度の評価項目を詳述すべきである。確率論的解析では、インプットパラメータの不確実性の評価に使った分布を文書化し、正当化すべきである。分析に用いたデータについては、可能な限り詳細に示すべきである。

### 費用対効果の期待値の提示

- 5.9.2 費用の各項目の期待値および期待合計費用を示すべきである。分析で比較した各選択肢の期待QALYも主な関連項目という観点から詳述すべきである。ICERを必要に応じて算出すべきである。
- 5.9.3 関心のある介入および比較対象の治療パスについて、費用およびQALY双方で構成される主なコンポーネントを表にすべきである。QALYについては、これに別途、生存年の提示が含まれる。疾患の異なるステージに関連した費用およびQALYを別途示すことを考慮すべきである。費用およびQALYを統合する際は、標準的な決定の規則に従うべきである。これらは劣位または拡張劣位である状況を反映すべきである。報告するICERは、代替治療と比較した期待追加QALYに対する期待追加合計費用の比でなければならない。ICERの他にも、1QALYあたり£20,000および£30,000として、期待される金銭的純便益または健康純便益を示すことができる。モデルがパラメータの非線形的結合から成る場合は、確率論的感度分析を使って費用およびQALYの平均を算出すべきである。そのようなモデルでは、パラメータを平均値に設定しても費用およびQALYの正確な平均推定値にならない。

### 費用効果分析での構造仮定の不確実性への対処

- 5.9.4 分析に使われる構造仮定の不確実性を調査するためには、感度分析を用いるべきである。代表的な範囲の妥当なシナリオの分析を提示し、別の各分析について個別に結果を示すべきである。
- 5.9.5 費用対効果の結果の不確実性に関する重要な要素は、決定モデルの構造における不確実性から生じる。決定の不確実性に関する全パラメータでの不確実性の分析では、モデルの構造やデータインプットなどの要因が適切とみなされていると仮定される。ただし、当該モデルのこれらの特性も不確実性の対象であり、感度分析を使って特定し、正式に調査すべきである。
- 5.9.6 このタイプの感度分析を実施すべき一般的な例を以下に示す。
- 試験のフォローアップ期間を超える費用およびアウトカムの外挿に使う最適な仮定に不確実性がある場合
  - 分析において治療パスを最適に示す方法に不確実性がある場合
  - 規模の経済性が考えられる場合（例：診断技術の評価において）

- 5.9.7 リファレンスケースで使われる方法の適切性に関する不確実性も感度分析を用いて扱われるが、これらの分析は別々に示さなければならない。

#### **費用効果分析におけるデータ源の選択にある不確実性への対処**

- 5.9.8 適切なデータ源の選択にある不確実性は感度分析を使って対処すべきである。これには、パラメータ値のデータ源選択にある不確実性が含まれる。不確実性の当該データ源は感度分析、望ましくは確率論的分析法を使って調査すべきである。
- 5.9.9 分析に含めるデータの情報源が明確でない場合がある。このような場合、別の情報源を使って、または疑いのある試験を除外して、再度分析を行い、その結果を別途報告すべきである。このようなシナリオ分析を実施すべき場合の例を以下に示す。
- 当該疾患/介入に関連する健康関連効用に関する妥当な別のデータがある場合
  - 特定の資源またはサービスの費用あるいは特定の技術の取得価格について、病院間でばらつきがある場合
  - メタアナリシスまたは混合治療比較における特定の試験の質または関連度に疑問がある場合

#### **費用効果分析におけるパラメータの不確実性への対処**

- 5.9.10 分析で使われる全てのインプットは、不正確の程度で推測される。全てのインプット変数にある不正確さを、比較している選択肢の費用対効果における決定の不確実性の評価項目に変換するには、確率論的感度分析が望ましい。
- 5.9.11 費用対効果データのパラメータの不確実性を示す適切な方法には、費用効果平面上の信頼楕円や散布図（比較が2つの選択肢に限定される場合）および費用効果受容曲線がある。費用効果受容曲線を示す際は、費用対効果受容フロンティアおよびその説明を含めるべきである。不確実性も、表形式で提示すべきである。予測される結果の平均（費用、アウトカムおよびICER）の詳述に加え、1QALYあたり£20,000～£30,000の閾値で治療が費用対効果に優れる確率および誤差確率（治療が費用対効果に優れない確率）も示すべきである。3つ以上選択肢がある場合は特にそのようにすべきである。
- 5.9.12 単変量解析およびベスト/ワーストケース感度分析は、費用対効果にかなりの影響を及ぼすと考えられるパラメータを特定し、当該モデルの主要な推進因子を説明する重要な方法である。ただし、この分析は、パラメータの数が増えるにつれて、不確実性の複数の情報源の複合効果を示す上で有用でなくなる。

