

還払いの基準や利用条件を遵守せず、その介入を臨床・人口統計的基準または条件に合致しない患者にまで拡大適用する場合がある（ただし、それでも患者は認定された適応症の範囲内にあることがある）。このため介入の過剰な処方や無制限の利用を引き起こす可能性がある。評価には、医療費償還払い当局が意図したものより大きいこの集団における介入の予想される使用の分析が含まれることがある。患者が拡大適用されなければそうした治療を受けなかったような場合、こうした状況はコンパレーターの選択に影響を与えるおそれがある。Coyle. *et al.*²⁵は、効率的な限定的利用基準を作り上げるために、層別分析をどのように実施したらいいかという例を提示している。

- ・ 処方集で第2選択として収載してある薬剤が、第1選択薬として使用されることがある。
- ・ 介入に適している患者の候補者を正確に判断することは難しいかもしれない。こうした困難さは臨床試験では明白にならないかもしれないが、臨床診療でははっきりする。適切であれば、誤診された患者が意図しない介入を受けることになっても、何の影響も出ないと仮定できるかもしれない。例えば、Husereau *et al.*²⁶が行った研究では、感度分析を使用して、誤診され、(別々に) 後で症状が出た患者がインフルエンザの抗ウイルス薬を処方されるという仮定の影響を示している。示したように、加重平均 ICER は、適切使用から不適切使用の変動率を使って、計算することができる。
- ・ カナダ保健省が(まだ)承認していない適応症に関する臨床試験データが入手できるため、あるいは製品が市販されている管轄区域から情報が入るため、介入の未承認の適応外使用が起こる可能性がある。可能であれば、分析者はこうした適応外使用が行われる様子を知らせなければならない。
- ・ 償還払い状況や使用に関する変化については、公的資金が出されている技術は、起こりうる変化について評価を行うことになっている(規制や、リストからの削除など)。

こうした分析を行うには、対象となる特定の母集団やグループに関する臨床データが必要になる(平均的リスクの患者対高リスクの患者など)。状況によっては、当該介入を利用することができる管轄区域から利用データを入手したり、すでに市場に参入している類似の介入の利用データを検査すると役に立つかもしれない。場合によっては、間接的な比較が必要かもしれない。

3.4 コンパレーター

3.4.1 全般的な考慮事項

分析のために適切なコンパレーターを選ぶことは極めて重要である。というのも、何を選ぶかということが、介入の費用効果や、その研究が意思決定者にとって関連性があるかどうかを判断するのに重要になってくるからである。原則として、コンパレーターは、その

介入が採用された場合、臨床慣行において代用される可能性が最も高い代替策である。

- まず、研究課題、介入の適応症あるいは目的、そして研究の対象読者を考える。また、認可され、受容されており、技術的にも実行可能な代替策で、コンパレーターとして当該の症状に適応されるものをすべて考える。とはいえ、必ずしもそうした代替策全部を分析の中でコンパレーターとして使用するというわけではない。不要なものを削除しながら選択する。代替策として妥当なものを特定していくプロセスの中で、臨床専門家や研究内容の専門家に意見を求めると参考になるかもしれない。
- コンパレーターの選択は、同じ適応症について認可された代替策が数多くある場合、あるいは管轄区域や患者の下位集団の間で臨床慣行に差異がある場合（老人ホームにいる患者対一般大衆など）、簡単ではないかもしれない。実際には、分析者は少数の最も適切なコンパレーターを分析のために特定しなければならないかもしれない。それをする際には、科学的精密さを持って、データの入手可能性、時間の制約、および多数のコンパレーターの分析可能性とのバランスをとることが必要である。
- 対象となる介入と関連するコンパレーターを直接比較した臨床試験がない場合、分析者は適切な手法に基づいて間接的な比較を行ってもよい。間接的な比較を統合するために使用する方法については、説明をして、それを選んだ理由を述べなければならない。用いた方法に何らかの限界があり、パラメーター推定値に潜在的な偏りがあり、結果の解釈に警告をつける場合には、それらを報告しなければならない。コンパレーターに関する仮定の影響を評価するとき、感度分析も使用することができる。

3.4.2 「通常ケア」

参考事例ではコンパレーターは「通常ケア」となるが、これは症状に対する臨床慣行の中で最も一般的に、あるいは最も広く用いられている治療のことである。これはまた、「既存の慣行」、「現在の慣行」、「典型的ケア」あるいは「現状」とも呼ばれる。介入が取って代わろうと意図している最も一般的に使用されている治療は、おそらく利用データや臨床専門家の意見に基づいて、最も多くの患者に使用されているものになるはずである。さらに、臨床慣行の中で優位を占めている最も一般的なタイプのケアがあるかもしれず、あるいは2つか3つの優勢な代替策があるかもしれない。その場合、すべての代替策は研究対象の介入と個別に比較する必要がある。

3.4.3 追加的考慮事項

通常ケアは必ずしも推奨される、あるいは臨床的に最も効果的なレベルのケアであるとは限らない。こうした状況下では、それが実行可能で適切な選択肢であると考えられる場合には、「推奨される」あるいは「適正な」ケアを（通常ケアの他に）含めるべきである。科学的根拠に基づく標準的治療法のガイドラインや臨床専門家の推奨を参考にして、高品質

で臨床的に適正なケアを決定することができる。こうした場合、代替策に関する証拠の強度を示さなければならない。望ましい治療がはっきりしない場合には、複数の代替策が適切かもしれない。

介入によっては、個別の製品を比べるというより、治療戦略を比較することになるかもしれない。こうした場合、それぞれの状況を、技術が戦略の付加的な要素である場合、異なった治療シーケンスである場合、あるいは、当該介入が採用されたら、戦略の他の要素に取って代わるような代替策の場合を区別する。例えば、ある研究においては、ある尿失禁の薬物治療の費用効果が、その薬物治療を行う場合によって変わってくるのが証明された（第1選択療法または第2選択療法として使用、既存の薬物治療が失敗したら使用など）²⁷。組織的な（多くの要素で構成されている）ケアの代替モデルや代替パッケージも比較することができる。例を挙げると、一般内科病棟での脳卒中のリハビリと脳卒中専門治療室でのリハビリ、および自宅で支援を受けながら行うケアを比較した研究がある²⁸。戦略は説明を行い（行う時期、行う状況、対象者など）、代替的な戦略の要素を定義しなければならない。

将来のコンパレーター、特に、分析に関連する時間枠内に市場に参入してくると思われる低価の技術を予想することは良い慣行である。それをしないと、新しい介入のICERを過小評価することになるおそれがある。一例としては、ある医薬品の特許がこの先切れることが分かっている場合、競合するジェネリック薬の登場を予想することである²⁹。また、低価の技術の市場参入を予想することも、非薬剤技術にとって社会的に意味のあることである（外科用器具で再利用できるもの対低価格の使い捨てタイプなど）。こうした分析は、ジェネリックテクノロジーの市場参入の時期や価格に関して新たな不確実性の原因を提起することになる。こうした状況は、実行可能なおきにはいつでも、感度分析や競合技術の低価格を使って処理することができる。

その他にも、妥当と思われる代替策を考える。場合によっては「何もしないこと」や「注意深く見守って待つ」ことが適切なコンパレーターになるかもしれない。

3.4.4 コンパレーターになる薬剤

- ・ コンパレーターは当該管轄区域の処方集に載っている代替薬でなくてもよい。出発点としては、同じ薬効分類の代替薬（適応症が同じ薬剤など）を考えるのが好都合かもしれない。これには新薬と同じ化学成分や作用機序を持つ薬剤を含めてもよい。例えば、ニューキノロン系抗生物質のコンパレーターとしては、他のキノロン系薬剤だけに限定しないで（すべてのキノロン系薬剤が同じ適応症を認められているわけではないので）、むしろニューキノロン系薬剤の用途である感染症の治療に臨床的に用いら

れている他のすべての抗菌薬を考えるほうがよい。新薬が新しい薬効分類に入るものである場合には、コンパレーターになるのは通常ケア（および、適切な場合は推奨ケア）であり、これには、利用可能な場合には、他の化学分類の薬剤を使用した治療や、非薬剤治療を含むことができる。

- コンパレーターとしては、低価で同じ適応症にしばしば用いられている入手可能な代替薬を選択する。この選択は、症状を治療するための薬剤の全推奨用量の費用と、薬剤の投与にかかる費用に基づいて行う（それぞれの費用を別個に報告する）。
- 費用計算のために使用する投薬計画は、評価で使用される薬剤の効力または効果に関するデータを支持する投与量と投与期間を反映したものでなければならない。臨床的に用いられている投与計画が効力を調べる臨床試験で用いられる投与計画と違うかどうかを明示すること。実際の（推奨されるのではなく）投与は、文献のレビュー、利用データの調査、あるいは臨床専門家を対象とした調査などによって決めることができる。適切な場合には、同じ治療効果をあげるためにそれぞれのコンパレーターが必要とする用量を使い、使用した線量当量の理由を説明する。
- 報告書には、薬剤の一般名と商品名、薬効分類、投与形態、投与経路、推奨される 1 日の投与量、治療期間、1 日薬価、および（薬剤およびその他のコンパレーターの）通常の治療単位費用を明示しなければならない。適応症や禁忌、使用上の注意、警告、および副作用に関して違いがある場合は、それを報告しなければならない。

