

2021 年度厚生労働行政推進調査事業費補助金(政策科学総合研究事業)  
「医薬品・医療機器等の 費用対効果評価における公的分析と公的意思決定方法に関する研究」

分担研究報告書

我が国における費用対効果評価分析ガイドラインの改定に関する研究

分担研究者 白岩 健(国立保健医療科学院)

費用対効果評価における標準的な分析手法を記述した「分析ガイドライン」は費用対効果評価を制度で活用している多くの国で作成されている。一般的に費用対効果評価の分析ガイドラインは、分析手法の標準化や分析の質を確保するなどの目的のために、各国の制度のあり方を反映して作成される。我が国における費用対効果評価の活用は、保険償還の可否には用いず、価格調整のみに活用する多くの諸外国とは異なる方法であることから、諸外国の分析ガイドラインを参照しつつも検討すべき課題が多い。

我が国における分析ガイドライン作成は、2012 年度頃までに研究者向け医療経済評価ガイドラインが厚生労働科学研究補助金によって作成され、2013 年 12 月には参考資料として左記ガイドラインが中央社会保険医療協議会(中医協)費用対効果評価専門部会に提出されている。費用対効果評価の制度化議論が進み、2016 年 4 月から試行的導入が開始されることが決まると、そのために使用される分析ガイドラインが作成され、2016 年 2 月に費用対効果評価専門部会によってガイドライン(「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」)が承認された。その後、2019 年に費用対効果評価の制度化が開始され、その直前の 2019 年 2 月に「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 2 版」について承認がなされ、現在に至っているところである。

しかし、2021 年度は費用対効果評価制度に関する各種の改正等が行われることとなっており、それとあわせて「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」も改定を行い、この間の科学的議論の進展や制度改正に対応することとした。

一方で、制度を運営していく仮定で、分析ガイドラインとしては十分使用に耐えうるものであることが明らかになっていることから、以下の 14 章からなるガイドラインの骨格は維持し、必要な箇所等について改定を行うこととした。「1 ガイドラインの目的」「2 分析の立場」「3 分析対象集団」「比較対照技術」「5 追加的有用性」「6 分析手法」「7 分析期間」「8 効果指標の選択」「9 データソース(費用を除く)」「10 費用の算出」「11 公的介護費・生産性損失の取り扱い」「12 割引」「13 モデル分析」「14 不確実性の取り扱い」

諸外国の分析ガイドラインや評価終了品目における課題などを参照しながら研究班員間で検討を行い、作成した「第 3 版」のガイドラインを次ページ以降に示す。

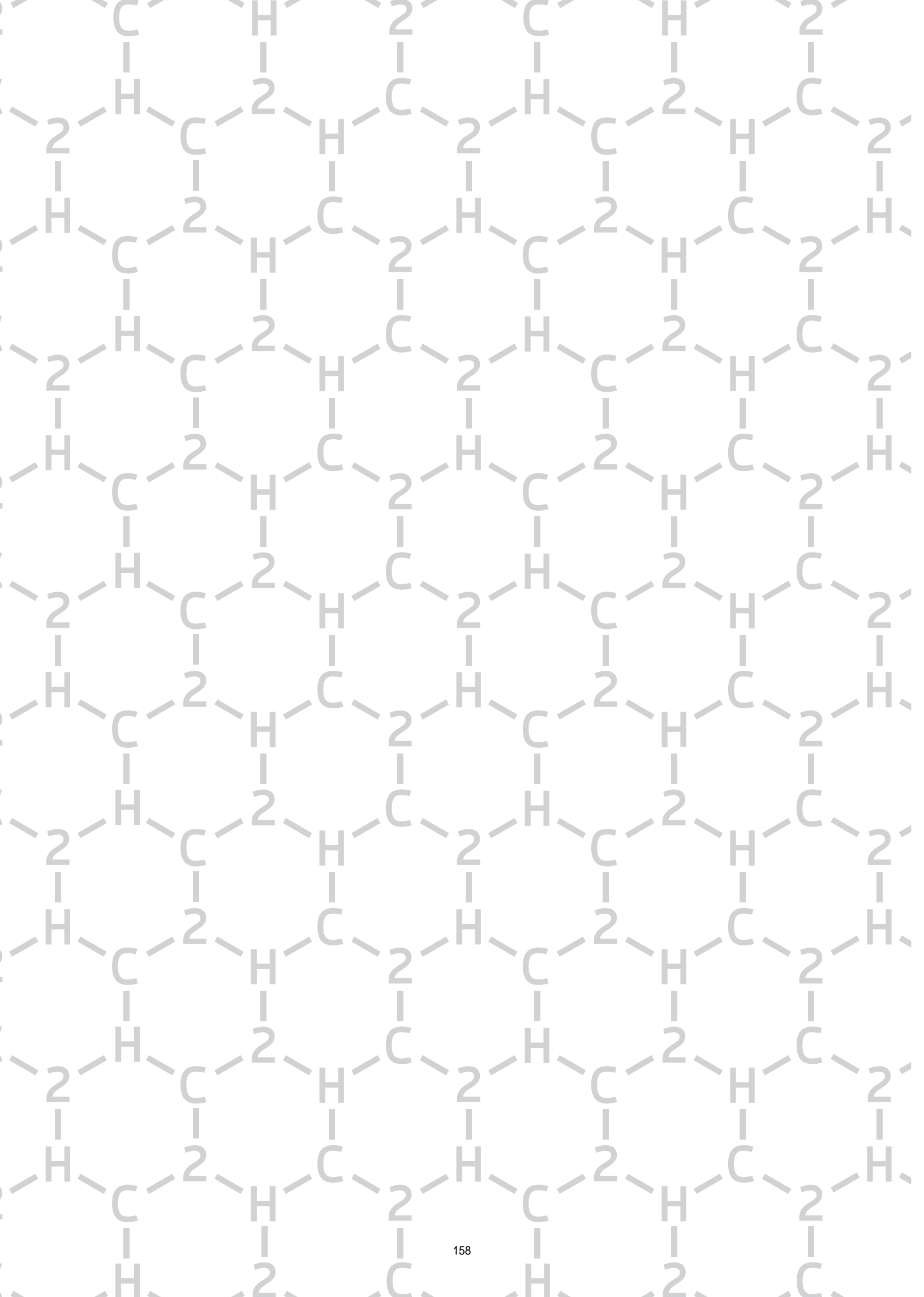


---

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

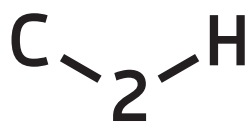
2022年1月19日 中医協総会了承

---



目次

1	ガイドラインの目的	4
2	分析の立場	5
3	分析対象集団	6
4	比較対照技術	7
5	追加的有用性	8
6	分析手法	11
7	分析期間	12
8	効果指標の選択	13
9	データソース(費用を除く)	15
10	費用の算出	16
11	公的介護費・生産性損失の取り扱い	19
12	割引	21
13	モデル分析	22
14	不確実性の取り扱い	23
	用語集	24
	略語一覧	29



---

## 1 ガイドラインの目的

- 1.1 本ガイドラインは、中央社会保険医療協議会において、評価対象として選定された医薬品・医療機器等(以下、評価対象技術)の費用対効果評価を実施するにあたって用いるべき分析方法を提示している。
- 1.2 本ガイドラインは製造販売業者により提出される分析と公的分析を対象としている。

---

## 2 分析の立場

- 2.1 分析を行う際には、分析の立場を明記し、それに応じた費用の範囲を決めなければならない。
- 2.2 費用や比較対照技術、対象集団などについて公的医療保険制度の範囲で実施する「公的医療の立場」を基本とする。
  - 2.2.1 「公的医療の立場」以外の立場から分析を実施する場合でも、「公的医療の立場」の分析を実施しなければならない。
  - 2.2.2 一部の予防技術(検診やワクチン等)など厳密には公的医療保険制度に含まれないものでも、それに準ずる医療技術であれば、公的医療保険制度に含まれるものと同様の取り扱いをした分析を実施する。
- 2.3 公的介護費へ与える影響が、評価対象技術にとって重要である場合には、「公的医療・介護の立場」の分析を行ってもよい。
- 2.4 評価対象技術の導入が生産性に直接の影響を与える場合には、より広範な費用を考慮する立場からの分析を行い、生産性損失を費用に含めてもよい。

### 3 分析対象集団

- 3.1 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を分析対象集団とする。
- 3.1.1 品目の指定から分析枠組み決定までの間に適応症が追加される場合は、それらの適応症についても原則として分析対象に含める。
- 3.1.2 「3.1.1」で定める期間以降に新たな適応症が追加され、評価結果に影響を与えると考えられる場合、当初の評価終了後に改めて評価を実施する。
- 3.2 複数の適応がある場合、あるいは同一疾患内においても治療成績や使用方法・用法用量、比較対照技術が異なる主要な集団がある場合は、各集団についてそれぞれ分析を実施することを原則とする。
- 3.2.1 ただし、「3.2」を実施することが困難な場合は、患者数や疾患の性質等を勘案して、協議における両者(製造販売業者と国立保健医療科学院/公的分析班:以下同様)の合意のもとで集団を選択することとする。
- 3.3 各集団の患者割合は、評価対象技術の対象患者に関する最新の臨床実態に基づくことを原則とする。

---

## 4 比較対照技術

- 4.1 評価を行う際の比較対照技術は、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ことが原則的な考え方である。
- 4.1.1 比較対照技術としては無治療や経過観察を用いることもできる。
- 4.1.2 「4.1.1」の場合を除いて、比較対照技術は原則として公的医療保険で使用が認められているものとする。
- 4.1.3 ただし、「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(Randomized controlled trial: RCT)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。
- 4.2 比較対照技術として選定した理由については十分に説明する。



## 5 追加的有用性

- 5.1** 費用対効果を検討するにあたっては、評価対象技術の比較対照技術に対する追加的な有用性がデータによって示されているか、示されていないかをまず評価する。
- 5.2** 追加的な有用性を検討する際に、「4.」で選定した比較対照技術に対するRCTのシステマティックレビュー (Systematic review: SR)を実施し、追加的有用性の有無を評価する。適切なものであれば公開されていない臨床研究や治験の結果等を含めてよい。
- 5.2.1** SRを実施する際には、リサーチクエスチョン (Research question: RQ) を明確に提示する。例えば PICO (P: 患者 (Patient), I: 介入 (Intervention), C: 比較対照 (Comparator), O: アウトカム (Outcome)) などで構造化されたRQを定義する。
- 5.2.2** 評価対象技術あるいは「4.」で選定した比較対照技術とアウトカムが同等であると考えられる同一作用機序あるいは同一機能区分等の類似技術について、協議の上で適切と判断されれば、それらをSRにおける介入(I)あるいは比較対照(C)に含めてもよい。
- 5.2.3** 「5.2.1」におけるアウトカム(O)指標は、臨床的な有効性・安全性・健康関連QOLの観点のうち、評価対象技術の特性を評価する上で、適切なもの(真のアウトカム指標など)を用いる。
- 5.2.4** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 声明の原則に従い、研究の組み入れ基準・除外基準、使用したデータベース、検索式、論文選択のプロセス(情報フロー図を含む)等を記載する。

