
Co-chairs:
Andrew Briggs, DPhil, MSc, Professor of Health Economics, University of Glasgow, Glasgow, UK
Sorrel E. Wolowacz, PhD, Head European Health Economics, RTI Health Solutions, Manchester, UK

Leadership Group:
Vasily Belozeroff, Health Economics Director, Global Health Economics, Amgen Inc., Thousand Oaks, CA, USA
Philip Clarke, PhD, MEd, Associate Professor of Health Economics, Centre for Health Policy, Programs and Economics, University of Melbourne, Melbourne, Australia
Lynda Doward, MRes, European Head, Patient-Reported Outcomes, RTI Health Solutions, Manchester, UK
Ron Goeree, BA, MA, Director, PATH Research Institute, St. Joseph's Hospital and Professor, Department CE&B, McMaster University, Hamilton, Ontario, Canada
Andrew Lloyd, PhD, DPhil, Director, Bladon Associates Ltd., Oxford, England, UK
Richard Norman, PhD, MSc, BA, Senior Research Fellow, School of Public Health, Curtin University, West Australia, Australia

YOUR feedback is very important. Every comment is read and discussed.
### 1 Introduction

Health-state utility values (HSUVs) are estimates of the preference for a given state of health on a cardinal numeric scale, where a value of 1.0 represents full health, 0.0 represents dead, and negative values represent states worse than death (Lenert and Kaplan, 2000, p 138; Feeny, 2000, p 151). HSUVs are used in cost-utility analysis, a special case of cost-effectiveness analysis in which health benefits are usually measured in terms of quality-adjusted life-years (QALYs) (Robinson, 1993, p 859). QALYs are calculated by multiplying the number of years lived in each state of health by the HSUV for each respective state (Kind et al., 2009, s 27). For example, if an intervention confers 2 extra years of life at an HSUV of 0.75, then it confers an additional 1.5 QALYs ($2 \times 0.75$) to the patient.

#### Table 1. Definition of terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRQL</td>
<td>Health-related quality of life</td>
</tr>
<tr>
<td></td>
<td>A multidimensional concept that represents the patient’s subjective, general perception of the impact of disease and its consequent therapy on daily life, physical, psychological and social functioning, and well-being (European Medicines Agency, 2005; US Food and Drug Administration, 2009; Schipper et al., 1996, p 11)</td>
</tr>
<tr>
<td>HSUV</td>
<td>Health-state utility value</td>
</tr>
<tr>
<td></td>
<td>The Estimate of the health utility value for a given economic model health state</td>
</tr>
<tr>
<td>MAUI</td>
<td>Multi-attribute utility instrument</td>
</tr>
<tr>
<td></td>
<td>An instrument for measuring health utility consisting of two components:</td>
</tr>
<tr>
<td></td>
<td>1. A standardized descriptive system for health or its impact on HRQL, composed of a number of multi-level dimensions that together describe a universe of health states; and</td>
</tr>
<tr>
<td></td>
<td>2. An algorithm for assigning values to each health state described by the system (often described as a value set)</td>
</tr>
<tr>
<td></td>
<td>Algorithms have been based on a variety of valuation methods. e.g., time trade-off, standard gamble, and visual analogue scales</td>
</tr>
<tr>
<td>Health Utility</td>
<td>A representation of strength of preference for a given health-related outcome on a cardinal numeric scale, where a value of 1.0 represents full health, 0.0 represents dead, and negative values represent states worse than death (Lenert and Kaplan, 2000, p 138; Feeny, 2000, p 151)</td>
</tr>
<tr>
<td>Clinical Study</td>
<td>For the purposes of this guideline, a clinical study is defined as a clinical trial investigating one or more interventions or an observational study (prospective or retrospective) in a routine clinical practice setting</td>
</tr>
</tbody>
</table>

US = United States.
1.1 The importance of health-state utility values for economic models

Cost-utility analyses, most often performed in economic models, are increasingly being used by health technology assessment (HTA), pricing, and reimbursement authorities in many countries to establish whether the increased cost of a new intervention can be justified in terms of expected health benefits. The decisions made by these authorities affect patients’ access to treatments, physicians’ ability to use them, the product’s price, and in turn, the return that manufacturers are able to realize on their investment in developing the product.

HSUVs are among the most important and uncertain data inputs in cost-utility models, and the accuracy and precision of the model results often depend heavily on the quality of the HSUV estimates. Since some HTA agencies expressed a preference for health-utility data collected from patients (National Institute for Health and Care Excellence, 2008; Pharmaceutical Benefits Advisory Committee, 2013), it has become increasingly common for such data to be collected in clinical trials. Clinical trials represent an important opportunity for collection of health utility data; however, trial design is devised primarily to evaluate efficacy and safety for market authorization by regulatory agencies and often presents challenges to the optimal collection of HSUV estimates for economic models. Careful planning is needed to define the HSUVs that will be needed for cost-utility analyses, determine which of the HSUVs may be measured in the planned trial or trials, establish the optimal health-utility instrument, timing and frequency of assessments, period of follow-up, and data analyses, as well as to plan additional studies to collect HSUVs that may not be adequately estimated within the trials.

1.2 Aim and scope

This guideline aims to provide a framework for researchers to plan the collection of health utility data in clinical studies in order to provide high quality HSUV estimates appropriate for economic modeling. A variety of study designs (Box 1) and measures (Box 2) have been used to estimate HSUVs. The focus of this guideline is on collecting data primarily within clinical trials because clinical trials often represent the best opportunity to collect high-quality data for a large patient sample; however other study types also are considered more briefly.

Box 1. Study design options for collection of HSUVs for economic models

- **Review of published estimates.** A review of existing estimates available in the published literature is an important starting point, to determine whether high-quality estimates relevant to the economic model are available and whether new research is needed. Such a review also can establish the range of available estimates that should be explored in sensitivity analyses. Formal systematic reviews should be conducted, to maximize the generalizability and representativeness of the values used in any economic model (Pappaioannou et al., 2013). Critical appraisal of studies should examine three broad issues: the relevance of the health utility data to the health states in the economic model, the potential sources of bias in a study design, and the extent to which data meet the needs of decision makers. For example, the UK’s National Institute for Health and Care Excellence prefers values estimated using the EQ-5D instrument that reflect the preferences of people living in England and Wales (National Institute for Health and Care Excellence, 2013). Pooling of HSUVs may be performed to improve the precision of the estimates (and estimates of uncertainty), provided the populations are sufficiently homogenous and the same method of elicitation and instrument was used (Pappaioannou et al., 2011; Peasgood and Brazier,
Prospective data collection in clinical trials. In most cases, with careful consideration of the optimum methodology, clinical trials represent an important opportunity and an efficient way to collect HSUVs. However, careful design of the health utility data collection is critical to the quality and usefulness of the data to the economic model. Additional work may be needed to provide a complete set of values for all the modeled health states. In particular, it is important to time the collection of data so that it captures profiles of change around the relevant health states. These issues are discussed in more detail in Section 5.1. The incremental cost of adding a utility measure to a regulatory trial is expected to be substantially less than the cost of performing a separate bespoke utility study.

