VALUE IN HEALTH 22 (2019) 267–275

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

ISPOR Report

Identification, Review, and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report

John Brazier, PhD [1,*], Roberta Ara, PhD [1], Ismail Azzabi, MSc, MAS [2], Jan Busschbach, PhD [3,4], Hélène Chevrou-Séverac, PhD [5], Bruce Crawford, MA, MPH [6], Luciane Cruz, ScD, PhD [7], John Karnon, MSc, PhD [8], Andrew Lloyd, DPhil [9], Suzy Paisley, MA, PhD [1], A. Simon Pickard, PhD [10]

[1]University of Sheffield, Sheffield, South Yorkshire, UK; [2]Takeda Pharmaceutical International AG, Zurich, Switzerland; [3]Erasmus University Medical Center, Rotterdam, The Netherlands; [4]Viersprong Institute for Studies of Personality Disorders, Halsteren, The Netherlands; [5]Celgene International Sàrl, Boudry, Switzerland; [6]APAC, Syneos Health, Tokyo, Japan; [7]Health Technology Assessment Institute, Federal University of Rio Grande du Sol, Porto Alegre, Brazil; [8]The University of Adelaide, Adelaide, Australia; [9]Acaster Lloyd Consulting Ltd, Oxford, UK; [10]Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA

ABSTRACT

Cost-effectiveness models that present results in terms of cost per quality-adjusted life-year for health technologies are used to inform policy decisions in many parts of the world. Health state utilities (HSUs) are required to calculate the quality-adjusted life-years. Even when clinical studies assessing the effectiveness of health technologies collect data on HSUs to populate a cost-effectiveness model, which rarely happens, analysts typically need to identify at least some additional HSUs from alternative sources. When possible, HSUs are identified by a systematic review of the literature, but, again, this rarely happens. In 2014, ISPOR established a Good Practices for Outcome Research Task Force to address the use of HSUs in cost-effectiveness models. This task force report provides recommendations for researchers who identify, review, and synthesize HSUs for use in cost-effectiveness models; analysts who use the results in models; and reviewers who critically appraise the suitability and validity of the HSUs selected for use in models. The associated Minimum Reporting Standards of Systematic Review of Utilities for Cost-Effectiveness checklist created by the task force provides criteria to judge the appropriateness of the HSUs selected for use in cost-effectiveness models and is suitable for use in different international settings.

Keywords: cost effectiveness, economic evaluation, health state utility, preference-based, quality of life, systematic reviews, utilities

Introduction

Cost-effectiveness models that present results in terms of cost per quality-adjusted life-year (QALY) for health technologies are used to inform policy decisions in many parts of the world. Health state utility (HSU) data are required to calculate QALYs. HSUs describe the value of a health state on a scale where 1 represents full health, 0 represents states deemed to be as bad as being dead, and negative values represent states deemed to be worse than being dead. The preference values are usually obtained by elicitation techniques such as standard gamble (SG) or time trade-off (TTO) from a sample of the general population (although preferences may be elicited from patient populations).

If HSUs are not available from clinical trial data and conducting a study to collect this evidence is not feasible, HSUs are often obtained from the literature. This approach can be problematic because analysts frequently cite outdated evidence used in previous evaluations without undertaking basic quality checks of the data in the original source material (eg, on the relevance of the patient population, utility measure, elicitation method, or sources of the preference weights used). Furthermore, systematic reviews of the literature are rarely undertaken for HSUs, and current reporting standards of HSUs used in cost-effectiveness models are often poor.[1]

Use of different samples, estimation methods, and preference weights can result in different HSUs for the same health state.[2–4]

* *Address correspondence to:* John Brazier, PhD, University of Sheffield, Regent Court, Sheffield, South Yorkshire S1 4DA, UK.
E-mail: j.e.brazier@sheffield.ac.uk

1098-3015/$36.00 - see front matter

https://doi.org/10.1016/j.jval.2019.01.004

Box: Background to the Task Force

The proposal to initiate an ISPOR Good Practices for Outcomes Research task force was evaluated by the ISPOR Health Science Policy Council then recommended to the ISPOR Board of Directors for approval.

The task force was comprised of international subject matter experts representing a diverse range of stakeholder perspectives (academia, research organizations, government, regulatory agencies and commercial entities). The task force met approximately every five weeks by teleconference and in person at ISPOR conferences. All task force members reviewed many drafts of the report and provided frequent feedback in both oral and written comments.

To ensure that ISPOR Good Practices Task Force Reports are consensus reports, findings and recommendations are presented and discussed at ISPOR conferences. In addition, the first and final draft reports are circulated to the task force's review group. All reviewer comments are considered. Comments are addressed as appropriate in subsequent versions of the report. Most are substantive and constructive improving the report

Selecting evidence in an ad hoc manner results in unjustifiable conclusions and raises the possibility of cherry picking. For consistency within a model, utility data on all health states should be informed by evidence obtained from studies that used the same preference-based measure and preference weights, although this may not always be possible.

When multiple appropriate HSUs are available for a particular health state or when it is not possible to identify all HSUs from the same measure, it might be reasonable to synthesize (ie, combine) the data. Furthermore, even the most appropriate HSUs might not exactly match the definitions of the health states within the model. Consequently, analysts frequently adjust the data in some way to account for age, concurrent clinical events, or adverse effects of treatment.[5]

To address these issues, this report provides recommendations for the identification, critical appraisal, and synthesis of HSUs from the literature; minimum reporting standards for these HSUs; and the use of these HSUs in cost-effectiveness models. It is the third ISPOR Good Practices for Outcomes Research Task Force Report on the topic of HSUs. Previous task force reports have addressed the collection of utility data in clinical studies and the derivation of mapping functions to estimate HSUs from non–preference-based outcome measures.[6,7]

Search Methods for Identifying HSUs in the Literature

HSUs are available from a wide range of study designs, including randomized controlled trials, observational studies, and economic evaluations.[8] Guidance on how to search for studies systematically and transparently is useful for informing reviewing in general (eg, which databases to search and how to devise search strategies).[9] Searches for HSUs from studies for cost-effectiveness models need to account for several requirements of the modeling process, including the iterative nature of model development, the scope of HSUs required, and the extensiveness of searches needed.[10]

Iterative Searching

Searches for HSUs are rarely discrete activities at the outset of model development because final requirements for the searches may not be fully defined at that time. Multiple searches are typically conducted iteratively to identify the full scope of evidence required.