- 5.2.5** 信頼できる既存のSRが存在する場合、それらを活用することは可能であるが、RQに合致するか、最新の文献まで含まれているか等を検討した上で、そのまま使用するあるいは追加的なSRを実施して研究を追加する等の検討を行う。
- 5.2.6** 実施することが適切な場合は、メタアナリシスの手法を用いて各試験を統合した結果を提示する。その場合、用いた統計手法、異質性の評価、フォレストプロット、統合した値とその信頼区間等を明らかにする。
- 5.2.7** RCTが存在しないことが明らかな場合、協議における両者の合意のもとで、「5.2」のプロセスは省略できるものとする。
- 5.2.8** SRにおける文献検索終了時点は、分析枠組みが決定された後から製造販売業者による分析提出までの一時点に決める。
- 5.2.9** 「5.2.8」の文献検索期間終了後に費用対効果評価において重要と考えられる臨床研究(症例数の多い主要な研究結果、現在の知見と異なる結果が得られた信頼性の高い研究等)が公表等された場合、それらを含めることの必要性について検討する。その際には追加的なSRを実施しなくてもよい。
- 5.3** 「5.2」のSRの結果、適切なものが存在しない場合、「5.2」のプロセスに基づき、アウトカムを比較した非RCT(観察研究等)のSRを実施し、追加的有用性を評価する。ただし、研究の質(研究デザイン、群間での患者背景の差異、統計解析手法、サンプル数や施設数等)について十分に説明しなければならない。

- 5.4 より信頼性の高い結果が得られると考えられる場合、協議の上で適切と判断されれば、既存の観察研究やレジストリーデータなどを再解析した結果をもって、追加的有用性の有無を評価してもよい。ただし、研究の質(研究デザイン、群間での患者背景の差異、統計解析手法、サンプル数や施設数等)について十分に説明しなければならない。
- 5.5 「4.」で選定した比較対照技術とのRCTは存在しないが、その他のRCTが存在する場合、SRの結果を用いて、間接比較により追加的有用性の評価をしてもよい。
- 5.6 単群試験しか存在しない場合は、評価対象技術と比較対照技術それぞれについてのSRに基づき、間接比較を実施する。
- 5.7 間接比較を行う場合、以下を原則とする。
- 5.7.1 個人レベルのデータを用いることができる場合、MAIC(Matched adjusted indirect comparison: マッチング調整された間接比較)等を用いて患者背景の違いなど適切な補正を行う必要がある。
- 5.7.2 個人レベルのデータを用いることができない場合、複数のRCTを用いたadjusted indirect comparison(調整された間接比較)あるいはネットワークメタアナリシスなどに基づくことが望ましい。
- 5.7.3 個人レベルデータも比較試験結果も利用できない場合、その他の代替手段が存在しなければnaïve indirect comparison(単純な間接比較)を用いて評価をすることもできる。この場合、結果の不確実性について慎重に検討する。

**5.7.4** 間接比較を行う場合は、間接比較を可能とする前提条件(疾患、重症度、患者背景等の異質性や試験の同質性など)についても十分に説明しなければならない。

**5.8** 「5.3」から「5.7」までの手法において得られた結果が、研究の質に課題があると判断されるものの治療効果が劣っているとは考えられない場合には、評価対象技術のアウトカムが比較対照技術と同等であるという前提で、「6.」の分析を実施する。

**5.9** ヒトを対象とした適切な臨床研究が存在しない場合、適切ならば、医薬品医療機器総合機構からの承認により、比較対照技術と同等であるという前提で、「6.」の分析を実施する。

**5.10** 「5.2」から「5.7」までの結果、アウトカムが比較対照技術と比べて劣ると判断される場合は、費用対効果の分析は実施しない。

## 6 分析手法

- 6.1 効果を金銭換算せず、費用と効果を別々に推計する費用効果分析を分析手法として用いることを原則とする。
- 6.2 「5.」の分析に基づき、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(Incremental cost-effectiveness ratio: ICER)を算出する。
- 6.3 ただし、以下の場合については、各群の期待費用と期待効果のみを提示をして、ICERは算出しないこととする。
  - 6.3.1 比較対照技術と比べて効果が高く、かつ費用が安い場合。このとき、ICERを算出せずに優位(dominant)であるとする。
  - 6.3.2 「5.」の分析により、追加的有用性が示されていないものの、アウトカムは同等と考えられる場合には、比較対照技術と費用を比較する(いわゆる「費用最小化分析(Cost-minimization analysis :CMA)」。このとき、結果は費用削減あるいは費用増加とする。
- 6.4 「3.2」あるいは「3.2.1」により、分析対象となる複数の疾患や、同一疾患内での複数の分析集団がある場合、ICERは疾患や集団ごとに算出する。
- 6.5 評価対象技術に関する既存の費用効果分析や主要な諸外国の医療技術評価機関において公表された評価結果が既に存在する場合は、それらもあわせて示す。

---

## 7 分析期間

- 7.1 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。
- 7.2 費用と効果は、原則として同じ分析期間を用いる。
- 7.3 分析期間については、そのように設定した理由を説明すること。

## 8 効果指標の選択

**8.1** 効果指標は質調整生存年(Quality-adjusted life year: QALY)を用いることを原則とする。

**8.1.1** QALYを算出することが困難であり、かつCMAを実施する場合、適切であれば、QALY以外の評価尺度を使用することもできる。

**8.2** 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。

**8.2.1** 費用効果分析を行うために、新たに日本国内でQOL値を収集する際には、EQ-5D-5Lを第一選択として推奨する。

**8.2.2** 「8.2」に該当するデータが存在しない場合、その他の適切な健康関連QOL(Health-related quality of life: HRQOL)データからQOL値へマッピングしたものを使用してもよい。マッピングにより得られた値を使用する場合、適切な手法を用いてQOL値に変換していることを説明しなければならない。

**8.3** PBMによりQOL値を測定する場合には、対象者本人が回答することが原則である。

**8.3.1** PBMを用いる場合、対象者本人からQOL値が得られない場合に限り、家族や介護者等による代理の回答を用いてもよい。

**8.3.2** PBMを用いる場合、医療関係者による代理回答は、対象者本人の回答と乖離する可能性があるため、避けることが望ましい。

- 8.3.3** 対象者本人からQOL値を得ることが困難な場合などには、一般の人々を対象に健康状態を想起させることによりSG(Standard gamble: 基準的賭け)法やTTO(Time trade-off: 時間得失)法、DCE(Discrete choice experiment: 離散選択実験)法などの手法を用いて測定してもよい。ただし、これらの手法で測定されたQOL値は、提示するシナリオにより結果に大きな影響を与えるため、同一の健康状態においては同一のQOL値を用いるなどの留意が必要である。
- 8.4** QOL値は、「8.2」および「8.3」を満たすものがある限り、国内での調査結果を優先的に使用することを推奨する。
- 8.4.1** ただし、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。
- 8.5** 「公的医療・介護の立場」からの分析の場合、実際のデータがあれば家族等の介護者や看護者に与えるQOL値への影響について考慮に入れてもよい。



## 9 データソース(費用を除く)

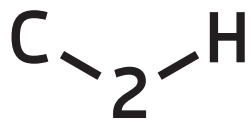
- 9.1 ICER等を算出するにあたって使用する有効性・安全性・QOL値等のデータ(モデル分析を実施するにあたって使用する推移確率等のパラメータも含める)については原則として、研究の質やエビデンスレベルが高く、かつ日本における現実の臨床成績を反映しているものを優先的に使用する。
- 9.1.1 有効性・安全性・QOL値等のデータ選定においては、国内外の臨床研究のSRに基づくことを推奨する。適切なものであれば公開されていない臨床研究や治験の結果等を含めてよい。
- 9.1.2 原則としてエビデンスレベルの高いデータの使用を優先すべきであるが、研究の質や分析における対象集団、結果の外的妥当性等を勘案して適切なものを使用することを推奨する。(例:RCTの結果が、実際の臨床成績と大きく乖離している可能性があるなど)
- 9.1.3 適切な場合は、既存の臨床研究やデータベース等を再解析してもよい。その場合は、患者背景や解析手法等について詳細を明らかにする。
- 9.2 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。
- 9.3 評価対象技術と比較対照技術で統計的に有意な差が認められない場合は、両群で同じ値を用いることを原則とする。両群で同じ値を用いない場合は、支持するデータや理由、治療効果の大きさ(臨床的に意味のあるものか)などについてあわせて検討を行い、それが妥当であることを説明しなければならない。

- 9.4 医療機器等の評価において、科学的に信頼できる定量的なデータがある場合は、協議における両者の合意のもとで、上記のデータに基づく分析とは別に、いわゆる習熟効果(経験の蓄積による治療効果等の改善)や製品改良による効果を反映した分析をあわせて提出してもよい。

## 10 費用の算出

- 10.1 「公的医療の立場」においては公的医療費のみを費用に含める。
- 10.2 各健康状態の費用は、評価対象技術によって直接影響を受ける関連医療費のみを含め、非関連医療費は含めないことを原則とする。
- 10.3 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。
- 10.4 各健康状態の費用の推計において、適切な場合には、「10.3」の観点から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。ただし、レセプト上で健康状態の定義が困難である、評価時点においてデータの十分な蓄積がないなど、推計の実施が困難な場合はその限りではない。
- 10.4.1 レセプトデータを用いて推計する場合、各健康状態の定義とその根拠を示さなければならない。
- 10.4.2 費用を推計するために用いた手法(外れ値処理や非関連医療費の除外などを含む)については、用いたものとその根拠を示さなければならない。
- 10.5 各健康状態の費用の推計において、レセプトデータベースを用いることが困難と判断される場合、あるいはより適切であると判断される場合は、標準的な診療プロセス等に基づき積み上げで算定してもよい。
- 10.5.1 積み上げで算定する場合は、「10.3」の観点から、その算定根拠について示すこと。関連する資源消費項目と消費量について、レセプト等のデータベースを用いて同定するなどしてもよい。
- 10.5.2 費用を積み上げで推計する場合は、医療資源消費量と単価を区分して集計、報告することを原則とする。

- 10.5.3** 「10.5.2」の資源消費量の集計において、注射剤のバイアルについては、残量廃棄を原則とする。
- 10.6** 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。
- 10.7** 公的医療費については、保険者負担分のみならず公費や患者負担分も含めて費用として取り扱う(公的医療費の全額)。
- 10.7.1** 「2.2.2」の原則に応じて、検診やワクチン等の公的医療費に準じる費用も費用として含める。
- 10.8** 単価は可能な限り最新時点の診療報酬点数表や薬価基準等を使用する。特に評価対象技術あるいは比較対照技術については必ず最新時点の価格を用いなければならない。
- 10.8.1** 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。
- 10.8.2** 結果に与える影響が無視できる程度である場合には、分析実施時点に調整しないことも許容する。
- 10.9** 比較対照技術に後発医薬品が存在する場合は、それらの価格を用いた分析もあわせて提出する。
- 10.10** 評価対象技術あるいは比較対照技術の費用が包括支払いの対象となっている場合は、出来高で費用を算出する。



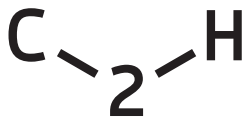
**10.11** 将来時点に発生する費用も、現時点における医療資源消費や単価に基づき推計したものを用いる。

**10.12** 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。単価は国内のものを反映させなければならない。

---

## 11 公的介護費・生産性損失の取り扱い

- 11.1 「公的医療・介護の立場」では、基本分析に加えて、公的介護費を含める追加的分析を実施することができる。なお、公的介護費は国内の知見に基づき推計されたものを用いる。
- 11.2 公的介護費を費用に含める場合は、要介護度・要支援度別に費用を集計することを推奨する。
- 11.3 要介護度ごとの公的介護保険の利用額は、対象疾患等における実際の資源消費量に基づくことが原則であるが、測定することが困難な場合は平均的な受給者1人当たり費用額等を用いてもよい。
- 11.4 生産性損失を含めた分析を行う場合、基本分析に加え、追加的分析として実施することができる。ただし、生産性損失を含めることができるかどうかは、疾患の特性等による就業可能性を考慮しなければならない。また、生産性損失は国内の知見に基づき推計されたものを用いる。
- 11.5 生産性損失の減少は、
- (A) 医療技術に直接起因するもの(治療にともなう入院期間の短縮等)
  - (B) アウトカムの改善(病態の改善や生存期間の延長等)を通じて間接的に生じるもの
- に分けて考えることができる。
- 生産性損失を分析に含める場合には、原則として(A)のみを費用に含めることとする。
- 11.6 生産性損失は、人的資本法を用いて推計することを基本とする。これは、当該疾患に罹患していなければ、本来得られたであろう賃金に基づき推計する方法である。



