

最初に分析者は、既存の証拠に基づいて介入の作用を修飾するかもしれない「現実」の要因を特定しなければならない。2番目に、実行可能で科学的に信頼できる場合は、「現実」の要因を参考事例に組み入れ、それを入手可能な最良の証拠に基づいた重要な患者アウトカムと結びつける。例えば、Husereau *et al.*²⁶が行った評価では、臨床試験での高めの診断精度を地域慣行での診断精度のよりよい推定値に置き換え、日常的診療慣行で使用されるインフルエンザ薬の効果をより反映させるようにした。評価において「現実」の要因を説明できないと、結局は次善の治療戦略を選ぶことになってしまう。

「現実」の要因を分析に組み入れるかどうかを判断するとき鍵となるのは、作用修飾因子を重要な患者アウトカムと結びつける入手可能な証拠の強さである。こうした証拠は臨床試験で得られた2次的なアウトカムではなく、1次的なアウトカムに基づいていることが望ましい。こうしたデータは介入が開始される前には無いのが普通であり、他の情報源からのデータが役に立つであろう。例えば、外国の「現実」での遵守パターンに関するデータや回顧的なデータベースからのデータなどがある。

3.6.4 不確実性と層別分析

介入の「現実」での体験が始まる前に、効果に影響を与える要因の推定値については高度な不確実性がある。データが入手できない場合、あるいは入手できるデータの質が低い場合、効果の推定値に関する不確実性は感度分析を通じて査定しなければならない。ただしこうした分析が質の高い証拠に基づいたものでない限り、それは推論的、あるいは探索的なものであると認識する必要があり、報告書には適切な警告も記載しなくてはならない。

例えば、臨床上の便益の持続期間と大きさについては、妥当と思われるシナリオに基づいてモデル化することができる。経時的な効果の大きさもシナリオの違いによって、臨床試験期間を過ぎると「効果がまったくない」から「効果が弱まる」、あるいは継続して「全面的に効果がある」まで変わってくる³⁰。さらに、測定対象期間がより長いシナリオの場合は、より重要な患者アウトカム（C型肝炎の感染による重度の肝臓合併症など）やQALYが含まれる可能性がある。代替アウトカムの場合は短めの期間のほうが適切であろう（C型肝炎感染の抗ウイルス療法に対する持続的ウイルス反応など）。

代替的な方法としては、証拠を集めて分析するのにベイズの反復的枠組みを使うことである。分析者は市場導入前臨床試験で得られた効力に関するデータを、事前の証拠（あるいは仮定）を使って効果に変換することができる。場合によっては、その介入がカナダの市場に出る前に他国の市場にすでに出回っているような場合、実証的証拠を使ってこれを行うことができる。また、カナダですでに使われている類似の介入の調査を行うことも理に適ったことである。その他の場合には、専門医やパラメーターについて知識を持っている

ような他の人々から得た情報に基づいた、主観的な事前予想になるかもしれない。介入が実際に開始された後、費用と効果に関する暫定推定値は、「現実」の費用と効果に関するデータを集めて、更新することができる。

層別分析は、ターゲット集団中の下位集団を対象に、介入の効果における変動の結果に対する影響を評価するときを使用すべきである。これは下位集団において、例えば、リスクプロフィール（高リスクの患者対平均的リスクの患者など）、発症率、病状の進行、あるいは医療の利用などにおける差異に起因して起こる効果の可変性を評価するのに使うことができる。下位集団におけるパラメーターの確実性のレベルは、母集団におけるそれよりも低い。したがって、層別分析を行う場合、感度分析は増加した不確実性を明白に考慮しなければならない。

3.7 測定対象期間

測定対象期間は、介入とコンパレーター間の費用とアウトカムにおけるすべての有意義な違いを捉えることができるほど長いものでなければならない。しかし、代替策の費用とアウトカムが収束するときのように、有意義な違いがまったく出ない期間を超えて測定対象期間を延長する必要はない。分析の一貫性を保つため、費用とアウトカムには同じ測定対象期間を適用すべきである。

分析者には、特に慢性症状（糖尿病、関節リウマチなど）の場合、あるいは代替策が死亡率に異なった作用を持つ場合には、初期設定として生涯の測定対象期間を考慮することが奨励される。それより短い期間を使用した場合（急性疾患についてなど）、その理由を説明しなければならない。場合によっては、外挿されたデータに対しては複数の測定対象期間が適切かもしれない。一部の慢性症状については、測定対象期間が1年と5年、およびそれ以上長い期間の代替シナリオが適切であろう。

長期の分析だからといって、それほど長い期間にわたって1次データを患者から収集しなければならないということではない。長期のデータ収集は実行できない、あるいは実際的でないことが多いため、中間アウトカム（あるいは代用アウトカム）に関して集められた実データに基づいて短期の分析を行い、次に外挿またはモデル化されたデータに基づき、症状にふさわしいより重要な患者アウトカムに関するもっと長期の分析を行うことにより、短期の分析を補完することもできる。例えば、HIV感染に対する非常に積極的な抗レトロウイルス療法の場合、短期の測定対象期間を使ってウイルス反応および免疫系代替マーカに関する試験データの分析を行い、それを補完する形で、数年後に起きるかもしれない重篤な日和見感染および早死にを考慮した長期（生涯）の測定対象期間の分析を行うこともできる。

分析で使用された外挿手法の影響を査定するために、代替的な測定対象期間の分析を別に提示しなければならない。入手可能な最良の証拠を使用し、データを外挿するために使用した因果関係、手法、および仮定について説明する必要がある。

3.8 モデリング

医療技術の経済的評価では一般的に、新しい治療法の長期の増分費用やアウトカムを推定するため、複数の情報源からの証拠と仮定を統合するためのモデルを構築してから使用する。アウトプット（結果）はモデル構造、データ、使用する仮定によって決まってくるため、モデルはできる限り透明性のあるものでなければならない。したがって、意思決定者はモデルをベースにした評価の結果を検討するときには、批判的になる必要がある。