Initial scoping searches can inform early conceptualizations of the cost-effectiveness model, and these early versions of the model will clarify the information needed for additional searches. For example, exploratory analysis may show that the model results are sensitive to certain HSUs and insensitive to others. Future searches can then focus on the HSUs that influence the results. Consequently, the modeler and the information specialist should consult each other to inform the evolving direction and scope of the iterative searches.

Iterative searching can combine more traditional in-depth search techniques with more efficient search techniques to explore a wide cross section of potentially relevant evidence. Techniques to increase the efficiency of searching include initial, focused searching to maximize the relevance of the search retrieval (eg, by searching for relevant terms in article titles only), followed by broader searches (eg, by extending the searches to abstracts). Guidance on iterative search techniques has been published by the National Institute for Health and Care Excellence (NICE) Decision Support Unit.[11]

Scope of Searches

The scope of evidence required should include all health states and all aspects of treatment and management that might affect health-related quality of life (HRQOL) or might be affected by the intervention and comparators under consideration within the model. For this reason, the use of multiple keyword search strategies may be required. For example, a cost-effectiveness model for hypercholesterolemia management required HSUs for downstream events, including stable and unstable angina, stroke, and myocardial infarction.[12]

A search approach that systematically addresses the full range of evidence required for the modeling framework is different from standard systematic review search methods. The latter are commonly used in reviews of clinical effectiveness that capture evidence using a single search strategy based on the population and intervention elements of the structured Population, Intervention, Comparator, and Outcome framework. A systematic account of the range of evidence to be retrieved should be determined by the requirements of the decision problem (see Fig. 1). Factors to consider when identifying possible search criteria are presented in Table 1.

Extent of Searching

Exhaustive searching is a fundamental methodological requirement of systematic reviews. Although this approach is recommended for parameters of treatment effect,[13] there is consensus that exhaustive searching for every model parameter is not an efficient use of resources.[14,15]

To the extent possible, the search process should be (1) systematic and (2) explicitly described to demonstrate that evidence has not been identified "serendipitously, opportunistically or preferentially."[11] Recommendations for completing at least a minimum amount of searching for all key model parameters have been published elsewhere.[16] These recommendations emphasize the need to undertake additional searching if required or to provide justification for the amount of searching completed if this process has provided sufficient evidence.

