SPINE An International Journal for the study of the spine, Publish Ahead of Print

DOI: 10.1097/BRS.0000000000002975

A GENERAL POPULATION UTILITY VALUATION STUDY FOR METASTATIC EPIDURAL SPINAL CORD COMPRESSION HEALTH STATES

Markian A. Pahuta, MD[1], Eugene K. Wai, MD, MSc[2], Joel Werier, MD[2], Carl van Walraven, MD, MSc[3], Doug Coyle, PhD[4]

[1]Department of Orthopaedic Surgery, Henry Ford Health System

[2]Division of Orthopaedic Surgery, The Ottawa Hospital

[3]Ottawa Hospital Research Institute, School of Epidemiology and Public Health, University of Ottawa Institute for Clinical Evaluative Sciences

[4]School of Epidemiology and Public Health, University of Ottawa

***Corresponding Author**

Markian A. Pahuta, MD *

Department of Orthopaedic Surgery

Henry Ford Health System

2799 West Grand Blvd, CFP-6

Detroit, MI 48202

248-661-6473

248-661-7155

mark@pahutamd.com

The manuscript submitted does not contain information about medical device(s)/drug(s).

The Hans K. Uthoff Fund grant funds were received in support of this work.

No relevant financial activities outside the submitted work.

Study Design
General population utility valuation study

Objectives
This study obtained utility valuations from a Canadian general population perspective for 31 unique Metastatic Epidural Spinal Cord Compression (MESCC) health states and determined the relative importance of MESCC-related consequences on quality-of-life.

Summary of Background Data
Few prospective studies on the treatment of MESCC have collected quality-adjusted-life-year weights (termed "utilities"). Utilities are an important summative measure which distills health outcomes to a single number that can assist healthcare providers, patients, and policy makers in decision making.

Methods
We recruited a sample of 1138 adult Canadians using a market research company. Quota sampling was used to ensure that the participants were representative of the Canadian population in terms of age, gender, and province of residence. Using the validated MESCC module for the "Self-administered Online Assessment of Preferences" (SOAP) tool, participants were asked to rate 6 of the 31 MESCC health states, each of which presented varying severities of 5 MESCC-related dysfunctions (dependent; non-ambulatory; incontinent; pain; other symptoms).

Results
Participants equally valued all MESCC-related dysfunctions which followed a pattern of diminishing marginal disutility (each additional consequence resulted in a smaller incremental decrease in utility than the previous). These results demonstrate that the general population values physical function equal to other facets of quality-of-life.

Conclusions
We provide a comprehensive set of *ex ante* utility estimates for MESCC health states that can be used to help inform decision making. This is the first study reporting direct utility valuation for a spinal disorder. Our methodology offers a feasible solution for obtaining quality-of-life data without collecting generic health status questionnaire responses from patients.

Key Words
Metastatic Epidural Spine Cord Compression, Heath Related Quality-of-Life, Quality-Adjusted Life Years, Utilities, Health Economics, Spinal Cord Injury, Spinal Neoplasms, Survey, Resource Allocation, Decision Making, Health Services Research

Level of Evidence 4

Key Points:

1. We report a comprehensive set of utility estimates for MESCC health states that can be used for decision making, such as allocation of limited resources for specific treatments
2. Regression model building revealed that members of the general population value all attributes characterizing MESCC health states equally
3. Direct utility valuation over the internet is a feasible solution for obtaining quality-of-life data when generic health status questionnaire data is lacking

INTRODUCTION

The decision to offer a patient with metastatic epidural spinal cord compression (MESCC) surgery followed by radiotherapy (S+RT) or radiotherapy alone (RT) is difficult because survival, physical function, and complications must be considered jointly. Quality-adjusted-life-year (QALY) analysis could allow patients and clinicians to determine the relative weight of these factors and reach an optimal decision. QALYs are calculated using “utilities,” or health-related quality of life weights, which are usually derived from patient responses to generic health status questionnaires (e.g. EuroQol-5D, Short Form-6D, or Health Utilities Index 3).(1) Unfortunately there is a paucity of quality-of-life data for MESCC as few high-quality studies compare interventions using generic health status measures.

When generic health status measures are not available, utilities can be derived by direct valuation. Direct valuation is the classical approach in which individuals rate hypothetical health state descriptions using the time-trade-off or standard gamble procedures.(2) These procedures can be used to measure utilities for very specific and uncommon health states. Best practices in economic evaluation are to recruit a sample of healthy individuals from the general population for direct utility valuation.(3,4)

Based on the work of the European Organisation for Research and Treatment of Cancer (EORTC),(5) our group has developed a comprehensive set of 31 unique MESCC health state descriptions.(6) We have also developed an online utility valuation module for MESCC using the “Self-administered Online Assessment of Preferences” (SOAP) tool. This module was found to be valid, reproducible and responsive in a sample of individuals how have not experienced MESCC.(6)

The primary objective of this study is to obtain general population utility valuations for the 31 unique MESCC health states. The secondary objective of this study is to determine the relative importance of various aspects to quality-of-life in MESCC.