Prospective longitudinal or cross-sectional observational studies. These studies may offer the greatest flexibility in terms of the data that can be collected. However, they may be time consuming and costly to perform and may not produce measures that are good proxies for measures of change associated with the occurrence of key health states. These studies are discussed in Section 5.2.

Early-access or compassionate-use-type programs, phase 4 studies, registries, and other post-licensing commitments. These types of studies can provide opportunities for including health utility assessment and they can be an efficient way to capture HSU data. Such studies may be performed too late in the product development program to provide HSUV estimates for HTA submissions of new technologies. However, data collected in these studies may be useful for re-assessments and ongoing cost-utility research for products.

Vignette studies. In this approach, detailed descriptions of each health state are developed from different sources of information (e.g., patient and physician interviews, trial data, literature review) (Lloyd et al., 2006; Lenert et al., 2004; Brazier and Rowen, 2011). Members of the public are asked to rate these states in a stated preference experiment (such as time trade-off or standard gamble). These methods are limited because the resulting values are entirely dependent on the validity of the vignette descriptions, do not reflect the variability in HRQL among patients within the health state, and do not offer the opportunity for patients in the health state to describe their own health (as is the case if a MAUI is used).

HRQL = health-related quality of life; HSUV = health-state utility value; HTA = health technology assessment; UK = United Kingdom.
Box 2. Measures for the estimation of HSUVs for economic models

- **Generic preference-based MAUIs.** Many HTA authorities require or prefer HSUVs estimated using a generic preference-based MAUI and a value set algorithm developed using data from the general population in their country. Examples of generic (condition-nonspecific) MAUIs include EQ-5D, HUI, 15D, AQoL, and SF-6D. These instruments differ considerably in the content and size of their descriptive system, the methods of valuation, and the populations used to values the health states (Brazier and Rowen, 2011).

- **Condition-specific, preference-based MAUIs.** These measures are similar in concept to the generic preference-based MAUIs but have been developed for a specific disease or condition. Examples include the AQL-5D for asthma (Young et al., 2011) and EORTC-8D for cancer (Rowen et al., 2011); other examples are summarized by Brazier and Rowen, (2011). Many condition-specific measures describe symptoms or symptom impact rather than HRQL (Brazier and Rowen, 2011).

- **Mapping of a condition-specific, patient-reported outcome measure onto a generic preference-based measure.** Mapping or “cross-walking” involves estimating a relationship between a condition-specific, patient-reported outcome measure and a generic, preference-based measure using a statistical model and making predictions from the estimates (Wailoo et al., 2015). These methods can be valuable in instances where there is no health-utility measure available but where some form of patient-reported outcome or clinical endpoint can be used as a basis for predicting HSUVs. Mapping also may be performed from one MAUI to another.

- **Direct valuation of patients’ own health state using preference-based methods such as time trade off or standard gamble.** This may be appropriate if no MAUI is available which is valid and responsive in the condition of interest and meets the needs of decision makers, and no means of mapping from a condition-specific measure to such a MAUI is available. Guidance on performing time trade-off studies is available in the published literature (Oppe et al., 2014, p 445; Norman et al., 2010, p 499). While this approach provides direct observations of health utility, there are technical and ethical obstacles to performing time trade-off and standard gamble experiments with patients. Further, the data represent patients’ values rather than those of the general population, which is the preference of many HTA authorities (Brazier and Rowen, 2011).

---

Note: the ordering of methods in this box is not intended to convey any order of preference; preferences for alternative methodologies vary among individual HTA authorities and other audiences. AQL-5D = Asthma Quality of Life Utility Index; AQoL = Assessment of Quality of Life; EORTC-8D = European Organization for Research and Treatment of Cancer cancer-specific instrument; HRQL = health-related quality of life; HSUV = health-state utility value; HTA = health technology assessment; HUI = Health Utility Index; MAUI = multi-attribute utility instrument; SF-6D = SF-6D Health Survey.

The guideline includes considerations for the following:

- Early planning of health-utility data collection within the research and development program
- Design of health-utility data collection during protocol development for a planned clinical trial
- Design of prospective and cross-sectional observational studies for estimating health utility
- Use of alternative study types for estimation of HSUVs
- Statistical analyses and reporting to maximize the value of patient-level health-utility data for economic models

Recommendations are discussed in sections 3 through 6 and are summarized in Boxes 3 – 6 and Figure 1. The focus of the guideline is on estimating HSUVs for cost-utility models. We do not consider direct...
longitudinal measurement and analysis to perform statistical comparisons of health utility between treatment arms, or for QALY estimation by treatment arm for cost-effectiveness analysis alongside clinical trials; a separate guideline from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) is available for this type of analysis (Ramsey et al., 2005, p 521). We focus on collection of HSUVs via preference-based measures in patients (particularly validated MAUIs with preference-based value sets) because these methods are favored by many HTA agencies at the time of writing (e.g. Canadian Agency for Drugs and Technologies in Health, 2006; Capri et al., 2001, p 189; Haute Autorité de Santé, 2012; Kennedy-Martin et al., 2014, p 108; Lopez-Bastida et al., 2010, p 513; National Institute for Health and Care Excellence, 2013; Pharmaceutical Benefits Advisory Committee, 2013; and other summarized in International Society for Pharmacoeconomics and Outcomes Research, 2011a-c). Other methods of estimation, including direct valuation of patients’ own health state and health-state descriptions (vignettes) valued by members of the general population, using methods such as standard gamble or time trade-off, are discussed briefly. Discussion of utility theory, development of instruments and measures, and details of how to conduct utility elicitation (e.g. time trade-off and standard gamble interviews) are beyond the scope of this guideline.

Methods for estimation of HSUVs by mapping from condition-specific measures are not considered in detail; separate guidelines for these methods are available (Longworth and Rowen, 2011) and are under development (Wailoo et al., 2015). Measurement of health-related quality of life (HRQL) using scales other than the health-utility scale (for example, using condition-specific or generic HRQL measures such as the SF-36 Health Survey) is a broad topic and is not covered here. Several guidelines have been developed by ISPOR’s Task Forces for the measurement of patient-reported outcomes (Wild et al., 2009, p 430; Coons et al., 2009, p 419; Eremenco et al., 2014, p 17; Zbrozek et al., 2013, p 480; Matza et al., 2013, p 461; Rothman et al., 2009, p 1075; Patrick et al., 2011, p967).

This article states the consensus position of the ISPOR Task Force on Good Practices for Outcome Research – Collecting Health-State Utility Values for Economic Models in Clinical Studies and provides recommendations for best practice. Recognizing that best practices may not always be feasible due to practical constraints, pragmatic approaches also are discussed. This position represents the best judgment of the Task Force at the time of writing and is subject to change as new methods for health-utility estimation and economic modeling emerge and the preferences of HTA, pricing, and reimbursement authorities change.

2 General Considerations for the Collection of Health-Utility Data for Economic Models

There are several key issues to consider when designing studies to collect HSUVs for economic models.