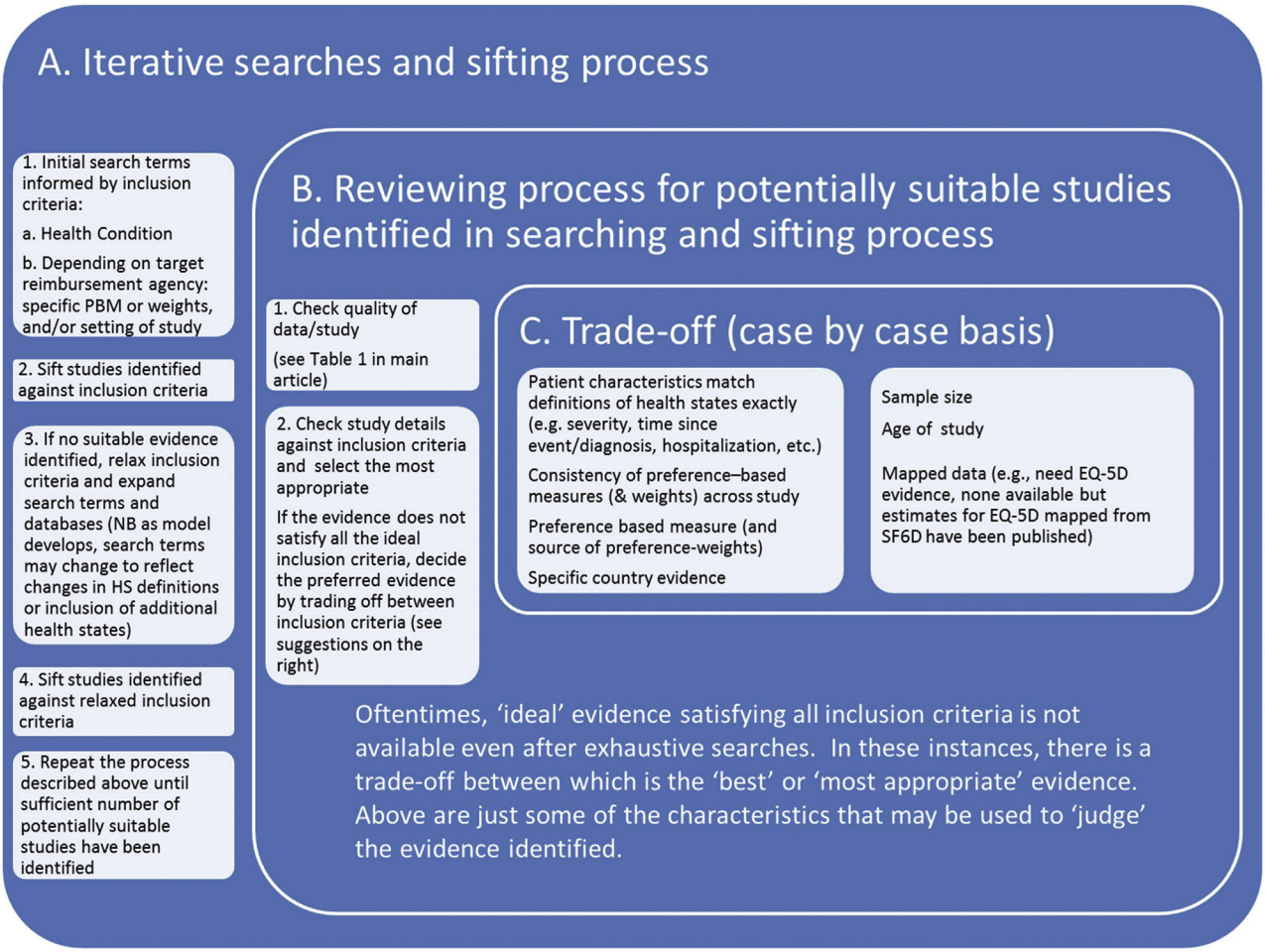

Fig. 1 – HSU identification and selection is often not a straightforward process.

Currently, no empirical definition exists of sufficient evidence or sufficient searching for HSUs to use in a cost-effectiveness model. In the absence of such definitions, the search objective should be "to identify the breadth of information needs relevant to a model and sufficient information such that efforts to identify further evidence would add nothing to the analysis."[17] This concept is useful for heuristic judgments about when to stop searching.[18] The sufficiency checks that should be conducted before concluding the searching process include:

- sensitivity analyses to understand the impact of HSUs on model outputs including cost-effectiveness ratios (search activities can prioritize HSUs to which outputs are most sensitive) and
- the availability of evidence because searching is not of value where there is a lack of appropriate evidence.

Table 1 – Factors to consider when defining search criteria to identify HSUs for a cost-effectiveness model.

Essential factors
• Health state descriptions to come from the cost-effectiveness model
• Treatment effects of interventions and comparators of interest (including utility gains from treatment benefits and utility losses from adverse effects)
• Treatment effects and management at all stages of the clinical pathway included in the model
• Caregiver health state utilities
• Comorbidities
• Concurrent clinical events and sequelae
• General population norms
• Moderator variables (eg, method of administration or treatment setting) that might affect HSUs
Additional factors that may be relevant
• Mapping functions for estimating preference-based utilities from other HRQOL measures or clinical variables
• The context within which the model will be used (eg, geographic location or reimbursement agency criteria)[23]

HRQOL indicates health-related quality of life; HSU, health state utility.

Search Tools

Guidance on how to search for studies to include in systematic reviews of HSUs provides details on how to search general biomedical databases, such as MEDLINE and specialist databases (including the TUFTS Database, ScHARRHUD, and HERC Mapping Database [Health Economics Research Centre, Oxford, UK]).[19] The InterTASC Information Specialist Subgroup Search Filter Resource[20] provides HSU filters for use with databases, such as MEDLINE.[21] Other sources of guidance on searching are useful for adapting the search process for HSU systematic reviews to the specific requirements of cost-effectiveness models.[11,16]

Process for Reviewing HSUs

The next step after the iterative literature searches are complete and articles have been identified that satisfy the inclusion criteria is to review the articles. Initial considerations in this review include the quality and appropriateness of the data in each article. Depending on the target reimbursement agency or audience for the cost-effectiveness model, additional considerations for the review might include the choice of preference-based measure and/or source of preference weights, the study setting, and whether to include evidence from other measures (Fig. 1).

Data Quality

Each study identified during the search process should be reviewed for evidence of methodological flaws, biases, and limitations using the following criteria as a minimum:

- *Precision of the evidence:* The precision of the data will be reflected in the variance of HSUs, which is related to sample size.
- *Response rate:* The generalizability and validity of the evidence may be compromised if a substantial proportion of eligible individuals declined to participate in the study.
- *Loss to follow-up and missing data:* The rates of loss to follow-up may compromise the representativeness of the final sample. The extent of missing data, whether these data are missing at random, and how researchers dealt with these limitations must be reported. This is particularly important for longitudinal data assessments at interim time points.

Data Appropriateness

First and foremost, the study population must be similar to the modeled population. The cost-effectiveness model's health states are often defined in terms of objective clinical measures. HSUs might be needed for health states defined by stage or severity of disease, comorbidities, age, sex, ethnicity, adverse events, or complications and sequelae. For chronic conditions characterized by symptom exacerbations (eg, Crohn disease or gout) or multiple discrete events (eg, transient ischemic attack or asthma attack), HSUs can fluctuate over time. Thus, it is important to consider timing of data collection including how close the timing of the event and data collection were, and whether this interval is likely to result in statistically different HSUs. The use of any medications that are likely to have independent effects (detrimental or beneficial) on HSUs should be considered and may need to be adjusted for in the final HSU estimate used in the cost-effectiveness model.