確率論的感度分析の使用によって、全てのインプットパラメータに関連する、パラメータ不確実性の評価が完全にできるようになる。確率論的分析において、情報の期待価値 (expected-value-of-information) 法を使って、各パラメータの不確実性が決定の不確実性全体に寄与する度合いおよびその影響をもたらすことができる。

## 5.10 患者サブグループのデータ分析

- 5.10.1 多くの技術について、患者ごとに特性が異なるため、治療で得られる効果のキャパシティも異なる。この点について、リファレンスケース分析の一貫として、関連する患者サブグループごとに、臨床的效果および費用対効果の推定値を示すことにより調査すべきである。サブグループ内の患者特性を明確に定義すべきである。望ましくは、既知の生物学的に妥当な機序、社会的特性または明確に正当化された因子による臨床的效果または費用対効果の差を事前に推測し、これに基づき、サブグループ内の患者特性を特定すべきである。可能な場合、スコوپング段階で、関連すると考えられるサブグループを特定し、サブグループ効果の予測の根拠を示すことを検討する。ただし、これによって、後のプロセス、特に評価委員会の審議におけるサブグループの特定が除外されるわけではない。
- 5.10.2 当機構の目的が限られた資源から最大の健康効果を得ることであることを考慮すれば、患者集団の特性の違いのために臨床的效果および費用対効果がいかに異なるかを検討することは重要である。一般に、治療で得る効果のキャパシティは患者間で異なるが、このことはその後の治療費にも影響を及ぼす可能性がある。明確な根拠を示すとともに、適切な場合、患者サブグループの定義および差異効果の予測に関する生物学的妥当性を示すべきである。サブグループの効果の研究において「明らかになった」事後データは避けるべきであり、懐疑的にレビューする。
- 5.10.3 介入の全体的正味効果の推定は、特定の疾患またはイベントのベースラインリスク、および/または関連する比較対象治療と比較した当該技術の相対的效果によって決まる。治療の全体的正味効果は、対象集団を形成する人のその他の特性によっても決まる。したがって、関連するサブグループは、絶対的治療効果に対する1つないし複数の寄与因子の相違で特定することができる。
- 5.10.4 特定の健康アウトカムベースラインリスクの差に基づくサブグループについては、それを定量化するデータの系統的な特定が必要である。サブグループ解析を目的として適切なベースラインデータを特定する方法を十分詳細に示し、反復および批判的吟味を可能にすることが重要である。

- 5.10.5 効果の変化を定義するスケールの選択など、サブグループ解析の実施方法を規定する際は注意すべきである。サブグループ推定値全ての統計的精度は、パラメータの不確実性の分析に反映すべきである。提示したサブグループに関連する患者特性を明記して、評価委員会が決定問題に関して分析の適切性を判断できるようにすべきである。
- 5.10.6 RCTまたはシステマティックレビューで実施される標準的なサブグループ解析は、相対的な治療効果の相違に基づくことが多い（技術の効果と患者特性の相互作用の分析による）。特に、多数のサブグループが報告される場合は、偶然相違が生じる可能性を調査すべきである。
- 5.10.7 評価委員会は、サブグループ解析の検討で、効果を支持するエビデンスの強さの他にも、サブグループ効果の生物学的または臨床的妥当性に特に注意する。サブグループの効果の生物学的または臨床的妥当性を裏付けるエビデンスは、統計的分析の詳細を含め、完全に文書化すべきである。
- 5.10.8 サブグループに固有のパラメータの推定には、利用可能な場合、個々の患者データの使用が望ましい。ただし、全てのエビデンスについて、当該データの適切性の評価では、分析の質、利用可能なエビデンスの代表性、および決定問題との関連度などの因子を必ず検討する。
- 5.10.9 特定の状況では、差額原価に基づくサブグループの検討が適切な場合がある。例えば、特定の複合治療を管理する費用は、特定のサブグループで異なることが知られている。
- 5.10.10 評価委員会は、サブグループを検討する際に、人権、差別および平等に関する法規上の責務に特段の注意を払う。
- 5.10.11 次の因子にのみ基づくサブグループは、関連があるサブグループとみなされない。
- 健康状態および患者の選好に関する個人的効用
  - 社会的特性に従った個人に対する治療費用の差異にのみ基づくサブグループ
  - イギリス国内の異なる地域で治療を提供する費用という点で決められたサブグループ（例：技術の提供に使える施設の費用が場所によって異なる場合）
- 5.10.12 任意の期間に規定の「レスポンス」を示した人々にのみ治療を継続することで費用対効果を最大にする「治療継続ルール」の分析は、別のサブグループとして分析しないこと。むしろ、「継続ルール」をリファレンスケースの介入および比較対象とともに、新たな治療戦略とみなして、「継続ルール」が関与する戦略を別のシナリオとして分析すべきである。これによって、継続ルールに関連する追加モニタリングなどの費用および健康効果の因子を経済分析に組み込めるようになる。継続ルールに関するその他の検討事項には次のようなものがある。



