VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Testing the External Validity of a Discrete Choice Experiment Method: An Application to Latent Tuberculosis Infection Treatment

Tima Mohammadi, MSc, MA[1,*], *Nick Bansback, PhD*[1,2], *Fawziah Marra, PharmD*[3], *Amir Khakban, MSc*[3,4], *Jonathon R. Campbell, BSc*[3], *J. Mark FitzGerald, MD*[5,6], *Larry D. Lynd, PhD*[2,3,4], *Carlo A. Marra, PharmD, PhD*[7]

[1]*Centre for Health Evaluation and Outcome Sciences, University of British Columbia, St Paul's Hospital, Vancouver, British Columbia, Canada;* [2]*School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada;* [3]*Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, British Columbia, Canada;* [4]*Collaboration for Outcomes Research and Evaluation, University of British Columbia, Vancouver, British Columbia, Canada;* [5]*Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada;* [6]*Centre for Heart and Lung Health, Vancouver Coastal Health Research Institute, University of British Columbia, Vancouver, British Columbia, Canada;* [7]*School of Pharmacy, University of Otago, Dunedin, New Zealand*

ABSTRACT

Objectives: To explore the external validity and predictive power of stated preferences obtained from a discrete choice experiment (DCE) by comparing the predicted behavior of respondents to their actual choices at an individual level. **Methods:** A DCE was performed in patients before being offered treatment for latent tuberculosis infection. A mixed logit model was estimated using hierarchical Bayes. The individual-specific preference coefficients were used to calculate the expected probability of choosing the treatment by each patient. The predicted choice using this probability was compared with their actual decision. We used a receiver-operating characteristic curve and different thresholds to convert probabilities into the predicted choices. The comparability of different distributions for the random parameters was also examined. **Results:** Our results identified significant heterogeneity in preferences for all attributes among respondents. The best model correctly predicted actual treatment decisions for 83% of the participants. The results from using different thresholds and a receiver-operating characteristic curve also confirmed the compatibility between predicted and actual choices. We showed that individual-specific coefficients reflected respondents' actual choices more closely compared with the aggregate-level estimates. **Conclusions:** The results of this study provided support for the external validity of DCEs on the basis of their power to predict actual behavior in this setting. Future investigations are, however, required to establish the external validity of DCEs in different settings.

Keywords: discrete choice experiment, external validity, hierarchical Bayes, mixed logit, revealed preferences, stated preferences.

Introduction

Understanding how the public and patients value different aspects of health care and outcomes can help decision makers provide more patient-centered care and improve health care system efficiency [1]. Consequently, studying patients' preferences for health care products and programs has become a common practice in health economics [2–4]. In the wider literature, individuals' preferences can typically be assessed by observing their choices in the real world [5]. Nevertheless, because of the lack of opportunities to make, observe, or record choices in many areas of the health care system, such actual choice data are more challenging to obtain [6,7].

Therefore, there has been an increasing interest in stated preference (SP) research in health care. The advantage of SP methods such as discrete choice experiments (DCEs) in eliciting preferences for health care services and products, particularly in situations in which no market data exist, makes these techniques powerful instruments. Results from SP methods are widely used to inform policy. They can be used at aggregate level in situations such as resource allocation decisions or at individual level to provide more information about patient decision and can also be used in improving treatment guidelines [8–10]. The hypothetical nature of SP methods, however, raises important questions about the external validity of results in characterizing and predicting consumer's actual behavior and choices [11,12]. As such, there is a need to investigate the validity of SP versus actual decision data in health care. Efforts to explore the hypothetical bias and to test the external validity of SP methods have

The study was conducted at Collaboration for Outcomes Research and Evaluation, University of British Columbia, Vancouver, British Columbia, Canada.

* *Address correspondence to:* Tima Mohammadi, Centre for Health Evaluation and Outcome Sciences, University of British Columbia, St Paul's Hospital, 570-1081 Burrard Street, Vancouver, British Columbia V6Z 1Y6, Canada.

E-mail: tmohammadi@cheos.ubc.ca.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.04.007

been performed predominantly in environmental [14–16,18,19,23], and transportation [11,20–22] economics and in marketing [13,17,24–26]. Different methods have been used to evaluate the validity of SP data. Some studies used SP data to derive preferences and then compared the predicted choices based on these preferences with the observed stated choices. But some studies had access to actual choices data and compared predicted choices with respondents' real decisions.

Despite the common use of SP studies, in particular DCEs, there has been more limited research on testing external validity in health economics with the focus instead on evaluating internal validity [27,28]. Some previous studies conducted to test the external validity of SP used preference data at the aggregate level and not at the level of the individual [29]. One study evaluated willingness to pay for a screening test derived from a contingent valuation exercise and a DCE and compared the data from these two SP methods with responses to a free-of-charge screening test [30]. From two recent studies conducted to evaluate the predictive value of DCE, one stated that results from a DCE can predict respondents' actual behavior [31] and the other reported a significant gap between predicted choices from a DCE and choices made using a decision board (DB) (as a measure of actual choice) [32]. Lancsar and Swait [28] reviewed the existing literature on testing the external validity of DCE and concluded that this issue is under-researched in health economics.

