METHODOLOGICAL ARTICLES

An Expert on Every Street Corner? Methods for Eliciting Distributions in Geographically Dispersed Opinion Pools

Daniel Sperber, MSc[1],*, Duncan Mortimer, PhD[1], Paula Lorgelly, PhD[1], David Berlowitz, PhD[2,3,4]

[1]Centre for Health Economics, Monash University, Melbourne, Australia; [2]Institute for Breathing & Sleep, Melbourne, Australia; [3]Austin Health, Melbourne, Australia; [4]University of Melbourne, Melbourne, Australia

A B S T R A C T

Recent publications outline developments in eliciting probabilistic opinions from clinical experts with which to inform structural assumptions and parameter estimates in health economic models. We outline approaches taken to date to elicit probabilistic distributions from experts within the health economic literature and outline the appropriate considerations and the resulting process in developing a new elicitation program with the aim of allowing low-cost elicitation of expert opinion from a heterogeneous and geographically dispersed opinion pool while preserving the essential features of good practice elicitation methods.

Keywords: clinical opinion, elicitation, expert opinion.

Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Recent "best-practice" guidelines in decision analytic modeling [1] suggest that clinical opinion may be included in economic models where necessary and well documented, despite broad concerns over its nonexperimental nature, potential overconfidence, and general susceptibility to biases and heuristics in memory and decision making. If clinical opinion is to be used influentially in decision models, further examination of elicitation methods may be appropriate, especially around the elicitation of uncertain probabilistic parameters. A small literature has developed within the field regarding the direct elicitation of probability distributions from clinicians and clinical researchers [2–5]). Existing examples of good practice such as the Sheffield Elicitation Framework (SHELF) [6] rely on personal interview and are therefore of limited assistance in eliciting opinion from geographically dispersed opinion pools. Our elicitation exercise seeks to build upon existing applied examples by attempting to emulate aspects of SHELF in a spreadsheet program, thereby allowing for wider dissemination and completion, to inform the economic evaluation of a multicenter clinical trial in a rare treatment area: sleep apnea related to acute quadriplegia.

Methods

A targeted review of elicitation methods in health economic applications was undertaken. We also reviewed guidelines for elicitation methodology including the psychology of elicitation, visual aids, and statistical methods for probabilistic elicitation and response aggregation. SHELF [6] was identified as an example of good practice, having been recently developed by an experienced group of elicitation methodologists in concert with an elicitation textbook [7] and a series of freely available research tools and programs.

It was not clear, however, that the personal interview program embodied in the SHELF recommendations was appropriate for our potential respondent pool. In an effort to design an elicitation framework suitable for the collection of opinions from a geographically dispersed, heterogeneous, and very busy group of respondents, we emulated aspects of the SHELF framework where possible and incorporated necessary departures in the SHELF framework with respect to elicitation survey design, expert selection, calibration, and weighting. An Excel spreadsheet was developed, pilot tested, and is currently being disseminated to a population of researchers and clinicians.

Results

Literature Review

The literature review identified a number of existing protocols for expert elicitation [2–6]. The method adopted by Stevenson et al. [2] conforms closely to good-practice guidelines arising from the general probability elicitation literature [7–10] in which in-person

* Address correspondence to: Daniel Sperber, Room 278, Centre for Health Economics, Building 75, Monash University, Clayton, Victoria, Australia, 3800.
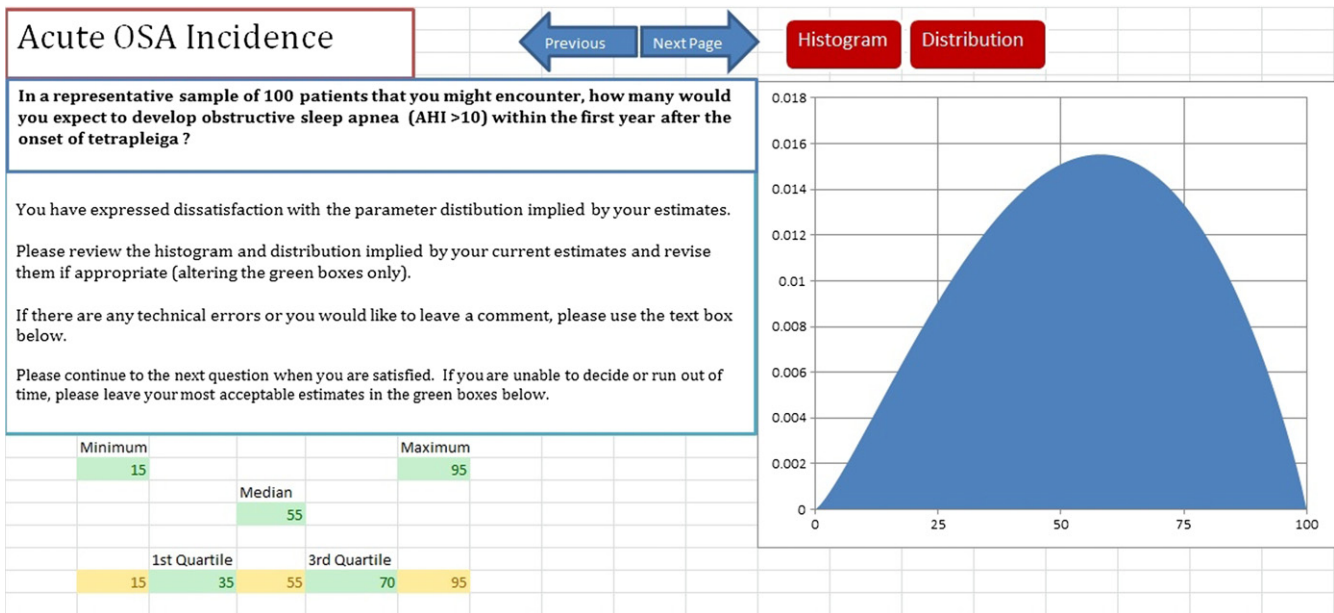
E-mail: Daniel.sperber@monash.edu.

