

# Labeled versus Unlabeled Discrete Choice Experiments in Health Economics: An Application to Colorectal Cancer Screening

Esther W. de Bekker-Grob, MSc,<sup>1</sup> Lieke Hol, MD,<sup>2</sup> Bas Donkers, PhD,<sup>3</sup> Leonie van Dam, MSc,<sup>2</sup>  
J Dik F. Habbema, PhD,<sup>1</sup> Monique E. van Leerdam, MD, PhD,<sup>2</sup> Ernst J. Kuipers, MD, PhD,<sup>2,4</sup>  
Marie-Louise Essink-Bot, MD, PhD,<sup>1,5</sup> Ewout W. Steyerberg, PhD<sup>1</sup>

<sup>1</sup>Department of Public Health, Erasmus MC—University Medical Centre Rotterdam, Rotterdam, The Netherlands; <sup>2</sup>Department of Gastroenterology and Hepatology, Erasmus MC—University Medical Centre Rotterdam, Rotterdam, The Netherlands; <sup>3</sup>Department of Business Economics, Erasmus University, Rotterdam, Rotterdam, The Netherlands; <sup>4</sup>Department of Internal Medicine, Erasmus MC—University Medical Centre Rotterdam, Rotterdam, The Netherlands; <sup>5</sup>Department of Social Medicine, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

## ABSTRACT

**Objectives:** Discrete choice experiments (DCEs) in health economics commonly present choice sets in an *unlabeled* form. *Labeled* choice sets are less abstract and may increase the validity of the results. We empirically compared the feasibility, respondents' trading behavior, and convergent validity between a labeled and an unlabeled DCE for colorectal cancer (CRC) screening programs in The Netherlands.

**Methods:** A labeled DCE version presented CRC screening test alternatives as "fecal occult blood test," "sigmoidoscopy," and "colonoscopy," whereas the unlabeled DCE version presented them as "screening test A" and "screening test B." Questionnaires were sent to participants and nonparticipants in CRC screening.

**Results:** Total response rate was 276 (39%) out of 712 and 1033 (46%) out of 2267 for unlabeled and labeled DCEs, respectively ( $P < 0.001$ ). The labels played a significant role in individual choices; approximately 22%

of subjects had dominant preferences for screening test labels. The convergent validity was modest to low (participants in CRC screening:  $r = 0.54$ ;  $P = 0.01$ ; nonparticipants:  $r = 0.17$ ;  $P = 0.45$ ) largely because of different preferences for screening frequency.

**Conclusion:** This study provides important insights in the feasibility and difference in results from labeled and unlabeled DCEs. The inclusion of labels appeared to play a significant role in individual choices but reduced the attention respondents give to the attributes. As a result, unlabeled DCEs may be more suitable to investigate trade-offs between attributes and for respondents who do not have familiarity with the alternative labels, whereas labeled DCEs may be more suitable to explain real-life choices such as uptake of cancer screening.

**Keywords:** colorectal cancer screening, discrete choice experiment, feasibility, labeled alternatives, unlabeled alternatives, validity.

## Introduction

Estimates of public and patients' preferences are of great importance in informing policy decision-making and improving adherence with public health-care interventions or programs [1]. Discrete choice experiments (DCEs) have become a commonly used technique in health economics to elicit preferences. The DCE is an attribute-based survey method for measuring benefits (utility) [2]. In a DCE, subjects are presented with a sequence of (hypothetical) scenarios (choice sets) and are asked to choose between two or more competing alternatives that vary along several characteristics or attributes of interest [2]. DCEs assume that subjects' preferences (as summarized by their utility function) are revealed through their choices [2] (for further details, see Bliemer and Rose [3], Hensher et al. [4], Louviere et al. [5], and Ryan et al. [2]).