3.8.1 モデリングに関する考慮事項

優れたモデル³⁵は次のように定義することができる。

- ・ 本来の目的にぴったり合わせて作ってある
- ・ 特定の対象とする意志決定の参考となる
- ・ 容易に伝達できる

分析者は、モデルおよび分析の質を確保するために優れた慣行に従わなくてはならない。ここで要約する優れたモデリング手法は、2つの指針文書から引き出したものである。指針の詳細については、Philips *et al.*³⁰、および Weinstein *et al.*³³による文献を参照されたい。さらに、モデルの設計については、経済ガイドラインのすべてのセクションを考慮すること。

モデルの範囲（すなわち境界線）は説明し、その理由を述べなければならない。モデル構築が可能かどうかを査定してから、それをコーディングする。研究課題はモデル構築になくなくてはならないものである。課題が定義されると、分析者はモデリングがその問題にとって最良の方法かどうかを決めることが可能となり、使用すべき最も適切な手法を決めることができる。

モデルの範囲、構造およびパラメーターは研究課題と対象読者のニーズに関連したものでなければならない。モデルには対象とする症状のすべての重要な様相と、考えられる介入の潜在的影響が盛り込まれている必要がある。モデルは柔軟なものでなければならず、そうすれば管轄区域や支払機関の状況に適応することができる（多様な治療パターンが可能になるなど）。

モデルの全体的な設計と構造が決まると、実施できる分析の範囲が決まってくる。モデルがあると研究課題に関連する経済的評価の種類を決めるのが容易になり、研究結果を取り巻く不確実性を適切に評価することができる。分析者は経済的評価の反復的性質を忘れず、より多くのデータを入手したとき、結果の更新ができるようにモデルを構築する必要がある。例えば、モデルは現在進行中の臨床試験から有害事象のデータを組み込めるような構造にしてもよい。

モデルの構造、入力パラメーターの数値および出所はその正当性を示さなければならない。仮定およびモデル構造に関する主観的な判断（関係、含まれる変数、分布など）は、利用者が受容性を評価できるように、その理由を説明しておく必要がある。モデルは研究課題に適切に取り組み、対象読者の情報に関するニーズに対応できるだけの複雑さがあればよい³⁰。

モデルの構造は症状や治療経過（treatment pathways）の明細事項、関連する臨床事象、および因果関係に対応したものである。

モデルの構造はその症状に関する基調となる理論と一致し、介入や代替策の影響を捕捉するもので、研究課題と関連するものでなければならない。その構造は、慣行パターンが変化してもそのモデルによって評価できるように、現在の慣行パターンで規定してはならない。論理的、あるいは理論的に介入と代替策の間に差があると予想できないような臨床事象は構造から排除することができる。報告書には症状や治療経過を示した図表（決定樹など）を含めることが推奨される。

データが入手できるかどうかということで、モデルの設計を過度に制限してはならない。なぜなら資金供給や償還払いの決定はデータなしで行われることが多いからである。データを制限するとモデルの範囲が限定されるので、モデルの詳細な特徴を設計するときにはデータの制限を考慮することができる。データの限界は調査や、専門家の意見、モデルの代替物の利用、あるいは感度分析などの手法を使って対処することができる。

状態遷移（マルコフ）モデルを使用するときは、サイクルの長さを規定し、その理由を述べなければならない。周期は、病状や症状が変わると予想される最短のインターバルにする³⁰。

データ外挿（短期アウトカムから長期アウトカムへ、など）は適切な方法に基づいて行う。外挿の関係、手法、およびモデルで使用される仮定は説明をして、裏付け証拠の強さに言及しながら、その理由を述べなければならない。費用データに直線外挿法を使用すること

は、例えばスケールメリットが出てくるため、適切とはいえない³⁶。

モデルは正式にその正当性を確認しなければならない^{33,35,37}。正当性の確認では、モデルを検証して、予定していることを確実にしていることを確認する。

内的妥当性確認ではモデルによって生じた結果が内的に一貫していることを確認する。これは、モデルを数理論理学で検証し、数学的コードのエラーと管理上のエラー（標識付けやスペリングのエラー）をチェックすることによって行う。この方法には極値やゼロ値のテスト、既知シナリオの結果の調査、コードの確認が含まれる。直感に反した結果が出た場合は、それを説明する。より綿密な検証では、2つのソフトウェアを使ってモデルを組み立て、結果を照合する³⁰。

外的妥当性確認では基本的なモデル構造、仮定、およびパラメーターが適切で、病状のプロセス、および介入とコンパレーターの影響を正確に反映していることを確認する。結果は文字通りつじつまが合うものでなければならず、そうでない場合はその理由を説明する。感度分析はモデルで使用された構造的な仮定に関して不確実性を査定することができる（データ外挿の手法、重要な臨床アウトカムの最終アウトカムへの転換、治療経過など）。結果は他のモデルから出された結果と比較することができる（すなわち、モデル間の確認）。モデルの中間結果（健康アウトカムなど）は信頼できる独立したデータセット（全国癌統計など）に照らして調整したり、比較をする。差異が出たら説明を加えるか、モデルを修正するための参考とする。

確認プロセスは文書化する必要がある、こうした文書は意思決定者が希望すれば、いつでも見られるようにしておかなければならない。理想的にはモデルの検証は公平な人物が行うべきである。分析者は報告書の中でモデルの限界について述べ、モデルの検証を実施したかどうか、実施したのであればどのように行ったのかを記載しなければならない。

3.8.2 データに関する考慮事項

多くの問題がデータの適切な確認、収集、および分析、またデータのモデルへの組み込みに関係している。こうした問題は経済ガイドラインの扱う範囲外のことなので、これに関する指針として、読者は Philips *et al.*³⁰、および Weinstein *et al.*³³ による文献を参照されたい。

モデルのデータソースは RCT、観察研究、管理データベース、疾病レジストリー、専門家の意見、標準費用リスト、それに分析者の仮定やガイドラインで指定した仮定（割引率など）である。データの選択は研究課題や対象読者のニーズに適したものでなければならない。データはモデル設計上の特徴（視点など）と一致したものであり、その介入によって

影響を受ける母集団に関連したものでなければならない。すべてのデータは報告し、出所を明らかにしておく必要がある。データの詳細、すなわちデータが引き出された母集団や結果が適用される母集団なども記載しなければならない。モデルの中のデータを直接比較することができない場合（すなわち、他の患者サンプルに関係している場合）、それを選択した理由および仮定を説明し、正当化する必要がある。