1. The HSUV estimates must fit the needs of the decision problem, so the study design must fit the needs of the economic model for which the data are to be used. The study should be designed (as far as is possible) to capture HSUVs for each health state likely to be in the economic model. The same criteria should be used to categorize patient assessments into health states in the study that are used to characterize health states in the economic model. The health-utility measure should be selected based on its acceptability to the economic model’s audience (e.g., HTA authorities), validity
2. The participants in the study should reflect the indication and the population that is considered in the economic model (usually the population who would receive the investigational intervention if it is adopted in routine clinical practice). Ideally, studies should recruit a sample of participants that will produce HSUVs representative of the entire population under consideration in the economic model. To try to achieve this, any sampling approach should be free from sources of bias. It is better to recruit from multiple centers and, if appropriate, to include patients on different types of treatment. Trial inclusion and exclusion criteria may result in populations which are younger and fitter than populations in routine practice, or exclude specific groups of patients. Participants in more severe health states can be less likely to complete assessments, which may be a significant source of bias (Bleehen et al., 1993, p 1157). Similarly, the participants recruited into any study should be representative of the types of patients being treated in the health care system that is requesting the cost-effectiveness analysis. For example, if the HSUVs are being collected for an economic model for the National Institute for Health and Care Excellence, then it is most appropriate to include participants from England and Wales. If it is not possible to recruit a representative patient sample, it is important to ensure adequate sampling of the model population and to adjust HSUV estimates in the analysis. However, it should be noted that this approach puts pressure on the validity of statistical adjustment. Protocols should be designed to minimize any missing data that can affect the representativeness of data.

3. Many economic models describe longitudinal changes in patients’ health utility as their disease progresses through different stages. Sampling needs to consider variability among patients within a health state (i.e. the study population should adequately reflect variability within the model population) and variability over time within a single health state (e.g. a decline in HRQL over time within a disease progression category). The order in which health states are experienced also can be important e.g. health utility in patients with low BMI who previously had high BMI might be expected to differ from patients who always had low BMI. The best way to understand how health utility changes over time is to capture data from patients at multiple time points as they progress through health states. There is emerging evidence that changes in health utility associated with experiencing health states may differ significantly from differences inferred from cross-sectional data (Alvar, 2013). This is not always practical, due to time and cost constraints. Where these cannot provide the necessary data, cross-sectional studies should be designed to ensure that each sample in each health state is representative.

4. There are valid arguments for collecting utilities in a real-world clinical practice rather than in a clinical trial. Trials contain many protocol-driven procedures and have very tight entry criteria, both of which may reduce the generalizability of the data. Collecting subjective data in a trial also may promote placebo-type effects that could inflate health utility scores (for example, regular monitoring or the possibility of receiving an investigational new treatment may inflate utility scores). This becomes important when we consider that the estimation of QALYs relies upon absolute (not relative) health utility scores; if utility scores are inflated, it will be a source of error. Data collection from routine clinical practice may produce more representative data for guiding decision making.
5. All studies will inevitably present a stylized representation of how patients’ HRQL is affected in the real world. Many factors influence HRQL; without random sampling, it is very difficult to control for these sources of heterogeneity. One longer-term solution is the use of routine outcome measurement embedded in clinical practice. For example, the Patient Reported Outcome Measures program in the United Kingdom attempts to capture data from all patients undergoing hip and knee replacement surgery, and part of the assessment includes the EQ-5D (Department of Health, 2012). However, data collected in routine clinical practice for patients receiving existing therapies may not fully reflect the HRQL of patients receiving a new treatment.

6. Many economic evaluations will include health states that occur infrequently but nevertheless are important to capture because of their impact on patients’ health, for example a severe adverse event or a complication related to a disease. If health states are rare, then it is often hard to capture enough data from patients in that health state. When making decisions regarding how data should be captured, it is worth reflecting on how important these states are for the outcome of the economic evaluation. If they are important for determining cost-effectiveness, then this may provide justification for investing in a prospective study. One way to address this problem is by purposively recruiting patients who are at increased risk for developing the infrequent health state. So, rather than recruiting a large number of patients and waiting for the health state to emerge, it may be better to specifically target patients at increased risk. However, this approach may not work for some disease areas. For some health states, it is extremely difficult to capture data outside of a clinical trial. For example, there are very few reports in the literature of health-utility data captured from patients experiencing severe adverse events. Further, published quantitative data regarding health utility in patients approaching the end of life in palliative care are rare. Such information can be important. Data collection after the end of the active treatment phase of the study would be very informative regarding how HRQL changes as people move through subsequent care and treatment (although may be more prone to missing data, see section 4.1.6).

7. Regarding the mode of administration of the assessment (e.g., paper instruments, online instruments, smartphone apps), while flexibility is often desired, standardization of the mode of administration across the patient sample is preferred, to limit the potential for bias or variance due to mixed modalities (Bowling, 2005, p 281). The mode of administration can affect response rates, task comprehension, response strategies, and the extent to which the sample is representative of the population (Norman et al., 2010, p 499). Mode may also impact responses— for example Clarke and Ryan (2006) showed that respondents were more likely indicate either the highest and lowest categories (i.e. excellent and poor) when asked an oral rather than a written self-reported health question. Furthermore, where electronic data capture is used, care must be taken to ensure that the questionnaires can be standardized across all potential devices. For example, it is essential that the length of a visual analogue scale is consistent for all participants on all devices at all time points (Critical Path Institute PRO Consortium, 2014).

8. The application of any research solution should be within the appropriate guidance for good clinical practice, patient consent, safety reporting protocols, and data transparency obligations appropriate to the study.
Box 3. General Considerations for the Collection of Health-Utility Data for Economic Models

- HSUV estimates must fit the needs of the decision problem evaluated in the economic model
  - Values for model health states and events should be captured
  - Criteria used to categorize patient assessments into health states should be aligned with model health-state definitions
  - Selection of the HSUV instrument or measure should take account of any needs / preferences of the economic model’s audience and suitability to health condition of interest
  - Study participants should reflect the indication and population considered in the economic model
- Changes in health utility over time should be captured
  - Sampling needs to consider variability among patients within a health state (i.e. the study population should adequately reflect variability within the model population) and variability over time within a single health state (e.g. a decline in HRQL over time within a disease progression category). The order in which health states are experienced also can be important e.g. health utility in patients with low BMI who previously had high BMI might be expected to differ from patients who always had low BMI.
  - In longitudinal studies, ensure that any changes in utility over time as patients progress through health states are captured (e.g. making multiple assessments over time within a given health state)
  - In cross-sectional studies, ensure that the data sample is fully representative of each health state
- There are valid arguments for collecting utilities in real-world clinical practice versus clinical trials, e.g. representativeness of respondents to the model population
- Where there are problems with collecting data for rare health states or events
  - Consider whether the estimates are important (do they influence incremental cost-effectiveness ratio)?
  - If they are, consider recruiting patients who are at increased risk or continued follow up after active treatment phase
- Mode of administration (e.g. paper instruments, electronic data capture systems via computer, dedicated device or, platform mobile apps)
  - Can affect response rates, task comprehension, response strategies, and representativeness of the sample
  - Standardization across patient sample and devices is preferred

3 Recommendations for Early Planning of Health-Utility Data Collection Within a Product’s Research and Development Program