- 11.6.1** 生産性損失を推計する際に単価として用いる賃金は、公平性等を考慮して、最新の「賃金構造基本統計調査」(賃金センサス)に基づき、全産業・全年齢・全性別の平均あるいは全産業・全性別の年齢階級別の平均を用いることとする。
- 11.6.2** 生産性損失を推計するにあたっては、対象となる集団において就業状況を調査し、実際に仕事等に従事できなかった日数や時間を測定する。これに全産業・全年齢・全性別の平均賃金を乗じて生産性損失を推計することが原則である。
- 11.6.3** 「11.6.2」の実施が困難な場合、対象集団において仕事等に従事できないと推計される日数(休日は除く)や時間に全産業・全年齢・全性別の平均賃金を乗じて生産性損失とする。ただし、この方法は18歳以上の就業率を100%と仮定するものであり、生産性損失を過大推計する可能性のあることに留意が必要である。
- 11.7** 家族等による看護や介護のために本人以外の生産性が失われることが明らかな場合は、本人の生産性損失と同じ条件・取り扱いのもとで費用として含めてもよい。
- 11.8** 仕事等の減少とは無関係な時間費用等については含めないこととする。

---

## 12 割引

**12.1** 将来に発生する費用と効果は割引を行って、現在価値に換算しなければならない。

**12.1.1** ただし、分析期間が1年以下、あるいは短期間でその影響が無視できる程度であるときは、割引を行わなくてもよい。

**12.2** 費用・効果ともに年率2%で割引を行うこととする。

**12.3** 割引率は、感度分析の対象とし、費用・効果を同率で年率0%から4%の範囲で変化させる。



## 13 モデル分析

- 13.1 「7.」の原則に基づき、予後や将来費用を予測するために決定樹モデル、マルコフモデル等を用いたモデル分析を行ってもよい。
- 13.2 モデル分析を行う際には、そのモデルの妥当性について示さなければならない。例えば、
- (A) 内的妥当性: なぜそのような構造のモデルを構築したのか、病態の自然経過を十分にとらえられているか、使用しているパラメータは適切なものか等
  - (B) 外的妥当性: 既存の臨床データ等と比較して、モデルから得られた推計が適切なものであるか等
- 13.3 モデルを構築する際に使用した仮定については明確に記述する。
- 13.4 モデルを構築する際に使用したパラメータとそのデータソースについてはすべて提示する。
- 13.5 使用したモデルや計算過程については電子ファイルの形式で、第三者の専門家が理解でき、かつ原則としてすべての主要なパラメータ(推移確率、QOL値、費用)を変更できる形で作成し、提出する。
- 13.5.1 費用については、その全体額のみでなく、それらの詳細(積み上げの場合は各資源消費量と単価など)を変更できるようにしておくことが望ましい。特に評価対象技術や比較対照技術の単価については、第三者が変更できるようにしなければならない。
- 13.6 マルコフモデルを用いて解析する場合、1サイクルあたりの期間が長く、結果に影響を与える場合は、半サイクル補正を実施する。

---

## 14 不確実性の取り扱い

- 14.1 分析の設定等が一意に定まらず、それらの違いが結果に影響を与える可能性がある場合は、複数のシナリオに基づいた分析を行う。
- 14.2 分析期間が長期にわたり不確実性の大きい状況では、臨床研究のデータが存在する期間を分析期間とするなど、より短期の分析もあわせて行う。
- 14.3 「5.」において比較対照技術との比較試験が存在しない場合、特に単群試験の結果同士を比較した場合は、不確実性が大きいので十分に広い範囲での感度分析を実施する。
- 14.4 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については、感度分析の対象とする。
- 14.5 推定値のばらつきのみが問題となる場合(パラメータの不確実性を取り扱う場合)、感度分析で動かす幅としては、95%信頼区間などを参考に設定する。
- 14.6 確率的感度分析(Probabilistic sensitivity analysis: PSA)もあわせて実施することが望ましい。その場合、使用したパラメータの分布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(Cost-effectiveness acceptability curve: CEAC)を提示する。

用語集

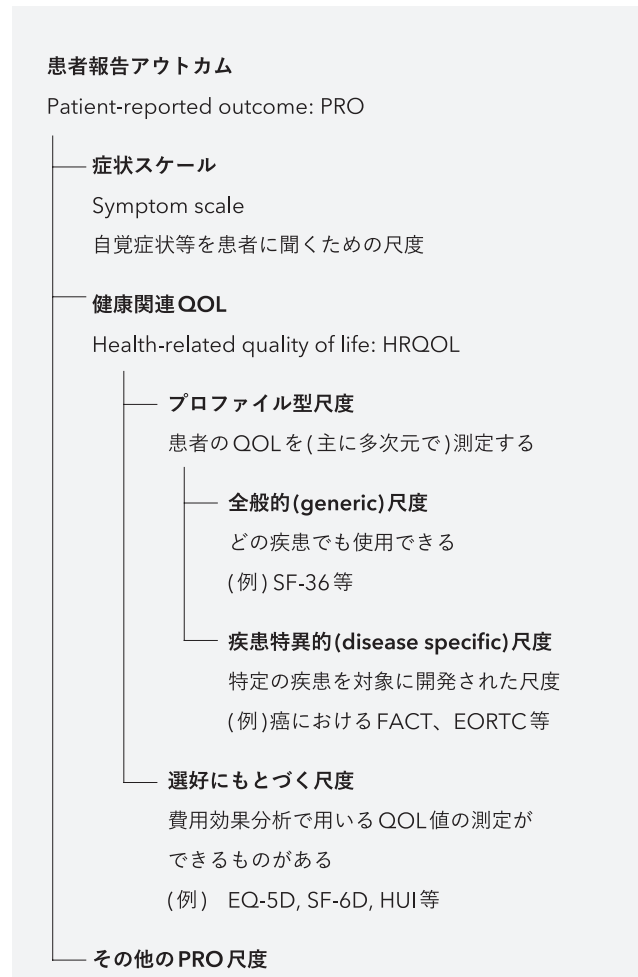
QOL値

人々の健康状態(から得られる価値)を0(死亡)から1(完全な健康状態)に基準化して、一次元で数値化したもの。ただし、「死ぬより悪い」健康状態として負のスコアも取りうる。

QOL値の測定方法には大別して、仮想的な(あるいは本人の)健康状態に対して、その状態のQOL値を一般の人々を対象に質問する「直接法」(基準的賭け(SG)法、時間得失(TTO)法など)と、QOL質問票により得られた回答からスコアリングアルゴリズムを用いてQOL値を算出する「間接法」が存在する。なお、TTOとSGで測定された値には、系統的な差が存在する可能性があるため留意が必要である。

すべての患者報告アウトカム(Patient-reported outcome: PRO)やQOL尺度での測定値から費用効果分析で使用されるQOL値が算出できるわけではないことに注意が必要である。費用効果分析で利用できるのは、下記のようにQALYを算出するために開発された選好にもとづく(preference-based)尺度で測定したもののみである。

現在のところ、日本でスコアリングアルゴリズムが開発されている尺度としてはEQ-5D(EuroQol 5 dimension)、HUI (Health Utilities Index)等がある。



---

## エビデンスレベル

エビデンスレベルには様々な分類法が存在するが、Minds (Medical Information Network Distribution Service) では以下のよう

I	システマティック・レビュー / RCTのメタアナリシス
II	1つ以上のRCTによる
III	非ランダム化比較試験による
IV a	分析疫学的研究 (コホート研究)
IV b	分析疫学的研究 (症例対照研究、横断研究)
V	記述研究 (症例報告やケース・シリーズ)
VI	患者データに基づかない、専門委員会や専門家個人の意見

ただし、RCTのような実験的研究が現実の臨床成績と乖離している可能性はしばしば指摘されている。エビデンスレベルの高いものの使用を原則としつつも、状況に応じた適切なデータによってを採用する必要がある。

---

## 確率的感度分析

確率的感度分析 (Probabilistic sensitivity analysis: PSA) は、モデルのパラメータに分布を当てはめること等により、増分費用や増分効果、ICERの分布を得るための手法である。確率的感度分析の結果は、費用効果平面上に散布図をプロットし、CEACとして  $f(\gamma) = \Pr(\gamma \cdot IE - IC > 0)$  を書くことが一般的である (IC: 増分費用、IE: 増分効果、 $\gamma$ : 支払意思額)。

---

## 間接比較

例えば臨床試験によって“A vs. B”と“A vs. C”の結果が得られているとき、これらの結果から直接比較のない“B vs. C”の結果を推測することを間接比較 (indirect comparison) と呼ぶ。適切な比較対照を用いて分析しようにも直接の比較試験がない場合、間接比較を適用することができる場合がある。

間接比較が成り立つためには“A vs. B”の結果が“A vs. C”の集団にも適応できること、逆に“A vs. C”の結果が“A vs. B”の集団にも適応できることが条件となる。このことを同質性 (similarity) の仮定と呼

ぶ。間接比較を行うにあたっては、このような仮定に関する検討や、適切な統計手法 (例えば、単純な (naïve) 間接比較ではなく調整された (adjusted) 間接比較) を使用することが必要である。また、ネットワーク・メタアナリシス (あるいは多群間治療比較 (multiple treatment comparison: MTC) とも呼ばれる) のようなより高度な手法を用いた分析についても検討する。

---

## 感度分析

不確実性が存在する場合等に、パラメータの値を変化させることにより、結果への影響を見ることを感度分析という。1つのパラメータを変化させる一次元感度分析、2つのパラメータを同時に動かす二次元感度分析、複数のパラメータの不確実性を同時に取り扱うPSA (確率的感度分析の項を参照) などがある。

---

## システマティックレビュー

システマティックレビュー (SR) とは特定の課題について文献等を網羅的に探索し、可能な限りバイアスなくその結果や評価報告する手法である。Mindsによれば『実際の作業面から定義すると、SRとは「クリニカル・エッセイに対して、研究を網羅的に調査し、研究デザインごとに同質の研究をまとめ、バイアスを評価しながら分析・統合を行うこと』とされる。

しばしばSRとメタアナリシスが混同して使用されることもあるが、SRにおいて得られた結果は必ずしも統計的に統合する必要はなく、このようなものを「定性的システマティックレビュー」と呼ぶこともある。結果を統合することが適切な場合は、SRの結果に基づき、メタアナリシスを実施することとしている。

なお、SR (メタアナリシス) の報告様式としては、国際的にPRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 声明が標準的に使用されている。

---

## 質調整生存年

質調整生存年 (QALY) は、生存年にQOL値を乗じることにより得られる。QOL値が1は完全な健康を、0は死亡を表す。QOL値0.6の健康状態で2年間生存した場合、生存年は2年だが、 $0.6 \times$

2=1.2QALY(完全に健康な状態で1.2年生きたのと同じ価値)と計算される。時間とともにQOL値が変化する場合、図のようにQOL値の経時変化をあらわす曲線下面積が獲得できるQALYとなる。



