

NICE TSD の紹介・QOL 値

－QOL 値の測定と費用効果分析への QOL 値の利用－

日本製薬工業協会

データサイエンス部会

2019 年度継続タスクフォース 5

「医療経済評価に関連する統計的手法の整理と紹介」

Ver 1.0

2021 年 10 月

本報告書の中で **Technical Support Document (Decision Support Unit, National Institute for Health and Care Excellence)** を引用／参考にした記載は、本タスクフォースの解釈であり、NICE には一切の責任はない。また TSD のすべてを紹介したわけではなく、一部を抜粋し紹介している。それらの紹介にあたり、タスクフォースによる解釈や追加の説明を加えている。これらは本タスクが独自にまとめたものであり、NICE から見解を得たものではない。

目次-

略語一覧.....	4
1 はじめに	5
2 QOL 値とは.....	7
2.1 QOL 値の使用：NICE 基本分析	8
2.2 効果の測定と価値づけ	9
2.3 NICE が EQ-5D を推奨する理由	16
2.4 日本では分析をする際はどうすればよいか？	17
2.5 EQ-5D の 3L と 5L.....	17
3 NICE Utilities TSD series	20
4 TSD8: NICE 申請のための各健康状態の QOL 値測定・価値づけの紹介	21
4.1 NICE 基本分析の概要	21
4.2 EQ-5D の適切性の評価	22
5 TSD 9：文献からの QOL 値の特定，レビュー，統合	25
5.1 EXECUTIVE SUMMARY の要約	25
5.1.1 背景	25
5.1.2 HSUV レビューの範囲	25
5.1.3 エビデンスの特定と選択	26
5.1.4 質及び適切性評価	26
5.1.5 データの抽出	26
5.1.6 データ提示	26
5.1.7 データの統合	27
5.1.8 HSUV レビュー実施における推奨	27
5.1.9 結論	28
5.2 HSUV レビュー：ケーススタディ	28
5.2.1 レビューの範囲	29
5.2.2 文献検索	30
5.2.3 骨粗鬆症での HSUV レビュー時の検索	33
5.2.4 骨粗鬆症ケーススタディでの選択/除外基準	38
5.2.5 レビュー結果の記述	39
5.2.6 個々の HSUV 研究の質と適切性評価	40
5.2.7 完全なデータ抽出	41
5.2.8 データの選択と統合	43
5.2.9 結果の提示	43

6	TSD10 : QOL 値の推定のためのマッピング法の利用	46
7	TSD11 : 各健康状態の QOL 値の EQ-5D 以外の方法での算出.....	49
8	TSD12 : QOL 値分析にあたっての留意点	53
8.1	ベースライン QOL 値.....	53
8.2	QOL 値の調整/統合	54
8.3	不確実性の考慮	56
9	NICE TA における QOL 値利用状況	57
9.1	背景	57
9.2	目的	57
9.3	方法	57
9.4	結果	57
9.5	考察	60
9.6	結論	61
10	まとめ	62
11	付録	63
12	参考文献	65

略語一覧

略語	正式名称（英語）	和訳
ASCOT	The new Adult Social Care Outcomes toolkit	—
BWS	Best-Worst Scaling	—
CMA	Cost-Miminisation Analysis	費用最小化分析
CUA	Cost-Utility Analysis	費用効用分析
EQ-5D	EuroQOL 5 Dimension	—
ERGs	Evidence Review Groups	—
HRQOL	Health-Related Quality Of Life	健康関連 QOL
HSUV	Health State Utility Values	各健康状態の QOL 値
HTA	Health Technology Assessment	医療技術評価
HUI	Health Utilities Index	—
MTA	Multiple Technology Appraisal	複数技術評価
NICE	The National Institute for Health and Care Excellence	英国立医療技術評価機構
NHS	National Health Service	英国民医療サービス
PBM	Preference-Based Measure	選好にもとづく尺度
PRO	Paient Reported Outcome	患者報告アウトカム
PSS	Personal Social Services	—
QALY	Quality-Adjusted Life Year	質調整生存年
QOL	Quality Of Life	生活の質
SF-6D	Short form 6 Dimension	—
STA	Single Technology Appraisal	単一技術評価
SG	Standard Gamble	基準的賭け法
TAs	Technology Appraisal guidance	技術評価ガイダンス
TSD	Technical Support Documnent	技術的文書
TTO	Time-Trade Off	時間得失法
VAS	Visual Analogue Scale	視覚的アナログ尺度
中医協	—	中央社会保険医療協議会

1 はじめに

2019年4月より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及び医療機器については、費用対効果評価の提出が求められることとなった。費用対効果評価を実施するにあたり、分析方法は「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）」（中医協ガイドライン）で示されている。この中で、分析に用いる効果指標は、原則、質調整生存年（Quality-Adjusted Life Year : QALY）を用いることとなっている。QALYとは、生存と健康関連の生活の質（Quality Of Life : QOL）の転帰を統合したものである。QALYを算出する際に使用される健康関連 QOL（Health-Related QOL: HRQOL）は、選好に基づく尺度で測定されたものに限定され、0が死亡、1が完全な健康を基準として一次元で数値化された値を効用値と呼ぶ。ただし、中医協においては費用対効果評価の制度化の議論の中で「QOL 値」という用語で統一されてきたため、以下、そのように示す。本資料では、これまで臨床試験等では馴染みの薄かった QOL 値とは何か、またそれはどのように取り扱われるものか、その技術的課題などを、費用対効果評価を早くから導入している英国 NICE が発行する評価ガイドライン「Guide to the methods of technology appraisal 2013」、および技術的な手順書（Technical Support Document, TSD）を参考にしつつ掘り下げていく。

2章では QOL 値について、NICE 基本分析及び中医協ガイドラインでの記載内容の概要を紹介する。3章から8章では”NICE Utilities TSD series”として、経済評価における健康利益の測定や価値づけに関する TSD5 報（TSD 8 から TSD 12）の概要を紹介する。5報の TSD はそれぞれ以下の内容について述べられている。TSD 8 は NICE 申請のための健康測定・価値の紹介、TSD 9 は文献からの各健康状態の QOL 値（Health State Utility Value: HSUV）の特定、レビュー、統合、TSD 10 は HSUV を推定するためのマッピング手法の利用、TSD 11 は HSUV を生み出すための EQ-5D の代替手法、TSD 12 は分析モデルでの HSUV の利用、についてである。NICE 技術評価（Technology Appraisal, TA）ガイダンスにおける QOL 値の利用状況を調査した内容を9章で紹介する。

本資料は、費用対効果評価について、ある程度予備知識のある企業のデータサイエンス担当者を対象に、費用効果分析の中でも QOL 値にフォーカスし、国内外でどのような論点・課題があるか理解することを目標として作成した。費用効果分析になじみのないデータサイエンス担当者であれば、例えば、本タスクフォースの過去の資料「医薬品の価値の科学的な評価—データサイエンス担当者のための費用対効果評価の現状と手法の解説—」は、費用効果分析の基本的な手順について説明しているため、一読をお勧めする（以下より入手可能 http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/2014ds_tf3.html）。

QOL 値は、費用効果分析を行う際の入力データの1つに過ぎないが、本資料からわかるように NICE においても企業と外部の評価グループ間で、その取扱い、引用文献の選択、算出方法などについて意見がわかれ、論点となりやすいテーマの1つである。自社での費

用効果分析に備えようと検討される担当者には是非，知識として有効活用され，QOL 値算出の際の足がかりになることを期待している．

2 QOL 値とは

医薬品を含む医療技術の有効性・安全性を評価する上で、検査などに基づく客観的な指標が多用されているが、近年、患者報告アウトカム（PRO）と呼ばれる患者自身による主観的な評価が着目されはじめ、医療技術の評価に用いられるようになってきた。

PRO とは、患者自身が疾患やその治療に対し、患者の認識に基づき直接評価した結果の総称であり、医師や第三者の解釈を必要としない(1), (2)。PRO は、自覚症状、健康関連 QOL（Health-related quality of life）、健康状態、治療に対する満足度、治療の順守（アドヒアランス）などの一次元的および多次元の尺度を包括的に網羅している(2)。中でも健康関連 QOL は、PRO の中でも複数の次元を有することが特徴であり、疾患とその治療が日常生活、身体的、心理的、社会的な機能と幸福に与える影響についての患者の主観的な認識と定義することができる(2)。これは、1948 年に WHO が発表した「健康」の定義である「健康とは、身体的、精神的、社会的に完全に良好な状態であり、単に病気がないことではない」を、共通の基盤としている。このように、多次元性という概念は、健康関連 QOL の定義の重要な要素であることから、患者自身が直接評価しても、疾患の中核的な症状（痛み、片頭痛、膿疱など）のような単一の領域（ドメイン）を評価したものと、複数領域を評価する健康関連 QOL とは明確に区別されるべきであるとしている(2)。

健康関連 QOL の尺度は、「選好にもとづく尺度（Preference-based measure: PBM）」と「プロフィール型尺度」に大きく分類される。PBM は、QOL を一次元で表し、QOL 値として費用効果分析で用いられる。一方、プロフィール型尺度は QOL に含まれる様々なドメインを 1 つにはまとめず、多次元（multi-dimension）のまま表現しようとするものである。プロフィール型尺度は、更に、全般的尺度（generic 尺度、包括的尺度ともいう）、疾患特異的（disease specific）尺度、症状インデックス尺度に分類することができる(3)。

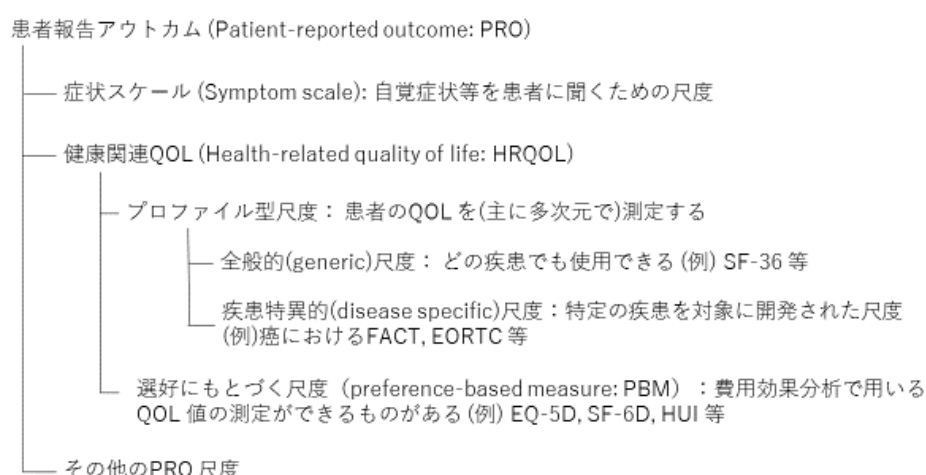


Figure 2-1. 選好に基づく尺度（PBM）及びプロフィール型尺度の分類 [43]

「QOL 値」は、死亡を 0、完全な健康を 1 とした間隔尺度であり、効用理論（von Neumann & Morgenstern の効用）に基づき一次元的にスコア化された指標である。死ぬより悪い健康状態として負のスコアも取りうる。英語では utility (value), health state utility などと表現される。

TF 注：もともとゲーム理論（経済や社会における複数主体がかかわる意思決定の問題や行動の相互依存的状况を数学的モデルを用いて研究する理論）のための基礎概念として提示された不確実性下の意思決定理論で Jeremy Bentham の功利主義（最大多数の最大幸福）を理論的基盤とする。行為の結果が不確実な状況下では、経済主体は効用を確率で加重平均した期待効用を最大化（ナッシュ均衡）するように選択する。John von Neumann は、原爆開発にも関与し、Alan Turing とともに現在のコンピュータの基礎を築いた一人でもある。

2.1 QOL 値の使用：NICE 基本分析

NICE の Guide to the methods of technology appraisal 2013(4)（以下、Methods guide）は、NICE が行う TA の原則と評価の方法について概要を示している。Methods guide はこれまで 2004 年、2008 年にも発行されてきたが、現在使用されているのは 2013 年版である。NICE の TA は、分析の枠組みの設定（スコーピング）、費用対効果の評価（アセスメント）、及び総合的評価（アプレイザル）の 3 つのフェーズに分かれており、中でもアセスメントにおいては、医療技術の臨床効果及び費用対効果評価について比較対照技術と比較し評価することを目的としている。アセスメントは、文献などのエビデンスのシステマティックレビューと、経済評価の 2 つの要素があり、得られたエビデンスの質、結果とそれが意味するところの分析で構成されている。Methods guide で示されているアセスメントの際の基本分析は、Table 2-1 のようにその概要が示されている（NICE methods guide p32-33 Table 5.1 Summary of Reference case を転載翻訳）。

Table 2-1. 基本分析の概要（NICE methods guide Table 5.1）

HTA の要素	基本分析	解説
意思決定の問題定義	スコープは NICE が作成する	5.1.4 ~ 5.1.6
比較対照	スコープの記載どおりとする	2.2.4 ~ 2.2.6, 5.1.6, 5.1.14
アウトカムの観点	患者あるいは、関連する場合、介護者に対するす	5.1.7,

	すべての直接的な健康に関する効果	5.1.8
費用の観点	NHS（National Health Service；英国民医療サービス）および PSS（Personal Social Services；対人社会サービス※福祉サービス）	5.1.9 5.1.10
経済評価の種類	完全な増分分析による費用効用分析（CUA）	5.1.11～ 5.1.14
分析期間	医療技術がもたらす費用とアウトカムの重要な差を反映するのに必要な期間	5.1.11～ 5.1.14
効果におけるエビデンスの合成	システマティックレビューに基づく	5.2
効果の測定と価値づけ	効果は QALY で表現すべき。尺度は EQ-5D が推奨される	5.3.1
健康関連 QOL の測定のためのデータソース	患者および／または介護者から直接報告されたものとする	5.3.3
健康関連 QOL の変化を価値づけるために好まれるデータソース	UK 国民を代表するサンプル	5.3.4
平等性の配慮	追加 QALY は、健康利益を受ける個人の他の特性にかかわらず同じ重みを有する	5.4.1
医療資源消費量と費用のエビデンス	NHS および PSS	5.5.1
割引	費用と効果で同じ年率（現在は 3.5%）	5.6.1

基本分析における QOL 値の使用については、NICE Methods guide の第 5 章基本分析の中の「5.3 効果の測定と価値づけ」で主に記載されている（2.2 に詳細を記載）。ここでは、NICE Methods guide の第 5 章基本分析の中の「5.3 効果の測定と価値づけ- Measuring and valuing health effects」に記載されている QOL 値の取り扱いについて整理する。特に、NICE Methods guide と中医協ガイドラインとの共通点・相違点を示す（

Table ）。なお、相違点がある場合であっても、日本の医療環境下における分析を行う場合は中医協ガイドラインに準じる必要がある。しかし、中医協ガイドラインも事例の蓄積に伴い将来的に変更もありうるため、参考として NICE Methods guide の該当部分を比較する取り組みを行った。

2.2 効果の測定と価値づけ

- 費用効果分析における効果指標の選択について

NICE Methods guide では、基本分析において、効果は QALY で表すこととしている。改訂された中医協ガイドライン（2019）も同様の方針で、効果指標は QALY を用いることを原則とし（中医協ガイドライン 8.1）、QALY を算出することが困難であり、かつ費用最小化分析（CMA）を実施する場合は、両者による合意のもとで QALY 以外の評価尺度を使用することもできる（Table 2-3 中医協ガイドライン 8.1.1）としている。

- 健康関連 QOL 測定の対象者

健康関連 QOL を測定する対象者は、両ガイドラインとも共通して対象者本人から測定することを原則としている（Table 2-3 中医協ガイドライン 8.3）。本人が回答できない場合は、代理回答として家族や介護者等による代理回答を用いてもよいとしている（中医協 8.3.1）が、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もあるのでその点について説明すること（中医協 8.3.2）、介護者による代理回答の方が医療関係者の回答よりも優先される（NICE 5.3.3）とし、両ガイドラインとも医療従事者よりは本人の介護者による代理回答を医療関係者の回答よりも優先して位置づけている。加えて、日本のガイドラインにおいては、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直接法（うち、時間得失（Time trade-off: TTO）法を推奨）を用いて測定してもよいとしている（中医協 8.3.3、付録参照）。

- 健康関連 QOL の測定尺度と価値づけ

健康関連 QOL の測定尺度について NICE は、直接法（Choice-based method）を用いて UK 国民（一般人）の選好に基づく尺度を用いるべきだとし、EQ-5D を推奨している（NICE5.3.1, 5.3.4）。EQ-5D は、標準化され検証（バリデート）された尺度であり、既に多くの患者で実証されていること、直接法（この場合 TTO 法）を用い、UK 国民の大規模調査で抽出された価値づけが利用可能であることが、使用推奨の理由である（NICE5.3.6）。更に、健康関連 QOL の測定は、使用する測定方法や測定尺度によって異なる QOL 値を生み出す。そのため、異なった方法・測定尺度を用いて得られた結果を比較することができない。アプレイザル間の一貫性を考慮すると、単一の尺度である EQ-5D による測定が推奨される（NICE5.3.5, 5.3.6）。