MATERIAL AND METHODS

Subjects

Recruiting a general population probability sample is a time-intensive and costly undertaking. Typically this is done by visiting homes, using random-digit-dialing, and selecting random phone numbers. In Canada, with over 80% of Canadians age 16 years and older having access to the internet,(7) web recruitment offers a practical alternative to traditional strategies. As has been done by several other investigators, we recruited participants from a proprietary market research panel (Toluna Group Ltd) for this utility elicitation study for MESCC.(8–10)

The market research panel was composed of over 80,000 individuals across Canada recruited by the company into a panel of potential survey participants through random-digit-dialing, internet banner advertisements, and partnerships with corporations to become panel members.(11) Panel members agree to be contacted about new surveys. We did not provide an incentive for participating in our study. However, the market research company managing the panel does award monthly prizes to panel members based on the number and length of surveys completed. Quota sampling was used to ensure that the sample of the market research panel represented the general Canadian population in terms of region of residence, gender, and age based on the 2016 Canadian Census (Supplemental Table 2.1, http://links.lww.com/BRS/B409). The market research company sent panel members an email invitation to participate in our study. Interested panel members were redirected to a secure website hosting the utility valuation exercise.

Survey Procedures

Participants were asked to value six health states in the online SOAP MESSC module. The SOAP MESCC module has previously been described.(6) Briefly, the first three pages of the module explained the utility valuation task, and provided an overview of MESCC. Following this, participants completed the standard gamble health state valuation exercises.(6)

Health states were derived from the EORTC item bank and were presented as a point-form description of five dysfunctional attributes: dependent (D), non-ambulatory (N), incontinent of urine (I), pain (P). EORTC items are restricted to health-related quality of life, and did not specifically consider the morbidity of treatment and adverse events Therefore, we reviewed prospective studies on MESCC to identify descriptors of these aspects.(6) To develop a manageable decision analytic model, the morbidity of treatment and adverse events were grouped as an "other symptoms"(S) attribute. Each attribute was characterized by the presence (+) or absence (-) of the dysfunctional attribute. Dependence was described as "you need help with eating, dressing, washing yourself or using the toilet. You are dependent on others." Other symptoms were described as "You have one or more of: nausea, vomiting, shortness of breath, lack of appetite, diarrhea." Valuations were obtained using standard gamble method.(12)

In this study, "perfect" health is represented by the fully functional health state (D-, N-, I-, P-, S-). To eliminate any bias introduced by respondents assuming different life expectancies for each scenario, all scenarios were framed as having a certain life expectancy of five years; that is, for both the probe health scenario, and perfect health, participants were told their life expectancy would certainly be five years.(13) Five years was the maximum survival reported in a randomized controlled trial on treatments for MESCC.(14)

Participants were asked to value six health states. Two health states formed a "test pair." These pairs shared non-dysfunctional attributes (eg. D-, N+, I-, P-, S- and D-, N+, I-, P- , S+). The other four health states were selected at random.

Statistical Analysis

We deemed participants to have misunderstood the task, or not engaged, if: they provided an illogical valuation for the test pair; or provided the same valuation for more than 4 health states. Such participants were excluded from the analysis. A sensitivity analysis including all participants is provided in Supplemental Text 3, http://links.lww.com/BRS/B409.

Utility values lie between 0 and 1 with 0 representing death and 1 perfect health. Therefore beta regression analysis was chosen to explore the relationship between health attributes and utility valuation (Supplemental Text 1, , http://links.lww.com/BRS/B409).(15) Regression was performed on the mean (μ) of the beta model. The dispersion parameter was treated as a constant.(15) To account for the presence of six observations per participant (indexed by i), we incorporated a random intercept term for each participant (ϵ_i) in the model for μ.(16)

The SOAP MESCC module is designed for *ex ante* (without experience) utility valuation from a general population sample. These utilities should reflect population preferences for the health states and were used to compute expected (mean) quality-of-life changes in economic analysis.