HSUs are often based on HRQOL measures completed by patients, with the general public providing weights for the measure using such techniques as TTO or SG.[22] In general, when the results of cost-effectiveness models are used to inform decisions about reimbursement or use of new technologies, societal weights are preferred over patients' weights.[23] Nevertheless, some decision makers (eg, the Dental and Pharmaceutical Benefits Agency in Sweden) prefer using weights from patients rather than from the public. In some cases, proxy assessments must be used and condition-specific measures may be acceptable. In addition, some agencies prefer using HSUs from a specific measure (eg, the UK's NICE and the Dutch National Health Care Institute prefer the EuroQol 5-dimensional questionnaire [EQ-5D]).[23]

An important consideration is the extent to which the measure used in a study is valid and sensitive to changes in the domains of health likely to be affected by the condition. For example, when interventions are evaluated for mental conditions that affect self-esteem or social relationships, the measure needs to capture changes in these outcomes.[24] Some measures are insensitive when used to measure the outcomes of certain conditions (eg, the EQ-5D is insensitive when used to measure the outcomes of hearing loss and some visual disorders).[25]

Using a single measure (and the same preference weights) for all HSUs within a model removes variability resulting from different valuation methods, populations, and other aspects of studies included in the systematic review. Using different measures can give spurious results arising from the impact of using different measures rather than differences between the options being evaluated. Different measures will give different results and therefore using different measures in 1 model introduces bias. It is, however, not always possible to identify all the HSUs from just 1 measure. Trade-offs might be needed between the desirable characteristics of the HSUs or of a specific measure and coverage of the most important health states in the model on the basis of their effect on the incremental cost-effectiveness ratio.

The final evidence used may be selected by trading off ideal characteristics (see Fig. 1) that are likely to differ across models. If all the needed HSUs are not available from a common source, consistency of the measure is a priority, subject to the robustness of the data. In some cases, it might be more important to find data on patients with certain characteristics or studies that collected data at certain time points than to find data on the same measure. The reasons for the final set of studies included in the review should be reported and justified using criteria established before the keyword searches. Any suitable alternative HSUs should be considered in sensitivity analyses.

Synthesis of HSUs

The Reasons to Synthesize HSU Data

Multiple published HSUs are often available for a given health state. Use of 1 source per state is not the best approach. The aim of synthesis is to generate a more accurate estimate of the mean HSU and the associated uncertainty by combining estimates from different sources and to improve the generalizability of the findings.

Requirements for HSU Synthesis

Before undertaking a formal synthesis of HSUs, it is important to consider whether enough HSUs are available and whether the studies with HSU data are sufficiently homogeneous for meaningful aggregation. A formal sample size calculation is rarely possible because it depends on too many factors, but a rule of thumb in the clinical context is that "when the sizes of the included studies are moderate or large, there should be at least 6 to 10 studies for a continuous study level variable; and for a (categorical) subgroup variable, each subgroup should have a minimum of 4 studies."[26]

Table 2 – ISPOR HSU Good Practices Task Force Minimum Reporting Standards of Systematic Review of Utilities for Cost-Effectiveness models (ISPOR SpRUCE checklist).

Criteria	Description
Search strategy	
Search terms and scope	Describe the final search strategy and ensure it covers appropriate databases.
Study selection criteria	Describe the criteria used to identify and select studies for the systematic review (eg, study sample, age range, and disease stage or severity).
Review process	
Quality check	Describe the quality criteria used during the review to decide whether to include or exclude studies from the analysis.
Assessment of HSU relevance	Describe the relevance of HSUs to the cost-effectiveness model and the target reimbursement agency if appropriate.
Data extracted and reported	
Population or patient characteristics	Report relevant patient characteristics, such as age, sex, comorbidities, diagnosis, and disease severity.*
Measure used	Provide the name of the measure used in the study.
Preference weights	Describe the technique used to value the health state (eg, TTO or SG) and the country in which the data were collected.
Descriptive statistics about HSUs	Include the mean and variance around all HSUs used in the model.
Response rate for the measure used†	Indicate whether the response rate is likely to jeopardize the validity of the measure.
Extent of missing data or data lost to follow-up†	Report rates of loss to follow-up and of missing data, especially if missing data could threaten the extent to which the HSUs are representative.
Original reference	Cite the original published study for the HSUs and not a previous economic study that used this evidence.
Selection and estimation of final HSUs for the cost-effectiveness model	
Basis for selecting HSUs	Provide the rationale for selecting the HSUs used in the model.
Method used to combine estimates	If HSUs were combined, describe the analytic method (eg, meta-analysis) used to combine them.
Methods used to apply the HSUs in the model	
Actual HSUs used	Report all HSUs used in the model as well as the measure from which the HSUs were calculated.
Adjustments or assumptions	Describe any adjustments or assumptions used about the HSUs in the cost-effectiveness model. Report both the raw and final HSU values used with examples, if required, to clarify the method used to adjust the data.

HSU indicates health state utility; SG, standard gamble; TTO, time trade-off.
* Check the primary source of the data rather than relying on data from a similar economic model without checking the relevance of these data.
† If this information is relevant.

Heterogeneity in the sources of HSUs can be a major challenge to HSU synthesis. Peasgood and Brazier[27] identified considerable variability in HSUs because of differences in measures (eg, EQ-5D vs the 6-dimensional health state short form), valuation method (eg, TTO vs SG), types of anchors used, country where the valuation was done, and who provided the preference weights (eg, patients with the health condition of interest vs general population). When many sources of variability exist, a formal synthesis may not be meaningful.

Role of Mapping in HSU Data Synthesis

Mapping can expand the number of relevant HSUs available for synthesis when the studies found use health or HRQOL measures that do not generate preference-based HSUs or when HSUs come from different preference-based measures or valuation techniques. In both cases, functions might be available to map or crosswalk the measures to a generic preference-based measure, such as the EQ-5D.[28] Nevertheless, mapping functions increase uncertainty and can produce systematic errors in estimates.[6] It is best to use mapping functions when patient-level data are available, although it is possible to map using mean patient HSUs where estimates are for appropriate groups of patients.

Synthesis Methods

Syntheses are conducted to estimate the absolute or relative impact of each health state on the corresponding HSU. Research into the application of synthesis methods for HSUs is at an early stage; therefore, the recommendations made here about which synthesis methods to use are limited.