## 人的資本法

生産性損失を、本来得られたであろう賃金に基づき推計する方法を人的資本法(human capital method)と呼ぶ。しかし、完全雇用が実現されていない状況下では、その人が働けなくてもかわりの誰かが働くはずであり、長期的には必ずしも生産性が失われるわけではない。そのため、社会における生産性がもとの水準に回復するまでにかかる摩擦費用(friction cost)のみを含めるべきという意見もある。賃金は、疾病により実際に労働できなかった期間を調査することにより推計すべきであるが、そのような作業が困難な場合は、家事労働を含めたデータが存在しないため就業率を100%として取り扱ってもよい。また、賃金の単価については、公平性の観点から、実際の単価にかかわらず全産業・全年齢・全性別の平均賃金を用いる。

## 生産性損失

病気が原因で仕事や家事ができなくなることによる社会的な損失(あるいは早期に回復できることによる社会的な便益)は生産性損失(productivity loss)として、分析の立場によっては費用に含めることができるが、基本分析では含めないこととしている。また、生産性損失の範囲としては、本人のみならず家族等による看護や介護(インフォーマルケア)について検討してもよい。ただし、本ガイドラ

インでは、効果との二重計上になる恐れがあること等から病態の改善や生存期間の延長などアウトカムの改善を通じた間接的な生産性損失は含めないこととし、入院期間の短縮など医療技術に直接起因するもののみを含めてもよいこととしている。

## 増分費用効果比

増分費用を増分効果で割ったものを増分費用効果比(ICER)という。以下の式により、治療Bと比較した場合の治療AのICERが算出される。

$$ICER = \frac{IC}{IE} = \frac{C_A - C_B}{E_A - E_B}$$

IC: 増分費用

C<sub>B</sub>: 治療Bの期待費用

IE: 増分効果

E<sub>A</sub>: 治療Aの期待効果

C<sub>A</sub>: 治療Aの期待費用

E<sub>B</sub>: 治療Bの期待効果

効果指標1単位獲得するあたり、いくらかかるかを表す指標であり、値が小さいほど費用対効果はよい。

## 追加的有用性

費用効果分析を実施するにあたっては、増分費用効果比を算出する前に、比較対照技術と比して追加的有用性等を有することを示す必要がある。追加的有用性を示す際の効果指標については、必ずしも費用効果分析における効果指標と同一である必要はないが、臨床的に意味のあるものを選択する必要がある。追加的有用性を有すると判断される場合は、費用効果分析によりICERを算出する。一方で、追加的有用性を有すると判断できない場合は、いわゆるCMAにより費用の比較を実施する。

## 非関連医療費

医療費は、評価対象技術によって直接影響を受ける関連医療費(related medical cost)と生命予後の延長等により間接的に影響されるもの、あるいは当該疾患と関連しないもの等の非関連医療費

(unrelated medical cost)とに分類できる。例えば、高血圧治療によって心血管疾患や脳卒中が減少すると、期待余命が延長して、非関連医療費(例えば認知症や糖尿病、腎透析など)が増大する可能性がある。このような非関連医療費は原則として費用に含めないこととしている。

---

## 費用効果分析

医療技術の経済評価は下記の4パターンに分類されることが多い。(a)アウトカムを同等とにおいて費用のみを検討する「費用最小化分析(Cost-minimization analysis: CMA)」、(b)QALY以外の種々のアウトカム指標(生存年、イベント回避など)を用いる「費用効果分析(Cost-effectiveness analysis: CEA)」、(c)QALYを用いる「費用効用分析(Cost-utility analysis: CUA)」、(d)アウトカムを金銭化して評価する「費用便益分析(Cost-benefit analysis: CBA)」。

しかし、CMA、CEA、CUAは費用とアウトカムを別々に推計するという点では、同種の分析であるとも考えられるので、本ガイドラインではこれらの手法をまとめて費用効果分析と呼んでいる。

---

## 不確実性

費用効果分析を行う上では、様々な分析の不確実性(uncertainty)がともなう。

異質性(heterogeneity)は、広義の不確実性の一種であり、比較対照技術や診療パターン、対象患者等が一意に定まらない状況を指す。これは、次に説明する狭義の不確実性とは異なり、統計学や医療経済学上の技術的な問題ではなく、現実が多様であることに起因する。このような異質性が存在する場合は、複数のシナリオ設定に基づいた感度分析を行うことを推奨している。

狭義の不確実性は、大きく(a)モデルの不確実性と(b)パラメータの不確実性に分けることができる。前者のモデルの不確実性は、さらに(a)-1方法論上の不確実性や(a)-2モデルの構造・仮定等に起因するものがある。

(a)-1方法論上の不確実性は、割引率や生産性損失の推計方法、QOL値の測定方法等が理論的には一意に定められないために生じる。これら为了避免するためには、標準的な共通の手法に従って分析を行うことが重要であるが、割引率など結果に大きな影響を与える場合には、一次元感度分析によってその不確実性の大きさを評価する。

(a)-2モデルの構造・仮定に起因する不確実性は、健康状態や治療プロセスのモデル化法、モデルに組み込むパラメータの選択、観察期間を超えて長期的な予後を予測するための仮定等によって生じる。感度分析等によって評価する。

(b)パラメータの不確実性は、パラメータの推定値が持つ不確実性によって生じる。例えば、ある臨床試験の中で100人中10人にイベントが起こったとしても、真のイベント発生率(母イベント発生率)は $10/100=0.1$ ではないかもしれない。このような統計的推測に起因する不確実性に対処するには、通常感度分析に加えてPSAを行うことも有用である。

---

## マッピング

選好にもとづく尺度による測定値が存在しない場合、患者報告アウトカム(PRO)での測定結果から、費用効果分析で使用するQOL値を算出することが有用な場面もある。このような尺度間のスコア変換をマッピング(mapping)と呼ぶ。他のデータが存在しないときなどにマッピングは次善の手法として許容されるものの、統計学的な妥当性などを十分に検討した上で実施すべきである。

---

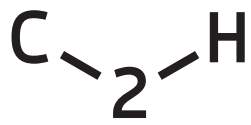
## メタアナリシス

システマティックレビューで得られた結果を統計的手法によって統合し、統合した値やその信頼区間を求める手法である。異質性が小さい場合は、固定効果モデル、異質性が大きい場合は変量効果モデルやベイズモデルを用いるのが一般的である。結果はフォレストプロットを用いて表記する場合が多い。治療間の対比較ではなく、複数の治療の比較を行う場合は、ネットワークメタアナリシスと呼ばれる手法が用いられる(→間接比較)。

---

## 優位・劣位

評価対象技術が比較対照技術と比して費用が安く効果も同等以上である場合、その医療技術は「優位(dominant)」であるという。一方、評価技術が比較対照技術と比較して費用が高いが効果は同等以下である場合、その医療技術は「劣位(dominated)」と呼ばれる。



---

## 割引

費用効果分析においては、将来に発生する(あるいは得られる)費用とアウトカムを現在価値に換算するため、一定の率で割引くことが一般的である。年単位で割引を行ったあとの現在価値に換算された費用  $C_p$  は、 $i$ 年後の費用  $C_i$  と割引率  $d$  を用いて