A fundamental question that arises in the application of DCE is whether to present the choice sets in a labeled or unlabeled form. The unlabeled form involves assigning unlabeled alternatives in the choice set, such as "alternative A," "alternative B," and so on. The labeled form involves assigning labels that communicate information regarding the alternative. In marketing applications, labels tend to consist of brand names and logos,

which consumers have learned to associate with different product characteristics and feelings. In the context of health economics, labels tend to consist of generic or brand-name medications, specific screening tests (e.g., colonoscopy, sigmoidoscopy), specific treatments (surgery vs. conservative), or other descriptors. An advantage of assigning labels is that alternatives will be more realistic and the choice task will be less abstract for the subject, which add to the validity of the results. Hence, the results may be better suitable to support decision-making at policy level. Nevertheless, by far, most commonly applied DCEs in health economics used unlabeled alternatives.

The aim of our study was to empirically compare the feasibility, respondents' trading behavior, and convergent validity between a labeled and an unlabeled DCE. All of these aspects were explored in the context of a DCE study directed at investigating population preferences for colorectal cancer (CRC) screening programs in The Netherlands. We were convinced that specific aspects of endoscopy (sigmoidoscopy, colonoscopy) or fecal occult blood test (FOBT) that determine its burden could not be totally captured by presenting an unlabeled "screening test A" variant to patients [6]. For that very reason, we expected differences between an unlabeled and a labeled DCE.

## Theoretical Basis of Labeled and Unlabeled DCEs

The aim of discrete choice modeling is to estimate the weights that respondents place on attributes of alternatives. An individual

*Address correspondence to:* Esther W. de Bekker-Grob, Department of Public Health, Erasmus MC—University Medical Centre Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands. E-mail: e.debekker@erasmusmc.nl  
10.1111/j.1524-4733.2009.00670.x

acting rationally is expected to evaluate the set of available alternatives and will choose that alternative that gives the greatest relative utility [4]. Thus, an individual will choose alternative A over B, if  $U(X_A, Z) > U(X_B, Z)$ , where  $U$  represents the individual's indirect utility function from certain alternatives,  $X_A$  represents the attributes of alternative A,  $X_B$  represents the attributes of alternative B, and  $Z$  represents the socioeconomic and other characteristics of the individual that influence his/her utility. Choices made in DCEs are analyzed by using random utility theory (i.e., an error term is included in the utility function to reflect the unobservable factors in the individual's utility function) [4]. Thus, an individual will choose alternative A over B, if  $V(X_A, Z) + \varepsilon_A > V(X_B, Z) + \varepsilon_B$ , where  $V$  is the measurable component of utility estimated empirically, and  $\varepsilon_A$  and  $\varepsilon_B$  reflect the unobservable factors in the individual's utility function of alternative A and B, respectively ( $X_A$ ,  $X_B$ , and  $Z$  defined as above).

There are two general types of DCEs: 1) unlabeled and 2) labeled DCEs [5]. Unlabeled DCEs use generic titles for the alternatives (e.g., radio-imaging "A" or "B"). Labeled DCEs use alternative-specific titles for the alternatives (e.g., "computer tomography" or "MRI-scan" [magnetic resonance imaging]). The number of alternatives (irrespective of whether labeled or unlabeled) in a choice set is unrestricted from a theoretical point of view [4]. The decision as to whether to use labeled or unlabeled DCEs is an important one [4]. The labeled alternative itself conveys information to respondents. This matters in choice and other decision tasks, because 1) respondents may use labeled alternatives to infer information that they perceive as missing; and 2) these inferences may be (and usually are) correlated with the random component [5]. Although we may not exactly know what respondents find relevant in the label for forecasting uptake of, for example, a health-care intervention, it may be worthwhile to find out if respondents prefer one alternative label to another. A labeled DCE can take effects into account, which respondents may have learned to associate with different health-care intervention characteristics and feelings, and, as a result, may be more suitable to predict [6]. Unlabeled and labeled DCEs both have their merits. If each of the labeled options has  $A$  attributes with  $L$  levels and the choice sets are of size  $M$ , then there are  $L^A M$  possible choice sets, assuming that all labels are presented in a choice set and that the same label does not appear more than once in a choice set. If the options are unlabeled, then there are  $L^A$  possible items that can be included in each position of each choice set. If the choice sets are of size  $M$  and we are not going to allow the same item to appear more than once in a choice set, then there are " $L^A$  choose  $M$ " possible choice sets of size  $M$ . Therefore, the designs of an unlabeled DCE can be much smaller. For example, two alternatives with four attributes and three levels yields  $6561$  (i.e.,  $3^{2 \cdot 4} = 3^8$ ) possible alternative combinations for a labeled DCE compared with "just"  $81$  (i.e.,  $3^4$ ) possible alternative combinations for an unlabeled design. Other merits of unlabeled DCEs include that 1) they do not require the identification and use of all alternatives within the universal set of alternatives, namely, the attribute levels are sufficiently broad to represent all alternatives; 2) they might be more robust in terms of not violating IID-assumption (i.e., error terms are independent and identically distributed), because the alternatives may be less correlated with the attributes as in labeled DCEs [4]; and 3) they encourage respondents to choose an alternative by trading-off attribute levels, which may be desirable from a nonmarket valuation perspective [7]. On the other hand, merits of using labeled DCEs include that 1) they will be more realistic and less abstract so that responses may better reflect the real preference structure; and 2) they can study the main effect of the labels.