One broad synthesis approach is to apply strict eligibility criteria to studies included in the analysis to reduce heterogeneity, such as limiting HSUs to those obtained from the same measure and population (eg, from patients with mild, moderate, or severe depression). This approach is appropriate when a sufficient number of HSUs meet the eligibility criteria. For example, Peasgood and Brazier[27] excluded all HSUs not collected with the EQ-5D (because the EQ-5D is NICE's preferred measure) and combined data from 9 studies to estimate mean HSUs. Considerable unexplained heterogeneity in HSUs remained even though the authors used the same measure, and this heterogeneity raised concerns about the usefulness of the estimates for use in cost-effectiveness models.

When not enough studies have used the same measure in a sufficiently homogeneous population, more sophisticated synthesis methods must be used. One such approach is to model the impact of heterogeneity on the HSUs using meta-regression. For example, Bremner et al[29] used a linear mixed-effects model to

estimate coefficients for prostate cancer stage, symptoms, severity, and valuation methods. The authors acknowledged that this approach overestimated HSUs at the lower end (ie, nearer 0) and resulted in HSUs greater than 1 at the upper end. A colorectal cancer study compared a similar linear mixed logit model with a Bayesian logit model–based model. The Bayesian model produced a better fit, although the coefficients had to be transformed for use in a cost-effectiveness model.[30] In both studies, considerable heterogeneity remained, partly because the cost-effectiveness models were limited by the variables in the published studies used and partly because the authors did not have access to individual-level data.

Meta-regression methods require a substantial amount of data to control for the different sources of variation among studies. Methodological research is needed on meta-regression methods for HSU syntheses and the types of settings in which each method is appropriate.

Another underexplored source of variation consists of differences in HSUs by country in which study participants were recruited. This variation could come from differences in patient characteristics that can be controlled for in meta-regression or from differences in country-specific preference weights for such measures as the EQ-5D. Nevertheless, studies in different countries can use the same preference weights (eg, the UK's EQ-5D preference weights in submissions to the US Institute for Clinical and Economic Review).

No standard way exists to adjust the weights of published values, which would require access to individual-level data. In addition, a country-specific effect might arise from the general health of patients or the healthcare system in general. The importance of these sources of variation and how to deal with them need to be further explored in future studies.

Experience using formal synthesis methods is limited for HSUs. For pragmatic reasons, many of the more complex techniques commonly used in the clinical efficacy literature may have little role in HSU synthesis because of the limited number of studies and the high degree of heterogeneity in the valuation methods and patient populations in these studies. Nevertheless, as the literature grows, opportunities to use meta-analysis with HSUs will increase.

Minimum Reporting Standards for Literature Reviews and Modeling Reports

We recognize that extensive documentation may be unrealistic when multiple literature reviews are necessary and cost-effectiveness models encompass multiple conditions and comorbidities. Nevertheless, the fundamental tenets of systematic reviews, such as use of systematic searches, critical appraisal of the literature, and transparent reporting, as described in the ISPOR Consolidated Health Economic Evaluation Reporting Standards report, are critical to the success of the review process.[31] The iterative nature of the search and review process is shown in Figure 1.

Table 2 presents a checklist of criteria for Minimum Reporting Standards of Systematic Review of Utilities for Cost-Effectiveness (SpRUCE) designed to help reviewers determine whether the process used to select HSUs for the cost-effectiveness model was transparent and appropriate. Herein we provide more details on what to report for each section. We would, however, point out that this ISPOR SpRUCE checklist provides only a minimum set of reporting standards for HSUs in models, because a greater level of detail is likely needed to proceed to peer-reviewed publication of a systematic review.[32]

Information on Search Strategies Used

Ideally, the search and selection methods used in the systematic review are described in a protocol before starting the review. Although initial searches may be somewhat cursory, HSUs determined to be important (eg, through sensitivity analysis) require a more comprehensive search strategy to be described.

The report should specify the terms used in searches and the databases that were searched. Additional, nonstandard search strategies (eg, hand searches of the non–peer-reviewed literature or searches for health technology assessment submissions) should be described. The search process and inclusion criteria used can be summarized in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.[33]

Review Process

The process for screening the studies identified in the searches and determining the eligibility of these studies should be described along with the number of reviewers who made these determinations and how disagreements among reviewers were resolved. Studies with measures from which HSUs were calculated that met the inclusion criteria but were not selected after the trade-off process described earlier should be listed.

Data Extraction

After studies that meet the inclusion criteria are identified, data used to assess the appropriateness of the HSUs should be extracted and summarized (Fig. 1). Data extraction could be limited to studies used in the cost-effectiveness model.

The review should describe characteristics of the studies with HSU data used in the model, such as whether the studies were observational or were clinical trials, and study limitations, such as aspects of the design that may result in placebo effects that could inflate the HSUs calculated from the study data. It is important to identify and cite the original source publication, and not rely on secondary references. If the authors of a study have been asked for clarification or original data, this should be documented in the report.

Selection and Estimation of Final HSUs for the Cost-Effectiveness Model

If a review identifies multiple appropriate sources of HSUs, the rationale for the selection of the best evidence or the methods (eg, meta-analysis) used for analyzing the evidence from multiple studies should be reported and justified. Tests for heterogeneity that support the appropriateness of combining the data should be described.

Use of HSUs in Cost-Effectiveness Models

This section describes issues related to the use of HSUs in cost-effectiveness models and associated recommendations (see Table 3).

Discrete Health States Versus Discrete Event Simulation

Cohort-based state transition cost-effectiveness models describe pathways through a set of discrete health states to which HSUs are assigned. Individual-level modeling techniques, such as discrete event simulation, follow individuals and can represent clinical status in more detail that requires the estimation of HSUs as a function of clinical status. When a model structure is

Table 3 – Methodological recommendations for using HSUs in cost-effectiveness models.