To successfully plan the collection of health-utility data for an economic model, it is important to perform early research to establish the HSUVs that will be required for the model, the data that are already available from the published literature, and, if there is any uncertainty, the validity of available instruments for the health condition of interest. The planned trial program may then be evaluated for opportunities to collect the required HSUV estimates and, if necessary, perform additional research to explore the validity and responsiveness of a particular health-utility instrument in the condition of interest. Any expected gaps in the data available from the clinical trial program should be identified, and additional studies to address these gaps should be planned. Engaging the teams responsible for economic modeling, health-utility assessment, patient-reported outcome assessment, and clinical trial design in strategic planning at an early stage within the product development lifecycle will maximize the potential for producing robust data that best support the health economic model. Recommendations for activities involved in collecting health utility data throughout the product development program are presented in Figure 1.
Figure 1. Recommendations for HSUV strategy during the product development process

**Phase 1**
- Determine HRQoL impact of disease, assess HRQoL benefits of intervention
- Identify and define health states (an early economic model can be helpful)
- Identify needs or preferences of key markets and model audience for HSUV data
- Review literature to identify HSUV estimates that have been published and assess their quality and appropriateness for the model and audience
- If there is doubt about the validity and responsiveness of possible utility instruments, evaluate this (e.g., by review of published qualitative and quantitative studies)
- It can be helpful to develop an HSUV Research Plan

**Phase 2**
- Perform any additional research needed to evaluate validity and responsiveness of utility instrument or select instrument
- If appropriate, use phase 2 trial with extended follow-up or other studies to collect long-term longitudinal data unlikely to be available from the phase 3 trial
- Perform any additional longitudinal studies necessary to supplement trial data

**Phase 3a**
- Include utility assessment in trials; prepare a utility data collection protocol and analysis plan
- Perform any additional cross-sectional or short-term longitudinal studies necessary to supplement trial data

**Phase 3b**
- Perform analyses of phase 3 trial or other longitudinal data for the economic model
- Conduct stand-alone, cross-sectional studies to collect HSUVs, if needed

**Phase 4**
- Continue to gather long-term outcomes data and compare with earlier data and model predictions

**HRQL** = health-related quality of life; **HSUV** = health-state utility value.
It is important first to describe the impact of the health condition or disease on HRQL and the benefits that the intervention is expected to provide in terms of HRQL, so that the research may focus on these aspects. As part of the planning process, it may be helpful to develop a preliminary economic model to identify and define the HSUV estimates that are likely to be required, and to perform exploratory analyses to determine the influence that individual parameter values have on the results. This can be helpful in determining how much resource to invest in the collection of specific estimates. It may be appropriate to consider whether it will be important to investigate the impact of the patient’s condition on the health utility of carers and/or family members or dependents. A literature review should be performed, to establish the availability of the required HSUV estimates, as well as the quality of the available data, the relevance of the data to the economic model’s health states, and the acceptability of the data to the model’s audience (e.g., specified HTA authorities).

Health-utility data may be generated using one of the following categories of instrument (see also Box 2):

- A generic, preference-based MAUI (e.g., EQ-5D, Health Utility Index, SF-6D Health Survey, 15D, Assessment of Quality of Life, or Quality of Well-being)
- A condition-specific preference-based MAUI (e.g., Asthma Quality of Life Utility Index AQL-5D or the European Organization for Research and Treatment of Cancer EORTC-8D)
- A non-preference-based, condition-specific patient-reported outcome measure mapped onto a generic, preference-based measure (Longworth and Rowen, 2011; Wailoo et al., 2015)

An instrument or instruments should be selected based on suitability for the disease or condition of interest and on its acceptability to the model’s audience (e.g., the HTA authorities to which the model is expected to be submitted). If there is any doubt about the appropriateness of utility instruments for the condition of interest, this should be evaluated in terms of content validity, face validity, construct validity, and responsiveness based on qualitative and quantitative empirical evidence (Brazier and Longworth, 2011). Some examples of this type of evaluation are available in the published literature (e.g., Brazier et al., 2014). The requirements and preferences of HTA authorities are evolving over time; researchers should refer to guidelines issued by individual authorities (links to many of these guidelines are available on the ISPOR website) and consult with the HTA authorities, if need be, regarding appropriate instruments in the condition of interest. If additional research is required to validate an instrument or develop an appropriate instrument, identification of this requirement early in the product’s development may make it possible to perform such research in parallel with the clinical research program.

These activities should establish what HSUVs need to be measured and the appropriate instrument for their measurement. Researchers should determine whether the planned trial program has the capacity to provide all HSUVs required for the economic model. The benefits of collecting all key health utility estimates for an economic model within a single study, or at least using a single methodology, should be recognized. Often researchers are faced with limitations on the number of measures that can be included in a clinical study (perhaps only one patient-reported outcome measure can be included, focusing on either HRQL or health utility). If it is not feasible to include all desired measures, researchers may need to consider whether a health utility measure such as the EQ-5D could be used as a substitute for a non-preference-based generic HRQL instrument; whether one instrument may be used to collect both HRQL and health-utility data, using an integrated HRQL and health-utility measure (e.g., SF-36/SF-6D,
EORTC QLQ-C30/EORTC-8D, or AQLQ/AQL-5D), or whether a HRQL measure could be included and mapped to health utility. Such decisions should recognize the importance of health-utility data and the requirements of the audience for the economic model.

As part of the phase 2 trial program, additional research may be performed either in the phase 2 trials or in parallel studies, to investigate the validity and responsiveness of a measure or measures in the population of interest, in order to inform the selection of a measure for phase 3 trials. There also may be the potential to use a phase 2 trial to collect coincident data using a HRQL and health utility measure which could be used to develop a mapping algorithm, or long-term longitudinal data to capture health states that may not be observed in the phase 3 trial due to limited follow-up.

Planning for HSUV collection in phase 3 trials should be performed carefully, after identification of potential issues and following the recommendations in Section 5.1. If early planning has not been performed prior to the phase 3 trial, the items highlighted with an asterisk in Box 4 are recommended prior to developing the trial protocol. If it is determined that the trial program will not be able to provide all utility data required for the model, early planning will allow time for any additional studies that may be required. Early planning can be the key to providing robust health-utility data for inclusion in the economic model.