Many studies investigating the external validity of DCE estimate preferences at the aggregate level, assuming homogeneous preferences among respondents [33]. In many choice situations, however, it is likely that individuals have different preferences, and preference heterogeneity may arise from unobserved sources. Existing heterogeneity implies that estimated preferences at the individual level would represent each decision maker's choice behavior more closely as compared with aggregate-level preferences and allows a more accurate comparison of the predicted and actual choices. The mixed logit (MXL) model, a flexible framework to model DCE data, can account for preference heterogeneity and can derive individual-specific utility weights [34,35].

This study aimed to test the validity of DCE by comparing predicted choices for latent tuberculosis infection (LTBI) preventive treatment derived from a DCE to the respondents' observed actual decisions.

Application: Patient's Preference for LTBI Preventive Treatment

Individuals with a diagnosis of LTBI are at higher risk of developing active tuberculosis (TB) disease during their lifetime [36]. Clinical evidence has shown the efficacy of isoniazid (INH) therapy as preventive treatment in reducing the risk of developing active TB in those with LTBI [37]. In British Columbia, TB clinics at the British Columbia Centre for Disease Control (BCCDC) offer publicly funded and optional preventive INH for 9 months to people with diagnosed LTBI. Patients can choose to have this preventive treatment, balancing the risk of developing active TB in the future against the risk of experiencing side effects including liver toxicity, rash, and fatigue. Individuals are provided with information about the treatment benefits and side effects, but the decision to take preventive therapy is left to the individual patient.

Methods

Study Design

Before being provided with information about treatment, we recruited individuals into this study and asked them to complete a DCE. The DCE comprised six key attributes associated with LTBI treatment: 1) length of treatment; 2) frequency of clinic visits; 3) risk of developing active TB after treatment (an indicator of the treatment's effectiveness); 4) chance of developing liver damage; 5) chance of developing skin rash; and (6) chance of developing fatigue. The attributes and levels are presented in Table 1. Details of the DCE design and survey administrating are provided elsewhere [38]. Sawtooth CBC/SSI V6.4.2 (Sawtooth Software, Inc., Sequim, WA) was used for design, and orthogonality, level balance, and minimal overlap were taken into account. Twelve versions of the questionnaire were generated. In each version, 12 choice tasks (10 different choice tasks plus 2 identical fixed to check the respondent's consistency) were presented to the patients. For each choice task, respondents were asked to choose between two hypothetical treatment options and one opt-out option (Fig. 1). The final questionnaire included the DCE questions, two further questions to assess consistency, and questions on sociodemographic characteristics and medical history. Respondents were recruited through BCCDC TB clinics if they: 1) had a diagnosis of LTBI; 2) were 19 years or older; and 3) were able to read and understand English. Ethics approval was obtained from the University of British Columbia Behavioural Research Ethics Board. After completing the survey, we asked whether patients would accept the preventive treatment of 9 months of INH offered by the TB clinic at the BCCDC and examined the dispensing patterns on the pharmacy system in which patients are required to fill their TB drug therapies to confirm that treatment was initiated and completed.

Table 1 – Attributes and levels used in discrete choice experiment.

Attribute	Level
Length of treatment	4, 6, 9, 12 months and none
Frequency of clinic visit	Every 2 months, every 1 month, every 2 weeks and none
Risk of active TB developing	0%, 1%, 2%, 4%, 10%
Chance of liver damage developing	0%, 1%, 3%, 5%, 10%
Chance of skin rash developing	0%, 5%, 10%
Chance of fatigue developing	0%, 5%, 10%

Econometric Model

We used an MXL model, also known as random parameter model, to overcome limitations in the conditional logit model, in particular the assumption of homogeneous utility weights across all individuals [34]. MXL also makes it possible to model repeated choices per respondent.

Following Train [35], we assume an individual n is deciding over J alternatives in S choice sets. The utility of individual n choosing alternative $j=1,\ldots,J$ is decomposed in two parts: one nonrandom (systematic) part that is observed by the researcher and one part that is unknown to the researcher and is treated like a stochastic error term. If utility is specified to be linear in parameters, it can be denoted as follows:

$$U_{njs} = \beta_n' x_{njs} + \varepsilon_{njs}, \quad (1)$$

where β_n is a vector of coefficients specific to individual n, x_{njs} is a vector of observed attributes for alternative j, and ε_{njs} is an error term that is assumed to have independent and identically distributed extreme value distribution. If β_ns are known, the