中医協ガイドラインは、QALY を算出する際の QOL 値は、一般人の価値を反映したもの（選好に基づく尺度（preference-based measure: PBM）で測定したもの、あるいは基準的賭け（Standard gamble: SG）法、時間得失（Time trade-off: TTO）法などの直接法で測定したもの）を用いるとしている（中医協 8.2、付録参照）。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、TTO 法を用いて国内データに基づき換算表が開発された（或いは TTO 法で測定されたスコアにマッピング等された）PBM を第一選択として推奨する（中医協 8.2.1、付録参照）ことから、新たに日本国内で QOL 値を収集する際に使用する尺度は、現時点において EQ-5D に限定される。ただし、介護を受けている方を対象とする場合は ASCOT もこの条件に該当する（Table 2-2. PBM の質問票

の日本語版の有無と価値づけ方法 Table 2-2.)。

Table 2-2. PBM の質問票の日本語版の有無と価値づけ方法

PBM の種類	EQ-5D-3L	EQ-5D-5L	HUI-3	SF-6D	ASCOT
質問票の日本語版の有無	あり	あり	あり	あり	あり
質問数	5 問	5 問	15 問	(SF-36 或いは SF-12 に準じる)	9 問 (介護を受けている方を対象)
質問項目	「移動の程度」 「身の回りの管理」 「ふだんの活動」 「痛み/不快感」 「不安/ふさぎ込み」	「移動の程度」 「身の回りの管理」 「ふだんの活動」 「痛み/不快感」 「不安/ふさぎ込み」	「視力」「聴力」 「会話」「歩行」 「器用さ」「感情」 「認知」「痛み」 の 8 領域	身体機能, 役割, 社会活動, 痛み, メンタルヘルス, 活動性の 6 領域	日常生活のセルフコントロール, 身だしなみ, 飲食, 安心・安全, 望む人との付き合い, 仕事, 住居の快適性, ケアのされ方と尊厳 (2 項目)
回答の水準	3	5	5-6	4-6	4
価値づけ方法	TTO	TTO	VAS, SG	SG	BWS, TTO
日本人集団における換算表	あり (5)	あり (6)	あり (7)	あり (8)	あり (9)
日本人集団における標準値	あり (10)	あり (10)	調査中	あり (10)	なし

国立保健医療科学院保健医療経済評価研究センター (C2H) ウェブサイト

(<https://c2h.niph.go.jp/tools/pbm/index.html>) を参考に作成

TTO: time-trade off, SG: standard gamble, VAS: visual analogue scale, BWS: best-worst scaling, HUI: Health Utilities Index

● 治験で QOL 値を測定していない場合

治験で QOL 値が得られていない場合, NICE Methods guide は, 文献から EQ-5D により測定された QOL 値の引用も可能であるとしている. その場合, データの特定方法は系統的吗かつ透明性を有してなければならず, 特定の QOL 値を選択した理由を明示する必要がある. EQ-5D により測定された QOL 値が複数あった場合, どれか一つを選び, 選ばれな

かった値は感度分析によってその影響を示す必要がある（NICE 5.3.8）。

中医協ガイドラインでは、一般人の価値を反映した QALY に換算可能な測定尺度（PBM, SG, TTO）を用いて対象者本人が回答していれば、国内での調査結果を優先（中医協 8.4）しつつも、海外で測定された結果を使用してもよいとしている（中医協 8.4.1）。QOL 値を文献から引用する際の文献の選び方、QOL 値を報告する文献が複数あった場合の対応方法などについては記載がない。

- マッピング

NICE Methods guide は、EQ-5D データが入手できない場合、他の健康関連 QOL データを EQ-5D にマッピングすることを許容する（NICE 5.3.9）。

中医協ガイドラインは「8.2」に該当するデータが存在しない場合、すなわち、一般人の価値を反映した QALY に換算可能な測定尺度（PBM, SG, TTO）を用いた本人による回答結果（QOL 値データ）がない場合は、その他の健康関連 QOL からマッピングしたものを使用してもよいとしている。マッピングされる測定尺度の種類は限定していない。

- NICE のガイダンスにはあるが中医協ガイドラインに記載がない項目

NICE Methods guide は、年齢や合併症の調整、EQ-5D の使用が不適切である場合、子供の健康関連 QOL の測定、EQ-5D-5L の位置づけ・取り扱い方法についても記載されている。一方、中医協ガイドラインではこれらに関し記載がない。

Table 2-3. NICE Methods guide 5.3 効果の測定と価値づけの各項目と、対応する中医協ガイドライン 2019（下線はガイドライン 2015 から追加・変更された点）

NICE Methods guide		中医協ガイドライン	
要訳	章	原文	章
● 費用効果分析における効果指標の選択			
効果指標には QALY を用いる.	5.3.1	効果指標は質調整生存年(Quality-adjusted life year: QALY)を用いることを <u>原則とする.</u>	8.1
		<u>QALY を算出することが困難であり, かつ CMA を実施する場合は, 協議における両者の合意のもとで, QALY 以外の評価尺度を使用することもできる.</u>	8.1.1
● 健康関連 QOL 測定の対象者			
健康関連 QOL の測定は, 患者本人から直接測定されたものであること.	5.3.1	選好に基づく尺度 (preference-based measure: 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

NICE Methods guide		中医協ガイドライン	
要訳	章	原文	章
健康関連 QOL は患者本人から直接測定されたものを使用することとなっているが、それができない場合は、医療専門家による代理の回答よりは、介護者による代理の回答を優先する。	5.3.3	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。	8.3.1
		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、その点について説明する。	8.3.2
		<u>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直接法を用いて測定してもよい。ただしその提示するシナリオの妥当性等については臨床家のチェックを受けることが望ましい。また、直接法を用いる場合、TTO 法により測定することを推奨する。</u>	8.3.3
● 健康関連 QOL の測定尺度と価値づけ			
健康関連 QOL の測定は、患者本人から直接測定し、choice-based method を用いた一般人の選好に基づいたものであること。成人の健康関連 QOL を測定する尺度として EQ-5D が推奨される。	5.3.1	QALY を算出する際の QOL 値は、一般の人々の価値を反映したもの（選好に基づく尺度（preference-based measure: PBM）で測定したもの、あるいは基準的賭け（Standard gamble: SG）法、時間得失（Time trade-off: TTO）法などの直接法で測定したもの）を用いる。 <u>ただし、TTO と SG での測定値には系統的な差がある可能性について留意すること。</u>	8.2

NICE Methods guide		中医協ガイドライン	
要訳	章	原文	章
健康関連 QOL の価値づけ（valuation）は、choice-based method を用い、英国人の代表サンプルから得られた一般人の選好の価値づけに基づくこと。	5.3.4	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、 <u>TTO 法を用いて国内データに基づき換算表が開発された（あるいは TTO 法で測定されたスコアにマッピング等された）PBM を第一選択として推奨する。</u>	8.2.1
健康関連 QOL の測定は、使用する測定方法によって異なる QOL 値を生み出す。それゆえ異なった方法・測定尺度を用いて得られた結果を比較することができない。アプレイザル間の評価の一貫性を考慮すると、成人の健康関連 QOL の測定には、単一の尺度、EQ-5D による測定が推奨される。	5.3.5		
EQ-5D は、標準化され検証（バリデート）された、どの疾患でも利用できる全般的尺度であり、多くの患者集団で検証され広く用いられている。EQ-5D においては、大規模調査で抽出された英国国民の選好の直接法に基づく価値づけ（TTO 法）が利用可能である。	5.3.6		
● 治験で QOL データを測定していない場合			
EQ-5D データを臨床試験で得られていない場合は、文献から EQ-5D により測定された QOL 値の引用も可能である。その場合、データの特定方法は系統的かつ透明性を有していなければならない。特定の EQ-5D データを選択した理由を明確に説明する必要がある。EQ-5D データが複数あった場合、どれか一つを選び、選ばれなかった値は感度分析によってその影響を示す必要がある。	5.3.8	QOL 値は、「8.2」及び「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
		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。	8.4.1
● 年齢や合併症の調整			

NICE Methods guide		中医協ガイドライン	
要訳	章	原文	章
場合によっては、年齢や合併症など、QOL 値の調整が必要な場合がある。	5.3.7	該当項目に関する記載なし	n/a
● マッピング			
EQ-5D データが入手できない場合、臨床試験等で観察された他の健康関連 QOL データや、健康関連便益を EQ-5D にマッピングすることによって推定することができる。選択されたマッピング関数は、健康関連 QOL 指標とその統計的特性の両方を含むデータセットに基づいていなければならない。その統計的特性は完全に記述され、その選択が正当であり、その関数がどれだけデータに適合するか十分に示されなくてはならない。また、マッピングアルゴリズムを使用した場合の出力のばらつきを調べるため感度分析を示す必要がある。	5.3.9	「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 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。	8.2.2
● EQ-5D が不適切である場合			
EQ-5D が不適切であると主張するためには、評価対象の疾患における EQ-5D の内容的妥当性が十分に得られておらず、健康の重要な側面を網羅できていないことを質的な実証的証拠を提示して示す必要がある。このような状況においては、代替の健康関連 QOL 尺度が使用してもよいが、その際にはデータを作成するために使用した方法、その妥当性、およびこれらの方法が QOL 値にどのように影響するかについて詳細な説明が必要である。	5.3.10	該当項目に関する記載なし	n/a

NICE Methods guide		中医協ガイドライン	
要訳	章	原文	章
● 子供の健康関連 QOL の測定について			
子供の健康関連 QOL の測定尺度について、選好に基づく測定尺度の標準化及び、その妥当性の検証が必要。EQ-5D の標準版は子供向けには設計されていない。7-12 歳については代替版が利用可能だが、英国の価値づけはまだ利用できない。	5.3.11	該当項目に関する記載なし	n/a
● EQ-5D-5L の位置づけ、取り扱い			
EQ-5D-5L が開発され基本分析で使用可能である。EQ-5D-5L の質問票は検証（バリデート）されているが、QOL 値を算出する換算表（バリューセット）は存在しない。EQ-5D-5L の換算表が利用可能になるまで、QOL 値算出のため既存の EQ-5D（-3L）から EQ-5D-5L へのマッピング関数が使用可能である（ www.euroqol.org ）（TF 注：現在は適切でないという評価もあるため、マッピング関数は使用できない）。 2017 年 8 月、NICE は EQ-5D-5L の換算表（バリューセット）の使用に関する立場声明を発表した。	5.3.12	（EQ-5D-5L に関する記載はなし）	n/a

2.3 NICE が EQ-5D を推奨する理由

NICE が EQ-5D を推奨している理由は、健康関連 QOL を測定する上で UK の一般人を対象とした直接法に基づいた価値づけ方法が採用されている（EQ-5D は TTO 法を用い UK 国民の大規模調査を行って QOL 値への換算表があるため該当）ためである。さらに、標準化され検証された尺度であり、既に多くの患者で実証されていること、健康関連 QOL は測定方法や測定尺度が異なると異なる QOL 値を生み出しかねないが、EQ-5D に統一すればその問題は生じないことなどが、推奨理由として挙げられる。これは他の健康関連 QOL からマッピングする際にも適用されており、EQ-5D データが入手できない場合、他の健康関連 QOL データを EQ-5D にマッピングした QOL 値を使用することが可能としている。

2.4 日本では分析をする際はどうすればよいか？

中医協ガイドラインの記載によると、日本人の QOL 値は優先するものの、日本人の QOL 値の研究がない、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としている（中医協 8.4, 8.4.1）。費用効果分析を行うために、新たに日本国内で QOL 値を収集する際に使用する尺度は、現時点において EQ-5D に限定される（中医協 8.2.1）。しかし、中医協ガイドライン「8.2」では、QALY を算出する際の QOL 値は一般人の価値を反映したもの、つまり PBM, SG 法, TTO 法を用いればよいとしており、更にそれに該当するデータがない場合はその他の健康関連 QOL データから QOL 値へマッピングしたものも使用可能としている（中医協 8.2.2）。また、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直接法を用いて測定してもよいとしており（中医協 8.3.3）直接法であれば、TTO 法による測定が推奨されている。これを図式すると次のようになる（Figure 2-2.-2）。

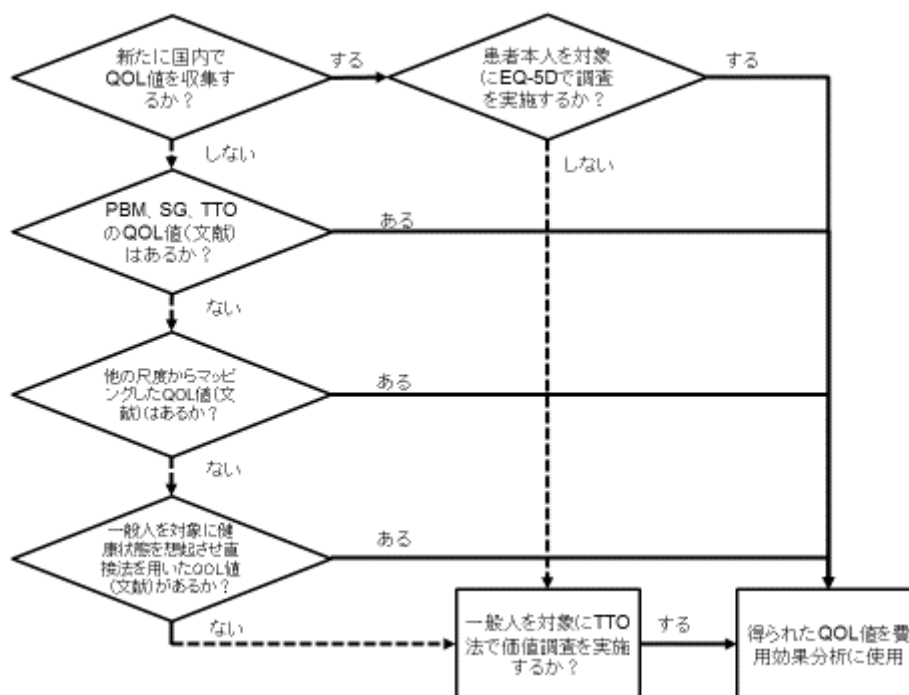


Figure 2-2. 日本における QOL 値収集のためのダイアグラム

2.5 EQ-5D の 3L と 5L

EQ-5D (-3L) に新しいバージョンの EQ-5D-5L が開発された。各 5 項目の質問について、従来は 3 段階の回答であったが、感度が高くなるようより細かい 5 段階の回答 (no

problem, slight problems, moderate problems, severe problems and unable to or extreme problems) が設定された。NICE Methods guide では、基本分析において 5L の質問票を使用するのは問題ないが、Methods guide 作成時点において、英国人の価値づけを反映した QOL 値を算出する換算表（バリューセット、タリフ）が開発されていなかったことから、換算表が開発されるまでは、既存の EQ-5D（-3L）から EQ-5D-5L の QOL 値を導出する検証済みのマッピング関数が存在するため、これを使用する、としている。その後 Devlin らによってイングランド向けの 5L の換算表が開発されたが、NICE は、EQ-5D-5L のバリューセットの使用に関するポジションステートメントを発表し、開発された 5L の換算表は更なる検討が必要であり、現時点では 5L の換算表の使用は推奨しないと発表した（2017 年 8 月）。更に 2019 年 10 月に次のようにステートメントが発表され、EQ-5D-5L の使用が行えるようデータの整備を行うとしている(11)。

1. EuroQol グループは、5 つの回答レベル（問題なし、軽度の問題、中等度の問題、重度の問題、極端な問題）を有する新しい質問票「EQ-5D-5L」を開発した。EQ-5D-5L は、EQ-5D-3L よりも感度が高いように設計されている。
2. 現在の NICE Methods guide（2013）では、EQ-5D-5L の質問票を用いて収集したデータを基本分析に使用してもよいとしている。ただし EQ-5D-5L の換算表（バリューセット）が入手可能になるまでは、既存の EQ-5D(-3L)から EQ-5D-5L へ QOL 値を導出するための有効なマッピング関数を使用すること。
3. イングランドの EQ-5D-5L の換算表が Devlin ら（2018）により開発・公表された。しかし、その開発方法、収集されたデータの信頼性等の懸念から、同バリューセットは使用しないこととなった。その後、国際標準の価値づけ研究プロトコルが改良され、追加の品質管理方法が導入された。
4. 以下のポジションステートメントは、CUA を使用する NICE のすべてのガイダンス作成プログラムに適用される。

ポジションステートメント

5. 我々（NICE）は、Devlin ら（2018）が発表したイングランドの EQ-5D-5L の換算表の使用を推奨しない。NICE へのエビデンス提出を準備している企業、学術団体、その他の者は、基本分析では 3L の換算表を使用すべきである。
6. EQ-5D-5L の質問票を用いてデータを収集した場合、基本分析における QOL 値は、5L の質問票で得られたデータを 3L の換算表にマッピングして算出する必要がある。また、EQ-5D-3L と EQ-5D-5L の両方の質問票を用いてデータが収集された場合、すべての QOL 値の導出に 3L の換算表を使用すべきであり、必要に応じて 5L を 3L にマッピングする必要がある。
7. 現行の Methods guide との整合性を保つため、EQ-5D-5L から EQ-5D-3L へのマッピ