Issue	Recommendation
Individual vs function-based HSUs	Decisions should be informed by the relevance of the data and the reliability of the reported analyses.
Comorbidity utility effects and general population norms	If age-specific HSUs are not available, they should be estimated using age-specific population norms. "Condition-free" HSUs should be estimated from the target population.
Treatment-related adverse events	The extent to which the utility effects of important adverse events are captured by the data used to estimate a model's non–adverse-event HSUs should be assessed.
Concurrent clinical events	The multiplicative method should be used to represent the utility effects of multiple concurrent clinical events.
Acute clinical events	In the absence of utility data collected during or immediately after an acute (temporary) event, plausible estimates of the QALY loss per event should be used in sensitivity analyses.
Sensitivity analysis	One-way and multiway sensitivity analyses of HSUs should be used. The appropriate order of HSUs should be maintained in all sensitivity analyses.

HSU indicates health state utility; QALY, quality-adjusted life-year.

conceptualized, the number of discrete health states required to capture changes in clinical status that result in important changes in utility should be carefully considered.[34]

No consensus exists on how to define important changes in HSUs. The definition is likely to vary by health condition, utility measure, incremental cost differences between treatment options, and decision context, but the basis on which important changes in HSUs inform the model structure should be stated explicitly. The number of discrete health states included in the model could lead to a decision to use discrete event simulation.[35] If a simpler model structure is used that does not represent all potentially important HSUs, the potential effects of such omissions should be examined and discussed.

Individual Versus Function-Based HSUs

Individual HSUs can be estimated by analyzing the data for each health state separately, or HSUs can be defined as a function of a relevant measure of clinical status. If both options are available, which data to use in the main (base-case) model analysis should be informed by the relevance of the data (see Synthesis of HSUs section) and the reliability of the analyses (eg, the precision of the mean HSUs and the validity of estimated functions).

Comorbidity Utility Effects and General Population Norms

HSUs should reflect HRQOL effects associated with the condition of interest as well as comorbidities. Comorbidities have real effects on HRQOL and should be represented in HSUs.[36,37]

It is reasonable to assume that mean HSUs represent comorbidity effects at the mean age of the population in studies used to calculate HSUs. Age-specific HSUs should be estimated to reflect age-related comorbidity utility effects, using the appropriate multiplier. For example, if a condition-specific HSU of 0.72 is derived from a study sample with a mean age of 70 years and the general population norm for people of that age is 0.80, the multiplier is $0.72/0.80 = 0.90$. Age-specific HSUs are then estimated for other ages using the multiplier (eg, if the general population norm at age 71 years is 0.79, the age-adjusted, condition-specific HSU at age 71 years is $0.79 \times 0.90 = 0.711$).

If the intervention is prophylactic and suitable data are available, it is preferable to use age-adjusted HSUs for the condition-free health state that are derived from the target population. These HSUs may be lower than those for the general population.[38]

Treatment-Related Adverse Events and Concurrent Conditions

The disutility associated with adverse events reflects the extent to which this information is already captured in the HSUs used for the model's health states. If individuals experiencing adverse events were less likely to provide utility data, the disutilities of adverse events are likely to be underestimated. Alternatively, few data may be available on high-impact, but uncommon, adverse events. In these cases, additional literature should be sought to estimate the disutility of adverse events, although the original HSUs may partially reflect adverse-event effects.

The choice of the adverse events for which HSUs are calculated should be justified by the incidence rates of these events in the treatment groups, their severity and duration, and the expected sensitivity of the cost-effectiveness results to the adverse-event HSUs. A wider range of adverse-event HSUs should be estimated when the expected impact of the HSUs on the cost-effectiveness results is higher. Estimated utility effects of adverse events should reflect the expected duration of their expected quality-of-life effects.

When study participants experience concurrent clinical events related to the condition of interest (eg, when patients with diabetes experience cardiovascular disease and retinopathy), the utility effects of these concurrent clinical events can be assessed by (1) subtracting the sum of the estimated utility decrements for overlapping events from the estimated HSU in the absence of an event (additive method), (2) multiplying the HSU in the absence of an event by the product of the ratios of the HSU for individuals with the clinical events to the HSU for individuals who do not experience the clinical events (multiplicative method), or (3) using the lowest HSU for all the clinical events (minimum method).[37]

A review of 11 studies that used HSUs for individual health conditions to estimate HSUs for concurrent health conditions found that the minimum method overestimated all observed HSUs and the multiplicative method was generally more accurate than the additive method.[37] The reviewers also described regression-based predictions of concurrent utility effects while recognizing the need for further research to validate regression approaches. On the basis of the existing evidence, we recommend the multiplicative method in this situation.

Effects of Acute Clinical Events on Utility

Acute clinical events, such as asthma exacerbations and bone fractures, may be associated with large utility decrements because they result in high levels of pain or discomfort. Nevertheless, respondents rarely complete the HRQOL measure during the period when they experience the effects of such acute events.

The impact of omitting these effects on utility is likely to be the greatest for acute events that occur on a regular basis, such as asthma exacerbations and angina attacks. To inform HSUs for acute events, utility study participants should be asked to complete a measure during or immediately after the events of interest.[7]

To assess the sensitivity of the model outputs to acute event effects on utility, analysts should generate plausible HSUs on the basis of the expected clinical effects of the event by, for example, determining the expected responses to a measure of health such as the EQ-5D. The estimated HSUs can be multiplied by the expected duration of the effects to estimate the QALY loss per acute event; for example, an acute event with an expected utility of 0.7 and an expected duration of 2 weeks would be associated with a QALY loss of $0.7 \times 2/52 = 0.027$ per event. The QALY loss per acute event can be applied to each occurrence of the event in the cost-effectiveness model.