**Box 4. Planning HSUV estimation**

Begin planning for HSUV estimation early in the product development process

Prepare a description of the HRQL impact of the disease condition and the potential HRQL benefits of the new intervention*

Prepare a list of HSUVs required for the model, appropriate measurement methods and data already available; define new data needs*

- Identify and clearly define the patient population of interest and the modeled health states and events for which HSUVs will be required. It may be appropriate to consider the need for health utility data for patients’ carers and/or family members or dependents. Development of an early economic model can be helpful in this process*

- Identify the requirements and preferences of key model audiences (e.g., HTA authorities) for HSUV data*

- Identify the most appropriate health-utility instrument, based on appropriateness for the condition of interest and acceptability to the audience(s) for the economic model(s). If there is any uncertainty about the appropriateness of HSU instruments in the disease condition, evaluate the validity and responsiveness of possible instruments (e.g., by review of published qualitative and quantitative studies)*

- Review the published literature to identify HSUV estimates that are already available; assess their quality, relevance to the model health states, and acceptability to the audience for the model (e.g., HTA authorities)*

- It can be helpful to determine the sensitivity of cost-effectiveness estimates to individual model HSUVs using an early economic model, to evaluate the importance of collecting high-quality data for each HSUV*

Evaluate the planned research and development program of the product for opportunities to collect the HSUV data defined in the previous steps

- Consider opportunities to use phase 2 trials or other clinical or observational studies to collect long-term longitudinal data and/or collect additional data unlikely to be available from the phase 3 trial

- Recognize the benefits of collecting all key HSUVs within a single study (or at least using single methodology)

It can be helpful to prepare an HSUV research plan*

**HRQL = health-related quality of life; HSUV = health-state utility value; HTA = health technology assessment.**
* If early planning has not been performed prior to the phase 3 trial, the items highlighted with an asterisk are recommended prior to developing the trial protocol.

The following content is suggested for the HSUV research plan: a definition of the patient population; the HRQL impact of the disease condition and potential HRQL benefits of the new intervention; the definitions of health states and events for which HSUVs are required for the economic model; a summary of validity and responsiveness of utility measures for the condition of interest and their acceptability to the economic model’s audience; the selection of an appropriate measure; a summary of HSUV estimates already available and assessment of their quality, relevance to the model’s health states, and acceptability to the model’s audience; an assessment of the importance of collecting high-quality data for each HSUV (based on analyses using the early economic model); an evaluation of the product’s planned research and development program for opportunities to collect the HSUV data; the identification of data gaps (HSUVs which are unlikely to be able to be estimated in planned studies); a plan for collection of health-utility data within the planned research program; and an outline for any additional studies to bridge data gaps.

4 Considerations for the Design of Studies to Collect HSUVs for Economic Models

4.1 Recommendations for the design of health-utility data collection during the protocol development for a planned clinical trial

Trials often represent an important opportunity to measure health utility as they provide a large sample of the patient population of interest within a study that is performed and monitored to a high standard, thus maximizing data quality and completeness. In addition, the health utility data may be linked to the endpoints which are used to estimate treatment effect. However, there are a number of pertinent issues that can determine the value of the health-utility data collected in this way.

4.1.1 Number of possible assessments for a given health state

There may be some HSUVs required for which it is not feasible to collect health-utility data in the trial or for which the number of assessments possible is likely to be low—for example, values for rare health states or health states that tend to occur after the end of the trial’s follow-up period. Researchers should identify the economic model’s health states and events for which it is feasible to estimate health utility in the trial and define these clearly using (as far as possible) the same criteria that will be used to characterize health states in the economic model. Any modeled health states or events for which data of sufficient quality may not be available from the trial should be identified, and alternative plans for their collection should be made.

In some cases, patients may enter the trial with different numbers of previous events experienced or different numbers of lines of treatment received. This may be problematic if the economic model requires HSUVs for a first, second, or third event or treatment line. There also may be differences in combinations of previous events or treatments or in time since the event or treatment. These issues should be considered at the design stage. Collection of appropriate data on disease and treatment history would allow separate estimates to be calculated as required.

Sample size considerations should take into account the need for precision of HSUV estimates rather than hypothesis testing, as the main objective is to estimate utility values for model health states and not to make statistical comparisons of utility between treatment arms.

At each health-utility assessment, it is important to capture any other variables that will be needed to connect the patient’s assessment with the economic model health state the patient is experiencing at the
time of the assessment. The nature of the data collection (e.g., categorical or continuous data, units, etc.) should be selected in view of the planned analyses of the data for the economic model.

4.1.2 Representativeness of the trial population to the population of interest in the economic model

Issues which may affect the representativeness of the trial population to the economic model population should be identified, for example trial inclusion and exclusion criteria. If the economic model population includes individuals who are excluded from the trial, it is important to consider whether they are likely to differ in terms of health utility, both in absolute and incremental terms. Researchers should evaluate the potential for bias arising from issues of generalizability and formulate plans to address any such issues. It may be possible to adjust trial selection criteria or collect utility data for patients excluded from the trial who would be eligible for treatment in routine practice. The extent and direction of any bias may be examined (e.g., by comparing baseline characteristics with characteristics of real-world populations) and/or corrected for in the analysis (see Section 6).

4.1.3 Special patient populations

Collecting health-utility data in certain patient populations poses significant issues (e.g., patients with dementia [Bryan et al., 2005], younger children and the cognitively impaired), although work is ongoing in this area (Ungar, 2011, p 641; Sonntag et al., 2013, p 1131). Also, patients who are severely ill may not be able to complete assessments. Using proxy respondents where the patients themselves are unable to fill out questionnaires is an acceptable practice; it is common practice to use a nominated informant for each patient to ensure consistency across multiple assessments for individual patients. For example, in cognitively impaired patients who have experienced stroke and cannot complete questionnaires, research suggests that family caregivers can reliably complete the utility assessment (Mathias et al., 1997, p 1888). On the other hand, health preference studies of children (including those with pediatric brain tumors or extremely low birth weight infants) have reported higher utilities for parent proxy-report compared with child self-report, whereas other studies have reported lower utilities (Lee et al., 2011, p 924). Therefore, when using proxy respondents, researchers should review available evidence examining inter-rater reliability or agreement between patient and proxy respondents for the instrument and condition in question, and should examine the potential for, direction, and extent of bias that might result; bias should be explored using sensitivity analyses in the economic model.

4.1.4 Timing of assessments

Clinical trial assessments often are made at regular scheduled intervals. The optimum timing of health-utility assessments may not coincide with the clinical endpoints. For example, health-utility assessments in cancer therapy trials often are made alongside clinical assessments at the chemotherapy administration visit (e.g., approximately every 3 weeks) and a short time after disease progression. This often results in numerous utility assessments prior to disease progression but few after progression and very few or none during terminal illness.

As a general rule, the number and timing of assessments should be planned to optimize the amount and quality of data collected and its relevance to the economic model, recognizing the health states and events for which HSUV estimates are required in the model. It is also important to recognize the instrument’s recall period, which can range widely. For example the EQ-5D describes health today, while the SF-6D
describes health over the last 4 weeks. If HRQL is expected to vary over time within an individual economic model health state and the HSUV is conceptually an average for all patients and all times in that state (e.g., as in the case of post-progression utility in many cancer models), care should be taken when scheduling assessments to ensure that the variability between patients and over time is adequately sampled. Data quality and relevance should take precedence over convenience, as far as is possible; and it may be necessary to schedule assessments specifically for health-utility measurement. For example, in cancer trials, it would be valuable to schedule assessments throughout the period after disease progression until death; this may be possible particularly in trials that include overall survival as an endpoint (use of proxy respondents may be considered if necessary). Economic models often are concerned with changes in HRQL as patients move between health states. Therefore, it is important to schedule assessments that allow these changes to be estimated (for example, at baseline and when certain clinical outcomes are reached, which define the economic model health states). Trials that involve a change from baseline (e.g., trials of surgical interventions) may benefit from multiple measures prior to baseline, to reduce the uncertainty around the treatment effect.