Within this context, the objective of the analysis was to estimate mean utilities for the population, not predict individual utility valuations. Consequently, goodness-of-fit measures such as the R^2 statistic or root mean squared (RMSE) are not appropriate because they quantify how well the model fits *individual* observations.(17) To instead quantify the performance of the regression model in estimating the mean utility for the 31 health states relative to direct estimation of mean utilities, we used the absolute agreement intraclass correlation coefficient (ICC). For this ICC calculation, health states were *targets* and utilities were *ratings*.(18) A two-way model was used because ratings can only be obtained by direct estimation or regression estimation.

In an effort to strengthen the generalizability of the regression analysis, we implemented internal validation by allocating participants to a test set and validation set in a 1:1 ratio.(19) Regression models were fit using the test set, and the optimal model was identified by jointly the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).(20) The optimal regression model was used to compute mean utilities for each health state (regression estimation of means). Mean utilities for each health state were then directly estimated using the validation set (direct estimation of means). The performance of the optimal regression model was quantified using the ICC by comparing mean utilities obtained from the regression model to mean utilities obtained by direct estimation.

Four regression models were considered. In Model 1, each dysfunctional attribute was coded as a categorical variable. In Model 2, the number of dysfunctional attributes was coded as nominal categorical variables. This strategy was used because it avoided assumptions of linear or extra-linear effects. Model 3 incorporated all first-order interactions in Model 1. Model 4 combined Model 1 and Model 2.

Before proceeding with model selection, we graphically checked all four models for misspecification of the variance and link function.(21) We attempted to simplify the optimal model using standard regression model building procedures.(22) These procedures ultimately led us to consider two additional models. In Model 5, the square root of the number dysfunctional attributes (num dys) was included as a continuous predictor. In Model 6, the natural logarithm of the number dysfunctional attributes (num dys) was included as a continuous predictor. The formula for the beta regression linear predictor for each model is outlined in Supplemental Text 3, http://links.lww.com/BRS/B409.

Sample Size Calculation

The sample size for the regression analysis was based on a commonly cited rule-of-thumb that linear regression requires 10 to 20 subjects per variable.(23) Our regression models considered a total of 21 variables: four numerical variables, five categorical variables, ten interaction terms, and two intercepts. Following this rule of thumb, we would require $15 \times 21 = 315$ subjects. As we planned to split into a test and validation set in a 1:1 ratio, we required 630 subjects. However, to ensure at least 2 participants in each quota, 650 participants were required (Supplemental Table 2.1, http://links.lww.com/BRS/B409).

RESULTS

1138 panel members logged into the SOAP MESCC module and provided consent. However, 488 were excluded for not valuing all health states, providing an illogical valuation for the test pair, or providing invariant responses (Figure 1). Of the 1029 participants who valued six health states, 379 (36.8%) were excluded for not having understood or engaged in the task.

The socio-demographic characteristics of 650 included and 488 excluded participants are shown in Table 1. There were no statistically or qualitatively significant differences between the groups in terms of sex, age, or province of residence.

Mean valuations for the 31 health states tended to decrease with an increasing number of dysfunctions (Table 2). For example, the mean utility valuation for 1, 2, 3, 4, and 5 dysfunctions was 0.691, 0.581, 0.471, 0.364, and 0.333 respectively (p-value for trend <0.0001).

The functional form of each model and example mean utility calculations are provided in Supplemental Text 3, http://links.lww.com/BRS/B409. Calibration and agreement parameters for all candidate models are provided in Supplemental Table 3.3, http://links.lww.com/BRS/B409. The model selection procedure identified the square root of the number of dysfunctions to be the optimal predictor of mean utility based on both the AIC and BIC criteria (Table 3).

The coefficients for Model 4 are given in Table 4. This model was well calibrated, estimating mean utilities in the validation set with a mean absolute error of 0.047 and mean signed difference of -0.02 (both of which are less than the MCID of 0.05 for utilities)(Figure 2).(24,25) There was excellent agreement between mean utilities obtained from this regression model and mean utilities obtained by direct estimation with an ICC of 0.936 (95% CI: 0.86, 0.97).(26)

Mean and median utility valuations were not sensitive to the inclusion of all participants (Supplemental Table 3.6, http://links.lww.com/BRS/B409).