Fig. 1 – Example elicitation page. OSA, obstructive sleep apnea; AHI, Apnea–Hypopnea Index.

elicitation by an experienced facilitator is considered to be a first-best approach in eliciting an expert's true parameter estimate.

The major applied challenges with intensive, in-person elicitation programs relate to expert recruitment, which may be constrained by geography, accessibility, and resources. Restricting the opinion pool to a small sample of geographically proximate and highly committed collaborating experts—as is typically necessary for in-person elicitation—may amplify biases arising from local paradigms, self-interest, or shared exposure to unrepresentative clinical experience or a selected body of research among a group of peers. An online tool for the SHELF program has recently become available, which may provide another means of elicitation at a distance [11].

The "Excel spreadsheet" approach of Leal et al. [3], Bojke et al. [4], and Soares et al. [5] is engineered to encourage ease of use and would allow dissemination to a wider opinion pool. For example, Leal et al. and Bojke et al. e-mailed their survey to a relatively wide opinion pool of approximately 15 identified experts, resulting in a response rate of approximately 30% to 50% of the sample frame. Soares et al. used the spreadsheet format as an aid to a large group elicitation of local nurses. We considered that each of these existing Excel-based protocols lacked one or more elements of best practice (e.g., provision of visual and quantitative feedback, anchoring and adjustment errors, and appropriate aggregation methods) as described by O'Hagan et al. [7] and reflected in the SHELF framework [6] and Stevenson et al. [2].

In practice, it may be the case that there exists a trade-off between the ability of an elicitation exercise to survey the "true" personal belief of an expert and to be a palatable and feasible undertaking for a wide cross section of respondents.

Program Development

An Excel workbook was programmed to create an elicitation experience evoking that of a structured interview. Respondents begin with reading approximately 10 minutes of context: the rationale for the elicitation, an introduction to the elicitation process, and a summary of common biases and heuristics in decision making. Following a stylized trial elicitation and a calibration question, they are asked to estimate the probability distribution describing their uncertainty regarding two structural assumptions in our decision model.

A quartile-bisection approach was adopted to elicit respondent distributions and attempt to mitigate the impact of anchoring effects. Instantaneous visual feedback with regard to the impact of quartile estimates was provided by a histogram and beta distribution fitted by least squares using Easyfit software (Figure 1) [12].

The prepared elicitation program departed markedly from the SHELF process in conducting elicitation at a distance, by foregoing a group dynamic, and by diluting the implicit "expert" standard. As such, the program is not a SHELF implementation but an effort to emulate selected points of emphasis in the elicitation recommendations within a spreadsheet program. Direct adaptations include the quartile elicitation procedure, the emphasis on visual feedback to the respondent, and the introductory expert training material [13] (see Appendix 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2012.10.011).

Our elicitation plan sought to avoid the various problems associated with restricting the opinion pool to local collaborating experts by distributing a spreadsheet program widely: to published authors in our research area, international trial collaborators, and unaffiliated physicians identified via the Australian Medical Register. Given the lack of an in-person interviewer to dissuade irrational responses, the handling of potential irrational or faulty elicitation responses was of particular concern. In light of this concern, we tested a variety of response aggregation methods, the most established of which were an unweighted linear opinion pool and a weighted linear opinion pool informed by "best estimate" fractions [4] arising from Monte Carlo sampling of the calibration distribution against a relatively known parameter distribution (in our case, informed by the sole relevant clinical publication in the area) [14].

We also tested the behavior of a couple of arbitrary weighting methods: one including only our clinical collaborator, to reflect the impact of not performing a broader elicitation exercise on our model, and another using only the median estimates provided by respondents (which we assume to be reflective of the deterministic estimates they might have provided to another research design, and to which we fitted a monomodal beta distribution with the specified mean, and a maximized SD).