When HSUV are collected during a visit involving other interventions and/or assessments, it may be important to standardize the timing of the utility data collection (e.g. at the beginning of the visit), bearing in mind whether and how any of the other interventions/assessments may affect HRQL.

4.1.5 Acute events

An additional challenge for collecting health-utility data in clinical trials is in circumstances where much of the quality-of-life gain from an intervention is in the reduction of acute episodes i.e. those with a transient effect on HRQL (e.g., heart attacks, asthma exacerbations, hemophilia bleeds) or of short-term (but often severe) treatment-related adverse events. These circumstances represent a particular challenge in health-utility assessment because of the practical and potential ethical considerations in requiring patients to complete a questionnaire during an acute episode (e.g., during a hospitalization, or disease exacerbation). The recall period of the utility measure should be carefully considered in relation to the timing of assessments in relation to the occurrence of acute events.

Another challenge facing researchers is in measuring the period of time over which the event impacts HRQL. From a practical standpoint, clinical trial personnel may not be aware of a patient’s acute event until the patient’s scheduled trial visit, by which time some recovery may have occurred. Researchers therefore should think about whether their economic model needs health-utility estimates for acute events or states; if the answer is yes, further consideration should be made as to whether these estimates can be realistically measured in the context of a clinical trial. As a practical consideration, the more the acute events are related to the intervention of interest and the more they are expected to affect the economics, the more accurately they should be described. Important acute events (e.g., heart attacks, asthma exacerbations, hemophilia bleeds) are commonly included among the trial endpoints that trigger collection of clinical data, and economic models often focus on grade ≥ 3 adverse events because these require treatment. Both types of event therefore require clinical contact and data collection that could include health-utility assessment.

When considering acute states, the following approaches may be helpful: 1) asking proxy respondents to complete the assessment, 2) asking patients to recall their experience with the acute events of interest.
from the recent past, or 3) conducting additional health-utility elicitation outside of the trial. Adverse events are often acute conditions, and the previously discussed approach to health-utility assessment for general acute events may apply to the measurement of health utility for individual adverse events. Alternatively, the impact of adverse events on HRQL may be measured in aggregate, i.e., by sampling the time during which patients are on treatment in such a way that the average effect of adverse events over time is captured. For example, in cancer trials, patients could be asked to complete the health-utility measure on randomly assigned days after chemotherapy administration (the random allocation could be performed as part of the process of random allocation to treatment arm).

4.1.6 Missing data

Where there is loss to follow-up from the trial, the patients remaining in the trial may represent a subpopulation that is distinct from that which was lost to follow-up (e.g., patients with less severe disease, response to treatment or fewer adverse events). Therefore, measurement of the health utilities in the patients who remain in trial may not accurately capture the health utilities of the target population. In these situations, standard imputation procedures such as last observation carrier forward may not be appropriate. Where deliberate plans are in place not to follow up certain patients (e.g., if it is no longer necessary for the purposes of the efficacy and safety assessments), the relevance of these patients from a health-utility perspective should be assessed and continuing follow-up performed as appropriate and when possible. If assessing those lost to follow-up is not practical, it may be valuable to contrast the baseline general characteristics and HSUVs of those who subsequently drop out and those who do not.

Plans should be put in place to minimize missing data resulting from loss to follow-up and other causes. Recommendations for minimizing missing data are available in the published literature (e.g., Little et al., 2012, p 1355).

4.1.7 Optimization of health utility data collection

Using good design and appropriate analysis should allow researchers to account for many of these issues. Health economists who understand the needs of the economic model should be involved in and able to influence the design of any trials in which health-utility data are collected, as well as of the analyses of the data. The aim of their participation should be to optimize the design of the health-utility data collection in the trial to meet the needs of the economic model and the model’s audience and to identify any HSUVs that may not be estimated within the trial so that other plans can be made for their estimation. It is important to plan any utility data collection during protocol development i.e. before the trial design has been finalized.

The objectives of health-utility measurement in trials should be clearly defined to aid in the selection of the instrument and measurement schedule. The economic model health states for which it is feasible to estimate health utility in the trial should be identified and clearly defined. Careful consideration should be given as to whether it is appropriate to reflect differences between treatment groups within health states—for example, due to different patterns of adverse events or different intensities of therapeutic response.
4.2 Recommendations for the design of prospective or cross-sectional observational studies for health-utility estimation

For a variety of reasons, it may not be possible to collect health utility data within a clinical trial program. In this section, we discuss the merits of conducting separate observational studies to collect these data. We also review factors to consider in the design and conduct of such research. Any planned studies should be designed to be representative of the population considered in the economic model, free from sources of bias and valid in their design and execution.

Bespoke prospective and cross-sectional health-utility studies can range in complexity, from complex studies performed at clinical sites to very simple online surveys. Many of the issues and recommendations discussed in Section 5.1 also apply to these studies and should be evaluated during the study’s design phase. This section considers additional issues that are specific to prospective studies in clinical centers and cross-sectional surveys. Other study designs also are possible for which similar considerations apply, for example cross-sectional studies performed at clinical centers may be appropriate if longitudinal data are not required or if there is insufficient time for a longitudinal study but detailed and verifiable clinical data are needed. Longitudinal patient surveys with repeated administrations over time may be appropriate if longitudinal data are required and access to patients via clinical centers is difficult (as may be the case for very rare diseases).

4.2.1 Cross-sectional surveys

Cross-sectional surveys can be set up and run quite rapidly. Most HRQL measures have validated web-based or tablet-based versions available for use and guidance exists regarding this process (Coons et al., 2009, p 419). Cross-sectional surveys have a number of important advantages. They are a relatively quick and cost-effective method for capturing HRQL data. Surveys can be conducted nationally or internationally, which can improve the representativeness of data. Online surveys may be considered more discreet and thus a better format for capturing sensitive data (e.g., Kerr et al., 2012). In the study by Kerr and colleagues, the impact of sexual dysfunction on HRQL was assessed solely via an online survey that consisted of a combination of self-complete clinical assessment scales and HRQL measures. Cross-sectional surveys are also often completed at sites that can be useful when more detailed clinical data are required about the participants. Online surveys can recruit through different sources such as patient advocacy groups, social media, and recruitment panels, as well as from health care providers.