ングに関し、基本分析では van Hout ら（2012）が開発したマッピング関数を使用すべきである（例えば Hernandez Alava ら（2017）のマッピング関数など複数あるがそれらは使用しない）。

8. 我々は、前向き臨床研究において、QOL のデータ収集のために EQ-5D-5L を継続して使用することを支持する。
9. 我々は、保健福祉省やその他の主要な利害関係者と協力し、我々の方法での採用を可能にする許容される品質の 5L の換算表が利用できるようにすることを約束する。EuroQol は、更新された国際標準プロトコルを用いて、イングランドでの新たな 5L 評価研究を委託している。この新しい研究が完了した時点で、EQ-5D-5L に関する方針を見直す予定である。

3 NICE Utilities TSD series

NICE Decision Support Unit では”NICE Utilities TSD series”として、経済評価における健康利益の測定や価値づけに関する TSD を 5 報公開している (<http://nicedsu.org.uk/technical-support-documents/utilities-tds-series/>). ここで留意すべき点は、これら 5 報の TSD は 2011 年に公開されているため、2008 年版の Methods Guide をベースにしていることである。Methods guide の最新は 2013 年であり、2008 年版と比べて EQ-5D 質問票が適切でない場合の詳細 (NICE5.3.10)、一般化された健康関連 QOL が臨床試験で測定できなかった場合にどうすべきか (NICE5.3.8) についての記載が追加されている。

1. TSD 8: NICE 申請のための健康測定・価値の紹介
2. TSD 9: 文献からの各健康状態の QOL 値 (Health State Utility Value: HSUV) の特定, レビュー, 統合
3. TSD 10: HSUV を推定するためのマッピング手法の利用
4. TSD 11: HSUV を生み出すための EQ-5D の代替手法
5. TSD 12: 分析モデルでの HSUV の利用

TSD 9, TSD 10 には関連文献として論文も紹介されている(12, 13)

TSD 8 では EQ-5D を推奨している NICE の基本分析の紹介を行った後に、EQ-5D が利用不可の場合、EQ-5D が適切ではない場合、EQ-5D が適切ではないかをどのように決定するのかについて、対応方法の概要が記されている。

他の 4 つの TSD は NICE 基本分析から生じる課題に対する対処方法を述べている。TSD9 は適切な各健康状態の QOL 値を得るための文献レビュー方法、TSD 10 は EQ-5D が利用できないときのマッピングの方法、TSD 11 は EQ-5D が利用できないときの代替手法、そして TSD 12 では分析モデルで HSUV を用いるときの特別な課題について紹介している。

次章以降の内容は TSD 8～TSD 12 の Executive summary と key table を中心に各 TSD で何が述べられているか、重要な点は何かについて紹介する。TSD の内容が Methods Guide 2013 と異なると思われる部分については「TF 注」としてコメントを記載した。詳細については各 TSD、TSD 内の引用文献を参照されたい。

4 TSD8: NICE 申請のための各健康状態の QOL 値測定・価値づけの紹介

NICE は意思決定において一貫性を要求している。そのため、NICE は各健康状態の QOL 値の測定と価値づけを含む技術評価手法のための基本分析の考え方を **Methods Guide** として提供している。TSD 8 では、NICE 基本分析の概要から始まり、次に NICE の推奨する尺度である EQ-5D に代わる方法を用いることを正当化するためにどのようなエビデンスが必要かを検討している。

4.1 NICE 基本分析の概要

2.1 節でも紹介しているが、TSD8 で述べられている NICE 基本分析の主要な構成要素は以下のとおりである。

1. QALY は介入効果の推奨尺度である

QALY は、生存と健康関連 QOL の転帰を統合したものであり、健康関連 QOL を 0 が死亡、1 が完全な健康であるという尺度に置き、生存期間を通しその総和をとることである。健康関連 QOL の状態を評価するための様々なアプローチと手法がある。

2. 成人患者本人から報告された EQ-5D を優先する

成人患者本人から報告された EQ-5D を優先するが、患者が報告できない場合は、患者の近親者の介護者からの報告でも良い。EQ-5D には、TTO 法を用いて英国の一般集団の代表サンプルから得られた既存の値が存在している。

3. HUI3 や SF-6D を感度分析として用いることが可能である

HUI3 および SF-6D のような他の全般的な選好に基づく尺度は感度分析として利用可能である。

TF 注: Methods Guide 2013 では、HUI3 や SF-6D を感度分析として利用可能とは記載されていない。EQ-5D の使用が不適切である場合には使用しても良いことが記載されている。

4. 小児については標準化と妥当性の検証が必要である

小児については、標準化され、妥当性が確認された選好に基づく尺度がより広く用いられるようになることを望んでいる。

2 章では NICE 基本分析の概要を述べている。Implications (要点) に記載されていることを以下にまとめた。

Implications（要点）：

- 健康上の転帰は、QALYs を用いて表現すべきである
 - 基本分析では QALYs を使用する
 - QALYs の基礎となる仮定（健康状態の QOL 値は評価期間と独立しており、かつ、健康状態の QOL 値は健康状態の発生順番と無関係である）がこの特定の疾患で成り立たないという懸念がある場合には、裏付けとなるエビデンスを提供し、代替分析を提示すること
- 各健康状態の QOL 値の測定と価値付け
 - 健康関連 QOL は通常 EQ-5D を用いて患者から報告されるべきである
 - QOL 値への換算表は直接法を用いて UK の一般集団から得られるべきである
 - EQ-5D は、ほとんどの状況において QOL 値を生成するための望ましい尺度と考えるべきである
- 患者が QOL を報告できない場合
 - 患者が自身の健康状態を報告できない場合、患者に近い介護者から得られるべきである
- 介護者の健康
 - 重要であると想定される場合は、EQ-5D を用いて介護者の健康に対する治療の影響を含める
- 臨床試験に組み入れられた選好に基づく他の健康尺度データ
 - EQ-5D 以外の選好に基づく尺度が有効性を推定するために用いた臨床試験に組み入れられている場合は、感度分析として提示する*
- 小児における健康関連 QOL 値測定
 - 幼児の健康状態を測定する際には、小児用に開発された質問票の適用を検討する

*TF 注：Methods Guide 2013 では、HUI3 や SF-6D を感度分析として利用可能とは記載されていない。EQ-5D の使用が不適切である場合には使用しても良いことが記載されている。

4.2 EQ-5D の適切性の評価

EQ-5D は試験で収集することができるが、多くの場合では日常診療下でデータを収集したり、既存の推定値を文献から利用したりの方が適切であろう。関連する EQ-5D データが入手できない場合、もう一つの解決策は、関連する研究で使用されてきた健康関連 QOL または疾患重症度の別の尺度からマッピングし、マッピング関数から EQ-5D を予測することであろう。これらは、両方の質問票を含む他のデータセットから推定することができる。

この戦略は、EQ-5D データがない場合には NICE によって受け入れられているが、EQ-5D を直接使用することに対する次善策であり、不確実性が増加するというペナルティを伴う可能性がある。

状況によっては、NICE は EQ-5D が適切でない可能性があることを認識している。しかし、ゴールドスタンダードな尺度がない場合、特定の患者群において健康関連 QOL の尺度が妥当であるか否かを証明することは困難である。EQ-5D が内容的妥当性 (content validity)、表面的妥当性 (face validity)、構成概念妥当性 (construct validity)、反応性 (responsiveness) の点で不適切であることを実証するための経験的エビデンスを NICE は要求している。内容的妥当性は、その質問票が患者にとって重要である健康関連 QOL のすべての次元を網羅しているかどうかに関わる。表面的妥当性は患者が質問票の記述内容を正しく理解しているかで評価できる。構成概念妥当性には、その尺度が群間の既知の差異を反映しているのか、他の関連する尺度 (例えば HUI3, SF-6D) と相関しているのかに関する定量的な証拠が必要である。

反応性とは、EQ-5D が健康関連 QOL の変化をどの程度反映しているかである。これらの基準は一般的な指標と同様に、測定の 5 次元にわたって評価されることが望ましいが、これはまれにしか行われない。

妥当性を検証するために用いられる変数の関連性は慎重に検討しなければならない。先述の通り、ゴールドスタンダードな尺度がない場合、健康関連 QOL の尺度が妥当であるか否かを証明することは困難であるため、どの患者群においても、NICE が行使すべき判断の程度は常に残るであろう。

EQ-5D が不適切かどうかを評価するチェックリストとして Table 4-1 に要約する。

Table 4-1. EQ-5D の不適切性を評価するためのチェックリスト

	構成要素
妥当性	● 内容的妥当性 ➤ その質問票は患者にとって重要な健康の次元を除外していないか？ ➤ 感度の低い項目が含まれていないか？ ● 表面的妥当性 ➤ 当該項目は、母集団にとって関連があり、適切であるか？ ● 構成概念妥当性 ➤ EQ-5D は群間の既知の差異を反映しているか？ ➤ EQ-5D とその関連する次元は同じ概念の他の尺度と相関しているか？ (例. HUI3 や SF-36 と相関しているか)
反応性	● この質問票は健康状態の既知の変化を反映していないのではないか？

[NICE DSU TSD8 の Table 1 を改変して掲載]

この TSD Utility シリーズの残りの部分では、文献 (TSD 9)、マッピング (TSD 10)、および EQ-5D が不適当である場合の代替手段 (TSD 11) における価値の使用に関する重要な問題について、より詳細に考察している。最後に、本シリーズの最後では、分析モデルにおける HSUV の使用 (TSD 12) について考察する。

小児における最も適切な嗜好に基づく尺度を確立するためにはさらなる作業が必要である。TSD8 の要点を以下にまとめた。

Implications (要点) :

- EQ-5D に関するデータは費用効果分析モデルの要件に応じて、臨床試験、観察研究や文献で収集することができる
- EQ-5D が利用できない場合には他の健康関連 QOL 尺度からのマッピングが解決策となり得るが、不確実性が増加するかもしれない
- EQ-5D は、特定の患者群に対する治療では評価できないことが経験的に実証されない限り、適切であると仮定している
 - これを支持する証拠は、特定の状態における EQ-5D データのエビデンスのシステムティックレビューに基づくべきである
 - EQ-5D の内容的妥当性、構成概念妥当性および反応性に関するデータを含むべきである
 - 症例数及び完了率を含め、試験の質を考慮すべきである
 - 変化が予想される健康の側面、および特定された研究での比較に用いられた尺度の妥当性を考慮すべきである
- 代替法はあるが、これらを用いることで NICE での評価における比較可能性が低下するため、その使用に関するエビデンスは説得力のあるものでなければならない

5 TSD9：文献からの QOL 値の特定，レビュー，統合

本章では，TSD9 を紹介する．EXECUTIVE SUMMARY の要約を紹介した後に，各健康状態の QOL 値（Health State Utility Values：HSUV）レビュープロセスの説明と具体事例として TSD9 で紹介されている骨粗鬆症関連の事例の一部を紹介する．なお，TSD9 の要約は Pappaioannou らの論文に示されていて，骨粗鬆症関連の HSUV レビュー事例が紹介されている(14)．

5.1 EXECUTIVE SUMMARY の要約

5.1.1 背景

HSUV は分析モデルでの重要なパラメータであり，NICE は文献からの HSUV 推定値が系統的に特定・選択されているということを要求している．NICE は基本分析において，どのように HSUV が導出されたかについて特別な要求を出していて，EQ-5D から導かれた HSUV を優先している．TSD9 では，NICE Methods Guide の要件を満たすために，文献から HSUV を系統的に特定し，選択する方法について考察している．さらに，質および関連性評価，データ抽出，値の選択および統合に関するガイダンスを提供する．2 つのケーススタディ，骨粗鬆症関連病態における HSUV のレビューおよび乳癌における HSUV のレビューが，TSD9 全体にわたって使用されている．5.1.2~5.1.7 で詳細を説明する．

TF 注:Methods Guide 2013 5.3.8 には，”EQ-5D データを臨床試験で得られていない場合は，文献から EQ-5D により測定された QOL 値の引用も可能である”と記載されているため，臨床試験で全ての QOL 値を入手できれば文献検索は不要である．しかしながら，費用対効果評価モデルで設定する全ての健康状態の QOL 値を臨床試験で取得するのは困難であるため，一般的には文献検索が必要になる．

5.1.2 HSUV レビューの範囲

レビュー範囲を決める段階の目的は，分析モデルのために必要な HSUV を特徴付けることである．2 つの重要な要素を定義する必要がある．①分析モデルに必要な，特定の健康状態，そして②NICE が要求している HSUV データのタイプ，である．レビューの範囲を決める段階では，利用可能な HSUV エビデンスの量・質が正確には分からないため，当初はレビューの適用範囲を広くしておくことが推奨される．エビデンスの選択段階では，レビュー範囲のさらなる改良が行われる．HSUV は，多くの健康状態，サブグループおよび異なる期間にわたって必要とされる可能性があるため，分析モデルで探索した一次条件を超えた健康状態にまでレビュー範囲が及ぶ可能性がある．NICE が要求している HSUV データのタイプは，EQ-5D である．しかし，EQ-5D が一部の患者集団に不適切であると考えられる場合もある．

5.1.3 エビデンスの特定と選択

HSUV レビューを行うためには、検索した研究の参考文献リストを精査するとともに、HSUV に関する検索用語の広範なリストを用いた電子データベース検索が必要である。理想的には、専門家の意見、引用、著者検索など、他の補足的検索技術を用いるべきである。レビューのためのエビデンスの選択には、レビュー範囲の改良や選択/除外基準の改良が含まれる。

これには、次の 3 つの重要な詳細についての予備的データ抽出が必要である。①健康状態を記述する集団の詳細（例：年齢、性別、疾患の重症度）、②健康状態を記述するために用いられたアプローチの詳細、③ HSUV 算出方法（例：TTO, SG, VAS）。この予備的データ抽出からの知見に基づいて、利用可能なエビデンスの性質と量についての全体像を構築することができ、含めるか除外するか決定が可能になるであろう。エビデンスを特定し、選択するためのプロセスは透明性が求められ、文書化が必要である。

5.1.4 質及び適切性評価

HSUV 研究のための合意された報告基準はない。提案されている HSUV 研究の質評価の主要な基準は、症例数、回答者の募集/登録方法、選択/除外基準、回答率、追跡不能例数（および理由）、欠測値に対する解析の方法である。質評価と同じくらい重要なのは、データが分析モデルおよび評価機関の要求に対して適切かということである。NICE はデータがどの程度 NICE の要求する基本分析と一致しているかを検討する。

5.1.5 データの抽出

データ抽出は、臨床的有効性のレビューとほぼ同じ原則に従う。したがって、著者または発表国などの一般情報、選択/除外基準などの研究特性、年齢、性別、疾患特性および研究の設定、質問表・価値付けの方法・誰から測定したかに関する情報は、結果に関する記述統計と同様に記録することが重要である。

5.1.6 データ提示

データ提示には、以下のものが含まれる。

- ① 選択した研究の特性
- ② 分析モデルで用いる HSUV（使用に対する正当化）
- ③ 選択した研究の質と妥当性評価
- ④ 分析モデルで使用する HSUV への修正
- ⑤ 実施した感度分析の詳細

5.1.7 データの統合

選択および統合の方法は、HSUV データの利用可能性に依る。以下の 3 つの状況が考えられる。

① 適切な HSUV が 1 セットのみ

モデルで使用する健康状態の推定値とすべきである

② レビューの基準を満たす HSUV が複数存在

モデルで使用する値の選択を正当化する必要がある。他の HSUV より適切性が低いと思われる HSUV は感度分析として使用すべきである。複数の関連 HSUV が十分に均質である場合（つまり、同じ質問票で同じ患者母集団から収集し、同じ英国の価値付け方法を用いている場合は）、HSUV の平均値と分散の推定値の精度を改善する方法として統合を考慮すべきである

③ 適切な値がない場合

例えば、基本分析に合致しない場合や患者集団がモデルに適切でない場合である。この場合においても、値を選択して正当化する必要がある。メタ回帰分析が役に立つ可能性がある

5.1.8 HSUV レビュー実施における推奨

- HSUV レビューのためのエビデンスの範囲と特定は、最初は広くする必要がある
- 電子データベースの検索、参考文献リストの確認、専門家との連絡など、関連する研究を特定するために、理想的には様々な方法を使用する
- エビデンス選択の段階で選択/除外基準を改良する
- HSUV レビューのためのエビデンス選択は反復的なプロセスであり、重要な特性（母集団の詳細、健康状態を記述するために使用するアプローチ、HSUV 算出方法）の予備的データ抽出を伴うことがある。そのデータに基づいて、選択基準をさらに修正し、改善することができる。
- 対象研究の選択は、十分に正当化され、明確でなければならない。研究選択/除外理由の記録を保管しなければならない
- 質と適切性の基準の提案（Table 5-6 および 5-7 を参照）。
- 基本分析に合致する値が 2 セット以上存在する場合は、最終選択の正当化の理由が必要であり、他の値については感度分析で提示する。値の統合も考慮する
- データ提示には、研究を特定するために実施した検索、選択した研究の特性、分析モデルで使った HSUV、質および適切評価、モデルで使った HSUV を修正した場合は、その修正に関する説明を含めなければならない

