

## Value in Health Volume 19, Issue 4

### The following articles will be included in the June 2016 issue of *Value in Health* (Volume 19, Issue 4):

Conjoint analysis methods, particularly discrete-choice experiments (DCEs), have been increasingly used to quantify preferences of patients, caregivers, physicians, and other stakeholders. Recent consensus-based guidance on good research practices, including two recent ISPOR Task Force reports, has aided in improving the quality of conjoint analyses and DCEs in outcomes research. However, uncertainty regarding good research practices for the statistical analysis of data from DCEs persists. The latest ISPOR Good Research Practices Task Force Report, published in this issue with an accompanying editorial, discusses the methods of analysing DCE data.

Best-worst scaling is gaining in popularity as a method of preference elicitation, but it is unclear whether the estimates it generates differ substantially from those produced by other commonly used methods such as discrete choice experiments. In this issue, the journal publishes a paper by Karin Groothuis-Oudshoorn et al. that compares best-worst scaling (BWS) with a discrete-choice experiment (DCE) to estimate stakeholders' risk tolerance for hip replacement surgery. The authors conclude that BWS and DCE are comparable in estimating attribute weights and level preferences, but that the risk tolerance threshold based on the estimation of maximum acceptable risk differs between the methods.

Also, as the roll out of the ED-5D-5L continues, the journal publishes a paper by Matthijs Michael Versteegh et al. reporting on the estimation of the Dutch tariff. In addition, with the growing concern over the potential budget impact of the new treatments for hepatitis C, the paper by Alexis Chidi et al. argues that restricting access to patients with advanced disease might be costlier and less cost-effective than a strategy of offering wider access.

Kind regards,

**Michael F. Drummond, MCom, DPhil** and **C. Daniel Mullins, PhD**, Co-Editors-in-Chief, *Value in Health*



### FEATURED ARTICLES:

#### **Statistical Methods for the Analysis of Discrete-Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force**

A. Brett Hauber, Juan Marcos Gonzalez, Catharina Groothuis-Oudshoorn, Thomas Prior, Deborah A. Marshall, Charles E. Cunningham, Maarten J. IJzerman, John Bridges

This report will assist researchers in evaluating and selecting among alternative approaches to conducting statistical analysis of discrete-choice experiments data. (See page 47 for summary)

#### **An Empirical Comparison of Discrete-Choice Experiment and Best Worst Scaling to Estimate Stakeholders' Risk Tolerance for Hip Replacement Surgery**

Karin G.M. Groothuis-Oudshoorn, Deborah A. Marshall, Maarten J. IJzerman, and Joris David van Dijk

The authors compare the metrics obtained from a discrete-choice experiment and profile case best-worst scaling with respect to hip replacement. (See page 47 for summary)

### ISSUE HIGHLIGHTS:

#### **DECISION-MAKER COMMENTARY**

##### **Developing a Measure of Value in Health Care**

This commentary describes how health care can use existing quality measures and cost accounting data to measure value.

#### **ECONOMIC EVALUATION**

##### **Economic and Public Health Impacts of Policies Restricting Access to Hepatitis C Treatment for Medicaid Patients**

In this study, the authors find that current Medicaid policies restricting hepatitis C treatment to patients with advanced disease are more costly and less effective than unrestricted, full access strategies.

**MORE ISSUE HIGHLIGHTS:****PATIENT-REPORTED OUTCOMES****Using the Evaluative Linguistic Framework for Questionnaires to Assess Comprehensibility of Self-Report Health Questionnaires**

This study uses an adaptation of the Evaluative Linguistic Framework to evaluate the quality of self-report questionnaires. (See below for summary)

**PREFERENCE-BASED ASSESSMENTS****Dutch Tariff for the 5-Level Version of EQ-5D**

In this article, the authors present the Dutch EQ-5D-5L valuation study. (See page 48 for summary)

**COMPARATIVE EFFECTIVENESS RESEARCH / HTA****The Impact of Patient-Centered versus Didactic Education Programmers in Chronic Patients by Severity. The Case of Type 2 Diabetes Mellitus**

This study finds that different patients are sensitive to different categories of education programs. (See page 48 for summary)

**SYSTEMATIC REVIEWS****Do Modelling Studies in Chronic Obstructive Pulmonary Disease Measure Correct Values of Utility? A Meta-Analysis**

Foruhar Moayeri, Philip Clarke, David Dunt, and Ya-Seng A. Hsueh

In this systematic review, the authors compare COPD-related utilities with alternative values used in decision models.

For all articles in this issue, and to see what services *Value in Health* provides for its authors see:

[http://www.ispor.org/valuehealth\\_index.asp](http://www.ispor.org/valuehealth_index.asp).

**Spotlight on *Value in Health***

*From Volume 19, Issue 4 (May 2016):*

**PATIENT-REPORTED OUTCOMES****Using the Evaluative Linguistic Framework for Questionnaires to Assess Comprehensibility of Self-Report Health Questionnaires** (pp. 335-342)