3.5 視点

評価のために選ぶ視点は、対象読者のニーズに合致するものでなければならない³⁰。参考事例の視点は公的医療制度の視点となる。管轄区域によっては、この視点に長期ケアや社会福祉サービス、あるいは地域ベースサービスの費用が含まれる。

幅広い視点を採用したことに伴う費用は、分析結果に影響を及ぼす可能性がある場合には、別個に報告しなければならない。こうした費用が生じるのは、その介入を使うと他の方法を使うよりも早く患者が職場に復帰できる、介入により費用が患者および患者の家族に転嫁される（病院ベースのケア対在宅ケア）、あるいは節約ができるようになる、他の公的部門の機関に追加費用が発生する（学習障害のある子どものための特別教育など）といった場合である。こうした費用は可能な場合、個別に定量化し、感度分析を行う必要がある。費用を定量化することが難しい場合は、考えるそれらの費用の大きさや、それらが分析結果に及ぼす影響について考慮すべきである。

個々の視点に関連する費用の種類の詳細を表 2 に示す。資源は特定し、測定し、評価しなければならない。

表 2：経済的評価の視点と関連費用			
視点	費用の種類	例	
社会的視点	公的支払機関	公的資金で行われている（ヘルスケア以外の）サービスにかかる直接費	ホームヘルパーや食事の宅配サービスなどの社会福祉サービス* 支払った所得移転支出（障害給付など） 特別教育
		公的医療制度への直接費	薬剤、医療機器 装置、場所、施設、および関連する間接費 政府が費用を支払う補助機械器具 医療提供者およびその他のスタッフ 処置を含む医療サービス 病院の諸サービス 救急外来 救急車サービス 診察、治療、検診 施設または自宅でのリハビリ* 在宅ケアや社会的支援などの地域ベースのサービス* 老人ホームでの長期ケア*
	公的医療制度	患者およびその家族にかかる直接費	薬剤、歯科治療、歩行器に関する現金支出費（自己負担を含む） 治療を受けるための旅費、有償介護者 民間保険会社への支払済み保険料と保険会社から受け取った給付金† 受け取った所得移転支払（障害給付など）
		患者およびその家族にかかる時間費用‡	患者が移動および治療を受けるために使った時間 患者および患者を介護する家族の無給労働（家事など）による損失時間
		生産性費用	低下した就業能力、あるいは短期または長期の欠勤（摩擦期間中）に起因する生産性損失 雇用者が患者に代わる労働者を雇用し、訓練するための経費

*これらの費用の一部は、その正確な性質や当該管轄区域により、公的医療制度が被ることがある。

†民間保険会社への費用（すなわち、患者から支払われた保険料や患者に支払われた給付金）は、患者の直接費として含まれている。これらの金額は、別な方法で行うもっともな理由がない限り、通常は相殺されるとみなすことができる。雇用者が従業員に対する給与の一部として支払う民間保険会社の保険料は、生産性損失費用の一部として含むことができる。

‡表中の分類システムでは患者および患者の介護をする家族に対する一部の（間接的な）時間費用を除外している。有給労働（摩擦期間を越えた）での損失時間の価値は、私的費用とみなされ、社会的視点の中では取り上げない。失われた余暇時間は QALY が健康アウトカム指標として使われる場合、少なくとも一部は選好測定基準の中で捕捉されるので、費用とはみなされない。

3.6 効果

3.6.1 効力対効果

効力指標と効果指標には違いがある。効力は管理された環境下、たいていの場合ランダム化比較試験（RCT）という状況下での医療技術の能力をいう。研究志向の臨床医によって厳密に書かれたプロトコルに従って管理される試験の参加者は、限定的な試験対象・除外基準に従って選ばれることが多く、その後治療に従うよう励まされ、注意深く監視される。こうした試験は、大学付属病院のような専門的施設で実施することができる。これに対して、効果は「現実」（すなわち、臨床慣行での日常的な使用）における技術の能力をいう。ここでは多様な医療提供者が、広く不均質な患者の集団に対して、適当であるとみなす技術を使っており、対象となる患者は通常あまり十分な説明を受けておらず、検査や診断が正確に行われる可能性が低く、患者の遵守度が低く、併存疾患や本来の効力試験では排除されている治療にさらされる。加えて、医療提供者のコンプライアンスと技能（一部の外科的処置の症例数とアウトカムの関係）、および対象者集団の規模（ワクチンなど）が一部の技術については低いため、アウトカムに悪影響を及ぼすおそれがある。概して「現実」の患者は、RCTの参加者と比べると、治療に応じる可能性が低いことが多い³¹。