Formulas for the linear predictor in the beta regression are provided in Supplemental Table 3.1, http://links.lww.com/BRS/B409. Formulas for mean utility valuation for a given MESCC health state using the fitted models are provided in Supplemental Table 3.5, http://links.lww.com/BRS/B409. To illustrate the use of these formulas, we will calculate mean utilities for the health state D-, N+, I-, P+, S- using the optimal Model 4. This health state is coded as: $\text{num dys} = 2$. Therefore using Model 4 the mean utility valuation is calculated as:

$$\text{inverse logit}\left(1.71 - 1.11 \times \sqrt{2}\right)$$

$$= 0.53.$$

DISCUSSION

In this paper we report directly valued utilities for MESCC health states using the validated SOAP MESCC module.(6) The study sample was representative of the population in all English-speaking Canadian provinces. Our findings are valid with 63.2% of participants having understood and engaged in the task. This rate is superior to validity rates reported in general population direct valuation studies for the EuroQol-5D.(27–30)

The regression model building exercise revealed that members of the general population value all dysfunctions characterizing MESCC health states equally. Furthermore, dysfunction follows a pattern of diminishing marginal disutility. That is, each additional dysfunction effects a smaller incremental change in utility than the previous dysfunction. These results demonstrate that from the societal perspective, physical function is valued equal to other facets of well-being.

It must be recognized that *ex post* utilities are not equivalent to *ex ante* utilities obtained from patients who have experienced the health states.(31) In part due to adaptation, patients tend to provide higher valuations for health states which predominantly affect physical health than the general population for the same health state. The *ex ante* utilities collected in this study are highly appropriate for facilitating healthcare decision making and can be used to conduct decision analysis and cost-utility analysis for MESCC.(32)

Although it may seem that applying lower *ex ante* utilities may infringe on patient autonomy and deny care, healthcare system decision making impacts patients with various conditions. If the objective of healthcare decision making is to maximize the benefit of all patients, utilities across different disease must be comparable in order to set priorities. Rawles argues that *ex ante* utilities can be used ethically if valued under a "veil of ignorance" (33). If we assume that the general population providing *ex ante* utility valuations may eventually develop the condition of interest, out of self-interest, they should provide fair valuations. Although *ex ante* utilities are theoretically restricted to system policy decisions, *ex ante* utilities have become the de facto standard for individual patient decisions. Utilities obtained from generic health surveys such as the EuroQol-5D, Short Form-6D, and Health Utilities Index 3 are actually *ex ante* valuations.(1) Therefore we have chosen to evaluate *ex ante* SOAP MESCC module to conform with conventions in the literature.

Utility valuations for a single health state were highly variable across participants. This was evidenced by wide Inter-quartile ranges. However, since our objective was to measure general population *ex ante* utilities for health policy decision making, the expected values and the underlying uncertainty represented by the 95% confidence intervals are pertinent.

Health policy decisions concern the allocation of scarce resources to health programs with superior cost-effectiveness ratios at the expense of those with inferior cost-effectiveness ratios.(1) If cost-effectiveness ratios are uncertain, there is a risk that a wrong policy decision will be made (ie. funding the program with an inferior cost-effectiveness ratio). Even if the risk of a wrong policy decision is large (ie. great uncertainty in cost-effectiveness ratio), from both ethical and economic perspective a decision should still be made.(34) From an ethical perspective, deferring a decision (ie. not funding any health program under consideration) denies care to several groups of patients. From the economic perspective, convincing arguments have been given to defend the practise of expected value decision making – i.e. the irrelevance of inference with respect to public decision making.(35) For these reasons, unbiased estimates of costs and effects are more important than precise estimates.

The results of the regression model building exercise have relevance for clinicians counselling MESCC patients regarding treatment. As dysfunctions follow a pattern of diminishing marginal disutility, each dysfunction is valued equal to the others – it is the total number of dysfunctions that drive quality-of-life. Surgeons should be cognizant that ambulation and continence, which are dysfunctions addressed by surgery, are no more important than other attributes (pain, other symptoms, and level of independence). Furthermore, attempting to reverse a single dysfunction in a patient with high functional status will lead to a greater increase in quality-of-life relative to a patient with low functional status.

To our knowledge, this is the first study reporting direct utility valuation for a spinal disorder. Utilities for spine disease specific instruments such as the NDI, ODI, and SRS-22 have been developed using an indirect "cross-walk" protocol.(36–39) Patient responses are collected using both the disease specific instrument and a generic health measure (eg. EuroQol-5D, Short Form-6D, and Health Utilities Index 3). Next, a regression model is developed to relate the disease specific score to the generic health measure. Then the regression model estimated generic health measure score can be used to compute a utility. However, utilities obtained from generic health measure scores are actually computed from another regression model relating the generic health measure score to directly valued utilities.(1) Our study demonstrates that disease specific direct utility valuation is feasible and valid. Investigators may consider applying our study protocol and the validated SOAP tool to other disease contexts. This approach may eliminate the potential error of the cross-walk approach introduced by the need for two regression models.

CONCLUSION

We provide a comprehensive set of *ex ante* utility estimates for MESCC health states. The utility values derived from this study can be used to help inform population level healthcare decision making, such as allocation of limited resources for specific treatments. The results of this study can also help clinicians counsel MESCC patients regarding treatment.