$$C_p = \frac{C_i}{(1+d)^{i-1}}$$

によって計算することができる。効果についても同様である。

---

## 略語一覧

**CBA:**

Cost-benefit analysis  
費用便益分析

**CEA:**

Cost-effectiveness analysis  
費用効果分析

**CEAC:**

Cost-effectiveness acceptability curve  
費用効果受容曲線

**CMA:**

Cost-minimization analysis  
費用最小化分析

**CUA:**

Cost-utility analysis  
費用効用分析

**DCE:**

Discrete choice experiment  
離散選択実験

**EQ-5D:**

EuroQol 5 dimension

**HRQOL:**

Health-related quality of life  
健康関連QOL

**ICER:**

Incremental cost-effectiveness ratio  
増分費用効果比

**MAIC:**

Matched adjusted indirect comparison,  
マッチング調整された間接比較

**MTC:**

Multiple treatment comparison  
多群間治療比較

**PBM:**

Preference-based measure  
選好に基づく尺度

**PRO:**

Patient-reported outcome  
患者報告アウトカム

**PRISMA:**

Preferred Reporting Items for  
Systematic Reviews and Meta-Analyses

**PSA:**

Probabilistic sensitivity analysis  
確率の感度分析

**QALY:**

Quality-adjusted life year  
質調整生存年

**RCT:**

Randomized controlled trial  
ランダム化比較試験

**RQ:**

Research question  
リサーチクエスション

**SG:**

Standard gamble  
基準的賭け法

**SR:**

Systematic review  
システマティックレビュー

**TTO:**

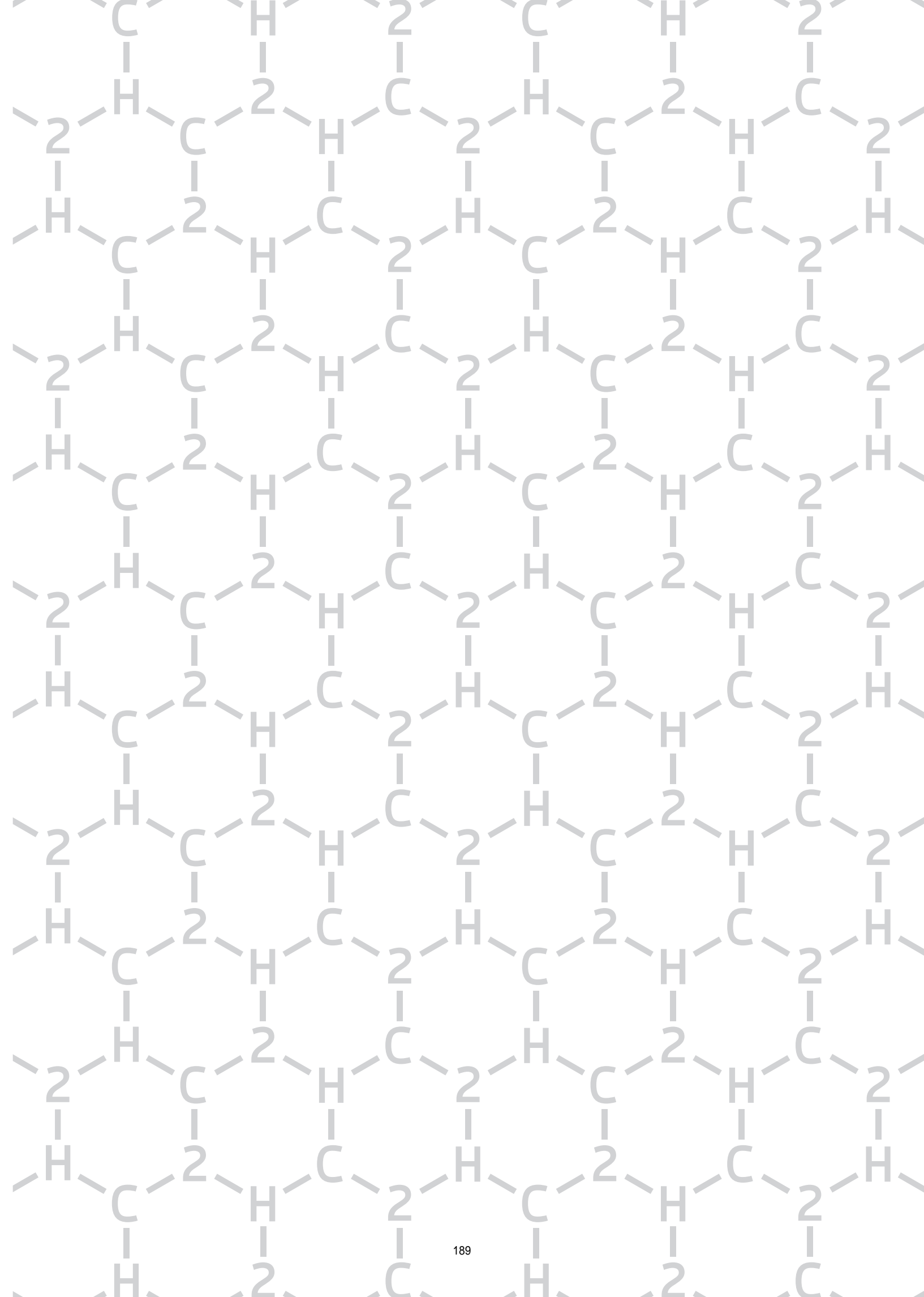
Time trade-off  
時間得失法



政策科学総合研究事業（政策科学推進研究事業）

「医薬品・医療機器等の費用対効果評価における公的分析と公的意思決定方法に関する研究」班

（研究代表者：福田 敬）







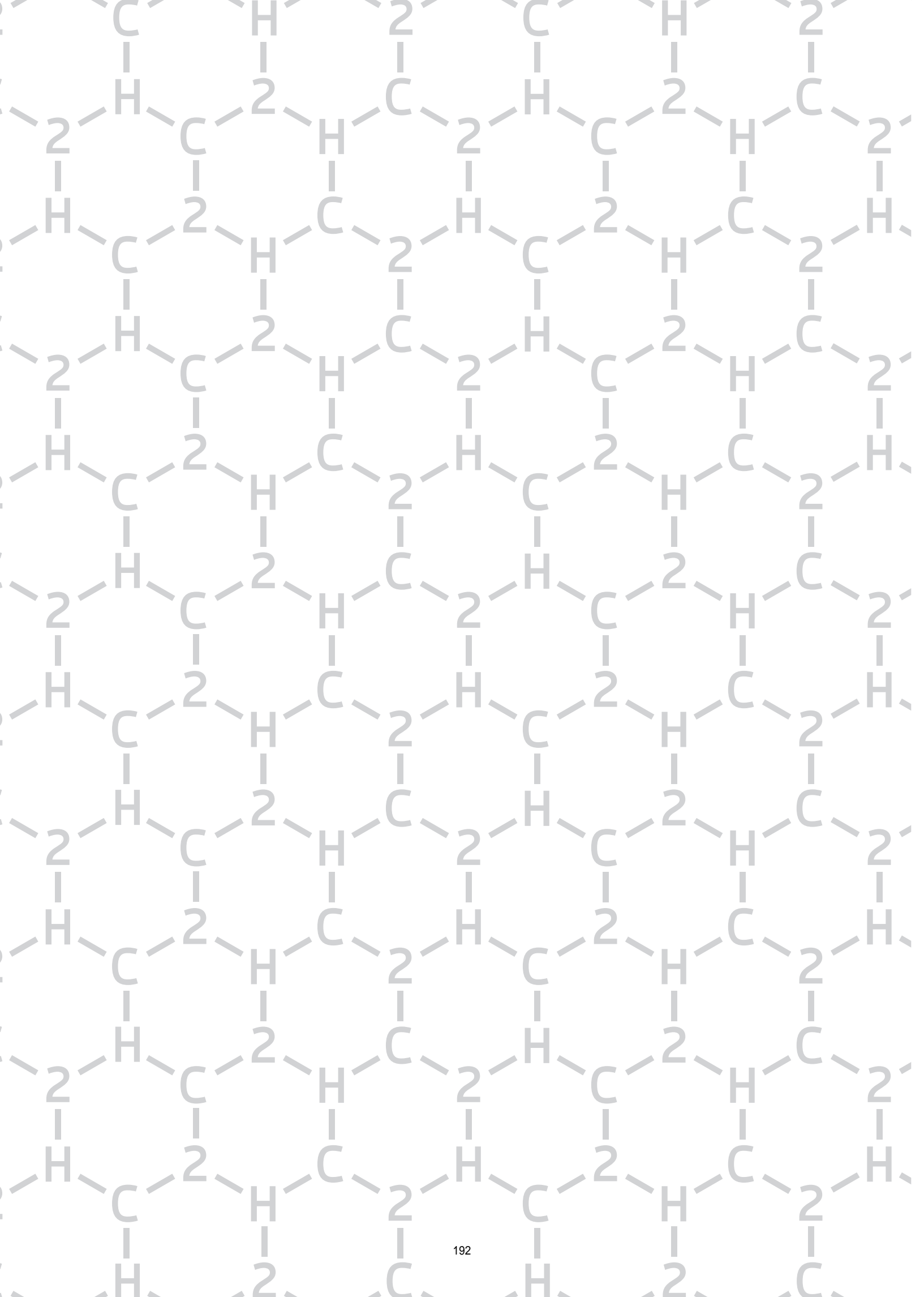
---

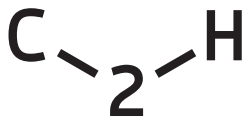
# Guideline for Preparing Cost-Effectiveness Evaluation to the Central Social Insurance Medical Council

—  
Version 3.0 approved by CSIMC on 19th January, 2022

---

The English version is a translation of the original version in Japanese. The Japanese version is preferentially applied in cases of discrepancy between the two versions.



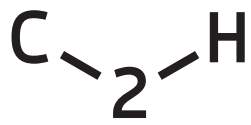


---

## Contents

---

1	Objectives	4
2	Analysis perspective	5
3	Target population	6
4	Comparator	7
5	Additional benefits	8
6	Methods of analysis	11
7	Time horizon	12
8	Choice of outcome measure	13
9	Sources of clinical data (excluding costs)	15
10	Calculation of healthcare costs	16
11	Public long-term care costs and productivity loss	19
12	Discounting	21
13	Modeling	22
14	Uncertainty	23
	Terminology	24
	Abbreviations	29



---

## 1 Objectives

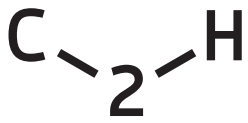
- 1.1 This guideline presents standard methods to perform cost-effectiveness evaluations of medicines and medical devices selected by the Central Social Insurance Medical Council (“selected technologies”).
- 1.2 This guideline is applied to manufacturers’ submissions and academic analysis (review and re-analysis).

---

## 2 Analysis perspective

- 2.1 The perspective of the analysis should be specified. In particular, the analysis should consider the range of costs corresponding to this perspective.
- 2.2 “Public healthcare payer’s perspective” is a standard. It uses costs, comparator, and target populations to reflect the situation of public healthcare insurance in Japan.
  - 2.2.1 Even when an analysis is conducted from a perspective other than the “public healthcare payer’s perspective,” an analysis from the “public healthcare payer’s perspective” should be submitted.
  - 2.2.2 There are some healthcare technologies that are not covered by public healthcare insurance but are publicly funded, such as some prophylactic procedures (e.g., health check-ups, vaccinations). Analyses including these technologies should be submitted from the “public healthcare payer’s perspective.”
- 2.3 If the effect on public long-term care costs is important with regard to the selected technology, it is acceptable to perform an analysis from the “public healthcare and long-term care payer’s perspective.”
- 2.4 If the introduction of a selected technology influences productivity, it is acceptable to perform an analysis that includes productivity loss from the broader perspective.





---

### 3 Target population

- 3.1** Patient populations that meet the indications when the target technology is selected should be considered as the target population of the cost-effectiveness evaluation.
  - 3.1.1** In the case that a new indication is approved between the selection of target technology and the determination of the framework of analysis, it is to be included in the target population.
  - 3.1.2** A new indication is added after the time defined by item 3.1.1 and a new evaluation is performed after the first evaluation is completed if it may influence the results.
- 3.2** An analysis should be conducted for each population if the technology has multiple indications or subpopulations that differ in outcome, application method/dose, and administration or comparator.
  - 3.2.1** However, if item 3.2 is difficult to achieve, it is acceptable to perform analyses of limited population(s) considering factors such as the number of patients or features of the illness. The exemption is determined based on agreement between the manufacturer and the National Institute of Public Health/public academic group in consultation.
- 3.3** The percentage of each patient group should be determined in a manner that reflects the recent clinical context of the target population.

---

## 4 Comparator

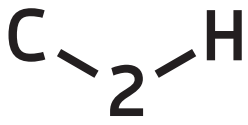
- 4.1** The comparator should be principally selected from technologies that are widely used in clinical practice and are expected to be replaced by the selected technology when it is introduced to treat the target population. Among them, technologies that result in better outcomes should be selected.
- 4.1.1** Non-treatment or watchful waiting can also be used as comparators.
- 4.1.2** Except for the cases described in item “4.1.1,” the comparator should be selected from technologies that can be used by public healthcare insurance.
- 4.1.3** If a single comparator cannot be determined based on item “4.1,” the comparator should be selected by considering the comparators in randomized controlled trials (RCTs), similar technologies for the official pricing, and cost-effectiveness based on agreement after consultation with C2H.
- 4.2** Sufficient explanation of the reasons underlying the selection of the comparator is needed.

---

## 5 Additional benefits

- 5.1** When a cost-effectiveness evaluation is conducted, whether the additional benefit of the selected technology to the comparator is proven should first be evaluated.
- 5.2** Evaluations of the additional benefit should be conducted on the basis of a systematic review (SR) of RCTs compared with a technology selected in section “4.” If appropriate, the results of unpublished clinical studies/trials may also be included in the SR.
- 5.2.1** When an SR is conducted, research questions (RQs) should be clearly presented. For example, a definition of structured RQs according to PICO (P: patient, I: intervention, C: comparator, O: outcome) may be provided.
- 5.2.2** There may be technologies with a similar action mechanism or function category to the selected technologies or comparator(s) determined in section “4” that are expected to show equivalent outcomes. These technologies can be included as an intervention (I) or comparator (C) in the SR if deemed appropriate after consultation with C2H.
- 5.2.3** For outcome (O) in item “5.2.1,” the most appropriate clinical outcomes (e.g., a “true outcome”) should be selected from among clinical effectiveness, safety, and health-related quality of life (HRQOL).
- 5.2.4** Descriptions of the inclusion/exclusion criteria, databases used, search algorithm, and research selection process (inclusion flow diagram) are required in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.
- 5.2.5** It is acceptable to utilize any existing reliable SR. In such cases, the existing review will be used solely or in combination with a new additional study. In this case, it should confirm the consistency of the existing review by considering the RQs and coverage of the most recent literature.

- 5.2.6** If appropriate, pooled results via a meta-analysis should be presented. In such cases, the required reporting factors include the statistical method, assessment of heterogeneity, forest plot, pooled results, and confidence interval.
- 5.2.7** When it is obvious that no RCTs have been performed, the process described in section “5.2” can be skipped upon agreement in consultation.
- 5.2.8** A time point between determining the framework of analysis and manufacturer’s submission can be used as a cut-off date for the literature search in the SR.
- 5.2.9** There may be cases in which the results of new clinical trials are published after the cut-off date defined in item “5.2.8” but are regarded as important information for the cost-effectiveness evaluation (e.g., clinical trials with a large sample size or reliable results that differ from current studies). Inclusion of these trials in the SR should be considered. In that case, additional SR is not required.
- 5.3** When no studies are available based on the result of SR described in section “5.2,” the additional benefit is evaluated by SR of comparative non-RCT (e.g., observational) studies based on section “5.2.” In that case, sufficient explanation regarding the research quality is needed (e.g., study design, differences in background factors between groups, statistical analysis, sample size, and number of institutions).
- 5.4** When more reliable results are obtained, the additional benefit can be evaluated via re-analysis of existing observational study or registry data, if agreed upon in consultation. In that case, sufficient explanation on the research quality is needed (e.g., study design, differences in background factors between groups, methods of statistical analysis, sample size, and number of institutions).
- 5.5** When there are no RCT studies using the same comparator selected in section “4” but there are RCT studies of the selected technology compared to others, the additional benefit is evaluated via an indirect comparison using the SR results.