モデルのパラメーターの推定に使用されるデータが信頼できるものであればあるほど、そのモデルの結果は確かなものになる³⁸。データソースの選択（メタ分析で使用された研究の選択など）、データ入力を分析するための方法の選択（包括解析での試験脱落者の処理など）、それにその後のデータのモデルへの組み込みなどは、評価結果の不確実性や、結果の一般化可能性に関係してくる³⁹。データの限界ははっきりさせ、それを処理するための方法についても説明しておく必要がある。こうしたデータの限界が評価結果の不確実性に与える影響を定量化する努力をすべきである。

研究の設計はデータの質や評価の結果に影響を与えることがある。介入のアウトカムや費用を推定するとき、好ましい（または好ましくない）臨床試験やデータばかり選ぶのは不適切なことである。モデルで使用したデータの質についても説明が必要である。さまざまな種類の研究から取り出した証拠の強さを評価するのに、異なった手段を使用することは可能である⁴⁰。それぞれの形態の証拠は、現地の母集団や費用、慣行パターンのデータを調整するのに役立つ、あるいは感度分析のベースとして役に立つ補足的な情報を追加する。

専門家の意見を使ってパラメーター値を規定するときには注意が必要である。専門家の意見を使うときにはその理由を明記しなければならず（他の情報源から適切なデータを引き出すことができないなど）、その意見の出所、引き出し方法、それを使用した結果について説明をする必要がある。こうした推定値によって生まれる不確実性は、感度分析をすることによって適切に査定しなければならない。

系統的レビューやメタ分析を行うと、モデルのパラメーターにとって高品質のデータを生み出すことができ、経済的評価に信頼性を与えることになる³⁸。系統的レビューはまた、関連する推定値を取り巻く不確実性を分析するのに役に立つ情報を提供する。こうした主要なモデル・パラメーターは、その結果が最も反応するので、注意が必要である^{30,33}。容易に入手したデータの適切性および一般化可能性に基づいて、主要なモデル・パラメーターを系統的にレビューすることができなかった場合には、その理由を記載する必要がある³³。

確率論的感度分析の場合、データをポイント推定値あるいは分布としてモデルの中に組み入れることができる。これを行うプロセスは説明する必要がある。データをポイント推定

値としてモデルの中に組み入れた場合、規範事例ではパラメーターの平均推定値を使用しなくてはならない⁴¹。変換が必要なデータは、生物統計学および疫学の一般に認められている方法に従う必要がある。こうしたデータ分析の方法や結果は記載しなくてはならない。

3.9 アウトカムの評価

3.9.1 健康関連の生活の質

「1つの構成概念として、健康関連の生活の質（HRQL）は、ある人の生活の健康的側面がその人の全体的な福利に及ぼす影響を指す。また個人の健康状態の価値を示すときにも使われる⁴²」HRQLを測定するために多くの方法が開発されてきた。それらの方法は、特異的尺度、包括的尺度、選好に基づく（効用）尺度に分けることができる⁴³⁻⁴⁵。特異的尺度（Western Ontario-McMaster Osteoarthritis Index など）と包括的尺度（Short Form 36 など）は一般に経済的評価にとってはそれほど価値のあるものではない。選好に基づく尺度はHRQLを数字的に示す総括的なスコアを提供するもので、費用効用分析（CUA）で使うのに適している唯一の方法である。

介入と代替策の間にHRQLに関して有意義な違いがあり、選好に関する適当なデータが入手できる場合には、適切な選好に基づく尺度を使う必要がある。プロスペクティブ研究でHRQLが測定されている場合には、経済的評価を行うことが目的なのであるから、選好に基づく尺度を含めるのが良い。これが行われていない場合でも、別方法で過去に遡って選好スコアを集めることが可能であり、その後で効力・効果の臨床試験のアウトカムの上にマッピングすることができる。もう1つの方法として、選好を2次文献から得ることも可能であるが、これにはその文献が対象母集団にとって適切なものであるという条件がつく。

「選好」と「効用」という用語は、経済ガイドラインのHRQLの測定基準として通常同義語として使われているが、厳密に言えば「効用」は不確実性を伴う方法（すなわち、基準的賭け法）によって得られる選好である¹。

3.9.2 質調整後余命年数（QALY）

CUAのための望ましい測定基準はQALYであるが、これは介入により獲得された余命年数に、その期間中のHRQLを反映する標準ウェイトを掛けて計算される^{1,46,47}。QALYは、その明瞭さ、単純さ、適用の容易さ、それに表面的な妥当性から、望ましい。

通常、QALYを算出しやすくするために、選好は基数（すなわち、間隔）尺度で測定され、スコアは、死が差し迫っている健康状態の場合は0.0、完全に健康な場合は1.0となる。死よりも悪い状態をこの尺度で表すことが可能で、その場合のスコアは0.0未満となる。選好

に基づくスコア（すなわち、CUA の場合は品質加重値）は直接的にも間接的にも測定することができる。CUA の中で選好を測定する代替的な尺度、例えば健康等価余命（HYE）や若年者救命等価生命（SAVE）を使用した場合には、その理由を説明する。

選好を直接測定することは、複雑で費用のかかる仕事である。選好の直接測定法としては3つの方法が使われる。すなわち基準的賭け法、時間得失法、それに視覚アナログ尺度である⁴⁸⁻⁵⁰。分析者は基準的賭け法を好むが、これはフォン・ノイマン＝モルゲンシュテルン効用理論に強力な規範的基盤があるからである^{1,51}。基準的賭け法の優位性に対しては反対意見がある⁵²。視覚アナログ尺度はよく知られた偏りがあるため単独で使うのは適切ではない¹。選好を直接測定したいと思っている分析者は、手持ちの問題に取り組むために、理論的特性と実証的特性を持っている方法を選ぶべきであり、選んだ理由を説明しなければならない。二重計算を避けるため、回答者には、実行可能な場合、選好の変化の観点から失われた余暇時間を評価し、医療費および所得損失は全額補償されるものと仮定するよう求める。