This is the first study reporting direct utility valuation for a spinal disorder. We demonstrate that direct utility valuation over the internet is a feasible solution for obtaining quality-of-life data when generic health status questionnaire data is lacking. Investigators may consider applying our approach to other disease contexts.

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes [Internet]. Oxford University Press, USA; 2005 [cited 2012 Apr 29]. 400 p. Available from: http://www.amazon.com/Methods-Economic-Evaluation-Health-Programmes/dp/0198529457
2. Furlong W, Feeny D, Torrance GW, Barr R, Horsman J. Guide to design and development of health-state utility instrumentation [Internet]. Hamilton; 1990. (CHEPA Working Paper Series). Report No.: 90–9. Available from: http://www.chepa.org/Files/Working Papers/WP 90-9.pdf
3. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. JAMA [Internet]. 1996 Oct 16 [cited 2012 Apr 15];276(15):1253–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8849754
4. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. JAMA - J Am Med Assoc. 2016;316(10):1093–103.
5. Mitera G, Loblaw A, Sahgal A, Danielson B. Canadian-led International Development of a European Organization for Research and Treatment of Cancer Quality of Life Module for Malignant Spinal Cord Compression: Results of Phase I. Int J Radiat Oncol [Internet]. 2010 Nov;78(3):S604. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0360301610023801
6. Pahuta M, Frombach A, Hashem E, Spence S, Sun C, Wai EK, et al. The Psychometric Properties of a Self-Administered, Open-Source Module for Valuing Metastatic

Epidural Spinal Cord Compression Utilities. PharmacoEconomics - open [Internet]. 2018 Sep 3; Available from: http://www.ncbi.nlm.nih.gov/pubmed/30178420

7. Haight M, Quan-Haase A, Corbett B a. Revisiting the digital divide in Canada: The impact of demographic factors on access to the Internet, level of online activity, and social networking site usage. Information, Commun Soc [Internet]. 2014;17(4):503–19. Available from: http://www.tandfonline.com/doi/abs/10.1080/1369118X.2014.891633%5Cnhttp://dx.doi.org/10.1080/1369118X.2014.891633

8. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian Valuation of EQ-5D Health States: Preliminary Value Set and Considerations for Future Valuation Studies. van Baal PHM, editor. PLoS One [Internet]. 2012 Feb 6 [cited 2012 Jul 5];7(2):e31115. Available from: http://dx.plos.org/10.1371/journal.pone.0031115

9. Augustovski F, Rey-Ares L, Irazola V, Oppe M, Devlin NJ. Lead versus lag-time trade-off variants: does it make any difference?. Eur J Health Econ [Internet]. 2013;14 Suppl 1:S25-31. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23900662

10. Augestad LA, Stavem K, Kristiansen IS, Samuelsen CH, Rand-Hendriksen K. Influenced from the start: anchoring bias in time trade-off valuations. Qual Life Res [Internet]. 2016;25(9):2179–91. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=27016943

11. Toluna Ltd. Esomar 26: 26 questions to help research buyers of online samples. 2013.

12. Von Neumann J, Morgenstern O. Theory of games and economic behavior. Princeton: Princeton University Press; 1953.

13. Stiggelbout AM, Kiebert GM, Kievit J, Leer JW, Stoter G, de Haes JC. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. Med Decis Making [Internet]. 1994;14(1):82–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8152360
14. Patchell R a., Tibbs P a., Regine WF, Payne R, Saris S, Kryscio RJ, et al. Direct decompressive surgical resection in the treatment of spinal cord compression caused by metastatic cancer: a randomised trial. Lancet (London, England) [Internet]. 2005 Aug 20 [cited 2012 Apr 15];366(9486):643–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16112300
15. Ferrari S, Cribari-Neto F. Beta Regression for Modelling Rates and Proportions. J Appl Stat [Internet]. 2004 Aug [cited 2013 Mar 8];31(7):799–815. Available from: http://www.tandfonline.com/doi/abs/10.1080/0266476042000214501
16. Gardiner JC, Luo Z, Roman LA. Fixed effects, random effects and GEE: what are the differences? Stat Med [Internet]. 2009 Jan 30;28(2):221–39. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19012297
17. Picard RR, Cook RD. Cross-Validation of Regression Models. J Am Stat Assoc [Internet]. 1984 Sep;79(387):575–83. Available from: http://www.tandfonline.com/doi/abs/10.1080/01621459.1984.10478083
18. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Vol. 86, Psychological bulletin. 1979. p. 420–8.
19. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol [Internet]. 2001 Aug;54(8):774–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11470385
20. Kuha J. AIC and BIC: Comparisons of Assumptions and Performance. Sociol Methods