5.1.9 結論

NICE は、公表文献からの HSUV 推定値は系統的に特定・選択されたエビデンスに基づくことを要求している。臨床的有効性レビューのためのシステマティックレビューの原則は、系統的に QOL 値を特定・選択する方法で参考となるが、HSUV レビューのエビデンスの範囲および特定には特有の課題が存在する。エビデンスの特定・選択のプロセスは、1つの文献検索ではなく、一連の検索がしばしば必要となる場合があるという点で異なる。また、研究選択は、利用可能なエビデンスに応じて HSUV データのタイプを改良するという観点から、選択基準にも影響を与える。エビデンス評価のプロセスには、質と関連性の両方の評価が含まれる。モデルで使用した値の最終的な選択は正当化が必要であり、他の値についても感度分析を実施する必要がある。

5.2 HSUV レビュー：ケーススタディ

ケーススタディとして骨粗鬆症関連疾患のシステマティックレビューを取り上げて、HSUV レビューの一連の流れを紹介する。ケーススタディの主要な特徴を Table 5-1 に示す。

TF 注：EQ-5D 以外も検索する理由は、EQ-5D が無いなら、その他のデータで利用可能なものがないか検討するためである。検索漏れを防ぐためにも、エビデンスの範囲と特定は、最初は広くする必要がある。

Table 5-1. ケーススタディの主要な特徴

	母集団	健康状態	特定された HSUV のタイプ
Peasgood, T et al (2009). 骨粗鬆症関連疾患に対する HSUV の最新のシステマティックレビュー(15)	骨粗鬆症関連疾患を合併する成人の男女	• 確定した骨粗鬆症 • 脊椎骨折 • 大腿骨頸部骨折 • 手首関節骨折 • 肩骨折	自身の健康状態: SG, TTO, VAS 選好に基づく尺度: EQ-5D, SF-6D, HUI3, QWB 専用の vignette (患者や専門家へのインタビュー)

SG; standard gamble, TTO; time-tradeoff, VAS; visual analogue scale, QWB; quality for wellbeing

[NICE DSU TSD9 の Table 1 を改変して掲載]

本ケーススタディは以下の Figure 5-1 に沿って行われた。

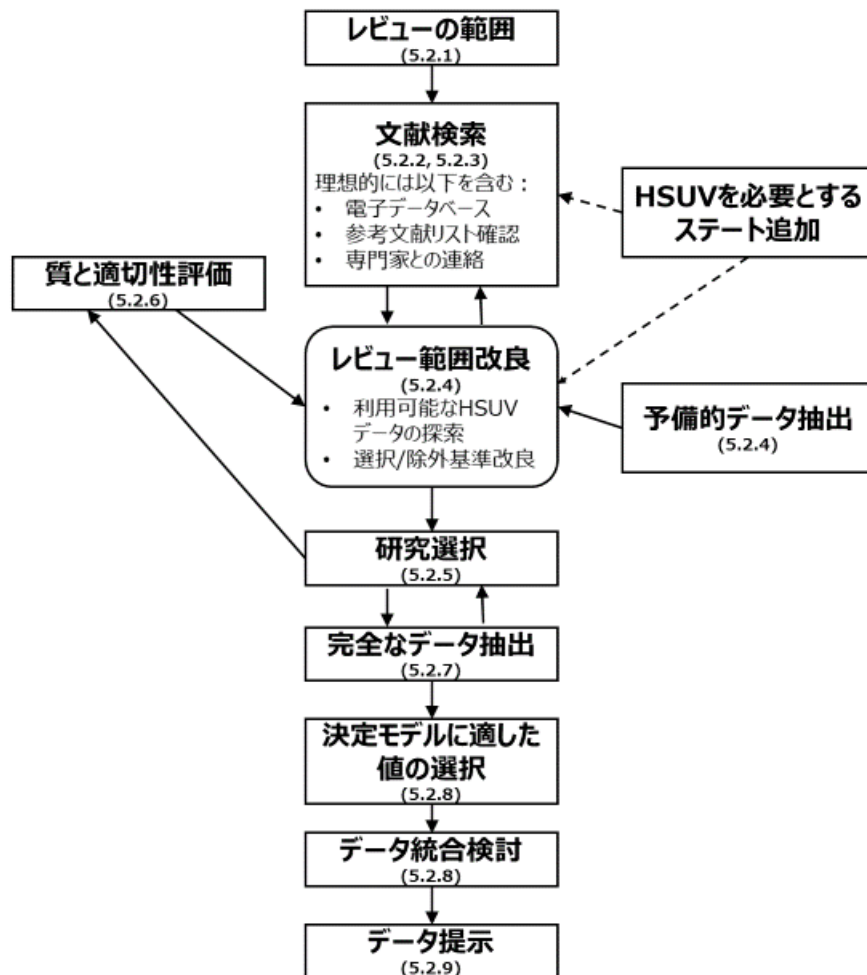


Figure 5-1. HSUV のレビュー

図の各ステップの数字は下記の 5.2.1～5.2.9 のどこに対応しているかを記載した

[NICE DSU TSD9 の Figure 1 を改変して掲載]

5.2.1 レビューの範囲

骨粗鬆症の費用効果モデルに必要な健康状態を Table 5-2 に示す。矢印で示した通り、骨折直後の HSUV は骨折直後から時間とともに変化する考えられるため、骨折に関する

る HSUV は複数時点での健康状態の考慮が必要であろう。

Table 5-2. 骨粗鬆症の治療と予防の分析モデルに必要な特定された HSUV データ

健康状態	病期 - 骨折前 - 確定した骨粗鬆症 → 脊椎変形あり → 脊椎変形なし 骨折 - 骨折既往あり - 骨折既往なし 骨折のタイプ - 脊椎骨折 - 股関節 - 肩 - 手首 - 多発性骨折 非骨粗鬆症の健康状態 - 乳がん - 心房細動 - 歯周病における骨喪失
母集団のサブグループ	年齢別 閉経状態（閉経前，閉経後）
他の検討事項	ナーシングホーム 自立生活

骨折後の時間
(複数時点での健康状態を考慮)

[NICE DSU TSD9 の Figure 2 を改変して掲載]

5.2.2 文献検索

Medline (MeSH : Medical Subject Heading) や Embase (EMTREE) のシソーラス検索では，共通 HSUV (EQ-5D, SF-6D など) のシソーラス用語はない。しかしながら，いくつかの用語，例えば Quality of Life や quality adjusted life year といった用語は，HSUV に直接は関連しないが，より広い概念の下でインデックス化される用語として考えられる。最も良く使われる MeSH 見出しを Table5-3 に記載した。

Table 5-3. HSUV 研究での頻出 MEDLINE MeSH 用語（300 件の検討に基づく）

Quality of life	
Questionnaires	
Psychology (subheading)*	
Health status	
Health status indicators	
Activities of daily living	
Health surveys	
Quality adjusted life years	
Treatment outcome	
Psychometrics	
*Subheadings とは、見出しに追加して意味を絞り込める用語である。	

[NICE DSU TSD9 の Box 2 を改変して掲載]

骨粗鬆症のケーススタディで選択した研究の MeSH をレビューした。選択した 28 研究のうち、Medline に存在した研究は 24 研究であった。24 研究のうち、1 研究を除くすべての研究に MeSH 用語'Quality of Life'が割り当てられていた。次に頻繁に割り当てられていた MeSH 用語は、'Questionnaires'であり、24 研究のうち 8 研究に割り当てられていた。しかしながら、Table5-3 に記載されている他の MeSH 用語の大部分は、骨粗鬆症のケーススタディで選択した研究に割り当てられていないか、2～3 件の研究のみであった。

MeSH 用語として'Quality of Life'や'Questionnaires'を用いて検索することは有用な方法であるかもしれないが、感度と特異度はトレードオフであり、これら 2 つの MeSH 用語のみでの検索は重要な研究を取りこぼすかもしれない点に留意すべきである。

フリーテキスト検索は 3 つのカテゴリーに分類される。一般的な用語（例えば QALY, HSUV）、質問票特有の用語（例えば EQ-5D, SF-6D など）、そして、HSUV 算出方法（SG, TTO）である。Table5-4 に電子データベースで HSUV を検索する際に用いるフリーテキスト用語を列挙した。

Table 5-4. HSUV を電子データベースで検索する際の共通のフリーテキスト用語

quality adjusted life
quality-adjust-life
(qaly\$ or qald\$ or qale\$ or qtime\$)

disability adjusted life
 daly\$
 (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six)
 (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six)
 (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve)
 (sf6D or sf 6D or short form 6D or shortform 6D or sf six D or sfsixD or shortform six D or short form six D)
 (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty)
 (euroqol or euro qol or eq5d or eq 5d)
 (hql or hqol or h qol or hrqol or hr qol)
 (hye or hyes)
 health\$ year\$ equivalent\$
 utilit\$
 (hui or hui1 or hui2 or hui3)
 disutili\$
 rosser
 quality adj2 wellbeing
 qwb
 standard gamble\$
 SG
 time trade off
 time tradeoff
 tto
Key
 \$ =truncation (*とするデータベースもある) 例 utilit\$ は utility や utilities を検索
 adj= 隣接演算子 例 adj2 = 2 語以内で隣接する単語

[NICE DSU TSD9 の Figure 4 を改変して掲載]

骨粗鬆症のケーススタディで選択した研究のフリーテキスト用語をレビューした。
 Medline に存在した 24 研究のうち、17 研究が'Quality of life'をタイトルに含んでいた。残

り 7 研究のうち 5 研究は'Quality of life'を抄録に記載していた. 残りの 2 研究はタイトルか抄録に HSUV を算出するために用いた特定の QOL 質問票に関する用語がフリーテキストとして含んでいた.

臨床的有用性のレビューと同様に, HSUV データの検索においても電子データベース検索は主要な方法である. MEDLINE 等のデータベース検索と同様に, 以下のウェブサイトも HSUV データ検索に有用である.

- Cost-effectiveness Analysis Registry(16) (CEA registry)
- Centre for Reviews and Dissemination (CRD) databases: NHS EED and HTA(17)
- Health Economics and Evaluation database (HEED) (18) *TF 注: 現在はサイトなし*
- The EQ-5D website (他の質問票のウェブサイト)(19)
- The MAPI Institute website(20) *TF 注: 現在は Mapi Research Trust と思われる*
- Submissions to NICE or other health care decision agencies
- Research Papers in Economics (RePEc)(21)
- Patient-reported outcome and quality of life instruments database (PROQOLID) (22)

TF 注: 2.2 節で記載している通り, 中医協ガイドラインでは, HSUV データ, つまり QOL 値は国内の調査結果を優先的に使用することが推奨されている. C2H のウェブサイトで紹介されているデータベースは上述の CEA registry, CRD database に加えて QOL データベースがある(23). これは日本国内で選好に基づく尺度で測定された研究のデータベースである. 例えば'general population'で検索すると 4 件ヒットし, EQ-5D-5L で測定された男女年齢別の QOL 値の研究結果を入手することができる (10).

5.2.3 骨粗鬆症での HSUV レビュー時の検索

以下に使用した検索用語と検索方法を, アプローチの長所と短所を強調して示す.

健康状態用語 (Table 5-5-1)

- 健康状態: 骨粗鬆症に関連する病態を合併する成人男女
- ステップ 1~4 は骨粗鬆症を記述する用語
- ステップ 5~14 は骨密度を表す用語
- ステップ 15~46 は骨折用語とホルモン補充療法または閉経を記述する用語の組み合わせ
- ステップ 47 は, 健康状態に関する用語を OR で組み合わせる.

Table 5-5-1. 骨粗鬆症ケーススタディ: 健康状態用語に関する Medline 検索*

No	Request
1	*Osteoporosis/
2	*Bone Diseases, Metabolic/
3	osteoporos\$.ti.
4	1 or 2 or 3
5	(bone adj6 densit\$).ti.
6	*Bone Density/
7	(bone or bones).ti.
8	*densitometry/
9	*Tomography, X-Ray Computed/
10	densit\$.ti.
11	9 and 10
12	8 or 11
13	7 and 12
14	5 or 6 or 13
15	*Colles' Fracture/
16	*hip fractures/
17	*Spinal Fractures/
18	15 or 16 or 17
19	*Fractures, Bone/
20	fractur\$.ti.
21	18 or 19 or 20
22	colles\$.ti.
23	(hip or hips).ti.
24	(femur adj6 neck).ti.
25	(femoral adj6 neck).ti.
26	(spine or spinal).ti.
27	vertebra\$.ti.
28	*Lumbar Vertebrae/
29	22 or 23 or 24 or 25 or 26 or 27 or 28

30	21 and 29
31	*Estrogen Replacement Therapy/
32	estrogen replacement therapy.ti.
33	oestrogen replacement therapy.ti.
34	hormone replacement therapy.ti.
35	ert.ti.
36	ort.ti.
37	hrt.ti.
38	31 or 32 or 33 or 34 or 35 or 36 or 37
39	*menopause/
40	*Climacteric/
41	menopaus\$.ti.
42	postmenopaus\$.ti.
43	climacteric.ti.
44	39 or 40 or 41 or 42 or 43
45	38 or 44
46	30 and 45
47	4 or 14 or 46

* TF 注：詳細については割愛するが、現在では *PubMed* で高度な検索機能を活用すると自動的に検索式を記載してくれる。詳細は成書(24)を参照。

[NICE DSU TSD9 の Appendix A.1 を改変して掲載]

HSUV 用語 (Table 5-5-2)

- HSUV：HSUV（典型的には VAS，SG または TTO）を得るための価値付け方法を用いて経験的に推定
- Quality of Life（ステップ 48 および 49），または quality of well-being（ステップ 61 および 62）などの幅広いフリーテキスト用語を含める。
- 特定の質問票の用語を含める：SF-36（ステップ 52），EQ-5D（ステップ 54），HUI 1-3（ステップ 60）および SF-6D（ステップ 71）
- QOL 値の算出方法に関連する用語を含める（ステップ 65-70）

- 他の一般的な用語の使用 ; health utility\$, heatlh state\$, preference\$, health state\$, utilit\$

Table 5-5-2. 骨粗鬆症ケーススタディ : HSUV 用語に関する Medline 検索

No	Request
48	exp quality of life/
49	quality of life.tw.
50	life quality.tw.
51	hql.tw.
52	(sf 36 or sf36 or sf thirtysix or sf thirty six or short form 36 or short form thirty six or short form thirty-six or short form36).tw.
53	qol.tw.
54	(euroqol or eq5d or eq 5d).tw.
55	qaly\$.tw.
56	quality adjusted life year\$.tw.
57	hye\$.tw.
58	health\$ year\$ equivalent\$.tw.
59	health utilit\$.tw.
60	hui.tw.
61	quality of wellbeing\$.tw.
62	quality of well being.tw.
63	qwb.tw.
64	(qald\$ or qale\$ or qtime\$).tw.
65	standard gamble\$.tw.
66	time trade off.tw.
67	time tradeoff.tw.
68	tto.tw.
69	visual analog\$ scale\$.tw.
70	discrete choice experiment\$.tw.