直接測定法を用いずに「既成の」手段を使って効用を求めることができる。このカテゴリーで広く使われている手段は、健康効用指標（HUI）⁵³⁻⁵⁵、EQ-5D⁵⁶⁻⁵⁹、SF-6D⁶⁰、それに15D⁶¹である。これらの手段は「情報に通じた」一般大衆から得た選好を用いているが、これは集合的に資源配分を行う目的には適した情報源である。その内の一部、例えば HUI では調査回答者に対して、健康状態を評価するときには所得効果を排除するように求める⁶²。これらの手段を使うために、分析者は患者の健康状態を提示されるシステムに分類し、公式により効用を算定する必要がある。スコアは平均的効用の推定値を表すが、これが一般大衆の無作為標本法による健康状態に対する推定値となる。ただし、この方法では直接測定法の感度が失われるおそれがある。

分析者は入手も比較も解釈も容易なことから、間接測定法を使うことが奨励される。選好の直接測定は容認できる方法であり、状況によっては、このほうが適している（機能の一分野が広く病気に冒されている症状など）。

分析者は推測的に代替的方法を研究して、最も症状に適しており、研究課題に最適な方法を事前を選んでおくことが推奨される。その選択と方法に関して理由を述べ、選好を測定するために実施するステップを説明する。さまざまな方法を試してみて、その介入が最高のものに見えるような方法を選ぶというやり方は不適當である。Brazier *et al.*⁶³ は健康の選好に基づく尺度のメリットを判断するために、その手段の実用性、信頼性、および妥当性を基準にした有用なチェックリスト（報告書ボックス3）を提示している。

使用する方法に関係なく、生命の長さとは品質加重値の変化は別々に報告し、それを組み合

わせた方法についても明らかにする。経時的な品質加重値の変化（直線的、曲線的など）に関する仮定、および QALY を評価するために使用した方法（基準スコアからの変化、曲線下にある全エリアなど）についても報告する。

QALY の使用で問題なのは、重症度が異なる症状を区別できないということである。従来の QALY 法では選好スコアの絶対的な変化に焦点が当てられていたが、さまざまな研究の結果、さまざまなグループのための介入に対する社会的評価も、初期症状の重症度によって左右されることが分かってきた²⁰。この懸念に部分的にでも対処するため、分析者は臨床的に重要とは考えられないアウトカム（短期的、自己制御的、重症でないアウトカムなど）や、意思決定者にとってあまり関心のないようなアウトカムの QALY を排除した感度分析を行うことができる。例えば、インフルエンザ治療の新薬の場合、感度分析では受診や入院を必要とするほど重症なインフルエンザ関連の合併症だけの影響を示すことが可能で、治療により予防されたインフルエンザは分析から除外される。

品質加重値を引き出すために、特別な健康状態を経験している患者、またはそうした経験を持たない一般大衆（地域社会）の代表標本のどちらの選好を使うべきかに関して論争が起こっている⁶³。主要な間接測定手段は一般大衆を対象とした選好の調査に基づいている。特定の症状に関する選好を測定する直接的な方法では、おそらく臨床試験的な環境下で調査対象になっている患者を使うことが多い。選好評価は症状経験によって変わってくることを示す証拠がある⁶³。

選好を直接的に測定する分析者は、評価対象の健康状態について適切に情報を与えられている一般大衆の代表標本を使用することが望まれる。その理由は、彼らが公的医療制度の最終的な支払人であり、潜在的な患者でもあるからである⁶⁴。該当する健康状態を直接体験した患者は、選好の受容できる情報源となるかもしれない。直接測定法を行う分析者は選好を引き出した母集団と測定方法について説明しなくてはならない。

参考事例で一般大衆の選好を使い、感度分析で患者の選好を使うことができれば理想的であるが、これは実際的でない、あるいは不要かもしれない。分析者はカナダの母集団に推定した選好が適用できるかどうか、検討する必要がある。

一部の研究では大規模な感度分析が品質加重値のデータソースであるという専門家の判断を使っている。このアプローチは支持されていない。研究結果が品質加重値に反応しないことを示せる場合には、概算で十分であろう。

介入が患者の介護者の生活の質に影響を与える場合、これは測定することができるが、ケ

アの費用を二重計算するのを避けるために、ケアは取り除くべきである。介護者の生活の質の変化は分析の中で別に報告すべきであり、ICER を計算するときには排除しなければならない。

3.9.3 費用便益分析におけるアウトカム

費用便益分析（CBA）の中で健康アウトカムに割り当てられる金銭的価値は通常、支払意志額（WTP）法を適用することにより求める。WTP 研究で使用される 2 つの方法は、仮想評価と結合分析（離散的選択実験としても知られている）である。仮想評価では仮想的調査を使い、通常は市場価格を持たない物やサービスに関して、個人の最大限の WTP を推定する⁶⁵。結合分析では順位付けや評点付けや比較をすることによって、物やサービス（ヘルスケアも含む）のさまざまな属性に人々が付ける相対的ウェイトを推定する^{66,67}。ヘルスケアにおいては、結合分析が非健康アウトカムやプロセス属性を評価する方法としてこれまで使われてきた。

アウトカムを評価する WTP 法は今も発展中であり、解決する必要があるいくつかの方法論上の問題や倫理的な問題がある^{22,68}。これらの評価方法はヘルスケア部門での資源配分に関する決定を必要とする状況下で広く使われてはこなかった。ヘルスケアに関する資金供給の決定に情報を与えるための、これらの方法の使用が有効なものになるためには、さらなる研究調査が必要である。

一般的に CBA を使用すること、特に WTP 法を使うことは、2 次的な分析とみなすべきである。CBA を使った場合、評価ではアウトカムを金銭的価値に置き換えるためのステップを説明しなければならない。この分析の主要な仮定は、感度分析をすることによって正当性が確認され、全面的に検証されなければならない。

3.10 資源利用と費用

3.10.1 資源の特定

費用計算プロセスのこのステップでは、分析者はスケジュールに沿って、それぞれの代替策で発生しそうな活動と資源を特定する。研究の視点によって、どの資源項目を分析対象に含め、どの項目を除外するかが決まる。そのような視点のいくつかは表 2 に概説されている。