Res [Internet]. 2004 Nov 1;33(2):188–229. Available from: http://smr.sagepub.com/cgi/doi/10.1177/0049124103262065

21. McCullagh P, Nelder J. Generalized Linear Models. 2nd ed. London: Chapman & Hall/CRC; 1989. 391-418 p.

22. Harrell FE. Regression Modeling Strategies [Internet]. 2nd ed. Cham: Springer International Publishing; 2015. (Springer Series in Statistics). Available from: http://link.springer.com/10.1007/978-3-319-19425-7

23. Green SB. How Many Subjects Does It Take To Do A Regression Analysis. Multivariate Behav Res [Internet]. 1991;26(3):499–510. Available from: http://www.tandfonline.com/doi/abs/10.1207/s15327906mbr2603_7?journalCode=hmbr20

24. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes [Internet]. 2007 Nov [cited 2012 Apr 1];5(1):70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9366889

25. Wee H-L, Li S-C, Xie F, Zhang X-H, Luo N, Feeny D, et al. Validity, feasibility and acceptability of time trade-off and standard gamble assessments in health valuation studies: a study in a multiethnic Asian population in Singapore. Value Health [Internet]. 2008;11 Suppl 1:S3-10. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=18387064

26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.

27. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian Valuation of EQ-5D Health States: Preliminary Value Set and Considerations for Future Valuation Studies. van

Baal PHM, editor. PLoS One [Internet]. 2012 Feb 6 [cited 2012 Jul 5];7(2):e31115. Available from: http://dx.plos.org/10.1371/journal.pone.0031115

28. Lamers LM, Stalmeier PFM, Krabbe PFM, Busschbach JJ V. Inconsistencies in TTO and VAS values for EQ-5D health states. Med Decis Mak. 2006;26(2):173–81.

29. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? Med Care. 2005;43(3):221–8.

30. Devlin NJ, Hansen P, Kind P, Williams A. Logical inconsistencies in survey respondents' health state valuations - A methodological challenge for estimating social tariffs. Health Econ. 2003;12(7):529–44.

31. Nord E, Daniels N, Kamlet M. QALYs: Some Challenges. Value Heal [Internet]. 2009 Mar [cited 2012 Apr 15];12(10):S10–5. Available from: http://dx.doi.org/10.1111/j.1524-4733.2009.00516.x

32. Drummond M, Brixner D, Gold M, Kind P, McGuire A, Nord E. Toward a Consensus on the QALY. Value Heal [Internet]. 2009 Mar [cited 2012 Apr 29];12:S31–5. Available from: http://dx.doi.org/10.1111/j.1524-4733.2009.00522.x

33. Rawls J. A Theory of Justice: Revised Edition [Internet]. 2nd ed. Cambridge, Massachusetts, United States: Belknap Press; 1999 [cited 2012 Apr 29]. 560 p. Available from: http://www.amazon.com/A-Theory-Justice-John-Rawls/dp/0674000781

34. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ [Internet]. 1999 Jun [cited 2012 Nov 15];18(3):341–64. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0167629698000393

35. Arrow KJ, Lind RC. Uncertainty and the Evaluation Decisions of Public Investment. Am Econ Rev. 1970;60(3):364–78.

36. Richardson SS, Berven S. The development of a model for translation of the Neck Disability Index to utility scores for cost-utility analysis in cervical disorders. Spine J [Internet]. 2012 Jan;12(1):55–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22209244

37. Carreon LY, Anderson P a, McDonough CM, Djurasovic M, Glassman SD. Predicting SF-6D utility scores from the neck disability index and numeric rating scales for neck and arm pain. Spine (Phila Pa 1976) [Internet]. 2011;36(6):490–4. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3005013&tool=pmcentrez&rendertype=abstract

38. Carreon LY, Glassman SD, McDonough CM, Rampersaud R, Berven S, Shainline M. Predicting SF-6D utility scores from the Oswestry disability index and numeric rating scales for back and leg pain. Spine (Phila Pa 1976) [Internet]. 2009 Sep 1;34(19):2085–9. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00007632-200909010-00020

39. Bridwell KH, Cats-Baril W, Harrast J, Berven S, Glassman S, Farcy J-P, et al. The validity of the SRS-22 instrument in an adult spinal deformity population compared with the Oswestry and SF-12: a study of response distribution, concurrent validity, internal consistency, and reliability. Spine (Phila Pa 1976) [Internet]. 2005 Feb 15;30(4):455–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15706344