71	(sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or short form six or short form six).tw.
72	health state\$ utilit\$.tw.
73	health state\$ value\$.tw.
74	health state\$ preference\$.tw.
75	or/48-74

[NICE DSU TSD9 の Appendix A.1 を改変して掲載]

制限 (Table 5-5-3)

- ステップ 76～79 で letter, editorials, comment を除外
- 過去にレビュー実施済みであったため、2000 年以降の文献に限定 (ステップ 82)

Table 5-5-3. 骨粗鬆症ケーススタディ：制限に関する Medline 検索

No	Request
76	letter.pt.
77	editorial.pt.
78	comment.pt.
79	or/76-78
80	75 not 79
81	47 and 80
82	limit 81 to yr="2000 - 2007"

[NICE DSU TSD9 の Appendix A.1 を改変して掲載]

情報源

- 10 の電子データベースを検索 (学会抄録検索を含む) : Cochrane controlled trials register (Central), Cochrane database のシステマティックレビュー, EMBASE, Science Citation Index, MEDLINE, MEDLINE in process, NHS Database of Reviews of Effectiveness (DARE), NHS Economic evaluation database (NHS EED), Index to Thesis, ISI 議事録
- 2 誌のハンドサーチ (Osteoporosis International 誌と Journal of Bone and Mineral Research 誌)

- 関連する研究の参考文献リストの確認
- 骨粗鬆症研究の臨床医師・経済専門家との連絡

Table 5-6 が検索結果である。電子データベース検索での広範な検索用語の使用だけでなく、参考文献リストの確認や専門家との連絡も必要であることが分かる。

Table 5-6. 骨粗鬆症レビューで選択された研究

特定方法	選択数
電子データベース検索	13
参考文献リスト	8
専門家との連絡	4
過去のレビュー	3

[NICE DSU TSD9 の Table 2 を改変して掲載]

5.2.4 骨粗鬆症ケーススタディでの選択/除外基準

選択基準を広めに保つと幅広いアプローチからの HSUV 報告と数多くの健康関連 QOL 調査票をレビューすることになる。ケーススタディでは、Table 5-5 のように母集団を特定したが、レビューのためのエビデンスの範囲を決めて検索・選択する前に個別の健康状態を特定しなかった（例：手首関節骨折）。また、健康状態別プロファイルを記述するために使用される個々の調査票（例：EQ-5D）を事前に規定しなかった。しかしながら、広く認められている価値付け方法（VAS, TTO または SG）を用いて HSUV を推定した研究であることをレビュー基準に含めた。このようなアプローチは、分析モデルに用いるには広すぎる可能性がある。エビデンスの性質が明らかになるにつれて、選択基準は狭くした方が良い。

Table 5-7. 骨粗鬆症ケーススタディでの選択基準

- | |
|---|
| <ul style="list-style-type: none"> • 17 歳以上の成人 • 原発性または二次性骨粗鬆症に罹患している男性および閉経後女性 • 広く認められている価値付け方法である、SG, TTO または VAS を用いて経験的に推定された HSUV • 英語または翻訳 |
|---|

[NICE DSU TSD9 の Table 3 を改変して掲載]

5.2.5 レビュー結果の記述

研究特定方法の透明性と再現性のために、検索方針の詳細を記録することは重要である。NICE の企業 HTA 申請様式の B3.4 節と補遺 H では選択/除外基準の検索方針で使った用語の妥当性を説明するよう要求している (25)。(TF 注：様式は最新の様式番号に置き換えて記載した。以下も同様である。詳細は様式のユーザーガイドを参照のこと)。

フローダイアグラムは検索結果の総数、タイトルレベルで選択/除外した記録、抄録レベルで選択/除外した記録などの選択プロセスを記述するのに有用である。Figure 5-2 に骨粗鬆症でのフローダイアグラムを示す。

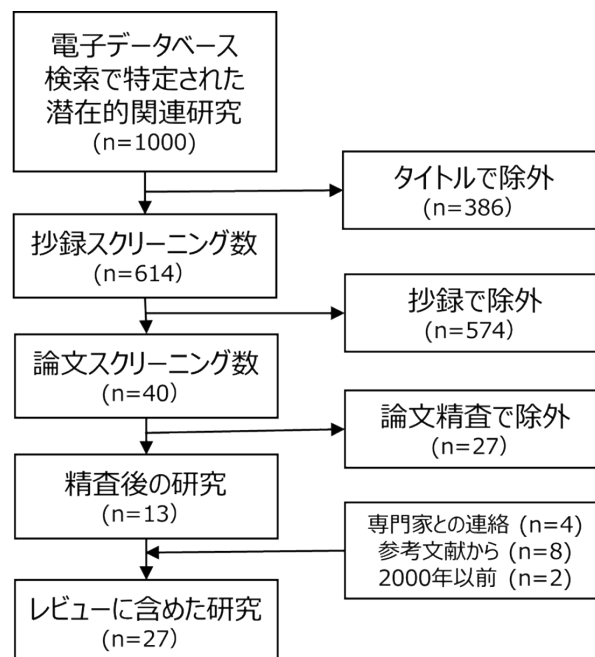


Figure 5-2. 研究選択/除外のフローダイアグラム

TF 注：中医協ガイドラインでは追加的な有用性を検討する際に実施するシステマティックレビューは PRISMA 声明の原則に従うことが記載されている (中医協 5.2.4)。

PRISMA 声明で用いられているフローダイアグラム(26) は上記フローダイアグラムに類似しており、HSUV データ検索においても参考になるであろう。

[NICE DSU TSD9 の Figure 5 を改変して掲載]

5.2.6 個々の HSUV 研究の質と適切性評価

HSUV 研究では合意された報告基準が存在しないため、HSUV レビューで選択した研究の質的評価は難しい。ケーススタディで用いたアプローチと経験則に基づき、Table 5-8 に HSUV 研究の質的評価で考えられる主要な基準を示す。

Table 5-8. HSUV 研究の質的評価で考えられる主要な基準

基準	考慮すべき点
サンプルサイズ	除外基準ではないが、推定値の精度はモデルで用いる推定値の分散で反映されるべきである
回答者の選択と登録	回答者の集団は分析モデルの対象集団の母集団とみなせるか？
選択/除外基準	除外されうる集団を除外しているか？（例：80 歳以上）
質問票の回答率	回答率が報告されているか、回答率は妥当性に影響を与える可能性が高いか？
追跡不能	追跡不能（例：骨折後 1 年）はどの程度か？理由は調査されているか？推定値の妥当性に影響を与える可能性があるか？可能性がある場合は、それは何か？
欠測値	欠測値のレベルと対処方法は？推定値の妥当性に影響を与える可能性はあるか？
研究に対する他の問題	例：実施地域（例：UK 以外で患者を募集した場合）

[NICE DSU TSD9 の Box 3 を改変して掲載]

分析モデルとモデル提出先の当局との関連性は質評価と同様に重要である。以下の Table 5-9 に NICE 基本分析との関連性を示す。

Table 5-9. NICE 基本分析に関連する評価基準

関連する質問	NICE での要求（TF 注：2008 年時点）
母集団の患者特性（例：年齢、性別、合併症、診断、疾患重篤度）は検討モデルでの対象患者と合致しているか？	答えがない場合、分析モデルで適切になるように値を調整する方法が利用可能である（TSD12）
どの質問票を健康状態の記述に用いるか？	一般的な選好に基づく質問票であり、EQ-5D を優先する
健康関連 QOL の変化をどの集団で調査するか？	患者から直接

患者の健康関連 QOL 変化の価値付けをどの集団で実施するか？	一般の集団
健康状態の価値付けを行うための方法は何か？	TTO のような choice-based method

[NICE DSU TSD9 の Box 4 を改変して掲載]

5.2.7 完全なデータ抽出

データ抽出は、臨床的有効性のレビューと同様に、どのようなデータを抽出するかを事前にデータ抽出フォームで作成しておき、必要なデータを収集することができるようにすることが理想である。例えば、著者または発表国などの一般情報、選択/除外基準などの研究特性、年齢、性別、疾患特性および研究の設定である。臨床的有効性のデータ抽出プロセスは独立した 2 名のレビューアで実施することが理想であるが、1 人のレビューアが実施したデータ抽出を 2 番目のレビューアが確認することも良く行われる。同様の方法を HSUV レビューでも適用することができる。

ケーススタディで抽出した経験に基づき、Table 5-10 にデータ抽出の見本を示す。本表は疾患特性など必要に応じて修正する必要がある。

Table 5-10. データ抽出フォームの見本

一般情報：

データ抽出者 _____

データ抽出日： _____

研究 参照 ID	研究詳細： 著者，タイ トル，年	回答者 の国	研究デ ザイン	選択/除外 基準	分析期間を 含む疾患健 康状態	健康関連 QOL 測定 時の患者特性	健康関連 QOL の価値 付けに用い た患者特性	健康状態記述シス テムの詳細（可能 であれば）	HSUV 算出 方法と回答 者	価値付け方 法（例： SG, TTO, VAS）
						性，年齢，疾患重症 度，他の関連する患 者特性		QOL 値の下限と 上限値について： 死亡か最悪の健 康，完全健康か通 常健康	例：直接法 か間接法 か，患者， 一般，臨床 家か	

回答者の選択と 登録方法	回答率	追跡不能理由	研究に関わる他 の問題点	HSUV 値の記述統計量：健康状態毎			
				サンプルサイズ/ 回答者数	平均（SD）	中央値	範囲

[NICE DSU TSD9 の Appendix A.3 を改変して掲載]

5.2.8 データの選択と統合

選択および統合の方法は、HSUV データの利用可能性に依る。以下の 3 つの状況が考えられる。

① 関連 HSUV が 1 セットのみ

モデルで使用する健康状態の推定値とすべきである

② レビューの基準を満たす HSUV が複数存在

モデルで使用する値の選択を正当化する必要がある。他の関連性の低い HSUV は感度分析として使用すべきである。複数の関連 HSUV 値が十分に均質である場合（つまり、同じ質問票で同じ患者母集団から収集し、同じ英国の価値付け方法を用いている場合）は、HSUV の平均値と分散の推定値の精度を改善する方法として統合を考慮すべきである。

骨粗鬆症の HSUV レビューでは股関節骨折において、EQ-5D を用いた HSUV 値の報告が多く選択された。股関節骨折後 1 年の QALY 損失は、分散の逆数を用いて 5 研究を重み付けした値とサンプルサイズで統合した値を統合 EQ-5D 値として推定した。これらの統合方法は単純であるが、集団が十分に均質であり、同じ質問票 (EQ-5D) から得られた値であれば許容可能であろう。

③ 直接関連する値がない場合

例えば、基本分析に合致しない場合や患者集団がモデルに適切でない場合である。この場合においても、値を選択して正当化する必要がある。メタ回帰分析が役に立つ可能性がある。

5.2.9 結果の提示

NICE が提供している企業 HTA 申請様式では、見出しを提示することにより HSUV レビュー結果を整理できるようにしている。（TF 注：詳細は様式のユーザーガイドを参照のこと）。

① 選択した研究の特性

選択した研究の特性について、NICE の企業 HTA 申請様式では、以下のものを含めることが求められている。

- ・ 測定対象集団
- ・ 対象者募集情報
- ・ 介入と比較対照
- ・ サンプルサイズ
- ・ 回答率

- ・ 健康状態の記述
- ・ 有害事象
- ・ 健康状態の適切性
- ・ 算出方法
- ・ 価値付けの方法
- ・ マッピング
- ・ 値の不確実性
- ・ 基本分析との整合性

② 分析モデルで用いる HSUV

NICE の企業 HTA 申請様式では、Table5-11 のような有用な表を提示している。また、セクション B.3.4 の”費用効果分析で用いた健康関連 QOL データ”の項で分析に用いた QOL 価の詳細と QOL 値を調整、バリデートしたかどうかについて記載することが求められている。

Table 5-11. 費用効果分析で用いる QOL 値の要約

健康状態	QOL 値：平均 (標準誤差)	95%信頼区間	申請参照番号 (セクションと ページ数)	正当化
健康状態 1				
健康状態 2				
等				
有害事象 1				
有害事象 2				

[NICE DSU TSD9 の Table 6 を改変して掲載]

骨粗鬆症のケーススタディでは、表形式と散布図を用いて視覚的に表示した (Figure 5-3)。複数の研究の HSUV を散布図で示すことは HSUV 値の類似性や異質性を視覚的に示す良い方法である。

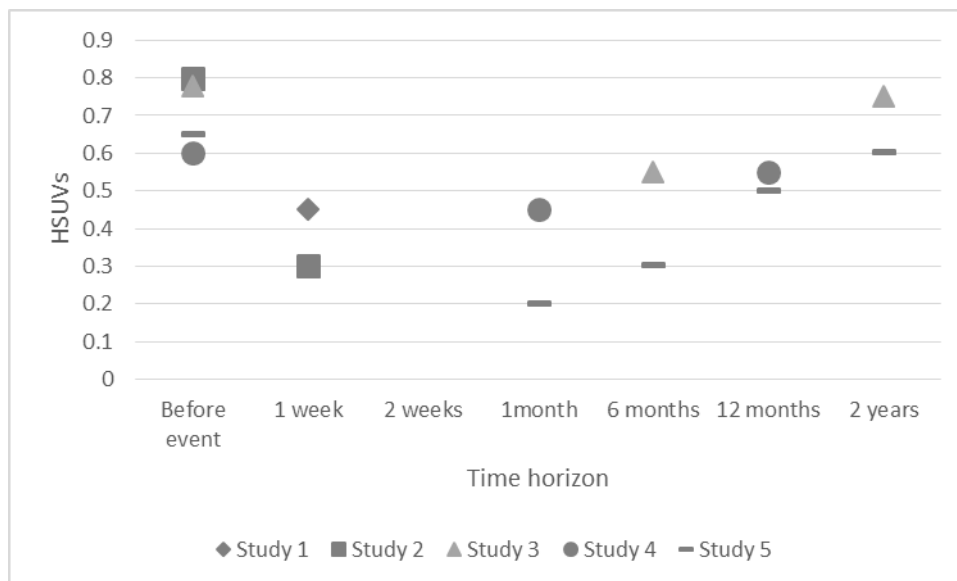


Figure 5-3. HSUV データの視覚的表示

[NICE DSU TSD9 の Figure 7 を改変して掲載]

③ 選択した研究の質と妥当性評価

現在の NICE Methods guide には、HSUV レビューの質および関連性評価の要件は含まれていない（TF 注：2013 年版も同様）。企業様式の補遺 H では、HSUV 研究の質及び関連性に関する項目（例：サンプルサイズ、回答率、基本分析との整合性、健康状態の適切性）が要求されているが、様式内でこれらの項目について詳細に考察する場所は存在しない。しかしながら、選択した研究の質と関連性を考察することは 5.2.6 節（Table 5-6, 5-7）に示された基準の実践であることを考慮すべきである。

④ 分析モデルで使用する HSUV への修正

TSD12 で述べているような分析モデルでの使用のために HSUV を適用あるいは修正した場合は詳細な理由を提示する必要がある。

⑤ 実施した感度分析の詳細

分析モデルの感度分析で HSUV の幅を用いた場合も、プロセスについて記載する必要がある。

6 TSD10 : QOL 値の推定のためのマッピング法の利用

NICE Methods Guide では、経済評価のアウトカム指標として QALY を使用すること、EQ-5D が QALY を計算するための QOL 値算出のために優先される尺度であることを推奨している。この NICE Methods Guide では、EQ-5D データが、報告書を作成するアナリストには常に利用可能というわけではないと認識している。そのようなデータが入手できない場合、NICE Methods Guide はマッピングを使用して EQ-5D データを予測することができると述べている。「マッピング」とは、他の指標または健康尺度に関するデータを使用して QOL 値を予測するためのアルゴリズムの開発および使用である。このアルゴリズムは、**target** である選好に基づく尺度、すなわち EQ-5D がオリジナルのデータ元である有効性試験に含まれていなくとも、QOL 値を予測するために、データ元となる予測尺度を含む臨床試験、他の研究、あるいは分析モデルからのデータに適用することができる。予測された QOL 値は、分析のための標準的な方法、または分析モデル内の各健康状態についての要約によって、分析することができる。

QOL 値を予測するためのマッピングの使用は、NICE Methods Guide で最近引用されているが、NICE への提出で長く使用されてきた。TA でのマッピングの使用の概要としては、マッピングが TA の 4 分の 1 以上に使用されているとのことであるが、我々の 2018 年から 2019 年の 1 年間の TA の調査でも 22% でマッピングが使用されていた（9 章参照）。

TSD10 は、EQ-5D データへのマッピングに使用できる方法を報告する既存の研究と、TA およびより幅広い文献に対して実施された既存の研究に基づいている。TA に含めるための QOL 値を推定するためのマッピングの使用を検討するアナリストに対する、一連の推奨事項を提供している。

- マッピングの使用は QOL 値の推定値の誤差と不確実性を増加させるので、ほとんどの場合、マッピングは EQ-5D により測定された QOL 値を直接収集することの次善の策と考えるべきである。
- マッピングは、専門家等の意見ではなく、データに基づいた統計的に関連付けられたマッピング関数に基づくべきである。これには 2 つの段階がある。まず、回帰モデルを使用して、**target** である EQ-5D スコアと他の **source** である健康尺度との関係を推定する。これらのモデルは、TA のために新たに推定してもよいし、またはマッピングに関する文献から推定してもよい。次に、これらのモデルの結果は、EQ-5D QOL 値の予測値を得るために、**source** となる予測尺度を含む臨床試験（単数または複数）または観察研究からの患者レベルまたは集団（平均）レベルデータに適用することができる。
- 推定サンプルの特徴はマッピング分析の **target** サンプルと同様であり、**target** サンプルまたは分析モデルに含まれる EQ-5D QOL 値に影響を及ぼすと考えられるすべての変数を含める必要がある。場合によっては、**target** サンプルが十分に表現されていれば、

推定サンプルにはより広い範囲の人を含めることが適切な場合がある。

- モデル選択と仕様の情報を提供するため、マッピング推定の前にデータを調べるために、標準的な統計的手法を使用する必要がある（例：頻度テーブルと相関）。マッピングの回帰モデルを推定するために使用されるデータセットは、EQ-5D QOL 値の範囲と分布を示すプロットの両方を含めて完全に記述されるべきである。
- データ元の EQ-5D QOL 値の範囲と費用効果分析モデルで使用された EQ-5D QOL 値の予測値は、予測値に外挿，すなわちデータ元の範囲外の QOL 値への予測の適用が含まれているかどうかの情報を提供するために十分に記述しなければならず，外挿は避けるべきである。
- 適切なモデルタイプは，データセットとその適用方法によって異なる。標準的な統計的手法と変数間の臨床的関連の事前知識に基づいた判断がモデルの選択と適用に用いられるべきである（統計的有意性，係数の符号と大きさ，R 二乗と調整済み R 二乗，AIC や BIC など）。モデルの選択と正当性の提示には，サンプルデータセットの特性を使用し，その最終モデルが選択された理由を説明する必要がある。
- マッピングアルゴリズムの統計的特性を明確に記述する必要がある。平均平方二乗誤差あるいは平均二乗誤差を報告する必要がある。また，EQ-5D QOL 値の範囲のサブセット（例：EQ-5D <0, $0 \leq \text{EQ-5D} < 0.25$, $0.25 \leq \text{EQ-5D} < 0.5$, $0.5 \leq \text{EQ-5D} < 0.75$, $0.75 \leq \text{EQ-5D} \leq 1$ ）にわたって誤差を報告し，観測値と予測値のプロットを使用する必要がある。
- モデルを検証する必要がある。理想的には，これは **target** サンプルと同様の外部サンプルを用いて行われる。しかし，これは多くの場合利用可能になることはほとんどないだろう。標本サイズが十分大きい場合，標本をランダムに分割し，推定用のサブ標本と検証用のサブ標本を作成することを推奨する。フルサンプルを使用して最終的なモデル仕様を再推定することができる。
- 関心のある測定値の間に内容の重複がない場合，マッピングは，QOL 値を推定するための関係を適切に規定することができない。これらの状況では，QOL 値を推定する代替方法を検討すべきである。