公的支払機関の視点に含まれる費用は、関連意思決定者に適したカテゴリーに分類することが推奨される。評価にあたっては、さまざまな公的支払機関部門（例：プライマリケア、病院、コミュニティ・ケア）別に発生費用を分類し、意思決定者が責任を負う費用を個

別に提示すべきである。例えば、州のドラッグ・プランの評価では、薬剤使用に伴う費用の内訳を示すべきである。公的支払機関部門間でコスト転嫁がある場合は、この評価で関連費用を定量化すべきである（例：入院日数の減少、在宅介護訪問回数の増加）。公的支払機関の視点では、すべての公的支払機関部門の費用を集計すべきである。

介入の結果である現在および将来の費用を評価に含めるべきである。非薬剤技術に伴う費用の特定は、薬剤関連費用の特定より複雑かもしれない。これらの費用には、開業費、資本費、営業費、維持修繕費、追加人員雇用費、間接費および職業訓練費が含まれる。これらの費用は必要に応じて含めるべきである。例えば、股関節補綴を評価する場合は、機器関連費用だけでなく、ケア全体にかかる費用と他のすべての関連費用（例：訓練費）を含めるべきである。介入と代替策とで使用状況がまったく同じである場合は、資源項目を分析対象から除外することができるが、分析者はこれが正当であることを証明すべきである。

臨床試験から生じるプロトコル連動費用は、評価対象から除外すべきである。所得移転支出（例：障害給付・雇用補償金の支払い）は、公的医療制度によって負担されず、社会にとっての実質費用ではない（相殺される）ため、分析対象から除外すべきである。分析者は、これらの費用が重要であれば報告してもよい。

ある費用項目は、評価対象の介入とは無関係の事象の影響を受けるため、関連性がないとみなされることがある（例：下肢骨折にかかる費用は通常、にきび治療薬の評価において考慮されない）。どの臨床事象が関連しているかを決定するための1つの代替措置は、裁定委員会（治療任務に左右されない）を通すことである。これにより分析者は、偏りのない方法で無関係の事象を取り除くことができる。

経済文献で議論されている未解決の問題は、介入によって獲得された余命年数の間に発生する無関係のヘルスケア費用や非医療費をめぐる問題である⁶⁹。分析者は独自の判断に基づき、そのような費用を含めるかどうか決めることができるが、これは感度分析においてのみ行うべきであり、そのような費用を特定すべきである。1つの選択肢は、これらの費用が全体的結果に与える影響が小さい場合は、分析対象から除外することである。データが入手可能で、かつ影響が大きい場合は、これらの費用を感度分析に含める。

3.10.2 資源の測定

透明性を確保するために、介入と代替策に関する資源利用データを物的単位で報告すべきである。

2つの費用計算方法があるが、多くの評価が両者の組み合わせを利用している⁶⁹。総費用計

算（トップダウン費用計算）は、入院 1 日当たり費用など、大きな要素を費用計算の基準として利用する。項目ごとの詳細なマイクロ費用計算（ボトムアップ費用計算）は、より正確な推定値を算出できるが、そのようなデータの収集に伴う時間と経費を考慮する必要がある。分析者は、利用した費用計算方法を明確にし、その方法が正当であることを証明しなければならない。費用計算方法の利用に関する指針は、Baladi⁷⁰、Oostenbrink *et al.*⁷¹ および Gold *et al.*⁴ など、いくつかの資料で入手することができる。

総費用と増分費用に最も寄与する資源項目は、より正確に測定・評価すべきである。これはモデルで資源利用と単位費用パラメーターの感度分析を実施し、総費用や増分費用および結果に対する予想される影響を決定することによって、推測的に行うことができる⁶⁹。

分析者は費用推定値を引き出すにあたり、以下の点に注意を払うべきである。

- ・ 可能であれば、間接費、共通労働コストおよび管理費の配分方法について説明する。
- ・ 学習曲線効果（例：外科的処置）など、資源の将来費用を削減する可能性のある利用資源量と単位費用推定値との関係や、新しいプログラムが既存のインフラストラクチャーに与える影響（プログラムの規模や範囲が要因となっている場合）、費用の推定に利用した設備稼働率（該当する場合、正常な作業能力の調整に利用した方法を説明）を考慮する。

さらに分析者は、推定の質を評価し、利用した推定値の質が低い場合は、感度分析を利用して費用に関する仮定の影響を決定する。質の特徴には、資源全体のうちどのくらいが推定に含まれているか、産出指標は資源利用を捕捉するために十分に精緻化されているか、データソースの質はどの程度か、費用の推定値をすべての医療機関に一般化できるか、といった問題が含まれる。

資源利用データは、いくつかの情報源から得ることができる。例えば、RCT、管理・会計データ、標準的治療法ガイドライン、専門家の意見、それに（さまざまな情報源からのデータを結合する）モデリング作業などである。データには、推定値の質やカナダの慣行への適用性に関して大きなばらつきがあるだろう。実験的慣行研究や国際的研究から得られた資源量をカナダの慣行に変換するにあたり、いくつか問題が生じる。得られたデータのカナダの慣行に対する適用性が正当であることを証明すべきである。

3.10.3 資源の評価

資源は、その機会費用、すなわち最善の代替利用の価値で評価すべきである⁶⁹。資源の単位費用の測定における指導原理は、これらの費用により、長期的に追加単位を生産するために必要なすべての資源を測定すべきだということである。これらの資源には、資本費、す

すべての営業費、配賦間接費、専門家費用が含まれるだろう。例えば病院費には、患者の入院に関連する診療料金を含めるべきである。実際には、この概念に近づくために多くの方法が利用されている⁷²。分析者は、単位費用測定基準として平均総費用（資本費や配賦間接費を含む）を利用することが推奨される⁷¹。