Table 1. Baseline characteristics of participants

		Included N = 650	Excluded N = 488
Sex – no. (%)			
	Female	331 (51)	234 (48)
	Male	319 (49)	254 (52)
Age – no. (%)			
	18 – 24 yrs	78 (12)	41 (8)
	25 – 34 yrs	113 (17)	76 (16)
	35 – 44 yrs	106 (16)	83 (17)
	45 – 54 yrs	118 (18)	91 (19)
	55 – 64 yrs	107 (17)	79 (16)
	≥ 65 yrs	128 (20)	118 (24)
Province – no. (%)			
	Maritimes	57 (9)	40 (8)
	Ontario	326 (50)	265 (54)
	Manitoba	30 (5)	27 (6)
	Saskatchewan	26 (4)	9 (2)
	Alberta	98 (15)	65 (13)
	British Columbia	113 (17)	82 (17)

Table 2. Utility valuations for all MESCC health states

Health State	Number of Dysfunctions	N	Mean (95% CI)	Median (IQR)
D+N-I-P-S-	1	181	0.676 (0.628, 0.725)	0.75 (0.45, 0.95)
D-N+I-P-S-	1	185	0.717 (0.671, 0.763)	0.85 (0.55, 0.95)
D-N-I+P-S-	1	187	0.681 (0.632, 0.73)	0.85 (0.45, 0.95)
D-N-I-P+S-	1	182	0.695 (0.647, 0.743)	0.85 (0.512, 0.95)
D-N-I-P-S+	1	202	0.685 (0.639, 0.732)	0.85 (0.462, 0.95)
D+N+I-P-S-	2	88	0.548 (0.474, 0.623)	0.65 (0.225, 0.9)
D+N-I+P-S-	2	108	0.619 (0.554, 0.684)	0.75 (0.4, 0.95)
D+N-I-P+S-	2	106	0.542 (0.475, 0.608)	0.575 (0.25, 0.85)
D+N-I-P-S+	2	112	0.6 (0.534, 0.666)	0.7 (0.35, 0.95)
D-N+I+P-S-	2	88	0.659 (0.589, 0.728)	0.75 (0.45, 0.95)
D-N+I-P+S-	2	107	0.57 (0.505, 0.635)	0.65 (0.325, 0.875)
D-N+I-P-S+	2	105	0.513 (0.44, 0.586)	0.6 (0.1, 0.9)
D-N-I+P+S-	2	100	0.583 (0.514, 0.652)	0.65 (0.35, 0.9)
D-N-I+P-S+	2	89	0.601 (0.525, 0.676)	0.65 (0.3, 0.95)
D-N-I-P+S+	2	113	0.582 (0.514, 0.65)	0.65 (0.25, 0.95)
D+N+I+P-S-	3	81	0.512 (0.428, 0.596)	0.55 (0.1, 0.9)
D+N+I-P+S-	3	76	0.443 (0.363, 0.524)	0.45 (0.1, 0.75)
D+N+I-P-S+	3	80	0.492 (0.413, 0.572)	0.5 (0.075, 0.75)
D+N-I+P+S-	3	88	0.445 (0.367, 0.524)	0.45 (0, 0.75)
D+N-I+P-S+	3	96	0.491 (0.415, 0.566)	0.5 (0.1, 0.85)
D+N-I-P+S+	3	88	0.479 (0.398, 0.56)	0.6 (0, 0.813)
D-N+I+P+S-	3	86	0.478 (0.401, 0.554)	0.525 (0.1, 0.75)
D-N+I+P-S+	3	81	0.407 (0.329, 0.485)	0.45 (0, 0.75)
D-N-I+P+S+	3	91	0.501 (0.426, 0.576)	0.5 (0.15, 0.85)
D-N+I-P+S+	3	81	0.457 (0.374,	0.45 (0, 0.8)

			0.539)	
D+N+I+P-S+	4	192	0.408 (0.357, 0.46)	0.35 (0, 0.75)
D+N+I+P+S-	4	199	0.359 (0.308, 0.41)	0.25 (0, 0.65)
D+N+I-P+S+	4	191	0.342 (0.292, 0.391)	0.25 (0, 0.65)
D+N-I+P+S+	4	173	0.357 (0.303, 0.411)	0.25 (0, 0.65)
D-N+I+P+S+	4	170	0.351 (0.297, 0.404)	0.25 (0, 0.65)
D+N+I+P+S+	5	174	0.333 (0.281, 0.384)	0.2 (0, 0.638)