マッピングの主な利点は，EQ-5D が含まれていない場合でも，治験または観察研究で収集されたアウトカムデータを経済的評価に使用して基本分析を満たすことができることである。しかし，依然としてほとんどの場合，有効性の研究の中で直接収集されたデータを得ることが望ましい。マッピングを使用したほとんどの NICE 提出物は，健康関連 QOL または疾患の重症度の臨床指標を使用して，患者の QOL 値を予測しているが，公表された文献の大部分は，一般的な人口サンプルを用いた測定間のマッピングに焦点を当てている。最も適切なマッピング方法には不確実な部分があるが，この分野では活発な研究が行われている。

以下にマッピングが使用されている旨の記載がある TA の一例を Table 6-1 に示す。後述の Table 9-1 に示されている 2018 年 4 月から 2019 年 3 月の間に発行された TA における SF-36 からのマッピングや疾患特異的な指標からのマッピングの事例となる。各項目の詳細については TA 本文を参照されたい。

Table 6-1. Overview of the use of mapping in TA

TA ID	Technology	Indication	Types of disease	Instrument	Source of utilities
TA537	Ixekizumab	Psoriatic arthritis after inadequate response to DMARDs	乾癬性関節炎	EQ-5D	HAQ, PASI
TA538	Dinutuximab	Neuroblastoma	癌	EQ-5D	SF-36
TA543	Tofacitinib	Psoriatic arthritis after inadequate response to DMARDs	乾癬性関節炎	EQ-5D	HAQ, PASI
TA545	Gemtuzumab ozogamicin with daunorubicin and cytarabine	Untreated de novo CD33-positive acute myeloid leukaemia except acute promyelocytic leukaemia for patients age 15 years and above	癌	EQ-5D	EORTC QLQ-C30
TA554	Tisagenlecleucel	Relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years	癌	EQ-5D, HUI2	CHRI, SF-36
TA563	Abemaciclib	Hormone receptor-positive, human epidermal growth factor receptor 2 (HER2)-negative, locally advanced or metastatic breast cancer in combination with an aromatase inhibitor, as initial endocrine-based therapy. In pre- or perimenopausal women, the endocrine therapy should be combined with a luteinising hormone-releasing hormone (LHRH) agonist	癌	EQ-5D	EORTC QLQ-C30
TA565	Benralizumab	Severe eosinophilic asthma	喘息	EQ-5D	AQLQ
TA571	Brigatinib	ALK-positive advanced non-small-cell lung cancer after crizotinib	癌	EQ-5D	EORTC QLQ-C30

7 TSD11：各健康状態の QOL 値の EQ-5D 以外の方法での算出

Methods Guide は、医療技術の QALY への影響を推定するために、EQ-5D を使用することを推奨している。EQ-5D が利用できないか不適切である場合に、QOL 値を算出する EQ-5D の代わりとなる方法が NICE によって認められている。利用不可能というのは、文献の系統的な検索から確認されるべきである。EQ-5D が利用できない場合でも、この TSD に記載されている代替方法ではなく、6 章で説明したマッピングを使用することができる(13)。EQ-5D が患者グループまたは特定の治療の健康関連 QOL を測定するには不適切であるとの主張は、証拠によって裏付けられなければならない。

EQ-5D 以外の方法で QOL 値を算出するための NICE ガイダンスは、以下のように要約することができる。

- 代替方法の選択に関する議論と証拠の提供
- 評価される健康状態の記述は、検証された患者報告の QOL 値に基づくべきである
- 評価方法は、EQ-5D を評価するために使用されたものと同等でなければならない
- 経済評価の結果に代替方法を用いることによる影響を提供し、可能であれば EQ-5D と比較すべきである

本 TSD では、NICE の推奨事項に準拠しているかどうかを判断するための代替方法が検討されている。QOL 値を生成する別の方法には、他の一般的な選好に基づく尺度、QOL 値の条件別選好に基づく尺度、健康状態の尺度、または自身の健康に関する直接法（TTO 法など）が含まれる。QOL 値および患者自身の健康状態評価に関する標準化された、検証された尺度に基づかない vignette（患者や専門家へのインタビュー）は、NICE Methods Guide を満たさず、限られた役割しか持たないだろう。これらの方法は、検証済みの QOL 値測定に基づく他のデータがない場合にのみ使用すべきである。

QOL 値の検証された尺度から得られた、代替的な選好に基づく全般的尺度および条件別の選好に基づく尺度の QOL 値を使用することができる。スコアリングは、選択に基づく手法を用いて抽出された英国の一般集団値に基づくべきであり、TTO を使用した英国の EQ-5D 評価と同じプロトコルが望ましい。

測定法の開発については、健康状態の記述が QOL 値の検証された尺度に基づいていることや使用された評価方法と EQ-5D に使用されているものとの比較可能性も含めて、十分に説明する必要がある。妥当な経験的証拠は、内容の妥当性、構築の妥当性、応答性および信頼性などの代替手段の特性を詳述すべきである。証拠は、EQ-5D よりもむしろ QOL 値を算出する代替方法を使用することによって得られた QALY への影響を実証すべきである。

TSD には一般的な選好に基づく指標と、条件別選好に基づく指標が評価手法やイギリスでの値の有無と共にまとめられている。各指標の注意すべき点などは TSD 本文を参照されたい。

Table 7-1. Summary of generic preference-based measures of health

<i>Instrument</i>	<i>Dimensions</i>	<i>Severity levels</i>	<i>Health states</i>	<i>UK values</i>	<i>Valuation technique</i>
AQoL	Independent living (self-care, household tasks, mobility) social relationships (intimacy, friendships, family role), physical senses (seeing, hearing, communication), psychological well-being (sleep, anxiety and depression, pain) 自立した生活（セルフケア、家事、移動）社会的関係（親密さ、友人関係、家族の役割）身体的感覚（見る、聞く、コミュニケーション）心理的幸福（睡眠、不安と抑うつ、痛み	4	16.8 million	No	TTO
AQoL 2	Social (including work, family and intimate relationships) independent Living, mental Health, illness (including pain), values and beliefs, sensory items. Obtained from 20 items 社会（仕事、家族、親密な関係を含む）自立した生活、精神的な健康、病気（痛みを含む）、価値観や信念、感覚的な項目。20項目から得られる	4-6 per item	64 billion	No	TTO
EQ-5D	Mobility, self-care, usual activities, pain/discomfort, anxiety/depression モビリティ、セルフケア、通常の活動、痛み・不快感、不安・抑うつ	3	243	Yes	TTO
HUI2	Sensory, mobility, emotion, cognition, self-care, pain (original version contained a fertility dimension for use in a specific patient group) 感覚、運動性、感情、認知、セルフケア、疼痛（オリジナル版では、特定の患者グループで使用するための豊饒度の次元を含む	4-5	8,000	Yes	VAS transformed into SG
HUI3	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain 視覚、聴覚、会話、歩行、手先の器用さ、感情、認知、痛み	5-6	972,000	No	VAS transformed into SG
QWB	Mobility, physical activity, social functioning 27 symptoms/problems モビリティ、身体活動、社会的機能 27 症状・問題点	3 2	945	No	VAS
SF-6D	Physical functioning, role limitation, social functioning, pain, energy, mental Health 身体機能、役割制限、社会的機能、痛み、エネルギー、精神的健康	4-6	18,000 (SF-36) 7,500 (SF-12)	Yes	SG
15D	Mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort/symptoms, depression, distress, vitality, sexual activity 移動性、視覚、聴覚、呼吸、睡眠、食事、会話、排泄、通常の活動、精神機能、不快感・症状、抑うつ、苦痛、活力、性的活動	4-5	31 billion	No	VAS

[NICE DSU TSD11 の Table 1 を改変して掲載]

Table 7-2. Selection of condition-specific preference-based measures of health

<i>Condition</i>	<i>Measure derived from</i>	<i>Dimensions</i>	<i>Severity levels</i>	<i>Health states</i>	<i>UK values</i>	<i>Valuation technique</i>
喘息 ^{(27), (28)}	Asthma Quality of Life Questionnaire	Concern, short of breath, weather and pollution, sleep, Activities	5	3125	Yes	TTO

	(AQLQ)	心配事, 息切れ, 天候や公害, 睡眠, 活動				
癌 ⁽²⁹⁾	EORTC QLQ-C30	Physical functioning, role functioning, pain, emotional functioning, social functioning, fatigue and sleep disturbance, nausea, constipation and diarrhea 身体機能, 役割機能, 疼痛, 情緒機能, 社会機能, 疲労と睡眠障害, 吐き気, 便秘, 下痢	4-5	81920	Yes	TTO
勃起不全 (機能不全) ⁽³⁰⁾	Index of Erectile Function (IIEF)	Ability to attain an erection sufficient for satisfactory sexual performance, ability to maintain an erection sufficient for satisfactory sexual performance 満足な性行為を行うのに十分な勃起能力, 満足な性行為を行うのに十分な勃起の維持能力	5	25	No	TTO
肺癌 ^{(31), (32)}	FACT-L	Physical, social/family, emotional, functional, symptoms - general: symptoms – specific 身体的, 社会的/家族的, 情緒的, 機能的, 症状-全般: 症状-特異的	2	64	Yes	VAS
更年期障害 ⁽³³⁾	Menopause-specific quality of life questionnaire	Hot flushes, aching joints/muscles, anxious/frightened feelings, breast tenderness, bleeding, vaginal dryness, undesirable androgenic signs ほてり, 関節・筋肉の痛み, 不安感・恐怖感, 乳房の圧痛, 出血, 膣の乾燥, 望ましくないアンドロゲン徴候	3-5	6075	Yes	TTO
過活動膀胱 ^{(34), (35)}	Overactive bladder questionnaire (OABq)	Urge to urinate, urine loss, sleep, coping, concern 尿意, 尿切れ, 睡眠, 対処法, 心配事	5	3125	Yes	TTO
パーキンソン病 ⁽³⁶⁾	N/A	Disease severity, proportion of the day with 'off-time' (impact on quality of life due to condition covering domains: social function, ability to carry out daily activities, psychological function) 疾患の重症度, 1日のうち「オフタイム」の割合 (疾患による QOL への影響: 社会的機能, 日常生活遂行能力, 心理的機能の領域を含む)	2-5	10	No	VAS, SG
肺高血圧症 ⁽³⁷⁾	Cambridge Pulmonary Hypertension Outcome Review (CAMPOR)	Social activities, travelling, dependence, communication 社会活動, 旅行, 依存性, コミュニケーション	2-3	36	Yes	TTO
尿失禁 ⁽⁸⁾	The King's Health Questionnaire	Role limitations, physical limitations, social limitations/family life, emotions, sleep/energy 役割の制限, 身体的制限, 社	4	1024	Yes	SG

		会的制限／家庭生活，感情，睡眠／エネルギー				
視力／視覚障害 ^{(38), (39), (40)}	N/A	Physical well-being, independence, social well-being, emotional well-being, self-actualization, planning and organization 身体的な幸福，自立，社会的な幸福，感情的な幸福，自己実現，計画と組織化	5-7	45,360	No	VAS , TTO

[NICE DSU TSD11 の Table 2 を改変して掲載]

8 TSD12 : QOL 値分析にあたっての留意点

治験時データや文献から QOL 値を入手し分析モデルで利用しようとする場合に注意すべき点がいくつか挙げられる。

NICE Methods Guide では介入および疾病領域全体にわたる結果の比較を容易にするため、EQ-5D データを可能な限り優先するとされ、これらのデータが入手不可能であるか不適切であると考えられる場合についての一般的なアドバイスが提供されている。

一方、次のような状況についてのガイダンスは十分ではなく、これらについては主に TSD12 において分析にあたっての留意点として提供されている。

- ・ ベースライン QOL 値
- ・ QOL 値の調整/統合
- ・ QOL 値の不確実性の考慮

以下に、事例の紹介と共に、TSD12 におけるアドバイスを紹介する。

8.1 ベースライン QOL 値

分析モデルにて評価する際には、臨床上のイベントを回避するあるいは特定の健康状態を緩和することによる QALY 利得の点で評価を行う。この時、事象または健康状態に関連する QOL 値に加えて、事象または健康状態を経験しないことに関する QOL 値も必要となる。例えば心血管疾患（CVD）の既往がある患者さんにおいて介入により脳卒中を回避できる可能性がある。この時、脳卒中の回避をモデル化する際には、理想的には脳卒中を経験したコホートからの平均 QOL 値および脳卒中を経験しておらず心血管疾患（CVD）の病歴があるコホートからの平均 QOL 値を必要とする。

条件特有の QOL 値は重要であるが、反事実、すなわち患者が事象または特定の健康状態を回避した場合の適切な QOL 値を得ることは困難かもしれない。これらのデータがない場合に、事象または特定の健康状態を回避したとして、QOL 値を 1 とするのは正しいアプローチではないだろう。ある事象を回避したからといって患者が併存疾患を持ったままである場合や、生涯の分析期間であれば、高齢になることによる合併症の有病率の増加によってベースライン QOL 値データは一定ではないと考えられる。

また、異なるデータソースでの比較では統合に適さない可能性もあるため注意が必要である。例えば臨床試験から得られたデータでは、心血管疾患を有するコホートが 0.75 の平均 EQ-5D スコアであり、異なるデータベースからのデータでは、心血管疾患のないコホートが 0.70 の平均 EQ-5D スコアとなることもあるだろう。このような問題は、様々なソースからのデータを統合するときに発生する傾向があり、次の原因が考えられる。

- ・ 異なる選好に基づく尺度の使用。
- ・ 重症度や病気の持続期間などの状態の定義における相違

・試験への組み入れに使用される厳密な選択/除外基準（例えば、合併症を有する被験者を除外した場合）。

・臨床的イベント後の年齢や時間の違いなどの他の要因

これらを考慮した上で、他の研究からのデータの使用がはっきりと正当化されることは非常に重要である。また不確実性を確率論的分析に組み込み、感度分析を実施する必要があるだろう。

QOL 値を複数のソースから引用している事例を見てみよう。

Table 8-1. QOL 値の直接的な推定値がない事例

TA ID	Year of Publication	Technology	Indication
TA550	2018/12	Vandetanib	Medullary thyroid cancer 甲状腺髄様がん

Table 8-1 の事例では企業は、甲状腺髄様がんに関する QOL 値の直接的な推定値がなく、増悪前の健康状態における QOL 値を他の指標から EQ-5D へのマッピングにより作成し、増悪後の健康状態における QOL 値はメラノーマの試験に関する文献値から入手していた。これに対し、ERG はどちらも同じソースから利用することを推奨し、分化型甲状腺がんの試験に関する文献値を利用した。Committee は分化型甲状腺がんとは甲状腺髄様がんの違いには留意したものの、利用可能なメラノーマの研究は一般化可能性が低いとし、ERG の推定値を受け入れている。

2つのソースがある場合についてはTSD12において以下のような助言がなされている。異なる情報源からのデータの組み合わせはバイアスによる誤差の原因を減らすのに役立つが、極度の注意が必要であり、メタアナリシスに含める臨床的証拠を検討する際に必要とされるのと同じ程度の厳格さが適用されるべきである。

最低限、以下を報告し、比較する必要がある

- ・重症度と期間を含む健康状態の詳細な説明
- ・年齢や性別などの人口統計学的特性
- ・元の研究とレビューの両方で使用された選択/除外基準

条件特有のデータが入手できない場合、一般的な集団からのデータを一連の妥当な値の一端として使用して、感度分析の範囲を作成する必要がある。

8.2 QOL 値の調整/統合

医療における分析モデルは、典型的な患者の臨床パスを反映するため、比較的複

雑になり、治療または介入に伴う合併症または有害事象を代表する複数の健康状態を含むことになる。理想的にはモデルで使用されている健康状態の定義を反映しているような患者からの QOL 値を使用すべきであるが、複合的な健康状態の QOL 値のデータは限定的である。

合併症や有害事象など、複数の健康状態を勘案した複合的な健康状態の QOL 値が利用できない場合、これらは、単一の健康状態のコホートから得られた平均 QOL 値を用いて推定される。しかし、これらの QOL 値を推定するために使用すべき方法において、異なるアプローチは大幅に異なる推定につながる可能性がある。有害事象を考慮するのであれば急性的な事象と慢性的な事象の区別も必要であろうし、有害事象に関連する QOL 値の減少を一定値として扱うことは不適切である可能性もある。