資源を評価する方法は、市場価格、管理費、直接測定、潜在価格の計算など、いくつかある⁷³。多くの資源の場合、単純明快な評価方法はない（例：インフォーマルな介護）。

最善の評価方法に関する合意はないが、以下の点を考慮すべきである。

- ・ 他の方法（例：超過利得）を採用する妥当な理由がない場合は、入手可能であれば市場価格を利用する。
- ・ 一貫性を確保するために、また便宜上、標準費用の利用を主張することができる。2004年、さまざまな州で比較的一般的に見られるサービスのいくつかについて、2002年の費用を利用して標準費用リストの改訂版⁷⁴が作成された。医療経済研究所ウェブサイト (<http://www.ihe.ca>) で入手可能である。このリストは多様な状況で利用されているが、経済的評価は多くの設定、さまざまな状況で行われており、そのすべてを1つの費用リストでカバーすることはできない。
- ・ 費用を直接計算するか帰属させる場合は、正常な業務レベルにおけるすべての関連資源の総経済費用を反映させるべきである。

対象読者が複数の管轄区域にまたがる場合は、最大管轄区域の費用またはカナダの平均費用（入手可能な場合）を参考事例で利用し、該当管轄区域の最高費用と最低費用を用いて感度分析を行うべきである。

公的支払機関の視点を評価する場合は、参考事例で介入とコンパレーターの総費用（すなわち、公的医療制度、民間保険会社および患者が支払う拠出金）を利用する。公的支払機関と患者との費用分担（例：病院外で処方される薬剤の自己負担）が関わる可能性がある介入については、分析者は感度分析を利用し、公的支払機関が支払う介入・コンパレーター費用の比率の変動の影響を評価すべきである。例えば分析者は、DSA で公的支払機関が負担する費用の典型的比率を利用したり、PSA で公的支払機関が負担しそうな費用の配分を指定したりすることができる。他の方法を採用する妥当な理由がない場合は、介入とコンパレーターに同じ比率を利用する。これによって意思決定者は、その管轄区域の保険規定に最も適した比率を利用することができる。費用分担取り決めは、患者が介入およびコンパレーターに支払う額が相殺される場合（例：患者が2つの代替的投薬計画に同額の定率料金を払う場合）か、ICERに与える影響が小さい場合は無視することができる。分析者は、これらの費用分担取り決めが感度分析でどのように取り扱われたかについて、また ICER

に対する影響について報告する際には、透明性を保つべきである。

単位価格は、過去の価格または他国の価格しか入手できないかもしれない。過去何年かに得られた価格は更新すべきである。カナダには病院サービスや医師サービスの価格指数がないため、消費者物価指数などの一般物価指数を利用することができる。

分析者は、カナダの設定に他国の価格や単位費用を適用するにあたって注意しなければならない。そのようなデータは、カナダに適していることを実証すべきである。そのようなデータを調整する場合は、適切な方法を利用し、説明・正当化すべきである。

3.10.4 損失時間

患者はその症状の結果、移動したり治療を受けたりするために費やされる時間、病気にかかっている時間、早死にが原因で、本来であれば実施していたであろう活動を放棄する場合がある。家族、友人またはボランティアが患者のケアに無給で従事するかもしれない。その結果、他の活動に費やすことができた時間が犠牲になる。より重度の慢性症状（例：片頭痛）の場合や、介入がフォーマル・ケアからインフォーマル・ケアへの移行を伴う場合、これは重要になる場合がある。介入が患者やインフォーマルな介護者の損失時間に与える影響は、分析の結果に影響を与えそうであれば、個別分析で定量化することができる。

評価に損失時間を含めるための適切な方法に関して議論が行われている。この問題は、実施された健康関連活動の点から、また放棄された活動（すなわち、機会費用）の点から時間をとらえることによって把握できる⁷¹。後者の方法では、活動を3つのカテゴリーに分けることができる。すなわち、有給の仕事、無給の仕事（例：家事）および余暇である。

患者にとって、病気の影響を受ける時間には、移動したり治療を受けたりするために費やされた時間や、放棄された活動に関する損失時間などの健康関連活動が含まれる^{71,75}。有給の仕事における損失時間には、工作中的の生産性の損失、短期または長期の欠勤、患者の早死による生産性の損失が含まれる。家族、友人またはボランティアによるインフォーマルな介護は、患者のケアに投入された時間や、他の活動に費やすことができた損失時間の点から測定することができる。

有給の仕事の損失時間を評価するための2つの方法は、人的資本法と摩擦費用法である。人的資本法では、病気が原因で失われた雇用期間が生産費の測定基準である。この方法では、完全（に近い）雇用があると仮定し、生産の損失の価値が過大評価される場合がある。摩擦費用法では、ある人が失業すれば最終的に失業者にとって代わられると仮定するため、社会にとっての生産性の損失は、その病人が別の労働者にとって代わられるまでの期間（摩

擦期間)に制限される。摩擦費用法では、早死による生産性の損失は、摩擦期間を超えて影響を及ぼさない。短期欠勤の場合、患者の生産の損失は、復職したときに本人によって、または会社の内部労働資源によって、ある程度取り戻される場合がある。有給の仕事の損失時間が短い場合は、2つの方法から得られる推定値に差がないかもしれない。より長期の場合は、摩擦費用法のほうが人的資本法よりも費用推定値が低くなる⁶⁸。

人的資本法と比べて、摩擦費用法のほうが労働市場の機能をより実践的にとらえているとする意見もある^{71,76}。すべてを考慮すると、有給の仕事における損失時間を評価するには摩擦費用法のほうが望ましい。この方法には、いくつか欠点がある⁷⁷。この方法を運用できるようにするには、摩擦期間の長さを指定する必要がある。摩擦期間は、産業、職業、マクロ経済環境(例：一般失業率)、それに求職者と欠員とのマッチング過程の効率によって異なる場合がある。カナダの摩擦期間の適切な長さに関するデータはない。摩擦期間の代用物として、仕事の欠員を埋めるために必要な時間の長さが提案されている。もう1つの方法は、カナダと同等の経済環境・労働市場の国々について推定された全国平均摩擦期間を利用し、その摩擦期間の長さを感度分析によって検証することだろう⁷⁶。この方法はGoeree *et al.*⁷⁸により、分裂病に伴う摩擦費用を推定するために利用された。

死亡や病的状態による長期の欠勤(すなわち、摩擦期間より長い欠勤)の単位費用は、年齢・性別調整賃金率に加えて、給付(年金や雇用保険など)の雇用者負担としての補助賃金を利用して評価することができる⁷⁶。この値は、各人口集団の被雇用率で調整することができる。これらの構成要素に関する人口レベル・データはカナダ統計局から入手可能である。摩擦費用は、代替労働者の採用・訓練に伴う費用を含む。