- 当該規則の根拠となるエンドポイントの頑健性および妥当性
- 当該規則で定義された「反応」基準が合理的に達成できるかどうか
- 反応を測定する時期の適切性および頑健性
- 当該規則を日常的な診療に組み込めるかどうか
- 当該規則により、当該技術が特に費用効果が高い患者を予測できる可能性があるか
- 効果が認められない患者の治療中止に関する問題およびその他の衡平に関する検討事項

## 5.11 エビデンスに基づく将来の研究ニーズの特定

- 5.11.1 *考えられる将来の研究のテーマは、システマティックレビューおよび費用効果分析によって特定されるエビデンスのギャップに基づき見出すことができる。決定の不確実性を抑制するという観点から新たな情報の価値を検討することにより、これらを最優先にすることができる。*
- 5.11.2 不確実性分析の一部は、意思決定の感度が最も高いパラメータおよび構造の不確実性を特定することである。そのような情報は将来の研究の優先順位に関する決定に取り入れることができる。費用効果分析の一環として、確率論的感度分析による正式な value-of-information（情報価値）法が利用可能であり、パラメータの不確実性を抑制するための新たな研究の金銭的価値やその研究に焦点を当てるべきかを判断することができる。
- 5.11.3 さらに研究の勧告が、当機構の研究開発諮問委員会（Research and Development Advisory Committee）が合意したプロセスおよび基準を使って優先される。また、NHS 研究開発プログラム（Research and Development Programme）などの研究に出資する機関に研究に関する勧告をする。

## 5.12 費用効果分析における衡平に関する検討事項の反映

- 5.12.1 *リファレンスケースでは、健康効果を得る人々のその他の特性にかかわらず、追加の QALYに同じ重み付けを行う。*
- 5.12.2 リファレンスケースで定義されたとおり、QALYの推定とは、個人間で獲得された健康効果の比較に関してある一種の見方を意味している。したがって、追加のQALYは、社会人口学的特性または健康の治療前後のレベルなど、個人のその他の特性にかかわらず同じ値である。QALYの算出に対する新たな重み付けの適用方法および条件に関しては方法論的問題がいくつか未解決であるため、リファレンスケースの一環として異なるQALYの重み付けを用いることは推奨されない。

## 5.13 NHSへの影響

### NICEガイドンスの実施

- 5.13.1 当該医療技術を実施することによるNHS（適切な場合、PSS）に対する最終的な影響の情報が必要である。
- 5.13.2 下に詳細を示したとおり、可能な場合、NHSの影響に関する情報には、主要な疫学および臨床的仮定の詳細、イングランドおよびウェールズの一般集団に関する資源単位および費用、並びに患者またはサービスの基準（例：100,000人あたり、プライマリケアの平均預託金あたり、または病棟あたり）を含めるべきである。

### 実施/取り込みおよび集団の健康への影響

- 5.13.3 現行のベースライン治療の割合に関するエビデンスベースの推定値、並びに評価対象技術およびNHSにおける比較対象技術の適切な実施/取り込み/治療の割合に関する期待値を示すべきである。また、任意の集団に結果的に生じる健康への影響の推定（例：獲得QALYまたは獲得生存年）を試算することが望ましい。その際、当該疾患の疫学およびNHSでの診断/治療の利用度に注意すべきである。また、主要な仮定または不確実性も取り上げるべきである。

### 資源への影響

- 5.13.4 新しい医療技術の実施は、評価中の技術および比較対象技術のNHSによる供給量に直接の影響がある（例：薬剤の投与量または手術時間（theatre hours））。また、当該技術は、代替治療または回避すべき治療、および新しい技術の使用をサポートするのに必要な資源など、その他のNHSおよびPSSの資源に連鎖的影響（増加/減少）を及ぼすことがある。それらには以下が含まれる。
- スタッフの人数および勤務時間
  - 教育訓練
  - サポートサービス（例：臨床検査）
  - 診療キャパシティ/施設（例：病床、診療時間、診断業務および滞在型の養護施設）。
- 5.13.5 評価する技術の実施をサポートするために必要な資源に関する制約の可能性があれば明確に示し、それが実施時間スケールに及ぼす影響についてコメントすべきである。

## 費用

- 5.13.6 国内および地方の有効な資金計画立案を可能にするため、予測される資源の影響に関するNHSの正味費用（適切な場合、PSS）の推定を提示すべきである。費用は、適切な包括的組織的（例：NHS、PSS、病院またはプライマリケア）および予算区分（例：医薬品、人件費、消耗品、および資本金）ごとに分類すべきである。可能な場合、これは資源単位情報に適用されているものと同じレベルおよび詳細にすべきである。節約が期待される場合は、実際に実現できる程度を特定すべきである。提供される費用は、VATの組み入れまたは除外を特定すべきである。費用の情報は、公表された費用分析または認められている一般に利用可能なデータベースまたは公定価格に基づく必要がある。
- 5.13.7 当該技術の実施が他のサービスにかなりの資源の影響を及ぼすような場合は、当該技術について提出された費用対効果のエビデンスを調査すべきである。
- 5.13.8 当機構は、NHSの個別の機関および地域の医療経済組織が、ガイダンスの地域予算への影響を迅速に評価できる原価計算ツールを作成している。原価計算ツールの開発方法に関する詳細については、当機構の文書「原価計算ツールの開発：方法ガイド（Developing costing tools: methods guide）」（付録C）に記載する。

## 6 エビデンスの評価および意思決定

### 6.1 緒言

- 6.1.1 本セクションの目的は、評価委員会がエビデンスを評価し最終結論に至る判断を行う方法を説明することである。
- 6.1.2 評価委員会は独立の諮問機関である。メンバーには、NHS職員、患者および介護者団体、関連する学術分野、並びに医薬品および医療機器業界が含まれる。評価委員会は、NHS内で使用される治療の臨床的効果および費用対効果について、当機構に勧告を行う。また、治療の患者への効果が証明されていない場合または費用対効果がない場合に、当該治療を使用しないよう勧告することも評価委員会の役割である。当機構はNHSへの最終ガイダンスの公布に責任を有する。
- 6.1.3 当機構への勧告を作成する際、評価委員会は各評価に最適とみなされている因子を検討する裁量がある。その際に、評価委員会はNICE設置令並びに人権、差別および平等に関する法規の条項を考慮する。医療技術の評価の実施において、当機構は次のように、保健大臣令 (Directions from the Secretary of State for Health) を考慮することを期待される (付録C参照)。
- 臨床的効果および費用の広範なバランス
  - 検討中の当該病態および疾患患者の臨床的ニーズの程度
  - 保健大臣がNHSに対して発行し、当機構に具体的に注意を促すガイダンス、および保健大臣によって発行されたあらゆるガイダンス
  - 技術革新がNHSの長期的な便益となる可能性
- 6.1.4 評価委員会は、科学的で社会的価値のある決定を行うための適切なアプローチについて当機構の助言を考慮する。社会的価値の判断に関する助言は、市民会議によって通知される。評価委員会が一般に検討するべき、社会的価値の判断を示すガイドラインを、当機構の文書「社会的価値の判断：NICEガイダンス作成の原則 (Social value judgements: principles for the development of NICE guidance)」に示す (付録C参照)。
- 6.1.5 当機構が作成するガイダンスの信用性は、評価委員会の意思決定プロセスの透明性に依存する。評価委員会の決定を明確に説明し、臨床専門家、専門的患者および評価期間中に協議に応じた人の見解の寄与度を検討することが不可欠である。検討された因子に関して、同委員会の決定の論拠をガイダンスの「検討事項」のセクションで説明する。