- 5.6** When only single-arm clinical studies of selected technologies are available, an indirect comparison should be performed based on the SR results of the selected technologies and the comparator.
- 5.7** When an indirect comparison is performed, the following items should be considered.
- 5.7.1** When individual patient data are available, the difference of background factors should be adjusted using an appropriate method such as MAIC (matched adjusted indirect comparison).
  - 5.7.2** When individual patient data are not available, an adjusted indirect comparison using RCT or network meta-analysis should be used.
  - 5.7.3** When neither individual patient data nor results of RCTs are available, a naïve indirect comparison may be acceptable if other methods cannot be used. In such a case, the uncertainty of the results should be carefully considered.
  - 5.7.4** If an indirect comparison is conducted, sufficient explanation on the prerequisites for the indirect comparison (e.g., heterogeneity of illness, severity, and patient background or similarity of the studies) is also needed.
- 5.8** There may be cases in which the results obtained by the methods in items “5.3” to “5.7” have serious problems regarding the quality of the studies. However, it is expected that the selected technology is not inferior to the comparator. In such cases, the analysis described in section “6” can be performed, assuming the outcome of the selected technology is equivalent to that of the comparator.
- 5.9** When there are no clinical data available on the selected technology in humans, the analysis described in section “6” can be performed, assuming the outcome of the selected technology is equivalent to that of the comparator(s), if appropriate. This is based on considering the approval of the Pharmaceuticals and Medical Devices Agency (PMDA).

**5.10** When the SR results obtained by the methods in items "5.2" to "5.7" show that the outcomes of the selected technology are inferior to those of the comparator(s), no cost-effectiveness analysis is performed.

---

## 6 Methods of analysis

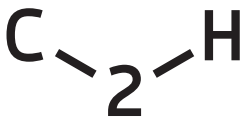
- 6.1** A cost-effectiveness analysis should be used. In this guideline, cost-effectiveness analysis is defined as an analysis that calculates cost and effectiveness separately without converting effectiveness into monetary units.
- 6.2** If the analysis described in section “5” reveals an additional benefit, the incremental cost-effectiveness ratio (ICER) should be calculated based on the expected cost and effectiveness in each treatment group.
- 6.3** In the following cases, only the expected cost and effectiveness in each group need to be presented and the ICER should not be calculated.
- 6.3.1** In cases where the technology has better effectiveness and lower costs relative to the comparator, the technology is considered “dominant” without a calculation of ICER.
- 6.3.2** A cost comparison with the comparator (a so-called “cost minimization analysis” [CMA]) should be performed if the analysis described in section “5” does not demonstrate an additional benefit but the outcome of the selected technology appears to be equivalent to that of the comparator. In such a case, the result is either “cost saving” or “cost increase.”
- 6.4** If the selected technology has multiple indications or sub-populations defined in item “3.2” and/or “3.2.1,” ICER should be calculated for each indication or sub-population.
- 6.5** If a cost-effectiveness analysis for a selected technology published in an academic journal or an evaluation from a health technology assessment (HTA) agency is available, these results should also be presented.

---

## 7 Time horizon

- 7.1 The time horizon should be sufficiently long to evaluate the influence of the technology on cost and effectiveness.
- 7.2 The same time horizon should be applied for both cost and effectiveness.
- 7.3 The reason for setting this time horizon should be specified.





---

## 8 Choice of outcome measure

- 8.1 Quality-adjusted life years (QALY) should be used in principle.
  - 8.1.1 When it is difficult to calculate QALY, and CMA is applied, other outcome measures can be used, if appropriate.
- 8.2 When QALY is calculated, the QOL score should be measured by a preference-based measure (PBM) in principal.
  - 8.2.1 If Japanese QOL scores are newly collected for a cost-effectiveness analysis, the use of PBMs with a value set developed in Japan using TTO (or mapped onto a TTO score) is recommended as the first choice.
  - 8.2.2 If data corresponding to item "8.2" are unavailable, it is acceptable to use mapping of other appropriate HRQOL data. When using a QOL score obtained from mapping, the conversion into a QOL score via an appropriate method should be explained.
- 8.3 When the QOL score is assessed by PBM, the subjects' own QOL responses should be used.
  - 8.3.1 In the case of using PBMs, responses from a proxy (e.g., family member or caregiver) may be used only when the subject cannot respond.
  - 8.3.2 In the case of using PBMs, proxy responses from a healthcare professional should be avoided due to possible discrepancies with subjects' own responses.
  - 8.3.3 If it is difficult to directly collect QOL scores from patients, it is acceptable for general people to evaluate the presented health scenario by standard gamble (SG), time trade-off (TTO), and discrete choice experiment (DCE). However, QOL scores measured via these methods are largely influenced by the presented health scenario. Therefore, the limitations of this method should be considered; for example, the same QOL score should be used for the same health state.

- 8.4** As long as a QOL score that satisfies items "8.2" and "8.3" is available, the use of Japanese results is preferentially recommended.
- 8.4.1** If Japanese research is absent or insufficient but high-quality research is available overseas, it is acceptable to use the data collected overseas.
- 8.5** In the case of the analysis from the public healthcare and long-term care payer's perspective, the QOL scores' influence on the informal caregiver may be considered if actual data exist.

---

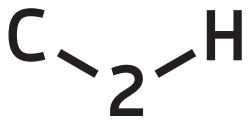
## 9 Sources of clinical data (excluding costs)

- 9.1** ICER calculations should preferentially use effectiveness, safety, and QOL data (inclusion of parameters such as transition probability for model analysis) derived from high-quality research with a high evidence level reflective of practical clinical results in Japan.
- 9.1.1** The selection of effectiveness, safety, and QOL data on the basis of an SR of all the clinical research is recommended. This review may also include unpublished clinical study/trial data if appropriate.
- 9.1.2** Data with a high evidence level should be used preferentially. The use of data deemed appropriate from the viewpoints of research quality, target population, and external validity is recommended (for example, it is possible that the results of an RCT may differ markedly from practical clinical results).
- 9.1.3** Data derived from re-analysis of existing study and/or registry data can be used if appropriate. In this case, detailed information on patient background, statistical methods, etc. must be provided.
- 9.2** Japanese data should be used preferentially if there is evident heterogeneity between Japanese and overseas data.
- 9.3** If the data do not differ statistically significantly between the selected technology and the comparator(s), pooled data of both groups should be applied. Otherwise, when considering factors such as data or rationale that support the difference, effect size (clinical meaning), it should be explained that the difference is interpretable.
- 9.4** Regarding evaluation of medical devices, if there are reliable and quantitative data, analysis reflecting “learning effect” (i.e., improvement of treatment effect via the accumulation of clinicians’ experience) or “product improvement effect” can be submitted in addition to analyses not considering the effects, upon agreement in consultation with C2H.

---

## 10 Calculation of healthcare costs

- 10.1 Only public healthcare costs should be included in the case of analysis from public healthcare payers' perspective.
- 10.2 The healthcare costs of each health state include only related costs that are directly affected by the selected technology, and do not include unrelated costs.
- 10.3 The healthcare costs of each health state should reflect the average resource consumption and standard clinical practices in Japan.
- 10.4 It is recommended that claims databases established in Japan, which reflect actual clinical practice from the viewpoint of item "10.3," should be used to estimate the costs of each health state if appropriate. However, this recommendation does not apply to cases in which it is difficult to define health states using only information from claims data or if there are insufficient data in the database.
  - 10.4.1 A definition of each health state and its rationale is required when claims data are used for cost estimation.
  - 10.4.2 The methods and rationale for the method to estimate costs (including handling outliers and unrelated costs) should be shown.
- 10.5 Micro-costing (by medical fee schedule etc.) based on the definitions of the standard clinical process can be used if it is difficult to estimate the costs of each health state via a claims database or if micro-costing is more appropriate.
  - 10.5.1 In the case of the application of micro-costing, the rationale for costing should be shown from the viewpoint of item "10.2." It may be better to identify relevant items and/or estimate the amount of medical resource consumption in the claims database.



- 10.5.2** When micro-costing is used, the medical resource consumption and unit costs should be reported separately.
- 10.5.3** In principle, for the estimation of resource consumption in item “10.5.2,” the resource consumption of injection products should be calculated by the number of vials rather than by patient dosage.
- 10.6** The estimation should include not only the costs of the selected technology and the comparator(s) but also the costs of adverse events and related future events etc.
- 10.7** An analysis of public healthcare costs should include not only the portion of costs paid by the insurer but also those paid by the government and patients as co-payment (i.e., the total public healthcare expense).
- 10.7.1** Based on the principal in item “2.2.2,” the analysis should include the costs of health check-ups, vaccinations, or similar procedures that are funded publicly but not reimbursed by Japan’s public healthcare insurance.
- 10.8** Unit costs should be derived from the latest medical fee schedule, drug price lists, or similar resources. It is particularly essential to use the latest unit costs for the selected technology or comparator(s).
- 10.8.1** Even if existing cost-of-illness studies or analyses of claims data are used, unit costs at the time of evaluation rather than when the medical resources are consumed should be applied. It is acceptable to make adjustments such as multiplication by the medical payment system revision rate.
- 10.8.2** Such adjustments may be omitted if the influence on results is minimal.

- 10.9** If generics of the comparator(s) are already on the market, an analysis using these costs should ] also be submitted.also submitted.
- 10.10** Even if the costs of selected technology and/or comparator(s) are included in a bundled payment, the estimation should be based on fee-for-service payment.
- 10.11** Future costs should also be estimated based on current medical resource consumption and unit costs.
- 10.12** Calculations of medical resource consumption based on other countries' data will require attention to possible differences in healthcare technology use between Japan and other countries. The unit costs in Japan should be applied in the analysis.

---

## 11 Public long-term care costs and productivity loss

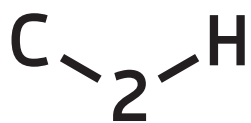
- 11.1** An analysis from the perspective of the public healthcare and long-term care payer may include public long-term care costs in addition to base-case analysis. It is acceptable to include public long-term care costs in additional analyses only if they can be estimated by Japanese data.
- 11.2** When public long-term care costs are included in the analysis, it is recommended that these costs be calculated based on the care level.
- 11.3** The amount utilized under public long-term care insurance should be based on the actual resource consumption. If this consumption is difficult to determine, it is acceptable to use the average amount utilized per beneficiary or similar data.
- 11.4** An analysis including productivity loss can be additionally performed on top of the base-case analysis. However, judgments regarding the appropriateness of including productivity losses should consider the possibility of working in the context of the illness characteristics. It is acceptable to include productivity losses in additional analyses only if they can be estimated by Japanese data.
- 11.5** Decreases in productivity losses may be classified as follows:
- (A) those arising directly from healthcare technology (e.g., treatment-related shortening of hospital stay) and
  - (B) those arising indirectly from outcome improvements (e.g., alleviation of illness, survival period extension).

When productivity loss is included in an analysis, only (A) should be included in the cost calculation.

- 11.6** Productivity losses should be estimated using the human capital method. This method estimates the loss via the expected earned wage in the absence of illness.