無給の仕事時間(インフォーマルな介護など)を評価するための選択肢はいくつかあるが、望ましい代替策はない⁶⁹。1つの選択肢は、自宅で患者に提供されるサービスの市場価値(すなわち総賃金)に基づく置換費用推定値を利用し、介護に費やされた時間か、専門家が職務の遂行に費やしていたであろう時間を用いることである。もう1つの選択肢は、インフォーマルな介護者が介護期間に有給の仕事に就いていれば受け取っていたであろう(純)賃金(すなわち留保賃金)を利用する。3番目の選択肢は、介護に費やされた時間を、有給の仕事、無給の仕事および余暇からの損失時間に分け、このセクションで概説された他の方法の1つを利用して各期間を評価することである。仮想評価法または結合分析を利用してインフォーマル・ケアを評価することも可能だが、これらの方法はそれほど一般的ではない⁷⁹。

何人かの著者^{69,71,80,81}が示唆しているように、失われた余暇時間は、可能であれば評価対象の費用から除外することが推奨される。失われた余暇時間は、QALYが健康アウトカム指標

として利用される場合は、選好測定基準によってある程度捕捉されるだろう。実行可能であれば、選好測定作業の被験者に、余暇の変化を選好の変化に関して評価し、医療費と所得損失が完全に補償されると仮定するよう求めると良い。

患者とインフォーマルな介護者の時間を測定・評価するにあたっては、注意を払うべきである。損失時間の測定・評価に利用した方法と仮定について説明しなければならない。損失時間の価値は利用する方法によって決まる可能性があるため、分析者は感度分析を利用し、代替的な方法と仮定を評価すべきである。

3.11 割引

費用と健康アウトカムが将来発生する場合は、現在価値に割り引き、社会の時間選好率を反映させるべきである。そこで、1年以上先に発生する費用やアウトカムは、標準方法を利用して割り引くべきである。

それぞれの評価の結果を比較できるようにするため、共通の割引率を利用することが重要である。参考事例の標準率は年間5%に設定されている。割引率0%を分析し、割引の影響を示すべきである。加えて、感度分析で3%の割引率を利用し、3%を標準割引率として利用する他の管轄区域で発表された評価と比較しなければならない。

参考事例と感度分析の割引率は実質（恒常価値）ベースで表示され、これは実質（すなわち、インフレ調整済みの不変）ドルによる資源の評価と調和する。

割引慣行に関しては Gold *et al.*⁴、West *et al.*⁸² および Smith *et al.*⁸³ を、健康結果の割引については Lazaro *et al.*⁸⁴ を参照されたい。

3.12 可変性と不確実性

可変性と不確実性を区別することが重要である。可変性は、状況における特定可能な違いに関連する、パラメーター値における既知の違いを表す⁶⁸。度数分布で表され、低下させることができない。不確実性はパラメーターの真の値が未知であるときに発生するため、知識や測定が不完全であるという事実を反映している。モデルのパラメーター値（例：資源利用、効用、作用）およびモデル設計上の特徴（例：モデル構造、分析方法、仮定）に関連している場合がある。これらは経済ガイドラインでモデル入力と呼ばれる。

3.12.1 可変性の取り扱い

慣行パターンの可変性は、さらなる分析によって取り扱わなければならない。これは感度

分析またはシナリオ分析を伴うことがある。この方法は一種の多元感度分析で、関連モデル・パラメーターの代替値を同時に置き換える。

患者の不均質性は、分析対象の患者のさまざまな個別的特徴（例：「リスクプロファイル」）に関係がある。例えば、治療効果は高齢の高リスク集団と中年の低リスク集団とで異なることがある。ある介入のターゲット集団において、患者群によって費用効果が異なるかもしれない。

モデリング研究における分析者の役割は、層別分析における費用、アウトカムおよび選好（効用）に関連するデータの不均質性を明らかにすることである。Coyle *et al.*²⁵ は、ターゲット集団における費用効果のばらつきに対処するための枠組みを提案している。この枠組みは、ターゲット集団をより均質な患者群に層別する。この分析では、各下位集団を個別にモデルにかける。したがって、治療の費用効果が高いグループ（例：純貨幣便益がプラスのグループ）を特定することができる。母集団が不均質である場合は、患者下位集団ごとにモデルまたはモデルのバージョンを分けるのが優れた方法である。

探索的データ解析の過程でグループ間に違いが現れる場合がある。できれば、アウトカム研究の企画段階で下位集団を定義することが望ましい。後知恵分析は、正当とする強い根拠が挙げられなければ受け入れられないことがある。情報から派生する後知恵は慎重に解釈すべきであり、現在の資源配分を促進するのではなく、さらなる調査課題を導入するものとみなすことができる。

3.12.2 不確実性の取り扱い

ほとんどの場合、経済的評価では決定分析モデルを利用して介入の費用効果を推定する。経済的評価の文脈における不確実性は、パラメーター不確実性とモデル不確実性に分けることができる。

さまざまな種類の変動や不確実性を扱うための推奨方法に関する要約を表3に示す。付録1は、可変性と不確実性の分析結果を図表によって提示するための提案である。

表 3：可変性と不確実性を取り扱うための推奨方法

カテゴリー	可変性または不確実性の種類	推奨方法
可変性	地理的地域間または設定間の臨床慣行パターンの違い	感度分析
	患者集団における可変性（患者の不均質性）	層別分析
モデルベースの不確実性	モデル不確実性 <ul style="list-style-type: none"> • 分析方法 • モデル構造 • 仮定 • データソース 	代替的仮定、一元分析、多元分析、限界値分析または極値解析を利用する DSA、およびモデル検証方法
	パラメーター不確実性	一元分析、多元分析、限界値分析または極値解析を利用する DSA モンテカルロ・シミュレーションを利用する PSA が奨励される