2004 年から 2007 年の間に発表された HTA 報告書に記述されている分析モデルに有害事象を組み込む際の手順をレビューした結果(41)、有害事象はしばしば考慮されておらず、使用された証拠の報告が貧弱であると結論付けている。QOL 値の評価において有害事象を考慮したのであれば、分析モデルにおいてどのように考慮されるのか明示的な報告を含めるべきである。

分析モデルにおいて有害事象が QOL 値に影響を与えるか否か、明確な正当性を提供すべきである。有害事象が QOL 値に影響を及ぼすことがわかっている場合、関連する費用と同様に扱われ、モデルに含まれなければならない、使用されるパラメータ値と方法の詳細な説明を提供すべきである (12)

QOL 値の検討において有害事象を取り扱っている事例を見てみよう。

Table 8-2. QOL 値の検討において有害事象を取り扱っている事例

TA ID	Year of Publication	Technology	Indication
TA546	2018/11	Padeliporfin	Adults with previously untreated, unilateral, low-risk prostate cancer in adults 前立腺がん

Table 8-2 の事例では企業は、文献値に基づいて QOL 値の減少を適用している。NICE DSU guidance では乗法の適用が推奨されている。前立腺がんを有する患者の文献値から腸の問題を有する男性の EQ-5D の値 0.84 と有しない男性の EQ-5D の値 0.94 の比は 0.8936(=0.84/0.94) であり、これを用いて年齢固有の英国人口基準に適用して前立腺がん患者における腸機能障害を有する患者の QOL 値を推定することができる。70 歳男性の QOL 値 0.813 に対して 0.8936 を乗ずることにより 0.727 が得られる。ただし、企業は乗数を適用する代わりにベースライン値 0.88 から 0.72 を減ずることにより腸機能障害の減分 0.16 を算出しており、元の

文献値で報告されている減少分 0.10 に対する過大評価であると指摘されている。

8.3 不確実性の考慮

QOL 値に関する不確実性は、平均のみが報告されていることもあり、医療経済の申請では十分な報告がされていない。NICE Methods Guide を参照するとモデルへの入力の不確実性、すなわちパラメータ精度については、モデルに含まれる各パラメータについて、平均値、平均値周辺の分布、およびソースと根拠を明確に記述する必要がある、と述べられている。また、主要パラメータの異なる推定値の影響は、感度分析（例えば、代替シナリオを含めることなど）に反映されなければならない、入力値は、代替入力値を用いた感度分析によって完全に正当化され、不確実性の影響を評価されなければならない、とされている。QOL 値データ、特に NICE が優先する選好に基づく尺度のデータ（すなわち EQ-5D）は、正規分布ではなく、負の値も含む。健康関連 QOL スコアに大きなばらつきがあり、平均スコアが比較的低い、あるいは高いコホートのデータを用いて、患者レベルのシミュレーションモデルのサンプリングを行う場合には、QOL 値を完全な健康状態からの減少分（すなわち 1 から HSUV を減算）とする必要がある。例えば、英国の EQ-5D 指標では -0.594 を取り得るが、減少分の分布としてベータ分布を考える際に、取り得る範囲は 0-1 であるが、1.594 倍することで減少幅の最大値を 1.594 に制限することが可能である。

分析モデルは、一般に、疾患の進行、有害事象または心臓発作および脳卒中などの別個の事象のような健康状態の変化を描写する複数の健康状態を組み込む。これらの健康状態のそれぞれは、異なる QOL 値を有することができ、これらの値の間の相関は、可能であれば、多変量分布を用いた確率的感度分析（probabilistic sensitivity analyses :PSA）において特徴付けられるべきである。

NICE Methods Guide には、QOL 値または臨床変数から生成されたマッピング関数を使用して、EQ-5D QOL 値のデータを推定できることが記載されている (4), (13)。QOL 値データを使用して機能をマッピングするための方法論に関しては TSD10 にて報告されており、本報告書でも紹介した。分析モデルでマッピングを使用する際には、モデルの不確実性分析を容易にするような出力を報告するようにすべきである。

9 NICE TAにおけるQOL値利用状況

9.1 背景

英国 NICE は、企業が提出した有効性・安全性・費用対効果評価の結果と、それに対する外部の評価グループ（Evidence Review Groups：ERGs）の批判的吟味等を参考に、総合的評価（アプレイザル）を行い、イングランドおよびウェールズの NHS における医療技術の使用を推奨するかについて判断する。この評価プロセスは単一技術評価（Single Technology Appraisal：STA）と呼ばれ、評価結果は NICE のウェブサイト公開されている。

費用効果分析で利用される QOL 値は、その測定尺度やデータ入手元などを考慮し、複数ある候補から選択される。NICE の Methods guide (4)では、費用対効果評価を行うにあたり、EQ-5D の利用を推奨している。

9.2 目的

本研究では、英国 NICE の TA おいて、製薬企業、ERGs がどのように QOL 値を選択し、アプレイザルで議論されたかを調査した。

9.3 方法

NICE のウェブサイト公開された NICE の TA とそのコミッティ・ペーパーを調査対象文書とした。調査対象期間は 2018 年 4 月から 2019 年 3 月の間に発行された TA とした。調査項目は、QOL 値の測定尺度、データ入手元、マッピングの使用を含む QOL 値の利用方法とし、企業と ERGs の提出書類および NICE による最終決定事項それぞれの差異を比較した。除外基準として、1) 複数の医療技術を同時に評価する MTA（Multiple Technology Appraisal）、2) 費用比較（cost-comparison studies）、3) 薬剤以外を評価対象とした TA、4) 製薬企業からの提出書類未提出による未評価の TA は除外した。

Table 9-1. 調査対象 TA と除外基準

分類		TA（数）
対象期間に発行された TA	TA517 – TA572	56
除外基準	複数の医療技術を同時に評価する MTA	3
	費用比較	2
	薬剤以外を対象	1
	提出書類未提出による未評価	7
調査対象 TA		43

9.4 結果

対象期間に発行された 56 TAs（TA517～TA572）のうち、除外基準の 3 MTAs、費用比較

の 2 TAs, 薬剤以外を対象とした 1 TA, 提出書類未提出による未評価 7 TAs を除き, 43 TAs に含まれた 150 の QOL 値について調査した (表 1). 43 TAs のうち, 35 TAs は抗がん剤を対象とした TA であった.

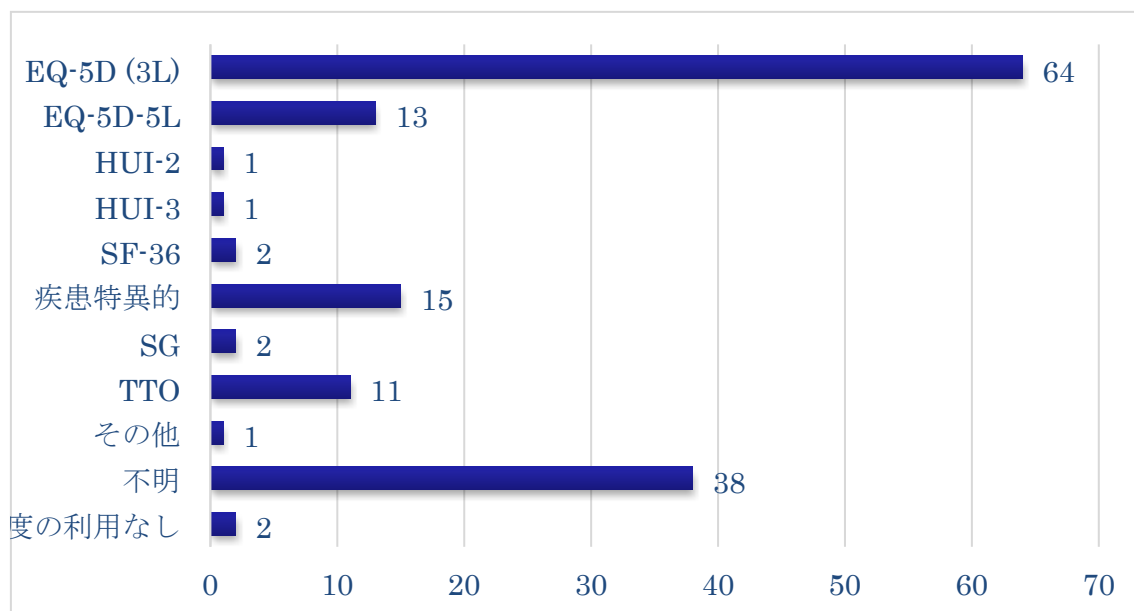


Figure 9-1 QOL 値の測定尺度 (N=150)

企業の提出

企業が提出した QOL 値の測定尺度は EQ-5D (3L) が最も多く 43% (64/150) を占めた. EQ-5D-5L は 9% (13/150) であった. データ入手元については, 臨床試験が最も多く 49% (73/150) を占めた. 過去の NICE TA での使用を含めた文献からの引用が 41% (62/150) とこれに続いた. マッピングは 22% (33/150) で用いられていた.

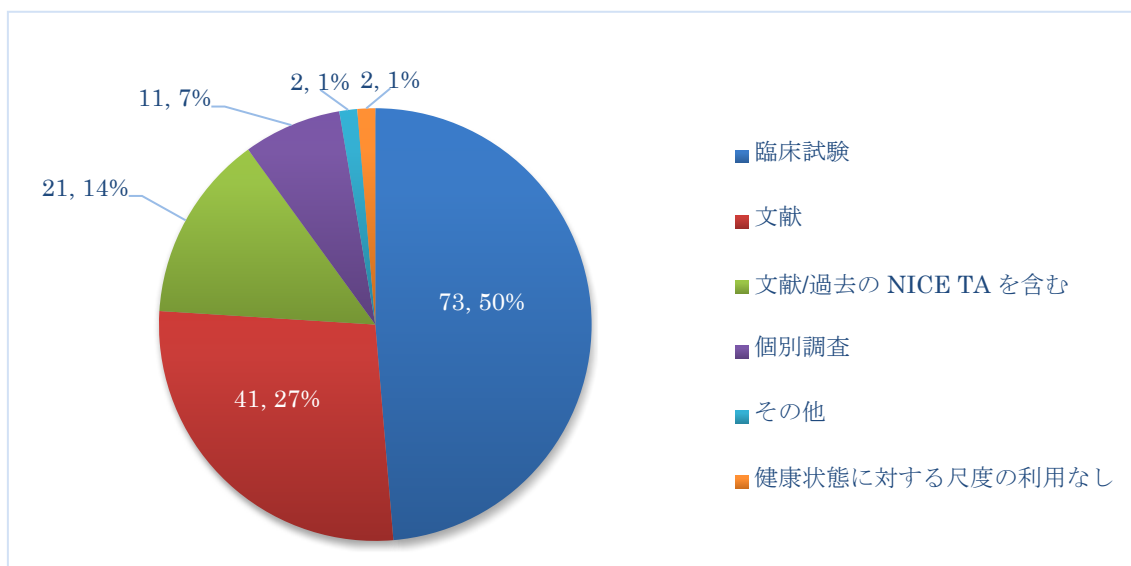


Figure 9-2. QOL 値のデータ入手元 (N=150)

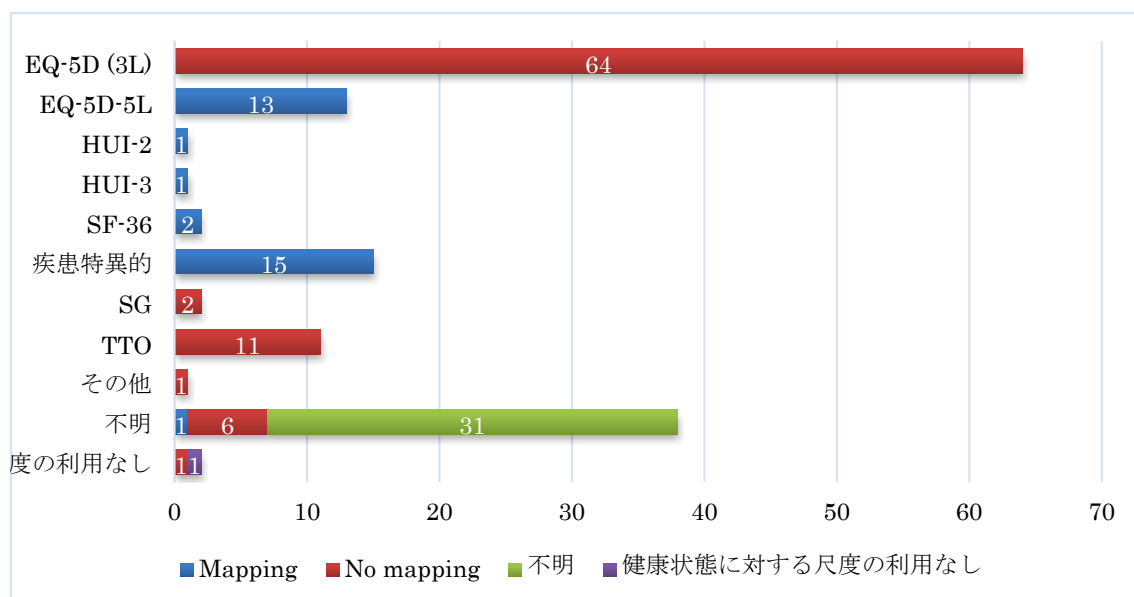


Figure 9-3 マッピング (N=150)

ERG の反応とアブレイザル

ERGs は、企業が選択した QOL 値の 49% (74/150) について同意したが、45% (67/150) に対して批判し、6% (9/150) は明確にならなかった。ERGs が批判した 67 の QOL 値のうち、アブレイザルにおいてその 70% (47/67) は最終的に ERGs が選択した QOL 値が採用された。19% (13/67) は置換されずに企業の意見が採用された。企業が選択した QOL 値のうち、最終的に採用されたのは 58% (87/150) であった。ERG の批判によ

りアプレイザルで置き換えられた QOL 値 44 について、置き換えられた理由は、次の通りであった。より適切な年齢調整方法を用いるため（14，32%）；健康状態の再分類（11，25%）；その他の理由（12，27%）。

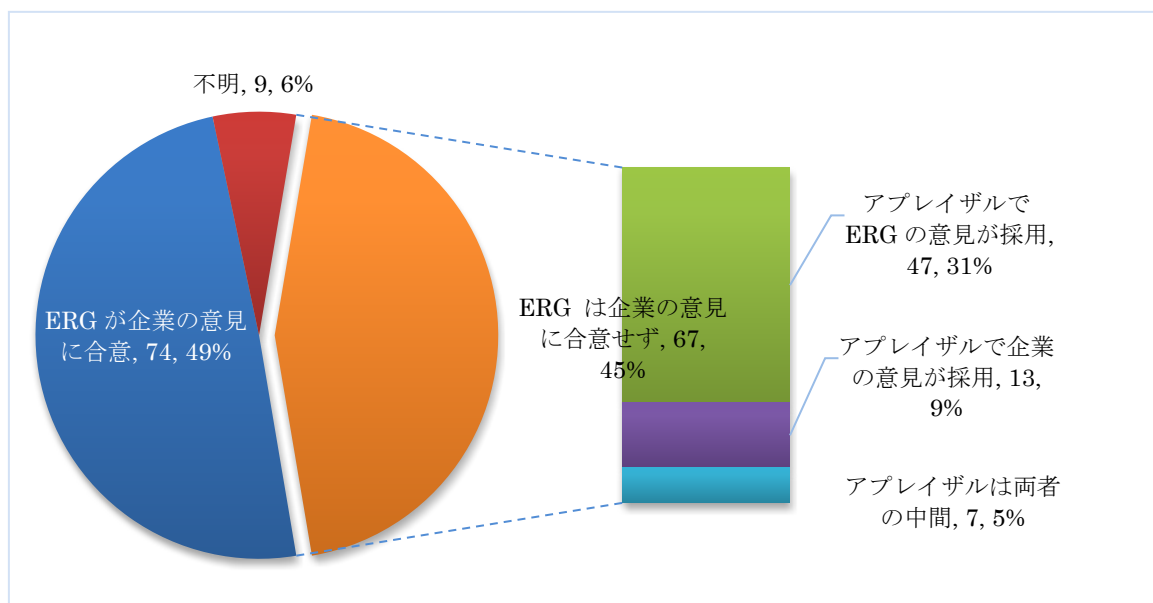


Figure 9-4 ERG が批判した企業提出に対するアプレイザルの結論（N=150）

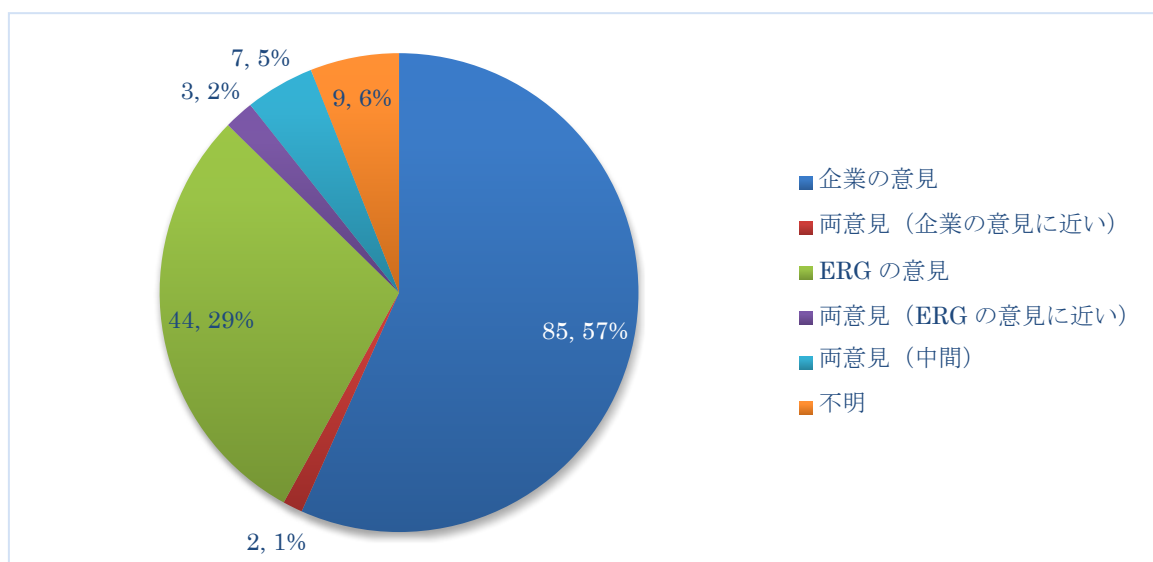


Figure 9-5 アプレイザルの結論（N=150）

9.5 考察

Tosh ら(42)は、2004 年の NICE Methods Guide の基本分析に準拠しているかの観点か

ら TA のレビューを実施した。一方、本研究は、企業と ERG 間の QOL 値の選択に着目し、TA のレビューを実施した。文献情報は元の情報源に遡らなかったため、文献に記載されている詳細な内容にまで十分確認することはできなかった（従って「不明」に分類された数が多くなった）。公開情報のみが使用されていたため、完全には理解されていない情報があった。また、企業、ERG 及び NICE の議論の詳細は、TA 及びコミッティ・ペーパーで十分に記載されていない場合もあり、最終的に判断がどのような意見に基づき行われたのか十分には理解できなかった。

9.6 結論

NICE Methods Guide の基本分析において QOL 値の評価尺度として推奨された EQ - 5D は、企業の提出の中で健康関連 QOL を測定する主要な尺度であった。多くの企業は、NICE の示す基本分析に従った。本研究では、NICE の評価過程において、企業と ERGs 間で QOL 値の選択に実質的な差異があることが分かった。ERGs のレビュー、アプレイザルでの議論という、NICE ガイダンスとして発行されるまでの評価プロセスで、企業が当初提出した QOL 値のうち約半分はそのまま採用された。企業は、評価プロセスの前に、両者意見の不一致を避けるために、事前に評価の範囲（案）について ERGs と更なる話し合いを行う必要があろう。

(ISPOR Europe 2019 で発表したポスター原稿を翻訳)

10 まとめ

本書では、最初に、英国 NICE が発行する医療技術評価の評価ガイドラインである「Guide to the methods of technology appraisal 2013（本文中においては「Methods Guide）」について紹介し、2019 年 4 月より日本で費用対効果評価が制度導入された際の分析ガイドライン「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（第 2 版，本文中においては「中医協ガイドライン）」との違いを比較し、健康関連 QOL の測定尺度について EQ-5D の推奨理由と共に解説した。また、NICE Decision Support Unit から公開されている、経済評価における健康利益の測定や価値づけに関する TSD を 5 報取り上げ、治験で QOL 値を測定していない場合の文献値の利用やマッピングについても紹介した。これらのことから EQ-5D が推奨されているものの、他の方策の利用可能性についてもご理解いただけたと思う。

今回我々は 2018-2019 年の 1 年間の TA とコミッティ・ペーパーを調査し QOL 値の測定尺度や企業と ERG，NICE による最終決定事項との差異を比較した。NICE Decision Support Unit からはこれらの TSD が示されているが、今回の TA 調査によると ERG が企業の選択した QOL 値と同意した割合は 49%のみであった。このような状況に陥らないようにするためには、事前に評価の範囲（案）について ERGs と更なる話し合いを行う必要があるだろう。

NICE の事例を踏まえた上で、中医協ガイドラインを見てみると、新たに QOL 値を収集する際の尺度としては EQ-5D とされているが、測定されていない場合に文献から引用する際の選択方法などについての記載はなく、年齢や合併症等の調整についても記載はない。これらの分析の枠組みも含め、日本においても分析実施中の協議が重要であろう。

なお、データサイエンス部会では引き続き費用対効果評価の手法を適切に理解いただけるよう、他の TSD についても解説を行っているので参考にされたい。

11 付録

日本製薬工業協会. 医薬品の価値の科学的な評価ーデータサイエンス担当者のための費用対効果評価の現状と手法の解説ー. 2016 (再掲) (43)

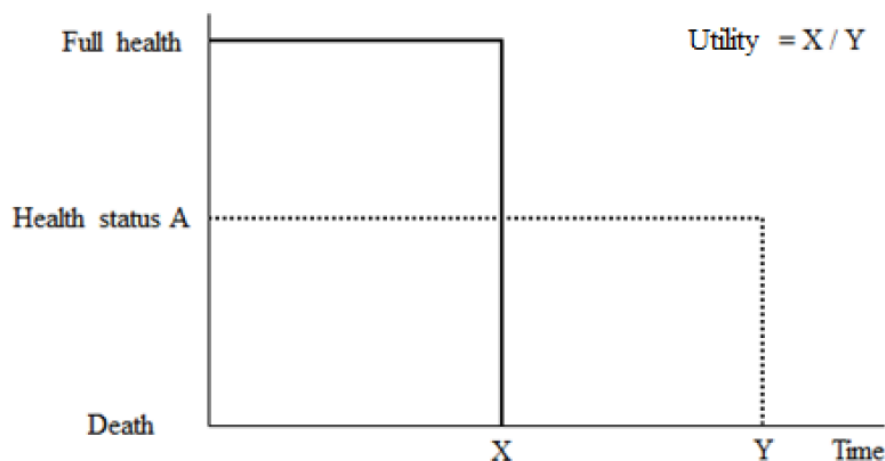
- QOL 値の測定方法 (直接法と間接法)

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」と, QOL 質問表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(44). 前者は, 時間得失法(Time Trade Off: TTO)や基準的賭け法(Standard Gamble: SG)等が該当し, 後者は, EQ-5D や HUI, SF-6D 等が該当する.

これらに該当するデータが存在しない場合, その他の適切な患者報告アウトカム (PRO) から QOL 値へマッピング (mapping) したものを使用できる. ただし, 適切な手法を用いて QOL 値に変換されていることの説明が必要となる. 海外の QOL 値と, 国内で PRO 尺度からマッピングした QOL 値の両方が利用可能である場合, どちらの測定値を使用することが望ましいかは, 今後の検討課題ではあるものの, 状況に応じて判断し, その根拠を明確にする必要がある.

- 時間得失法(Time Trade Off: TTO) (45)

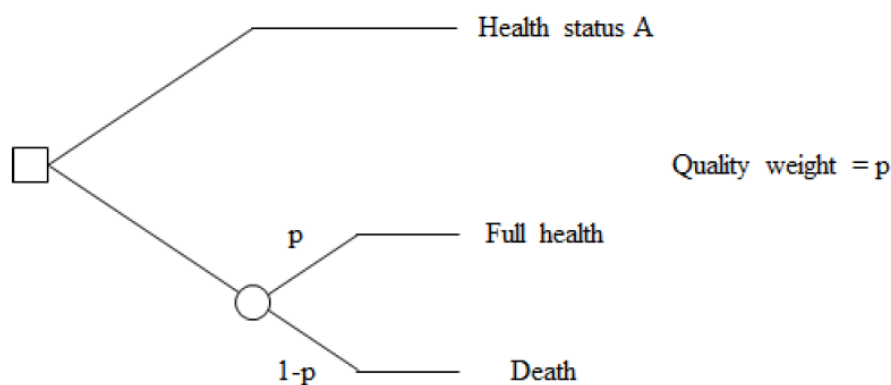
ある健康状態に対して, 生存期間が短くなる代わりに QOL を上げることができて完全な健康になれるとしたらどの程度の生存期間の短縮を容認できるかという質問により QOL 値を測定する方法. もとの状態での生存年数とそれと無差別な完全な健康での生存年数の比をもって, もとの状態の QOL 値とする.



- 基準的賭け法(Standard Gamble: SG)(45)

ある健康状態に対して, 仮想的な治療法を想定し, その治療法によりある確率 (p) で完全な健康になれるとした質問をする方法である. ただし 1-p の確率ですぐに亡くなるとする. そこである健康状態でいることと確率 p で成功する治療法とが無差別になる p を尋ね, こ

の p を QOL 評価値とする。もしある健康状態があまり悪いものでなければ、恐らく p は 1 に近いものが無差別になる、すなわち成功確率が高くなければ治療を受けようと思わないであろうし、逆にある健康状態が極めて悪いものであれば、 p は 0 に近づく、すなわちわずかな成功確率でも治療を受けようと思うであろうということである。



● マッピング (mapping) (46)

先行する研究で、選好にもとづく尺度の QOL 値による測定値 (EQ-5D, SF-6D, HUI 等) が存在しない場合、患者報告アウトカム (PRO) が測定され、線形回帰分析などにより 2 つの尺度間の対応付け (スコア変換) がある程度確立されている場合には、選好にもとづく尺度の QOL 値が存在しない場合でも、PRO の測定結果から、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はマッピング(mapping)といい、他のデータが存在しないときなどに次善の手法として許容されるが、上記の統計学的な妥当性などを十分に検討した上で、実施する必要がある。

12 参考文献

1. FDA. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
2. Agency EM. Regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. 2005.
3. 池上直己ら. 臨床のための QOL ハンドブック. 2001.
4. Guide to the Methods of Technology Appraisal 2013. London: National Institute for Health and Care Excellence (NICE).
5. Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. *Health Econ.* 2002;11(4):341-53.
6. Shiroiwa T, Ikeda S, Noto S, Igarashi A, Fukuda T, Saito S, et al. Comparison of Value Set Based on DCE and/or TTO Data: Scoring for EQ-5D-5L Health States in Japan. *Value Health.* 2016;19(5):648-54.
7. Noto S. Development of a multiplicative, multiattribute utility function and eight single attribute utility functions for the Health Utilities Index Mark 3 in Japan. *Journal of Patient-Reported Outcomes.* 2020;4(12).
8. Brazier J, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. *Med Decis Making.* 2008;28(1):113-26.
9. Shiroiwa T, Moriyama Y, Nakamura-Thomas H, Morikawa M, Fukuda T, Batchelder L, et al. Development of Japanese utility weights for the Adult Social Care Outcomes Toolkit (ASCOT) SCT4. *Qual Life Res.* 2020;29(1):253-63.
10. Shiroiwa T, Fukuda T, Ikeda S, Igarashi A, Noto S, Saito S, et al. Japanese population norms for preference-based measures: EQ-5D-3L, EQ-5D-5L, and SF-6D. *Qual Life Res.* 2016;25(3):707-19.
11. NICE. [Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/Nice-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf.
12. Papaioannou D, Brazier J, Paisley S. NICE Decision Support Unit Technical Support Documents 9: The Identification, Review and Synthesis of Health State Utility Values from the Literature. London: National Institute for Health and Care Excellence (NICE); 2010.
13. Longworth L, Rowen D. NICE Decision Support Unit Technical Support Documents 10: The Use of Mapping Methods to Estimate Health State Utility Values. London: National Institute for Health and Care Excellence (NICE); 2011.
14. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health

state utility values from the literature. *Value Health*. 2013;16(4):686-95.

15. Peasgood T, Herrmann K, Kanis JA, Brazier JE. An updated systematic review of Health State Utility Values for osteoporosis related conditions. *Osteoporos Int*. 2009;20(6):853-68.
16. CEAREgistry. [Available from: <https://cevr.tuftsmedicalcenter.org/databases/cea-registry>.
17. CRDdatabase. [Available from: <https://www.crd.york.ac.uk/CRDWeb/>.
18. HEED. [Available from: <https://health.ebsco.com/products/health-policy-reference-center>.
19. EuroQOL. [Available from: <https://euroqol.org/>.
20. MAPI. [Available from: <https://www.mapi-institute.com/>.
21. RePEc. [Available from: <http://repec.org/>.
22. PROQOLID. [Available from: <https://eprovide.mapi-trust.org/>.
23. QOLdatabase. [Available from: <http://cheqol.com/database/index.php>.
24. 岩下愛, 山下ユミ. 図解 PubMed の使い方-インターネットで医学文献を探す. 紀伊國屋書店; 第 7 版. 2016.
25. NICE-TA-guidance. [Available from: <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance>.
26. Taku K, Yoshida Y, Omori T. Practice guideline of evidence-based medicine : Preferred Reporting Items for Systematic Reviews and Meta-analyses (the PRISMA statement). *Journal of Information Processing and Management*. 2011;54(5):254-66.
27. Young TA, Yang Y, Brazier JE, Tsuchiya A. The use of rasch analysis in reducing a large condition-specific instrument for preference valuation: the case of moving from AQLQ to AQL-5D. *Med Decis Making*. 2011;31(1):195-210.
28. Yang Y, Brazier JE, Tsuchiya A, Young TA. Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the asthma quality of life questionnaire. *Med Decis Making*. 2011;31(2):281-91.
29. Rowen D, Brazier J, Young T, Gaugris S, Craig BM, King MT, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value Health*. 2011;14(5):721-31.
30. Stolk EA, Busschbach JJ. Validity and feasibility of the use of condition-specific outcome measures in economic evaluation. *Qual Life Res*. 2003;12(4):363-71.
31. Kind P, Macran S. Eliciting social preference weights for Functional Assessment of Cancer Therapy-Lung health states. *Pharmacoeconomics*. 2005;23(11):1143-53.
32. Lamers LM, Uyl-de Groot CA, Buijt I. The use of disease-specific outcome measures

in cost-utility analysis: the development of Dutch societal preference weights for the FACT-L scale. *Pharmacoeconomics*. 2007;25(7):591-603.

33. Brazier JE, Roberts J, Platts M, Zoellner YF. Estimating a preference-based index for a menopause specific health quality of life questionnaire. *Health Qual Life Outcomes*. 2005;3:13.

34. Young T, Yang Y, Brazier JE, Tsuchiya A, Coyne K. The first stage of developing preference-based measures: constructing a health-state classification using Rasch analysis. *Qual Life Res*. 2009;18(2):253-65.

35. Yang Y, Brazier J, Tsuchiya A, Coyne K. Estimating a preference-based single index from the Overactive Bladder Questionnaire. *Value Health*. 2009;12(1):159-66.

36. Palmer CS, Schmier JK, Snyder E, Scott B. Patient preferences and utilities for 'off-time' outcomes in the treatment of Parkinson's disease. *Qual Life Res*. 2000;9(7):819-27.

37. McKenna SP, Ratcliffe J, Meads DM, Brazier JE. Development and validation of a preference based measure derived from the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) for use in cost utility analyses. *Health Qual Life Outcomes*. 2008;6:65.

38. Misajon R, Hawthorne G, Richardson J, Barton J, Peacock S, Iezzi A, et al. Vision and quality of life: the development of a utility measure. *Invest Ophthalmol Vis Sci*. 2005;46(11):4007-15.

39. Peacock S, Misajon R, Iezzi A, Richardson J, Hawthorne G, Keeffe J. Vision and quality of life: development of methods for the VisQoL vision-related utility instrument. *Ophthalmic Epidemiol*. 2008;15(4):218-23.

40. Richardson J, Iezzi A, Sinha K, Khan M, Peacock S, Hawthorne G, et al. AQoL-7D (Vision) Instrument: Overview, survey results and utility algorithms.: Centre for Health Economics Research Paper 45, Monash University, Australia; 2009.

41. Drummond MF, Iglesias CP, Cooper NJ. Systematic reviews and economic evaluations conducted for the National Institute for Health and Clinical Excellence in the United Kingdom: a game of two halves? *Int J Technol Assess Health Care*. 2008;24(2):146-50.

42. Tosh JC, Longworth LJ, George E. Utility values in National Institute for Health and Clinical Excellence (NICE) Technology Appraisals. *Value Health*. 2011;14(1):102-9.

43. 日本製薬工業協会. 医薬品の価値の科学的な評価ーデータサイエンス担当者のための費用対効果評価の現状と手法の解説ー. 2016.

44. 福田敬. 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. 2019.

45. 池上 直己 西周. 医療技術・医薬品. 勁草書房. 2005.

46. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross

walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ. 2010;11(2):215-25.

2019 年度データサイエンス部会 KT5

作成担当者

東 美恵	エーザイ株式会社
奥山 ことば	MSD 株式会社*
野島 俊秋	興和株式会社
松下 泰之	第一三共株式会社*

レビュー担当者

荒西 利彦	日本イーライリリー株式会社
渥美 淳	東レ株式会社
河田 祐一	中外製薬株式会社*
中島 章博	帝人ファーマ株式会社
直井 一郎	大日本住友製薬株式会社
町田 光陽	塩野義製薬株式会社
吉田 瑞樹	ファイザーR&D 合同会社

*TF リーダー

担当副部長

酒井 弘憲	エーザイ株式会社
-------	----------