Table 3. Selection parameters for all models

	Description	AIC	BIC
Model 1	Attributes	-2750.242	-899.148
Model 2	Number of dysfunctions (nominal variable)	-2751.778	-906.2593
Model 3	Attributes and second-order interactions	-2747.422	-840.5717
Model 4	Square root of number of dysfunctions (continuous variable)	-2756.853	-928.0612
Model 5	Logarithm of number of dysfunctions (continuous variable)	-2754.902	-926.1101

Table 4. Summary of regression analysis using Model 4

	Variable	Coefficient	SE	t-value	p-value
μ coefficients	Intercept	1.70580	0.11594	14.71	<0.0001
	Square root of number of dysfunctions	-1.11247	0.07175	-15.51	<0.0001
σ coefficient	Intercept	0.74727	0.02139	34.94	<0.0001

Figure 1. Flow of participants

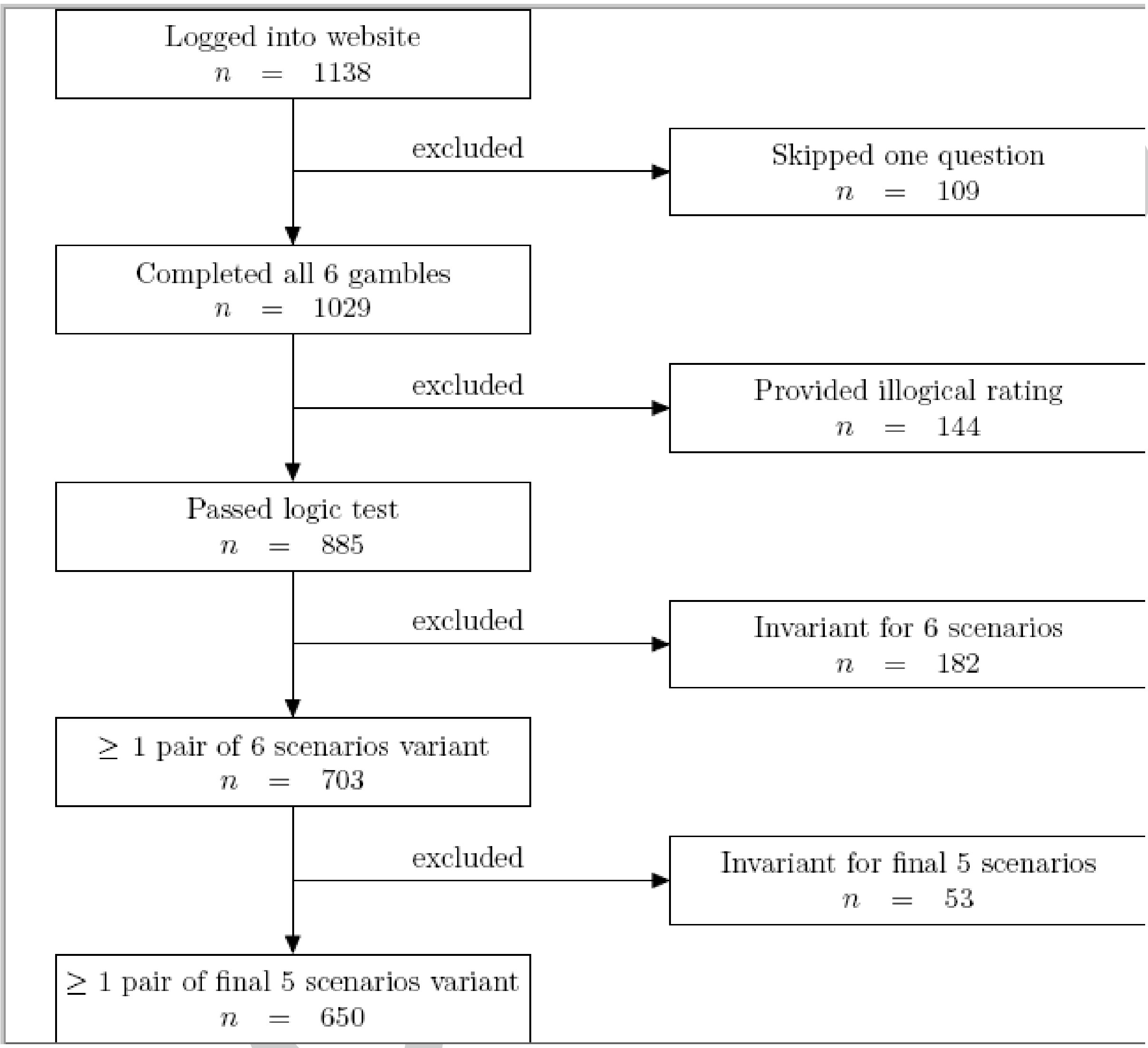

Figure 2. Calibration Plot for Model 4