Sensitivity Analysis

Uncertainty around the mean HSUs (including population norms) should be represented by parametric probability distributions.[39] Lower and upper confidence limits can be used in deterministic sensitivity analyses, and random samples can be generated from the distributions for probabilistic sensitivity analyses (PSAs). Uncertainty around HSUs should generally be represented by a standard beta distribution that is bounded by 0 and 1. Nevertheless, alternative lower and upper limits should be used if a negative HSU is possible.[40]

One-way sensitivity analyses should be used to identify the HSUs to which the model results are most sensitive. Relevant multiway sensitivity analyses include combined analyses of all HSUs using combinations of lower and upper HSUs that move the cost-effectiveness results in the same direction (ie, so that all selected HSUs either increase or reduce cost effectiveness).

Ordered HSUs are pairs of HSUs for which it is reasonable to assume that the true expected HSU for 1 state is higher than the true expected HSU for another state (eg, that a prediabetes HSU is higher than a diabetes HSU). In PSA, inconsistent HSUs can be sampled if probability distributions for ordered HSUs overlap (eg, in any iteration of the PSA, a higher utility value could be sampled for a prediabetes state than for a diabetes state).

To avoid sampling inconsistent HSUs, the difference method should be used.[41] This method involves generating a probability distribution of the difference in the HSUs of 2 ordered parameters. In PSA, one of the ordered parameters is sampled, and the difference between the 2 HSUs is then added to the sampled value to generate the second HSU.

Conclusions

This report provides guidance for identifying, reviewing, and synthesizing HSUs from the literature and using HSUs in cost-effectiveness models. In the past, analysts have paid insufficient attention to this parameter, often simply using evidence from previous models or from a known source without justifying this choice. Although the time and resources available for populating cost-effectiveness models are always limited, the HSUs can be just as important as other parameters used.

Comprehensive literature searches of HSUs are not always feasible or necessary, but the search methods used to identify studies and the criteria used to choose studies to include must be described in a report. The processes for searching and reviewing are iterative because the scope of the searches depends on the literature available. It may be necessary to broaden the search terms and inclusion criteria to identify a larger amount of appropriate evidence.

Any review criteria used should be identified a priori because trade-offs must often be made between the criteria considered. If some of the HSUs needed are not available from a single source, calculating HSUs using the same measure from different studies would be appropriate. If some HSUs are not available from the same measure (from any study) in the same patient populations as those in the cost-effectiveness model, then it must be decided whether using the same measure in the wrong population is more important than using a different measure but in the right patient population. The decision will be context-specific and will need to be justified.

The literature should be searched and reviewed as part of the cost-effectiveness model development. The results can influence the structure of the model. At the same time, the sensitivity of the incremental cost-effectiveness ratio can inform which searches are needed to calculate the most influential HSUs.

The evidence base for HSUs is increasing over time, and so there will be more utility values available per state. Analysts should consider using meta-analyses to generate representative estimates (as for all other model parameters) or meta-regressions to use the full range of evidence from heterogeneous studies.

Although the literature from which HSUs can be calculated is growing, many gaps in the evidence remain. Analysts frequently adjust HSUs for adverse events, comorbidities, and age not accounted for in the published evidence as a way to fill these gaps. Analysts should report any gaps with the evidence sources, methods used to adjust the data when appropriate, and HSUs used in the model to enable readers to review the implications of the decisions made. The uncertainty in the HSUs should be captured appropriately in the report of the cost-effectiveness model.

These task force recommendations and the ISPOR SpRUCE checklist offer a structured and more transparent basis for identifying and reporting the HSUs used in a cost-effectiveness model.

Acknowledgments

We thank all those who commented orally during 4 task force workshop and forum presentations at ISPOR conferences in the United States and Europe in 2017 and 2018. We also gratefully acknowledge the following 34 reviewers who generously contributed their time and expertise through submission of written comments on the task force report: Shohei Akita, Nermeen Ashoush, Suresh Bastakoti, Jenny Berg, Evelina Bertranou, Tânia Maria Beume, Kathleen Beustrien, Alan Brennan, Mike Colopy, Warren Cowell, Victor Dansk, Inigo Gorostiza, Parul Gupta, Jennifer Hinkel, Christine Huttin, Rosanne Janssens, Mehdi Javanbakht, Cicely Kerr, Samer Kharroubi, Stella Lartey, Karen Lee, Ka Keat Lim, Cathrine Mihalopoulos, Nicholas Mitsakakis, Guilhem Pietri, Julian Prescod, Gerry Richardson, Jing Shen, Gemma Shields, Daniel Touchette, Luke Vale, Allan Wailoo, Leslie Wilson, and Feng Xie.

Source of financial support: None of the authors received financial support for their participation in this task force. All authors volunteered their time for discussion and writing of this report.

REFERENCES

1. Ara R, Brazier J, Lloyd A, Chevrou-Severac H. How health state utilities used in cost-effectiveness models are currently identified, reviewed and reported. Value Outcomes Spotlight 2018;4(5):31–3.
2. Janssen B, Szende A. Population norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, editors. Self-Reported Population Health: An

