# Determining Criteria and Weights for Prioritizing Health Technologies Based on the Preferences of the General Population: A New Zealand Pilot Study

Trudy Sullivan, PhD[1],*, Paul Hansen, PhD[2]

[1]Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand; [2]Department of Economics, University of Otago, Dunedin, New Zealand

A B S T R A C T

**Objectives:** The use of multicriteria decision analysis for health technology prioritization depends on decision-making criteria and weights according to their relative importance. We report on a methodology for determining criteria and weights that was developed and piloted in New Zealand and enables extensive participation by members of the general population. **Methods:** Stimulated by a preliminary ranking exercise that involved prioritizing 14 diverse technologies, six focus groups discussed what matters to people when thinking about technologies that should be funded. These discussions informed the specification of criteria related to technologies' benefits for use in a discrete choice survey designed to generate weights for each individual participant as well as mean weights. A random sample of 3218 adults was invited to participate. To check test-retest reliability, a subsample completed the survey twice. Cluster analysis was performed to identify participants with similar patterns of weights. **Results:** Six benefits-related criteria were distilled from the focus group discussions and included in the discrete choice survey, which was completed by 322 adults (10% response rate). Most participants (85%) found the survey easy to understand, and the survey exhibited test-retest reliability. The cluster analysis revealed that participant weights are related more to idiosyncratic personal preferences than to demographic and background characteristics. **Conclusions:** The methodology enables extensive participation by members of the general population, for whom it is both acceptable and reliable. Generating weights for each participant allows the heterogeneity of individual preferences, and the extent to which they are related to demographic and background characteristics, to be tested.

*Keywords:* discrete choice experiment, health technology prioritization, multicriteria decision analysis, public consultation.

Copyright © 2016, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

## Introduction

All health systems have to grapple with how best to allocate new public funding across the myriad health "technologies" (drugs, devices, equipment, procedures, etc.) that are potentially available. Decision makers must choose between entirely new technologies under consideration and existing technologies that are already funded to some extent. Such decision making involves confronting trade-offs between multiple, conflicting objectives or criteria. In the last few years, methods based on multicriteria decision analysis (MCDA) have become increasingly popular [1,2]. In general terms, MCDA, a subdiscipline of operations research, is concerned with decision-making situations in which alternatives are to be ranked based on a variety of criteria [3].

Fundamental to the use of MCDA for health technology prioritization is the need to determine decision-making criteria and weights that represent their relative importance [4]. The reports of the MCDA Emerging Good Practices Task Force of the

International Society for Pharmacoeconomics and Outcomes Research (ISPOR) [5,6], recently published in this journal, emphasize that prioritization decisions depend on a broad set of criteria: "Criteria can be identified in a number of ways, from reviews of previous decisions to focus groups and facilitated workshops," and "weights could come from committee members, from patients, or from the general public" [5]. Methods for determining criteria and weights are surveyed in the reports.

In this article, we report on a methodology for determining criteria and weights that was developed and piloted in New Zealand and enables extensive participation by members of the general population. The methodology is consistent with the prioritization framework proposed by Golan and Hansen [7]. In that framework, for each quantum of potential new spending on a particular technology at the overall intervention level, its incremental costs and benefits are compared. Benefits are viewed in terms of multiple dimensions combined into a single benefits-related variable; the quality of clinical evidence for the technology is also considered.

It is axiomatic that people care about the cost of technologies and the quality of clinical evidence. What do people care about with respect to the multiple dimensions (criteria) underpinning technologies' incremental benefits? What is the relative importance (weights) of these benefits-related criteria? Rather than addressing these two questions by asking people directly, the earlier study, in effect, "made do" by first searching the literature for "criteria and 'other' considerations related to prioritization decision-making" and then administering a discrete choice experiment (DCE) to a small convenience sample mostly comprising health care professionals or researchers from Israel and Canada [7].

Relative to the earlier study, the main contribution of the present article is addressing the two questions just given and determining criteria and weights by consulting members of the general population (applied in New Zealand). Involving members of the general population is appropriate given that everyone consumes health care during their lifetime and most people pay taxes to fund the health system. Understanding "ordinary" people's preferences is important to prioritization agencies, such as New Zealand's Pharmaceutical Management Agency (PHARMAC), which recently undertook a major public consultation exercise with the aim of, in the words of Chief Executive Steffan Crausaz, ensuring that "the criteria we use to help us make those decisions … mean our funding decisions continue to reflect the things New Zealanders … value." [8].

Another strength of our methodology is that the DCE we use to derive the weights of the criteria generates a set of weights for each individual participant, in contrast to earlier DCE studies, which mostly produced aggregated data only [9,10]. Individual-level data enable cluster analysis [11] to be performed to identify any "clusters" of participants with similar patterns of weights.

## Methods

### Health Technology Vignettes

The first step of our study was to identify a diverse group of technologies to be ranked by the focus groups in the next step of the process for the purpose of stimulating discussion. Technology diversity was with respect to characteristics that, informed by the literature, we expected would be relevant for prioritization—ranging from simple preventive interventions to lifesaving ones to treatments for illnesses caused by lifestyle factors, and from treatments involving tens of patients to others involving hundreds of thousands. Mindful of the responder burden on focus-group members, we restricted the number of technologies to 14, which we selected in consultation with health professionals and academics in our professional networks and from PHARMAC.

For each technology, we wrote a "vignette," or short description, at the overall intervention level across the relevant patient group for the New Zealand health system. Based on the medical literature and with advice from clinical experts, each vignette included information about the technology's clinical indications, treatment and side effects, patient characteristics, and numbers treated. Costs and the quality of clinical evidence were not included because, as discussed in the Introduction, we accept, axiomatically, that these two considerations are important and would later be incorporated in the prioritization framework. The vignettes were pilot tested with respect to their clarity and, where necessary, refined by having our family members, friends, and colleagues read and prioritize them.

### Focus Groups

Six focus groups were convened to discuss what matters to people when thinking about technologies that should be funded.

Participants included health care consumers, providers, and academics. They were recruited through our professional and personal networks with the intention of representing a broad cross section of the adult population. We grouped participants according to things they had in common, such as their profession or workplace, to promote an environment in which people felt comfortable sharing and potentially challenging each other's opinions [12].

Before attending the meeting, participants were asked to complete an online "ranking survey" that involved using their unaided judgment to prioritize the 14 technology vignettes (discussed above) with respect to their benefits to society. As explained earlier, the vignettes did not mention the technologies' costs, and participants were instructed not to consider them. The survey was implemented using 1000Minds software [13], which was coinvented by the second author and is available for free for academic and noncommercial use from him or via the above reference.

The meetings were facilitated by the first author and followed an identical format. With participants' permission, discussions were recorded and later transcribed. After introductions, each meeting started with the individual rankings of the 14 technologies being presented to the group as a whole. Ranking variability was used to provoke discussion about why participants ranked the technologies the way they did. As the reasons were teased out, the main things people think about when prioritizing technologies in terms of their potential benefits to society were written on a whiteboard. As this process unfolded, the group also arrived at an overall ranking of the 14 vignettes by majority consensus. When the meeting ended, everyone was given the opportunity to change what was written on the whiteboard.

### Criteria for the DCE

Using the data from the focus-group meetings—the considerations written on the whiteboard, the transcriptions, and the facilitator's notes and impressions—we identified the main themes with respect to the potential benefits that matter to people when thinking about technologies that should be funded. In essence, we followed the process for qualitative data analysis recommended by Warden and Wong [14] involving data collection, note-taking, and the highlighting of relationships and themes.

To implement the DCE at the next step, the potential benefits related to the technologies needed to be expressed in a form suitable for the DCE in terms of benefit-related criteria, whereby each criterion has two or more mutually exclusive levels of performance or severity that are a priori ranked with respect to their relative importance. Ideally, these rankings of levels within each criterion are inherent and incontrovertible; for example, everyone would agree that a "large" health benefit is better (higher ranked) than a "medium" benefit, which in turn is better than a "small" benefit.

If such a ranking is not possible, then one must be imposed—preferably, one with which most people are likely to agree. The criteria and levels also need to be expressed succinctly and in simple terms so they can be easily understood by the DCE's intended participants: members of the general population. This requirement rules out using quality-adjusted life-years because most people without a background in health economics do not understand them (unless they are carefully explained).

Although the vignettes were described at an overall intervention level—across the relevant patient group for the New Zealand health system—the DCE was represented in terms of individual (hypothetical) patients who could be treated by the technology. This "narrowing" of the DCE's choice objects is for practical reasons: We observed that most people in the focus groups favor