## Methods

### Case Study

CRC is the most frequently occurring malignancy within the European Union and the second leading cause of cancer related death in the Western world [8,9]. Various countries have implemented a national screening program for CRC screening to detect CRC in an early stage or are investigating prerequisites for implementation [10,11]. There are several screening tests eligible for use as a population-based screening program, such as fecal occult blood tests (FOBTs), sigmoidoscopy, or colonoscopy. This study aimed to investigate individual preferences for CRC screening using a DCE.

### DCE Design

The questionnaire design phase involved extensive background research, expert opinions, and interviews with screened individuals. Experts ( $n = 3$ ) were asked to comment on a list of test characteristics derived from our extensive literature review. Potential screenees ( $n = 40$ ), both participants of a CRC screening program ( $n = 20$ ) and screening naive individuals ( $n = 20$ ), could also comment on the list of test characteristics and rank them in order of importance. Based on these data, we selected the most important test characteristics. The levels for each test characteristic incorporated the range of possible test outcomes based on the current literature (for more detail on how the qualitative data were used to select the final test labels, attributes, and levels, see work of Hol et al. [L. Hol, E.W. de Bekker-Grob, L. van Dam, et al., unpubl. ms] and van Dam et al. [12]). Table 1 lists the labels, attributes, and attribute levels chosen. The labeled CRC screening tests ("FOBT," "sigmoidoscopy," and "colonoscopy") may evoke individual feelings, which may not be captured in the unlabeled CRC screening tests ("CRC screening test A" and "CRC screening test B"). Notably, the invasiveness of the alternative test was (indirectly) described by the levels of five attributes: "side effects of the test," "complication risk of the test," "preparation for the patient," "location of screening," and "the duration of screening." Giving directly the information "how a sample is taken" is, in our case, totally equal to the screening test label, "taking a sample from your motion" is equal to FOBT, and "tube into your back passage throughout your colon" is equal to colonoscopy. If the unlabeled DCE would include directly this information about "how the sample is taken" (thus, actually naming the test), then the unlabeled DCE will be a labeled DCE as well; the attribute "how the sample is taken" will have an interaction with all other attributes, and a restricted design is needed to avoid implausible combinations of attribute levels (i.e., the attribute levels are alternative specific and, thus, a labeled DCE). Another point of notice is that the unlabeled experiment had, for some attributes, a smaller-level range than the feasible options in the labeled experiment. As a result, we avoided some extreme combinations of 30 times a screening test, resulting in a reduction in mortality from 3.0% to 2.7% in the unlabeled DCE, which added to utility balance.

The combination of the attributes and attribute levels of the unlabeled design resulted in 2048 CRC screening test alternatives ( $4^4 \cdot 2^3$ ). A fractional factorial design was used based on a Web site, which contained a library of more than 200 orthogonal arrays [13], to reduce the number of alternatives to a manageable level of 16 alternatives in which orthogonality and level balance were fulfilled. These 16 alternatives were paired up with another orthogonal array by using the fold-over technique (i.e., cyclic design), which caused minimal overlap between attribute levels [14]. Each choice set (i.e., a set of available alternatives) con-

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات