Volume 13 • Number 6 • 2010
VALUE IN HEALTH

Clustering Rasch Results: A Novel Method for Developing Rheumatoid Arthritis States for Use in Valuation Studies

Helen M. McTaggart-Cowan, PhD, John E. Brazier, PhD, Aki Tsuchiya, PhD

University of Sheffield, Sheffield, UK

ABSTRACT

Purpose: Health states that describe an investigated condition are a crucial component of valuation studies. The health states need to be distinct and comprehensible to those who appraise them. The objective of this study was to describe a novel application of Rasch and cluster analyses in the development of three rheumatoid arthritis health states.

Methods: The Stanford Health Assessment Questionnaire (HAQ) was subjected to Rasch analysis to select the items that best represent disability. *K*-means cluster analysis produced health states with the levels of the selected items. The pain and discomfort dimension from the EuroQol-5D was also incorporated.

Results: The results demonstrate a methodology for reducing a dataset containing individual disease-specific scores to generate health states. The four selected HAQ items were bending down, climbing steps, lifting a cup to your mouth, and standing up from a chair.

Conclusions: The combined use of Rasch and *k*-means cluster analysis has proved to be an effective technique for identifying the most important items and levels for the construction of health states.

Keywords: cluster analysis, health state, quality of life, Rasch analysis, rheumatoid arthritis.

Introduction

Economic evaluation plays a significant role in guiding health-care resource allocation decisions. For this reason, many health technology assessment organizations, including the National Institute for Health and Clinical Excellence (NICE), have proposed guidelines for the proper conduct of economic evaluations. Specifically, NICE guidelines state that the preferred methodology is cost-effectiveness analysis, in which the benefits of health interventions are quantified using quality-adjusted life-years (QALYs) rather than units specific to the condition under investigation [1]. The QALY describes an individual's preference for a health state by capturing quantity and quality of life (QOL) into a single summary measure.

In making decisions regarding the allocation of health-care resources, there is a normative debate as to who should be providing values for health states. The current recommendation is that these values be obtained from the general population rather than patients [2]. This follows the concept that in a publicly funded health-care system, the main objective is to meet societal preferences for maximizing health. The drawback, however, is that the general population, when appraising the impaired health state, do not consider that they can adapt to the condition over time [3]; this can result in significant ramifications when these values are incorporated into a cost-effectiveness analysis [4]. Therefore, to test the effect of providing disease adaptation information to general population respondents, health state descriptions need to be developed. The states need to both accurately portray symptoms a patient experiences and be comprehensible for respondents who may not be knowledgeable about the condition under investigation. In addition, the health states need to demonstrate detriment in some of the item levels to enable evaluation of the influence of disease adaptation information on general population values. If the health states do not demonstrate any detriment in at least one item level, then the descriptions are effectively synonymous to full health. Respondents will not need to consider the concept of adaptation in their valuation of this "full health" state. In this case, the impact of disease adaptation information on the value of this particular health state is non-existent.

Health states can be developed using various methods. Expert judgments, such as those from physicians, permit a range of patient experiences to be elicited; however, these opinions are subjected to biases. Physicians' viewpoints of health states may be distorted if their patients exaggerate their QOL in an attempt to please their doctor or family members. Patient responses, on the other hand, provide direct information about how the investigated health state impacts their lives. The use of interviews and focus groups allows researchers to gain in-depth knowledge about life in different health states, but the results may be subject to volunteer bias, as the small number of participants might not be representative of most patients. As such, using responses on condition-specific instruments from a large sample of patients may be a better alternative in the development of health states.

Condition-specific instruments are comprised of numerous items addressing different aspects of the disease under investigation. Nevertheless, for patients' responses on condition-specific instruments to be of use in the construction of health states, the number of items in the instrument needs to be minimized; this reduces the burden placed on the respondents when completing the valuation exercise. In addition, responses on condition-specific instruments need to be combined in such a way that patients with similar QOL are grouped together so that distinct health states are formed.

One method that identifies the most representative items of condition-specific instruments is Rasch analysis. It is a technique that converts categorical responses into a continuous latent scale [5,6]. It has been employed in the development of QOL instruments [7], and, more recently, in the construction of a health state classification systems [8,9]. Although Rasch analysis is not the

Address correspondence to: Helen M. McTaggart-Cowan, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK. E-mail: h.m.cowan@sheffield.ac.uk; hmctc@hotmail.com

10.1111/j.1524-4733.2010.00757.x

only method for selecting maximally discriminatory items, this technique was chosen for this study because it offers a systematic approach to identify items while making the best use of the richness and sensitivity of the original instrument. As Rasch analysis identifies the instrument's most meaningful items, the cognitive and time constraints placed on respondents are minimized when they are appraising the health states. Furthermore, items can be ranked from easiest to the most difficult; this ensures that the health states capture the widest range of severity. Rasch analysis aims to create a system that is ordered, additive, and of specific objectivity.