International Perspective Based on EQ-5D. Dordrecht, The Netherlands: Springer; 2014.
3. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ 2004;13(9):873–84.
4. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility measure produce different utilities: the relative importance of the descriptive systems, scale and "micro-utility" effects. Qual Life Res 2015;24:2045–53.
5. 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.
6. Wailoo AJ, Hernandez-Alava M, Manca A, et al. Mapping to estimate health-state utility from non–preference-based outcome measures: an ISPOR Good Practices for Outcomes Research Task Force Report. Value Health 2017;20(1):18–27.
7. Wolowacz SE, Briggs A, Belozeroff V, et al. Estimating health-state utility for economic models in clinical studies: an ISPOR Good Research Practices Task Force Report. Value Health 2016;19(6):704–19.
8. Paisley S. Classification of evidence in decision-analytic models of cost-effectiveness: a content analysis of published reports. Int J Technol Assess Health Care 2010;26(4):1–5.
9. Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0. http://handbook.cochrane.org. Accessed November 12, 2017.
10. Paisley S. Identifying Evidence to Inform Decision-Analytic Models of Cost-Effectiveness: A Qualitative Study of Information Seeking Processes and Behaviour [PhD thesis]. Sheffield, UK: University of Sheffield; 2012.
11. Kaltenthaler E, Tappenden P, Paisley S, Squires H. Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. NICE DSU Technical Support Document 13. http://www.nicedsu.org.uk. Accessed October 27, 2018.
12. Ara R, Tumur I, Pandor A, et al. Ezetimibe for the treatment of hypercholesterolaemia: a systematic review and economic evaluation. Health Technol Assess 2008;12(21):1–212.
13. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision-analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value Health 2003;6(1):9–17.
14. Cooper NJ, Sutton AJ, Ades AE, Paisley S, Jones DR. Use of evidence in economic decision models: practical issues and methodological challenges. Health Econ 2007;16(12):1277–86.
15. Golder S, Glanville J, Ginnelly L. Populating decision-analytic models: the feasibility and efficiency of database searching for individual parameters. Int J Technol Assess Health Care 2005;21(3):305–11.
16. Paisley S. Identification of evidence for key parameters in decision-analytic models of cost effectiveness: a description of sources and a recommended minimum search requirement. Pharmacoeconomics 2016;34(6):597–608.
17. Glanville J, Paisley S. Searching for evidence for cost-effectiveness decisions. In: Shemilt I, Mugford M, Vale L, Donaldson C, Marsh K, editors. Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice. 2nd ed. Oxford, UK: Wiley-Blackwell; 2010.
18. Fenwick E, Knies S, Rothery C, Koffijberg E. Value of information analysis for research decisions: emerging good practice recommendations from the ISPOR VOI Task Force. Paper presented at: ISPOR 20th Annual European Congress; November 4-8, 2017; Glasgow, Scotland. https://www.ispor.org/docs/default-source/task-forces/value-of-information-workshop-glasgow-2017.pdf?sfvrsn=88157374_2. Accessed October 7, 2018.
19. Papaioannou D, Brazier J, Paisley S. The identification, review and synthesis of health state utility values from the literature. NICE DSU Technical Support Document 9. http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD9-HSUV-values_FINAL.pdf. Accessed October 7, 2018.
20. InterTASC Information Specialists' Subgroup. ISSG search filter resource. https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home. Accessed November 9, 2018.
21. Arber M, Garcia S, Veale T, et al. Performance of Ovid Medline search filters to identify health state utility studies. Int J Technol Assess Health Care 2017;33(4):472–80.
22. Brazier JE, Longworth L. An introduction to the measurement and valuation of health for NICE submissions. NICE DSU Technical Support Document 8. http://www.nicedsu.org.uk. Accessed April 25, 2018.
23. Rowen D, Zouraq IA, Chevrou-Severac H, van Hout B. International regulations and recommendations for utility data for health technology assessment. Pharmacoeconomics 2017;35(1):11–9.
24. Brazier JE, Connell J, Papaioannou D, et al. Validating generic preference-based measures of health in mental health populations and estimating mapping functions for widely used specific measures. Health Technol Assess 2014;18(34):1–88.
25. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: systematic review, statistical modelling and survey. Health Technol Assess 2014;18(9):1–224.
26. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol 2011;64(11):1187–97.
27. Peasgood T, Brazier J. Is meta-analysis for utility values appropriate given the potential impact different elicitation methods have on values? Pharmacoeconomics 2015;33:1101–5.
28. 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.
29. Bremner KE, Chong CA, Tomlinson G, Alibhai SM, Krahn MD. A review and meta-analysis of prostate cancer utilities. Med Decis Making 2007;27(3):288–98.
30. Djalalov S, Rabeneck L, Tomlinson G, et al. A review and meta-analysis of colorectal cancer utilities. Med Decis Making 2014;34(6):809–18.
31. Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—explanation and elaboration: a report of the ISPOR Health Economic Evaluations Publication Guidelines Good Reporting Practices Task Force. Value Health 2013;16:231–50.
32. Santos J, Palumbo F, Molsen-David E, et al. ISPOR Code of Ethics 2017 (4th ed.). Value Health 2017;20:1227–42.
33. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009;62:e1–34.
34. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Value Health 2012;15:804–11.
35. Karnon J, Afzali HH. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. Pharmacoeconomics 2014;32(6):547–58.
36. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health 2010;13(5):509–18.
37. Ara R, Wailoo AJ. Estimating health state utility values for joint health conditions: a conceptual review and critique of the current evidence. Med Decis Making 2013;33(2):139–53.
38. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. Value Health 2011;14(4):539–45.
39. Briggs AH, Weinstein MC, Fenwick E, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value Health 2012;15:835–42.
40. Filliben J. Exploratory data analysis. In: Croarkin C, Tobias P, editors. Engineering Statistics Handbook. Gaithersburg, MD: National Institute of Standards and Technology and SEMATECH; 2012. https://www.itl.nist.gov/div898/handbook/eda/section3/eda366h.htm. Accessed November 7, 2018.
41. Ren S, Minton J, Whyte S, Latimer NR, Stevenson M. A new approach for sampling ordered parameters in probabilistic sensitivity analysis. Pharmacoeconomics 2017;28:1–7.