意思決定者は主として、当該介入が日常的臨床慣行において治療を受ける患者に対してもたらず「実生活」での影響に関心を持っている。重要な問題は、RCTで得られた効力に関するデータが「現実」の設定下で実現できる可能性のある効果を反映するかどうかということである（すなわち、臨床試験の外的妥当性）。実行可能な場合、経済的評価におけるアウトカムと費用は、評価が管轄区域に関連したものであるべきことから、その介入の効力よりむしろ効果に基づいたものにすべきである²。

3.6.2 最良の証拠の収集

介入の信頼できる臨床的検討が評価の基盤となるべきである。そのためには介入の効力および効果に関して入手できる証拠を系統的にレビューしなければならない。介入の効力および効果に関する系統的なレビューをしなかった場合には、その理由を説明する必要がある。

分析者は、系統的レビューに含まれる研究と、レビューを行うときに使用する方法について説明しなければならない。レビューにはさまざまな設計の研究が含まれることになり、それは多様なレベルの内的および外的妥当性を反映したものになる。個々の研究の実施および報告は重要である。なぜなら実施方法や報告がよい加減な研究（欠測値のある小規模な研究で、参加者の配置についても細かい説明がないなど）では多くの証拠を提示することはできないからである。各形態の証拠によって提示される補足的な情報は、感度分析を

使ってさらに深く分析することができる。データの分析または組み合わせに使用する方法（メタ分析、間接的比較など）については説明をして、選んだ理由を述べ、分析結果は報告しなければならない（ガイドラインステートメント 8.10 参照）。

非薬剤技術はしばしば経済的評価に難題をもたらすことがある。こうした技術の多くはあまり厳密でない臨床試験を受けただけで認可されており、市販の認可や市販後のサーベイランスでも薬剤と比べると余り厳しくない規制の対象になることが多い。臨床研究は設計が貧弱であり（ランダム化比較試験の欠如など）、また研究参加者に対し長期の追跡調査を行っていない可能性がある。したがって、効力の科学的根拠は、薬剤に比べると低い基準のものになりうる³²。最良の証拠を入手するためには、観察研究をもっと行ったり、レジストリーデータベースをもっと利用することが必要になるであろう。

内科的・外科的処置のアウトカム（および時には費用）や診断技術は、技術施行者の技能や経験によって決まってくる可能性がある。アウトカムに与える影響を測定することは難しいかもしれない。医療機器の評価をする場合、その機器の技術的な能力だけでなく、むしろそれを使ったケア全体のエピソードに焦点を当てるべきである。例えば、診断用機器の評価には、機器の感度や特異性がフォローアップケアや健康アウトカムに与える影響を査定することが含まれるかもしれない。効力や効果の証拠が弱い状況を評価するために、大々的に感度分析を行うことになる。こうした分析の推論的性質については、適切な警告を載せるべきである。

介入に関連する有害事象は、その介入と代替策との間の有意義な違いに基づいて、それが臨床的あるいは経済的に重要である場合、評価に盛り込むべきである。分析者はそうした判断をする前に、その性質、頻度、持続期間、および重症度について考慮する必要がある。有害事象は以下のようなものに影響を与える可能性がある。

- ・ 治療の継続や持続（治療の切り替え、投薬量の削減、あるいは中断など）
- ・ 患者の治療に対するコンプライアンス
- ・ 死亡率
- ・ 罹患率
- ・ HRQL（効用）
- ・ 資源利用（受診、入院、あるいは入院の延長など）

こうした事象は、経済ガイドラインに示されている原則に沿った形で評価しなければならない。改善した有害事象や副作用の特徴を表す証拠は、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。臨床試験で 2 次的なアウトカムとして現れた軽微な副作用に見られる重要でない違いは、臨床的には関連性が

ない。多くの軽微な副作用を同時に体験し、効果に影響を及ぼすおそれがある場合には、分析者は、それらが治療の遵守と関係があるかどうかについて論評しなくてはならない。さらに、ある介入に関連する最大の潜在的損害が、市場導入前の臨床試験では出てこないこともある。したがって、通常、事例の自発報告や市販後の臨床試験といった他の研究設計を使用する必要がある。

薬剤については、有害事象はその薬剤を使っている患者の予想投与量の範囲に関連するものでなければならない。すべての研究について、有害事象による薬剤の全面的使用中止（理由が内訳けされているもの）、および薬剤不耐性のために用量の削減を要する患者数は、当該薬剤に対する患者の耐性指標と考えるべきである。これは感度分析を通して調査することができる。

3.6.3 モデリング

参考事例は、感度分析を用いて行った推定に不確実性があったとしても、「現実」の効果に関して最良の量的推定に基づいたものにするのが望ましい。介入が市場で使われる前に「現実」の良好な証拠が入手できることはほとんどないため、分析者は、実行可能で科学的に信頼できる場合には、入手できる最良の証拠と適切なモデリング手法を使って、効力のデータを効果の推定値に変換することが奨励される。

効力のデータを効果の推定値に変換するときには使用した方法や仮定については、その説明とそれを用いた理由を述べなければならない。患者の健康にとって最も実現可能で（データがある場合）、最も重要なアウトカムを特定し、評価する。該当する場合、分析者はデータを外挿、または調整する理由となった証拠の強さについて説明しなければならない。これには、代用アウトカムに関する試験データを最終アウトカムへ外挿する、あるいは短期のアウトカムに関するデータを臨床試験期間を超えて外挿するといったことが含まれる。また、臨床試験の治療作用を修飾して、効力の臨床試験の因子とは違う「現実」の因子を説明することも、これに含まれる（患者の治療に対する遵守など）。

入手できるデータの性質にもよるが、モデリングをするには観察されたデータを適切なアウトカムに変換させる必要がある。アウトカムに関する短期のデータしか手に入らない場合には、そのデータを長期のアウトカムに外挿するのも適切であろう。データを外挿する理由となった証拠の強さについて説明し、使用するモデリング方法や仮定についてそれを用いた理由を述べる。

短期のデータを外挿するのに時々使用される方法の 1 つは、ベースライン・アウトカムの推定値を重ね合わせるというものである（臨床試験の治療効果の推定値に関する観察研究

により引き出された自然な生長による生存確率など)³³。外挿が妥当であるかどうかは、疫学的データの質、および介入によって修飾可能な危険因子と長期のアウトカムとの因果関係によって決まることが多い。臨床試験後の臨床上の便益の持続期間と大きさは、しばしば外挿に関して行う重要な判断となる。

分析者は代用アウトカムをより重要な患者アウトカムに結びつけるためにモデリングを使うことが奨励される。代用アウトカムを使用した研究の場合、その代用アウトカムは重要な患者アウトカムを大いに予示するものでなければならない。例を挙げると、心疾患の場合、血圧やコレステロールレベルといった代用エンドポイントを臨床エンドポイント（冠状動脈性心臓病の発症など）や、その後の最終アウトカム（全原因心血管系疾患死亡率および罹患率など）と結びつけるためにモデリングが使われるが、こうした結びつきの証拠があるかどうかによって決まってくる。こうした分析は感度分析をすることによって適切に検証しなければならない。

効力のデータを使って効果を推定するとき、いくつかの要素が考慮されることが多い。

- ・ 診断や検診の精度
- ・ 患者の治療に対する遵守
- ・ 医療提供者のコンプライアンスや技能
- ・ 併存疾患のための治療と研究では許可されない併用治療の使用に対する、下位集団の反応に見られる有意義な違い

著者によっては「遵守」と「コンプライアンス」を同義語として使用している人もいる。「遵守」という用語は経済ガイドラインにおいては、治療を受ける患者に関して3つの構成要素を包含するために使われている。すなわち、受容（治療を受けることを決めた患者の最初の決断）、持続（治療を長期間継続する）、およびコンプライアンス（患者が推奨された治療計画に従うときの一貫性と正確さ）である³⁴。患者がこの3つの構成要素を満たしたとき、遵守が達成される。

正確な診断（または検査選択）、患者の遵守、それに医療提供者のコンプライアンスや技能は、RCTにおいてよりも「実生活」で低いことが多い。医療機器や外科的処置といった技術に関していえば、そのアウトカムは技術施行者の経験や技能により決まってくる。患者の不遵守により、治療効果が臨床検査で観察されたものよりも低くなり、治療費用は高くなり、患者の職場での生産性が低下し、介護者の負担が大きくなり、薬剤耐性が生じる可能性がある。不遵守原因は患者が薬剤に対する有害作用を経験したから、あるいは薬剤投与の回数または容易さといったその他の要因である。

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に使用すべきである。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（あるいは代替アウトカム）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ー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で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。データの限界をはっきりさせ、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に、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を数値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-6 D⁶⁰、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

寄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、これは実際のでない、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、支払意思額（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