APPENDIX A – Methodologies Used to Elicit Utility Values

The three most frequently used methods for directly eliciting utility scores includes: rating scales (RS), standard gamble (SG), and time trade-off (TTO) methodologies. Additionally, a widely used alternative to directly measuring preferences for health outcomes is to use a pre-scored multi-attribute health status classification system such as the Quality of Well-being Index (QWB), Health Utilities Index (HUI), or the Euro-Qol (EQ-5D) [1].

Rating scales are the simplest approach to directly measuring patient preferences. Individuals are generally asked to rank specific health outcomes from most to least preferred and then to place the outcomes on a scale so that the intervals between each outcome corresponds to differences in the individual’s preferences for each outcome [1]. There are many variations of rating scales with visual analog scales (VAS) being the simplest. VAS involve asking an individual to indicate on a line ranging from 0 (death) to 1 (perfect health) where his or her preference for a particular health state falls. While the VAS is efficient and easy to use, one of its potential limitations is that it does not directly provide cardinal utility values like the SG or TTO [2]. Standard gamble is considered the classic method for obtaining utilities [3-6] and is based on paired comparisons from which an individual chooses one of two health states: 1) the first choice will involve a lottery with two possible outcomes, an immediate return to perfect health with a probability (p) or immediate death (1-p) and 2) the other choice will involve experiencing a specific health state of intermediate desirability with certainty for the rest of the individual’s life. The probability of experiencing either of these two choices (p) is varied until the individual is indifferent between the lottery and the choice with the certain outcome. The utility value is the value of p at the point the individual becomes indifferent between the two choices [7,8]. The SG involves decision-making under uncertainty or risk, which reflects the uncertainty patients may have when making decisions about health care.

Time trade-off asks an individual to trade off years of life for gains in health status [7] by converting the remaining years of life in a specific health state to its equivalent time in a state of perfect health [8].
Individuals are asked to assume they are living a specific health state (usually with the disease of interest) for the remainder of their lifetime. They are then asked to indicate how many years of life in their current health state they would be willing to give up to be returned to perfect health for their remaining years of life.

SG and TTO are considered to directly measure the relative preferences for different health states [7], while a RS asks an individual to provide specific preference values for a given health state. One important difference between SG and TTO is that the SG is framed in terms of risk, while TTO is considered to be riskless [9]. Utility values obtained using SG are typically greater than those obtained from using TTO, which are greater than those using RS [8-10]. Bleichrodt [9] mentions the effect of utility curvature, probability weighting, loss aversion and scale compatibility as contributing factors for upward biases in SG utility estimates. Additionally, the SG procedure requires patients to rate health states relative to death. When death becomes a consideration, many individuals become risk averse, which leads to higher utilities. In multi-attribute classification systems, individuals are not asked to consider his or her health status relative to death and thus, there is no bias against reporting health-related problems of greater severity [11]. While SG and TTO enable the user to directly measure preferences for various health states, key challenges with these techniques include a high level of respondent burden and time involvement needed to obtain estimates.

Because measuring preferences for health outcomes can be complex and time consuming, multi-attribute health status classification systems are frequently used. The QWB, HUI, or EQ-5D assess from 4 to 8 attributes. A respondent completes a questionnaire about his or her health status, which is then scored using a multi-attribute scoring function derived from community preference measures for health states [12]. Scores are presented on a scale from 0.0 (death) to 1.0 (full or perfect health) [1]. These instruments allow for easy incorporation of HRQOL evaluation and utility assessment into clinical trial protocols. However, these systems do not directly assess utility but rather rely on regression analysis to estimate the functional relationship between the overall utility weight and the individual health domains.
from the instrument. The utilities provided are only as accurate as the scoring function used to generate the estimates.

REFERENCES