- 11.6.1** The unit wage used for estimations of productivity loss should be the average wage across all industries, all ages, and both genders or the average wage for each age group in all industries and both genders derived from the latest “Basic Survey on Wage Structure” (Wage Census) and not discriminate by income.
- 11.6.2** Estimations of productivity loss require an actual investigation into the employment status in the target population (i.e., a measure of the days or hours of work missed). The actual measured number of days or hours should then be multiplied by the average wage across all industries, all ages, and both genders to accurately estimate the productivity loss.
- 11.6.3** If the item described in item “11.6.2” is difficult to perform, productivity loss should be calculated by multiplying the expected number of days (excluding holidays) or hours of work missed multiplied by the average wage across all industries, all ages, and both genders. A 100% employment rate should be assumed for those aged 18 years and older. However, note that this method may overestimate productivity losses.
- 11.7** If other individuals (e.g., family members) experience productivity losses due to the provision of nursing or informal patient care, it is acceptable to count these productivity losses as costs under the same conditions and using the same methods as those used to calculate the patient’s productivity loss.
- 11.7** Time costs that are unrelated to a decrease in work should not be included in the cost estimations.





---

## 12 Discounting

**12.1** Future costs and effectiveness must be discounted and converted into present values.

**12.1.1** Discounting is not needed if the time horizon is one year or less or is otherwise sufficiently short to ignore the influence of discounting.

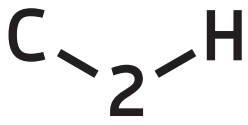
**12.2** Both cost and effectiveness should be discounted at a rate of 2% per year.

**12.3** The discount rate should be subjected to sensitivity analysis and be changed at the same rate of 0-4% per year for both cost and effectiveness.

---

## 13 Modeling

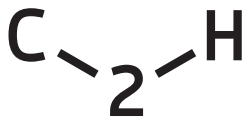
- 13.1** To predict prognosis and future expenses, it is acceptable to conduct a model analysis using a decision analytic model, the Markov model, and/or other models in accordance with the principle described in section "7."
- 13.2** The model analysis should present the validity of the model, including the following.
- (A) Internal validity: this addresses why a model with a given structure has been created, whether the natural history of illness has been sufficiently evaluated, and whether the parameters used are appropriate.
  - (B) External validity: this addresses whether the estimation yielded from the model is appropriate in comparison to other existing clinical data.
- 13.3** The assumption used to create the model should be clearly described.
- 13.4** All parameters and data sources used for model creation should be presented.
- 13.5** The model should be submitted in the form of electronic files. The model must be easily understood by third-party experts and all main parameters (transition probability, QOL score, and healthcare costs) must be alterable.
- 13.5.1** It is ideal that not only total costs but also the breakdown (in the case of micro-costing, the medical resource consumption and unit costs of each item) can be changed. Especially, the unit costs of the selected technology and comparator(s) must be adjusted for each academic analysis group.
- 13.6** Half-cycle correction should be used in the Markov model if the length of the Markov cycle is long and its influence on the results is not negligible.



---

## 14 Uncertainty

- 14.1 If the analysis setting has multiple scenarios and this could affect the results, a scenario analysis should be conducted.
- 14.2 For situations in which the uncertainty is high because of a long time horizon, a shorter-term analysis is necessary, such as an analysis of the period for which clinical study data are available.
- 14.3 If no available studies involve a comparison with the comparator according to section “5,” particularly when a comparison has been made concerning results between single-arm studies, a sensitivity analysis with a sufficiently wide range is required because of the large degree of uncertainty.
- 14.4 Sensitivity analyses are needed for parameters with large variances, those based on assumptions rather than actual data, those with possible heterogeneity between domestic and other countries’ data, and others.
- 14.5 When the variance of the estimator should be considered (parametric uncertainty), the range moving parameter in the sensitivity analysis can refer to the 95% confidence interval of the estimator.
- 14.6 A probabilistic sensitivity analysis (PSA) is also desirable. In such cases, the distribution used for analysis, scatter plots of the cost-effectiveness plane, and cost-effectiveness acceptability curves (CEAC) must be presented.



---

## Terminology

---

### Additional benefit

In a cost-effectiveness analysis, the additional benefit relative to the comparator should be demonstrated before calculating the ICER. The endpoint of effectiveness used to demonstrate the additional benefit does not always need to be equal to the outcome used for the cost-effectiveness analysis but should be clinically significant. If additional benefit is judged to be shown, cost-effectiveness analysis should be performed. On the other hand, if no additional benefit is shown, cost of both treatment should be compared by so called "CMA".

---

### Cost-effectiveness analysis

Economic evaluations of healthcare technologies are often divided into the following patterns: (a) cost-minimization analysis (CMA), in which the outcome is deemed equivalent and only cost is analyzed; (b) cost-effectiveness analysis (CEA), which uses outcome units other than QALY (LY, event avoidance, etc.); (c) cost-utility analysis (CUA), which uses QALY; and (d) cost-benefit analysis (CBA), which involves an evaluation of outcomes after conversion into monetary units.

However, CMA, CEA, and CUA can all be considered analogous in situations where the cost and outcome are estimated in different units. For this reason, these types of analysis are collectively termed "cost-effectiveness analyses" in this guideline.

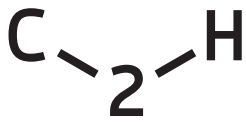
---

### Discounting

In a cost-effectiveness analysis, a discount at a constant rate is usually made to convert future costs and arising (or obtained) outcomes to current values. Costs converted to the current value after applying yearly discounts ( $C_p$ ) can be calculated from the cost at  $i$  years later ( $C_i$ ) and the discount rate ( $d$ ) using the following equation:

$$C_p = \frac{C_i}{(1+d)^{i-1}}$$

The same calculation can be used for effectiveness.



## Dominant/dominated

If a technology is lower in cost and equivalent or higher in effectiveness than the comparator is, the technology is called “dominant.” If the technology is higher in cost but equivalent or lower in effectiveness relative to the comparator, the technology is called “dominated.”

## Evidence level

Diverse classification methods for evidence levels are available. MINDS (Medical Information Network Distribution Service) set forth the following classification:

- I Systematic review/meta-analysis of RCTs
- II From one or more RCTs
- III From a non-randomized controlled study
- IVa Analytical epidemiological study (cohort study)
- IVb Analytical epidemiological study (case-control or cross-sectional studies)
- V Descriptive study (case reports or series)
- VI Views of an expert committee or individual experts that are not based on patient data

However, it has been often noted that the results of experimental studies such as randomized controlled trials (RCTs) can differ from real-world clinical data. Economic evaluations of healthcare technologies should primarily use data with a high level of evidence, although consideration should be given to appropriate clinical data.

## Human capital method

The “human capital method” is used to estimate productivity loss based on the wages originally expected to be earned. However, when viewed from a long-term standpoints, the inability of an individual to work does not always lead to a productivity loss because in a situation with an employment rate less than 100%, as other individuals are likely to work in place of the affected

individual who is unable to work. For this reason, one view suggests that productivity losses should include only friction costs (e.g., based on the period needed to restore the initial production level). Wages should be originally estimated through an investigation of the period for which an individual was actually unable to work because of illness. If this estimation is difficult due to lack of data including housework, it is acceptable to set the employment rate at 100%. From the viewpoint of fairness, the mean wage across all industries, all ages, and both genders should be used as the unit wage, regardless of the actual unit wage for individuals.

## Incremental cost-effectiveness ratio

The incremental cost-effectiveness ratio (ICER) is the incremental cost divided by the incremental effectiveness. ICER of treatment A compared with B is calculated using the following equation:

$$ICER = \frac{IC}{IE} = \frac{C_A - C_B}{E_A - E_B}$$

IC: incremental cost

$E_A$ : expected effectiveness

IE: incremental effectiveness

of treatment A

$C_A$ : expected cost of treatment A

$E_B$ : expected effectiveness

$C_B$ : expected cost of treatment B

of treatment B

ICER is an indicator of the cost to acquire one unit of effectiveness. A lower ICER indicates higher cost-effectiveness.

## Indirect comparison

When clinical studies yield results for “A vs. B” and “A vs. C,” an estimation of the results for “B vs. C” in which no direct comparison is available from the head-to-head results is called an “indirect comparison.” If no head-to-head study involving an appropriate comparator is available, an indirect comparison may occasionally be used.

The following conditions must be satisfied to enable indirect comparison: the results for “A vs. B” must also be applicable to the population “A vs. C” and the results for “A vs. C” must also be applicable to the population “A vs. B.” This is called an “assumption

---

of similarity.” When an indirect comparison is performed, it is necessary to test this assumption and to use appropriate statistical methods (for example, adjusted indirect comparison rather than naïve indirect comparison). This approach also enables analyses based on more advanced methods such as network meta-analyses (or multiple treatment comparisons; MTCs).

---

## Mapping

When preference-based measure-determined QOL scores are unavailable, it is sometimes advantageous to use PRO data to calculate the QOL score used for cost-effectiveness analysis. The conversion of scores between measures is called “mapping.” Mapping is acceptable as a second-best method when no other data are available but should be performed only after sufficient assessment of the statistical validity.

---

## Meta-analysis

Meta-analysis is a method by which the results from a systematic review are integrated statistically to yield integrated values or their confidence intervals. If the heterogeneity is small, a fixed-effect model is usually used. If the heterogeneity is large, random-effect or Bayesian models are usually employed. The results are often depicted as forest plots. If a comparison is made among multiple treatments rather than between two treatments (pairwise comparison), a “network meta-analysis” is used, employing different methods (ref. “Indirect comparison”).

---

## Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) is a technique used to determine the distributions of incremental cost, incremental effectiveness, and ICER by applying model parameters to the distribution. The results of a PSA are usually shown as a scatter plot on the cost-effectiveness plane and as a cost-effectiveness acceptability

---

curve (CEAC), defined as  $f(\gamma) = \Pr(\gamma \cdot IE - IC > 0)$  (IC: incremental cost, IE: incremental effectiveness,  $\gamma$ : willingness to pay).

---

## Productivity loss

Depending on the perspective, a loss resulting from the inability to perform work/housework because of illness (or benefit from early recovery) may be counted as a cost (i.e., productivity loss) but is not included in the base-case analysis. It is acceptable to consider not only the loss experienced directly by the patient but also losses experienced by family members or others arising from the need to provide nursing or informal care. According to this guideline, however, an indirect productivity loss resulting from an improvement in the patient’s health states (e.g., survival period extension) is not included in productivity loss to avoid double counting (i.e., counting a factor as both effectiveness and costs). Only a productivity loss directly attributable to the healthcare technology (e.g., shortened hospital stay) is permitted for inclusion.

---

## Quality-adjusted life year

A quality-adjusted life year (QALY) value is calculated by multiplying the life years (LYs) by the QOL score. A QOL score of 1 indicates full health, whereas 0 indicates death. If an individual has survived for two years under a health states with a QOL = 0.6, the LY is two years and the QALY is  $0.6 \times 2 = 1.2$  (equivalent to 1.2 years survival under full health). If the QOL score changes over time, the QALY is represented by the area under the curve of the QOL score over time, as illustrated in the figure below.



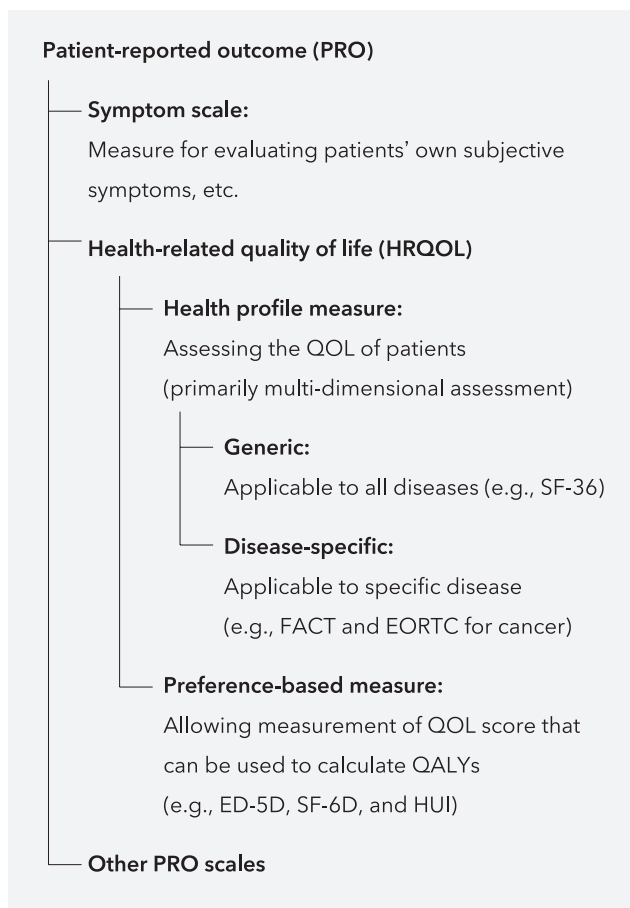
## Quality of life (QOL) score (utility)

The health states (i.e., value obtained from the health states) is scored using a one-dimensional scale ranging from 0 (death) to 1 (full health). Negative scores, reflective of a health states “worse than death,” are also possible.

QOL scoring methods can be categorized as follows: (1) direct methods that evaluate health states under a hypothetical situation (or about himself/herself), including the standard gamble (SG) and time trade-off (TTO) methods, and (2) indirect methods that calculate QOL scores from patients’ responses to QOL questionnaires using a scoring algorithm.

The QOL score used for cost-effectiveness analysis cannot always be calculated from any patient-reported outcome (PRO) or QOL data. Cost-effectiveness analysis can utilize only QOL scores determined using a preference-based measure developed for QALY calculation, as described below.

The EQ-5D (EuroQol 5 dimension) is one currently available measure for which a scoring algorithm has been developed in Japan.



## Sensitivity analysis

When uncertainty is present, its influence on the results can be evaluated by changing the parameter in a “sensitivity analysis.” Sensitivity analyses can be further classified as a one- (only one parameter is changed) and two-dimensional (two parameters are simultaneously changed) sensitivity analyses, as well as PSA (simultaneous uncertainty in multiple parameters; see “Probabilistic sensitivity analysis”).

## Systematic review

A systematic review is a method by which the literature is comprehensively searched regarding a specific topic and the results are evaluated/reported without bias if at all possible. This method was defined by MINDS as follows: “When defined from the aspects of practical actions, systematic review means ‘searching studies on a given clinical question comprehensively, grouping studies of identical quality on each research design and analyzing/integrating them being accompanied by evaluation of biases.’”

A systematic review is often confused with meta-analysis. The results of a systematic review do not always require statistical integration; this type of systematic review is also known as a “qualitative systematic review.” In cases where the integration of results is deemed appropriate, a meta-analysis of the systematic review results is needed.

Regarding the reporting style for a systematic review (meta-analysis), the style presented in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement has been used as a standard and can be used as a reference.

## Uncertainty

Various types of uncertainty accompany cost-effectiveness analyses.

Broadly, heterogeneity is a type of uncertainty that indicates a situation lacking uniformity in terms of the comparator, healthcare

---

patterns, targeted patients, and other factors. This differs from the uncertainty in the narrow sense, as explained below. This is not a technical problem related to statistics or health economics but rather arises from real-world variety. If such heterogeneity is present, a sensitivity analysis based on multiple scenarios is recommended.

Uncertainty in the narrow sense can be divided into (a) model and (b) parameter uncertainties. Model uncertainty can result from (a)-1 methodological uncertainty and (a)-2 model structure/assumptions.

Methodological uncertainty, mentioned in (a)-1, arises from the theoretical impossibility of setting uniform methods for the estimation of the discount rate and productivity loss, measuring the QOL score, and other parameters. To avoid this type of uncertainty, it is important to conduct an analysis in accordance with common and standard procedures. If results such as the discount rate are markedly affected, uncertainty should be evaluated through one-way sensitivity analysis.

Uncertainty arising from the model structure/assumption, as mentioned in (a)-2, is caused by the method used to model the health states and treatment processes, selection of parameters for incorporation into the model, assumptions regarding predictions of long-term prognosis beyond the observation period, and other factors. This uncertainty should be evaluated in a sensitivity analysis.

Parameter uncertainty, as mentioned in (b), arises from uncertainty inherent in the parameter estimation. For example, if 10 of 100 subjects develop events during a clinical study, the true incidence rate might not be  $10/100 = 0.1$  in the whole population. To deal with this type of uncertainty, which is attributable to statistical inference, it is useful to conduct a PSA in addition to a deterministic sensitivity analysis.

---

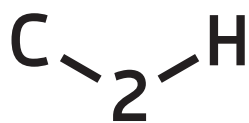
These unrelated costs are not included in the cost.

---

## Unrelated medical costs

Medical costs can be divided into related (i.e., those directly affected by the selected technology) and unrelated (i.e., those affected indirectly through survival extension or those not related to the illness). For example, a hypertension treatment that reduces the incidence of cardiovascular disease and stroke will extend life expectancy, possibly leading to an increase in unrelated medical costs (e.g., costs related to dementia, diabetes, and hemodialysis).



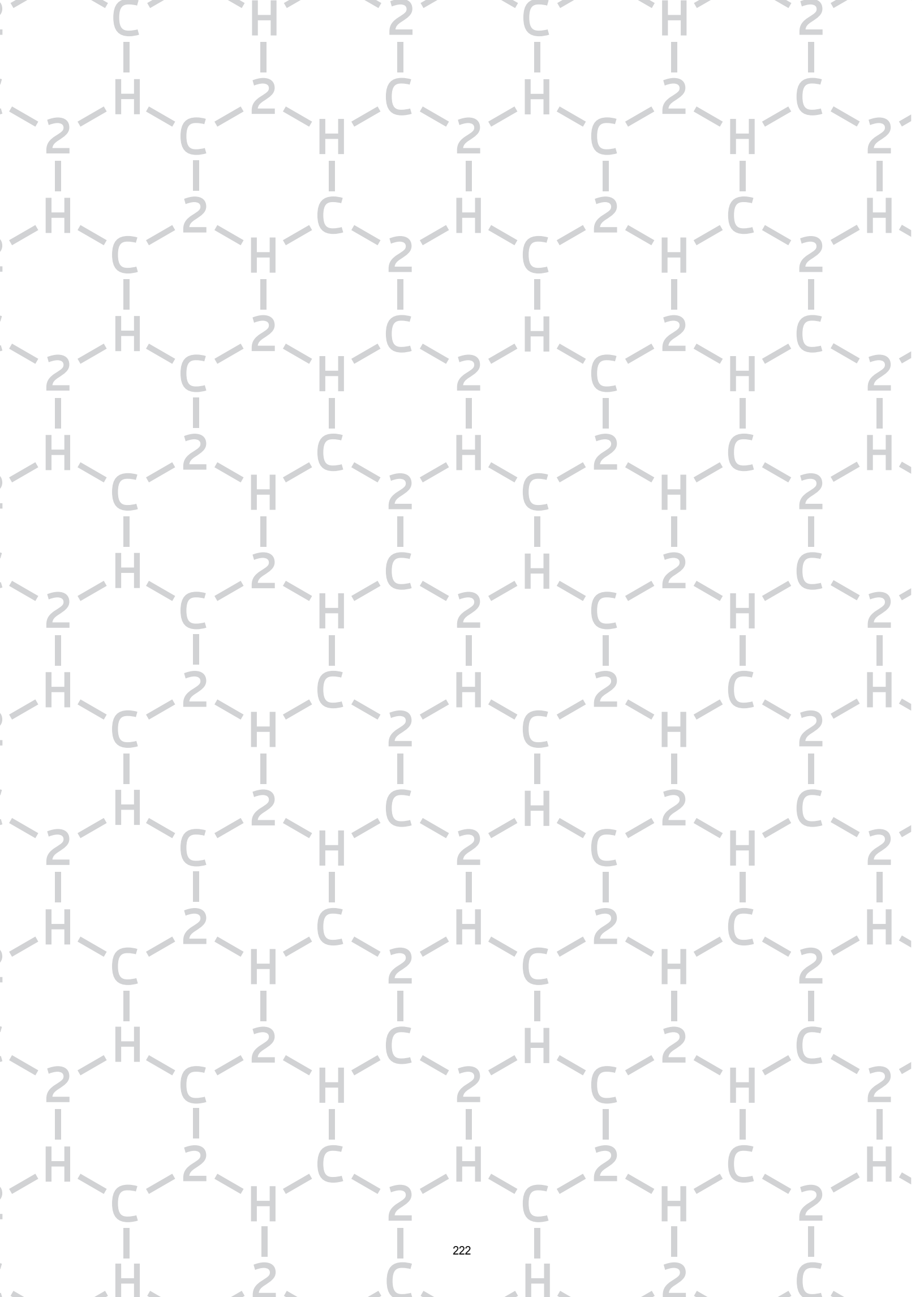


---

## Abbreviations

<b>CBA:</b> Cost-benefit analysis	<b>MTC:</b> Multiple treatment comparison
<b>CEA:</b> Cost-effectiveness analysis	<b>PBM:</b> Preference-based measure
<b>CEAC:</b> Cost-effectiveness acceptability curve	<b>PRO:</b> Patient-reported outcome
<b>CMA:</b> Cost-minimization analysis	<b>PRISMA:</b> Preferred Reporting Items for Systematic Reviews and Meta-Analyses
<b>CSIMC:</b> Central Social Insurance Medical Council (Chuikyo)	<b>PSA:</b> Probabilistic sensitivity analysis
<b>CUA:</b> Cost-utility analysis	<b>QALY:</b> Quality-adjusted life year
<b>DCE:</b> Discrete choice experiment	<b>RCT:</b> Randomized controlled trial
<b>EQ-5D:</b> EuroQol 5 dimension	<b>RQ:</b> Research question
<b>HRQOL:</b> Health-related quality of life	<b>SG:</b> Standard gamble
<b>ICER:</b> Incremental cost-effectiveness ratio	<b>SR:</b> Systematic review
<b>MAIC:</b> Matched adjusted indirect comparison	<b>TTO:</b> Time trade-off





**QOL/PRO 研究会 研究セミナー**  
**「選好に基づく尺度の世界：健康の価値を測定するとは」**

2022年3月12日(土) 14:00-17:00

共催：厚労科研「医薬品・医療機器等の費用対効果評価における公的分析と公的意思決定方法に関する研究」

・ハイブリット形式(参加申し込みは以下からお願いします)

オンライン <https://online-pbm.peatix.com/>

会場 <https://on-site-pbm.peatix.com/>

AP 品川 ルーム A (〒108-0075 東京都港区港南 1-6-31 品川東急ビル 8F 「品川駅」港南口より徒歩 6 分)

1. 選好に基づく尺度とは? 30分 池田 俊也 (国際医療福祉大学)
  
2. 各尺度の特徴など (各 15 分)
  - 2.1 EQ-5D-5L/Y 白岩 健 (国立保健医療科学院)
  - 2.2 HUI 能登 真一 (新潟医療福祉大学)
  - 2.3 EORTC/FACT 村田 達教 (クレコンメディカルアセスメント)
  - 2.4 ASCOT 白岩 健 (国立保健医療科学院)
  - 2.5 15D Michael Iopresti (インテージヘルスケア)
  
3. PBM を用いた日本の研究の現状 30分 能登 真一 (新潟医療福祉大学)
  
4. 直接法を用いた取り組みについて 30分 村田 達教 (クレコンメディカルアセスメント)

Q & A 15 分