There are certain limitations to cross-section surveys, and these limitations also should be considered. As well as capturing HRQL data, any survey needs to capture the correct variables for categorizing participant assessments into health states. Studies where the health state is defined by a laboratory or radiological marker would most likely need to be conducted at a clinical site. In studies where participants are recruited through advocacy groups or recruitment panels, it is only possible to capture basic clinical information (diagnosis, time since diagnosis, treatments, etc.). Depending on the recruitment method, it may not even be possible to accurately verify diagnosis of participants. Some self-complete tools are available that can be used to measure severity, but this is not always possible. Lastly, studies that are conducted exclusively online may be criticized for excluding participants who do not have access to the Internet, although with ever increasing access to such technology this has become less of a concern.
4.2.2 **Prospective studies in clinical sites**

More formal studies can be conducted at clinical sites to capture HRQL data. This is recommended when the potential participants can be considered a hard to reach group because of prevalence or other reasons. For example, an observational study of patients receiving treatment for a life-threatening condition like cancer is probably best conducted through clinical sites because it is usually important to have detailed diagnostic, staging, and treatment history information that requires access to medical records. Lloyd et al. (2007, p 22) report a prospective HRQL study conducted at four clinical sites, of people with moderate to severe allergic asthma which necessitated access to clinical data. However, recent experience from Sheffield, United Kingdom, shows how challenging this type of research can be (Dixon et al., 2014). In the Dixon study, the authors attempted to recruit a sample of women with breast cancer in order to measure the disease’s impact on HRQL. Despite the fact that the study was cross-sectional only, the authors were not able to recruit sufficient numbers of patients in the more advanced stages of the disease.

Prospective studies are often complex to undertake. Clinical sites may not see the value of this research, compared with an investigational trial, and thus it can be difficult to recruit participants. Usually patients cannot be in two studies simultaneously (even if one is an observational study), and patients may be more likely to be invited into a clinical trial. Thus, observational studies can be slow to recruit and more importantly may not be representative of patients generally (Lloyd et al., 2010, poster). Careful study site selection and engagement with medical specialists and advocacy groups may help to avoid these problems.

Lastly, it is also possible to undertake longitudinal research without recruiting participants through sites but instead by recruiting them through other sources (e.g. patient advocacy groups). This is particularly relevant if detailed clinical data are not required. Monthly or even annual data collection could be undertaken to understand changes over time. Endorsement from patient charities or medical specialists may help to minimize patient drop-out in such studies.

4.3 **Alternative study types for estimation of HSUVs**

In situations in which HSUVs may not be estimated using methods described in previous sections, alternative methods may be considered.

Early-access or compassionate-use-type programs, phase 4 studies, registries, and other post-licensing research activities can provide opportunities for including HRQL assessment. The design issues and recommendations discussed in the previous sections also apply to these studies. These types of studies can be an efficient way to capture HRQL data but these may be performed too late in the product development program to provide HSUV estimates for HTA submissions.

Valuation of health-state descriptions (vignettes) by the general population may be an option when other methods are not possible, for example if administration of an existing instrument to a cohort of patients is not feasible or impractical. The quality of the health-state descriptions is critical in vignette studies. Recent guidelines recommend extensive qualitative work with patients, independent psychometric validation of the vignette descriptions, and use of quantitative HRQL data to inform the content (Brazier and Rowen, 2011).
Estimation of HSUVs by health care professionals should be avoided. However, the opinion of clinical experts can be helpful in evaluating the plausibility of alternative available estimates and/or the expected rank order of HSUVs, from best to worst health state, on the health utility scale.

Box 5. Design of health utility data collection in a clinical trial

Design any health utility data collection in a clinical trial during the development of the protocol with reference to the anticipated needs of the economic model. Health economists who understand the needs of the economic model should be involved in and able to influence the design of any trials in which health-utility data are collected and the analyses of such data. Identify issues associated with collecting HSUVs for the economic model in the planned trial and make plans to address them. The following issues may arise for example:

- The timing of clinical assessments may not be optimal for utility measurement
  - These should be optimized to capture data for model health states and/or events and to maximize the amount and quality of data
  - Consider the recall period of the utility instrument
  - Consider acute events and their duration of HRQL impact
  - Consider changes in HRQL over time within a single model health state (e.g., "post progression" health states in cancer)
- Consider the number of assessments which are likely to be feasible for each model health state and/or event and the likely precision of the HSUV estimates.
- It may not be feasible to capture data for some model health states and/or events within the trial (e.g. rare events, events occurring after trial follow-up). Make alternative plans to collect these data.
- Consider the importance of the acute events (i.e. those with a transient effect on HRQL) for the economic model; if the collection of accurate estimates is important, consider whether measurement is feasible within the trial, or plan a separate study.
- The trial population may not be fully representative of model population e.g., due to trial exclusion criteria or geographic footprint. Evaluate the potential for bias, consider adjusting trial selection criteria, collecting data for excluded patients, and/or adjusting estimates in the analyses.
- Respondents may be unable to complete utility assessments e.g. young children, cognitively impaired, severely ill. Examine the relevant literature for any relevant research findings; if use of proxy respondent is expected to be the best solution, review relevant literature and examine the potential for bias.
- Identify potential causes of missing data (e.g., reasons for planned or unplanned loss of follow-up, types of patients less likely to complete assessments). Address these as far as possible by adjusting the study design, and formulate plans to adjust for missing data in the analysis.

Development of a utility data collection protocol (or a section of the trial protocol) is recommended:

- Clearly define the objectives of the health-utility measurement in the trial
  - Identify and define model health states and/or events for which it is planned to estimate health utility in the trial (considering the number of assessments feasible for each health state and whether HSUVs for acute events are important for the economic model and feasible to estimate in the trial)
  - Specify whether statistical comparisons between treatment groups in overall utility over time also will (or will not) be performed
- The following design features should be defined and justified in terms of the needs of the economic model:
  - Choice of instrument
  - Timing and frequency of assessments and period of follow-up
  - Variables that should be collected at baseline and at each health-utility assessment to determine health state at the time of assessment and to allow adjustment for covariates
  - Choice of respondents
5 Recommendations for Data Analysis and Reporting

This section presents recommendations for statistical analyses and reporting of the results of health-utility analyses conducted alongside clinical studies in order to maximize the potential of the HSUV estimates for current and future economic models.

5.1 Estimation of HSUVs for economic models

When clinical trials are used to collect HSUV data, it is often the case that the approach to the analysis is strongly influenced by the standard, between-arm comparisons that are used for regulatory submissions. However, if reimbursement submissions are to be made using an economic model, then it is more appropriate that the health-utility data in a clinical trial be analyzed to inform that model, rather than be analyzed by treatment arm.

Consider the following example of health-utility data collected as part of the UK Prospective Diabetes Study (UKPDS). When the study started in 1977, economic evaluation was barely used and health-utility measures had not been invented. However, toward the end of this trial, which reported in 1996, there was much interest in performing economic evaluations alongside clinical trials. Therefore, in 1996, as the study was drawing to a close, the EQ-5D was administered cross-sectionally to all 3,667 patients remaining in the study. When the data were analyzed by treatment arm, no difference in EQ-5D utility could be detected, despite the fact that intensive blood glucose control had been demonstrated to significantly reduce the long-term sequela of diabetes in this landmark trial. A subsequent analysis therefore regressed EQ-5D scores against the long-term clinical outcomes that formed the primary endpoint of the UKPDS study. As expected, the results showed very clearly that these long-term outcomes had a significant detrimental effect on HRQL as measured by the EQ-5D. Figure 5 shows the estimated health-utility decrements associated with the different clinical events and even differentiates the "acute" impact (event occurrence in the year of EQ-5D measurement) from the "chronic" impact (event occurrence one year or more before EQ-5D assessment).
The UKPDS example is compelling. The “story” told by the model, that treatment impacts the long-term risk of events, and those events impact HRQL, is confirmed by the analysis of the data. Yet with all the “noise” in the observed data, a between-arms difference in HRQL could not be detected at conventional significance levels. Of course there are problems with using cross-sectional data in this way. One issue is that those experiencing an event could have had a lower starting health utility than those who did not experience an event, and this was shown in an analysis of the extended follow-up data (Alva et al., 2014). However, and in contrast to the UKPDS, most studies collect health utility at randomization and at intervals throughout the study. With longitudinal data, more sophisticated analyses are possible that control for baseline health utility. The lesson is that the approach to analyzing data for reimbursement outcomes should reflect the needs of the economic analysis and should not be constrained by the traditional regulatory approach to analyzing data.