Rachelle Buchbinder, Rosemary Clerehan, Francis Guillemin, and Jonathan Epstein

This paper proposes a new framework for assessing the comprehensibility of self-administered health questionnaires. Health-related questionnaires are typically constructed by drawing on some combination of patient feedback, physician observation, expert opinion, theory and/or other research. While the instrument's construct validity and many other psychometric properties can then be determined empirically, there is less guidance for how to judge whether or not it will be clear to users or suitable for other contexts and cultures. The Evaluative Linguistic Framework (the ELF-Q) developed for this paper, based on systemic functional linguistic theory, provides a principled and usable approach to evaluating the linguistic quality of a health questionnaire, adapted from a tool (the ELF) previously used successfully to evaluate a range of different health care texts. The beta version of the ELF-Q was derived by applying the ELF to a health literacy questionnaire, then using this initial tool to evaluate 2 further self-report questionnaires in their English versions and in 4 other languages. The ELF-Q contains 9 items with descriptions and 34 assessment probes. These items are: the overall (generic) structure; the metadiscourse (or explanatory words about the questionnaire); the headings; the rhetorical elements – the function of each stage of the document (e.g. to inform, to instruct); the expressed writer-reader relationship; the technicality of the vocabulary; the (measured) content density; and the quality of the format. The overall judgement of the questionnaire developer can then be quantified at the end. There exists 'common-sense' approaches to appraising questionnaires, these do not appear to be grounded in any theory of language, and thus lack the robust explanatory power of the ELF-Q. It is now designed for self-training and, in conjunction with psychometric testing, can be used for choosing the most suitable questionnaire for observational research and clinical trials, or for assessing the comprehensibility of cross-cultural adaptations of questionnaires.

**PREFERENCE-BASED ASSESSMENTS****An Empirical Comparison of Discrete-Choice Experiment and Best-Worst Scaling to Estimate Stakeholders' Risk Tolerance for Hip Replacement Surgery** (pp. 316-322)

Karin G.M. Groothuis-Oudshoorn, Deborah A. Marshall, Maarten J. IJzerman, and Joris David van Dijk

This paper provides a comparison between outcomes from a discrete-choice experiment (DCE) and a profile-case best-worst scaling (BWS) task and it is evaluated whether the metrics obtained by those methods could lead to different regulatory decisions. As case, we considered the choice between two types of hip replacement procedures, where the new procedure is associated with additional

clinical benefits at the expense of higher risk on revision procedures. An online survey containing eight DCE and twelve BWS questions was distributed. Next, the estimated preference weights for the five attributes and the maximum acceptable risk (MAR) for various scenarios were determined using both methods. The MAR expresses the risk on revision that patients were willing to take to gain a certain amount of benefit such as pain relief. DCE and BWS provided comparable attribute weights and level preferences, which is in agreement with previous studies. However, although the estimated MARs for a revision procedure followed the same trend for both methods, the MARs were systematically higher in five of the six scenarios using DCE. So, although both methods estimated comparable attribute weights and level preferences, the risk tolerance threshold based on the estimation of MAR differed between both methods, possibly leading to inconsistencies when comparing treatment scenarios.

#### **Dutch Tariff for the 5 Level Version of EQ-5D** (pp. 343-352)

Matthijs Michael Versteegh, Karin Vermeulen, Silvia Evers, G. Ardine de Wit, Rilana Prenger, and Elly A. Stolk

In 2009, a new version of the EQ-5D was introduced with 5 rather than 3 answer levels per dimension, which is now known as the EQ-5D-5L. This paper describes the preference-elicitation study and subsequent estimation of a preference-based tariff based on time trade-off (TTO) results for EQ-5D-5L in The Netherlands. The study followed the TTO protocol for EQ-5D-5L valuation studies, which was developed to increase inter-country comparability of both valuation studies and outcome studies and to increase the likelihood that observed differences between EQ-5D-5L values collected in different countries reflect population preference differences rather than method heterogeneity. The new tariff is timely, since the new 2016 Dutch Guidelines for economic evaluations prescribe the use of EQ-5D-5L with a Dutch tariff as a reference case. If in certain disease areas EQ-5D-5L is not considered sufficiently sensitive, alternative measures may be presented alongside EQ-5D-5L. The new tariff for EQ-5D-5L estimated in this study produces values between -0.446 for the worst health state and 1 for the best health state, and is based on Tobit regression model with data of a face-to-face elicitation study throughout The Netherlands with 1,003 respondents. The new tariff was applied in a large dataset containing responses from 6 different patient groups who filled out both EQ-5D-3L and EQ-5D-5L; the distribution of responses suggests that the new EQ-5D-5L is better able to assign utility values to different mild health states as well as to states valued in the range between 0.3 and 0.5, bringing about the promise of increased sensitivity.

#### **COMPARATIVE EFFECTIVENESS RESEARCH / HTA**

#### **The Impact of Patient-Centered Versus Didactic Education Programs in Chronic Patients by Severity: The Case of Type 2 Diabetes Mellitus** (pp. 353-352)

Manuel García-Goñi, Paul Windrum, and Holly Coad

This paper analyzes the impact of two different types of education programs in patients recently diagnosed with type 2 diabetes mellitus. It is already known that education leads to better health-related decisions and protective behaviours. However, there are different types of education programs. Patient self-management education programs have been shown to be beneficial for patients with different chronic conditions and to have a higher impact on health outcomes than didactic education. This is especially important for chronic (such as diabetic) patients because they need to learn to live with their conditions, are the main users of health services, and account for most health expenditures. However, different to prior studies in the literature, we examine whether the initial severity of the condition is a factor affecting the relative impact of patients attending self-management and didactic education programmes. In this paper, we take advantage of a comparative trial in which we compare the glycaemic control (measured by HbA1c) of patients attending both education programs when they are initially diagnosed and one year later, using OLS regressions with fixed effects and quantile regressions techniques. Our results provide evidence of better mean glycaemic control in patients receiving the patient-centered programme, which engages better patients compared with the didactic programme, but the differential impact is non-monotonic. The patient-centered education program is significantly more beneficial for patients initially at the healthy range and also for patients with very high glycaemic level. Differently, other patients as those with the worst initial glycaemic control (far from the health range) improve equally their diabetes condition, regardless of which program they attended. Our results point to a relevant policy implication: different patients are sensitive to different categories of education programs and hence, there should be some patient stratification in the optimal design of preventive programs.