パラメーター不確実性とは、モデルのパラメーター（健康アウトカムや効用、資源利用など）の真の値に関する確率論的な意味での不確実性を指す。パラメーター値が標本に基づいている場合は、偶然によって推定に不確実性が伴う。ほとんどの評価において、費用とアウトカムに関するデータが患者レベルで直接観察されることはない。その代わりに分析者は、一般的に統合情報を利用し、調査対象の治療の費用とアウトカムに関連するパラメーターを推定する。一般に、この情報は多様な出所から得られる。例えば、前向き の 1 次データ、発表された研究および未発表の研究からの推定データ、専門家パネル、判断を反映する仮定などである。

モデル不確実性は、モデルにおける他の種類の入力（その値は確率論的な意味では必ずしも不確実ではないが、異なることが分かっているもの）に関連する不確実性である。モデル不確実性はモデル作成者の選択や仮定に依存する。モデル不確実性には以下が含まれる。

- 分析者が参考事例のために行う方法・仮定を選択（例：測定対象期間、割引率、生産の損失の評価方法）
- モデルにおける構造仮定（例：試験の追跡調査や治療効果の予測持続期間を超えてアウトカムを外挿するための手法）
- データソースの選択（例：メタ分析で利用する研究の選択、データ・プーリングの利用、データの層別）

不確実性が結果に与える影響を徹底的かつ体系的に評価すべきである。モデル構造、分析

方法、入力のパラメーターおよび仮定、出力など、あらゆるレベルでモデルの信頼性を検討しなければならない。不確実性を適切に考慮し、結論が有意義かつ確固不動であるかどうかを読者が判断できるようにする必要がある。

モデルに多くの仮定があり、どの入力（または入力の組み合わせ）が結果に最も大きな影響を与えるかを推測的に知るのが難しいことを考えれば、少なくとも、すべてのモデル入力について DSA を実施すべきである。感度分析からいずれかのモデル入力を省略する場合は、（例えば予備分析に基づいて）正当性を示すべきである。この分析の目的は、関連モデル入力の妥当と思われる範囲の値に全体的結果が反応し、結果の解釈が変化するかどうか調べることである。この分析を行うための方法として、代替的仮定、一元・多元感度分析、限界値（損益分岐点）分析、極値（すなわち、最高のシナリオと最悪のシナリオ）の分析が挙げられる。

分析者は、各モデル入力の妥当と思われる範囲の値を定義し、選ばれた範囲の正当性を示すべきである。これらの範囲は、信頼できる適切な限界に関する証拠に基づき、それぞれのモデル入力に関連する可変性または不確実性の全範囲を表すべきである。Drummond *et al.*¹ は、妥当と思われる範囲は文献調査、専門家の意見、（確率データの）平均の指定信頼区間の利用によって決定できると示唆している。

以下に示す要点は、DSA 実施の指針とするうえで役立つ。

- ・ すべてのモデル・パラメーターに極値を利用して一元感度分析を行う。
- ・ パラメーターが仮定または専門家の意見に基づいている場合は、一元感度分析で広範囲にわたって検証する。
- ・ 一元分析で反応性があることを示されたモデル・パラメーターの多元感度分析を行う。
- ・ 2 つの誤差原因が相互依存性である場合は、それぞれの誤差原因に関して二元感度分析を行う。
- ・ シナリオ分析を利用し、方法論的不確実性を分析する（例：異なる入力を利用する代替的バージョンのモデルの実施）³⁰。
- ・ 感度分析を利用し、構造仮定に関する不確実性を検証する（例：データを外挿するための手法、中間アウトカムを最終アウトカムに変換するための手法）。
- ・ 入力に小成分が含まれる場合は、モデルで小成分を分けて個別に検証する。

モデルの内的・外的妥当性を正式に確認すべきである。

パラメーター不確実性は DSA または PSA によって取り扱うことができる。この 2 つの分析方法には、どちらも長所・短所がある。分析者には PSA の利用を奨励する。PSA のほうが、

モデルのすべての入力に伴う不確実性を完全に評価することができる。この種の分析は入力間の相互作用を取り扱い、説明できる定量的な結果を提供する。平均的な費用と効果の不偏推定値を導き出し、非直線関係（例：割引）のもとで最適決定を下せる可能性が高い。さらに、ベイズ法による結果の解釈を可能にし、研究資金をどこに割り当てるべきかを決定するために利用できる。

PSA はモンテカルロ・シミュレーションの形態を取るべきである。重要なパラメーターの値の範囲、確率分布の形態、モンテカルロ反復の回数に関する仮定を明示し、正当性の根拠を挙げるべきである。シミュレーション・ソフトウェア（例：ExcelアドオンソフトのCrystal Ball[®]）を利用できるため、そのような方法が実施しやすくなっている。

PSA はパラメーター不確実性を評価するためにのみ利用でき、モデル不確実性の評価に利用してはならない。モデルの構造、方法および仮定は、DSA とモデル検証方法によって評価すべきである。

確率論的感度分析はベイズ解析に拡張することができる。ベイズ枠組みを利用すれば、さらに調査を行うための費用と不確実性の低下から得られる便益との比較に基づき、追加情報収集の価値を評価することができる。この枠組みは、意思決定者が直面する決定の二元性（受容または拒絶）を認識し、PSA であらゆる形態のパラメーター不確実性とモデル不確実性を定量化する。この枠組みの主眼は、リスクを抑えるためにより多くの標本情報を得る価値があるパラメーターの特定である。この情報は、調査の優先順位を決めたり、研究設計の効率を高めたりするために利用できる⁸⁵。

費用効果分析・計算のためのベイズ法、および完全情報の期待価値（EVPI）に関する情報については、Heitjan and Li⁸⁶ および Claxton *et al.*⁸⁷ が参考になる。新しいデータが利用可能になった場合に過去の証拠を「更新」という概念の概観については、Briggs⁸⁸ を参照されたい。

3.13 公平

公平は、意思決定者がしばしば取り組まなければならない関心事である。受益者が誰であるか（例：高齢者、低所得層、遠隔地コミュニティ）は、介入やプログラムの提供の社会的望ましさを決定するうえで重要な要素となることがある。公平に関する網羅的な議論は経済ガイドラインの範囲を超えている。とはいえ、ヘルスケア資源の配分に関する意思決定における公平問題の重要性を考えれば、いくつかの一般的事項を指摘しておく必要がある。