5.2 Distributional considerations

When examining health-utility data, it is clear that they are often subject to left skewedness and can often have a point probability mass at 1. This is because 1 represents perfect health and generates a ceiling effect in that values above 1 are not possible. By contrast, while 0 represents dead, states worse than death are allowable and therefore there is no theoretical lower limit for the utility scale. From a statistical
analysis perspective, left-skewed data are more problematic than right-skewed data. However, a very simple solution exists. A simple linear transformation of \( X = 1 - U \) facilitates a move to the “health-utility decrement” scale. Health-utility decrements are right skewed with a possible point probability mass at 0. The key is that, being right skewed, health-utility decrements are similarly constrained as cost data, and all the usual methods analysts will be familiar with in analyzing right-skewed cost data can be employed: namely, generalized linear modeling or generalized estimating equations for longitudinal data, with the possibility of a two-part model to handle excess zeros. Returning to the original health-utility scale after estimation is trivial; since the original transformation was linear, the back transformation of \( U = 1 - X \) does not impact any of the estimated quantities, such as standard errors.

### 5.3 Modeling heterogeneity and increasing generalizability

It has long been recognized by statisticians that statistical modeling of the effect of prognostic variables on the outcome of interest in clinical trials can improve the precision of estimated treatment effects, even where randomization is used to control for unobserved heterogeneity (Altman, 1985; Pocock, 1985). In nonrandomized clinical studies, the ability to control for observed prognostic variables is critical for the estimation of treatment effects. Furthermore, modeling health utility as a function of clinical endpoints offers the potential to describe a causal pathway of how treatment effects intermediate endpoints that themselves impact on health utility. It is well known that clinical trials trade external validity for internal validity. However, the explicit modeling of prognostic factors offers the potential at least to increase the generalizability of the results by offering a method of adjusting clinical trial data to the characteristics of real-world populations.

### 5.4 Transferability in multinational studies

In multinational studies, analyses of data collected via MAUIs should use the country tariff which is most appropriate for audience or jurisdiction for the economic model. A particular form of heterogeneity relates to the multinational aspect of many clinical studies. For example, a recent analysis of variations the ADVANCE study which was conducted in 20 countries showed substantial variation across regions in reporting on functional health problems, which cannot be explained by differences in demographic variables, clinical risk factors, or rates of complications (Salomon et al., 2011). This can manifest itself in predictable differences in the way in which patients map to the index score of the health-utility instrument. To this end, country- or region-level effects can be handled using a covariate in a standard statistical modeling approach. However, multinational aspects of transferability are further compounded by the existence of jurisdiction-specific tariff values for health-utility instruments. This means that even where there are no estimable country- or region-specific influences on the index measure, jurisdiction-specific differences nevertheless may exist through appropriate use of jurisdiction-specific tariffs, just as unit costs may drive differences between jurisdictions even when resource use differences do not exist.

### 5.5 Reporting to maximize value for future economic models

When reporting the analysis of health-utility values, care should be taken to always include variance and covariance estimates from the model. Alongside point estimates, these values will be important for other analysts seeking to apply the results of the reported health-utility analysis to their own economic models and model settings, including the appropriate assessment of uncertainty. In particular, when health utility
is regressed against ordered levels of health states representing disease severity it would be inconsistent to have higher utility values placed on more severe health states. Inclusion of covariance information can be crucial to ensure the integrity of the ordering of health-utility values under conditions of uncertainty. Alternatively, consideration should be given to specifying a functional form for the modeling that maintains the logical ordering of health-state utilities – perhaps by parameterizing the more severe health states in terms of how much worse they are than a less severe health state and choosing appropriate distributional assumptions.

It should be recognized that this covariance information described above relates to the covariance between model parameters. For some applications of individual simulation models it may be important to also have estimates of the level of variation due to the individual variation. In a standard regression model this is given by the error term, and in a GLM/GEE framework this is given by the residual deviance. Consideration should be given to also reporting these measures – especially when it is known that the model to be informed will be designed as an individual simulation model.

Box 6. Analysis and Reporting

The data analysis approach should reflect the needs of the economic analysis and should not be constrained by traditional regulatory approach to analysing data (e.g., comparison between treatment arms).

Data are constrained to values less than 1 and are left skewed. Consider a simple linear transformation \(X = 1 - U\) and then use familiar methods for right-skewed data, e.g., generalized linear modelling or generalized estimating equations for longitudinal data, with the possibility of a two-part model to handle excess zeros.

Consider explicit modeling of prognostic factors. This has the potential to increase the generalizability of results by adjusting clinical trial data to the characteristics of populations in routine clinical practice.

Regarding the transferability of data in multinational studies:

- There are predictable differences in the way patients map to the index score of the health-utility instrument.
- Country or region effects can be handled as a covariate in statistical modeling.
- Use the country value set / tariff appropriate for the economic model audience (separate analyses may be needed for each country-specific economic model).
- Careful reporting can maximize the value of research for future economic evaluations. E.g., the effect of treatment, independent of health state, can be summarized for use in future indirect comparisons with the current drug.

Covariance can be used to retain the integrity of the logical ordering of health-utility values under conditions of uncertainty. Alternatively, consider specifying a functional form that maintains the logical ordering of health-state utilities. This is important for other analysts seeking to apply the results in future models, including appropriate assessment of uncertainty.

6 Conclusions

The quality of HSUVs is critical to the decisions being made by HTA, pricing, and reimbursement authorities that affect patients’ access to new treatments, physicians’ ability to use them, and the return that manufacturers are able to realize on their investment in developing new products. Health economists with an understanding of the economic model should be influential in the design of any studies in which
health-utility data are collected, to ensure that methods are optimized to meet the needs of economic models.

Careful planning is needed, beginning early in the product development process, to define the HSUVs that will be needed for economic models, determine which of these may be measured in the planned trial or trials, establish the optimal design for the data collection and analyses, and plan any appropriate additional health-utility research.

The design of health-utility studies should start with a clear statement of the objectives, framed in terms of the needs of the economic model. The following aspects should be carefully considered: the choice of instrument and respondents, the mode of administration, the timing and frequency of assessments, the data that should be collected at each assessment, the follow-up period, and any issues related to heterogeneity of the sample and generalizability of results. Analyses should be designed to provide HSUVs for model health states, making any appropriate adjustments to generalize the results to the population of interest in the economic model and to preserve valuable information available in the patient-level data.

7 REFERENCES