Although Rasch analysis selects the representative items of the condition-specific instrument, there is still a need to identify different combinations of levels of the selected items to form distinct health states. A technique that meets this objective is cluster analysis. The main purpose of this statistical approach is to group similar objects together; for example, individuals within the same cluster are more similar to each other than individuals from different clusters. One of the first studies that employed cluster analysis to develop health states did so by identifying patterns in the physical and mental health domains of the Medical Outcomes SF-12 questionnaire [10]; other studies using cluster analysis have followed (e.g., [11,12]). By applying cluster analysis to the Rasch-reduced instrument, natural groupings in the dataset can be identified, thereby forming plausible health states.

This article aims to describe the combination of Rasch and cluster analyses alongside classical psychometric methods to develop health states. In this work, the Stanford Health Assessment Questionnaire (HAQ) [13], a condition-specific measure, and the EuroQol (EQ-5D) [14] were used to construct rheumatoid arthritis (RA) health states. This process is the first stage of a two-part procedure in the assessment of informed general population values. We used Rasch analysis to select items from responses on the condition-specific HAQ instrument, and then used cluster analysis to group the identified item levels to form health states. In this article, we do not claim that the joint use of Rasch and cluster analyses is the definitive method for developing health states; rather, we attempt to shed light on its potential as a novel approach to doing so. The second stage of this project is a quasi-experiment to determine whether individuals alter their values for the constructed health states after being given information about the ability of patients to adapt to their disease. The results from the second stage are beyond the scope of the current work and will be presented in a future article.

Methods

The HAQ

RA afflicts 0.8% of the United Kingdom population [15]. It is a chronic autoimmune inflammatory disorder, which results in upper- and lower-limb disability and discomfort. Although treatments are improving, RA can pose a burden on its patients by reducing their QOL. As such, physicians often use condition-specific instruments such as the HAQ to assess their patients [14].

Although there are other condition-specific instruments that also assess the disability that a RA patient faces, the HAQ was chosen because it has been administered in various rheumatic populations for nearly three decades. The instrument is widely used in clinical trials to evaluate RA outcomes [16] and is shown to be valid [17] and responsive [18]. Despite its popularity, the complete HAQ—which also includes questions on symptoms, medication use, and medical history—is lengthy. The HAQ is also known to have a floor effect, such that severely disabled individuals could be represented by scores implying minimal disability [19].

The component of the HAQ of particular relevance to this study contains 20 items that assesses an individual's ability to complete daily tasks in the following domains: dressing and grooming, arising, eating, walking, personal hygiene, reach, grip, and other activities. Two or three items comprise each domain (Table 1). Each item has four levels: no, some, or much difficulty performing the task, or an inability to perform the task. Respondents can select a score between zero and three, with higher scores implying a greater disability. The score on an individual item is increased by another point when the respondent requires assistive devices or additional help. The greatest item score yields the domain score. The averaged scores from the eight domains form an overall HAQ score.

Dataset

Rasch analysis is sensitive to large sample sizes; using too large a sample generates a greater frequency of statistically significant items [6], making item reduction difficult. Nevertheless, the literature reports that sample sizes in the range of 400 to 700 have been used successfully [20,21]. For this study, an anonymized dataset from the National Data Bank for Rheumatic Diseases (NDB) was used. The NDB is a non-profit organization that contains longitudinal outcomes research data from rheumatology patients [22]. The dataset contained information from a stratified sample of 600 RA patients. An equal number of individuals in three severity ranges (i.e., overall HAQ scores of <1, 1–2, >2) was used to ensure that each disability level was well represented.

Initial Criteria for Reducing the HAQ

Although Rasch analysis was the main approach in the selection of most representative HAQ items, psychometric methods were conducted simultaneously to ensure greater strength in the results [8]. The frequency and the internal consistency—the correlation between item and domain scores—were evaluated for the HAQ responses [10]. If some items elicited poor responses (i.e., low frequency) or poor internal consistency (i.e., weak correlation), they were considered to be less representative of disability than other HAQ items. As has been done in previous work, questions pertaining to the use of assistive devices were excluded from the analysis [19,23].

The research team decided a priori that three states composed of five items needed to be constructed. The developed states had to crudely define a form of mild, moderate, and extreme RA that patients may experience. The decision to have five items to describe the states was based on the fact that previous studies have shown that five-itemed instruments do not overburden respondents [14,24]. Although it may seem like a limitation to not have the chosen statistical methods identify the ideal number of health states and the number of items to be included in the descriptions, restricting the number of items and levels is normal practice in valuation studies. For this study, we wanted to use the health states in our subsequent valuation study; therefore, specific requirements as to what a successful product needed to look like had to be met. Nevertheless, if rigid criteria were not needed in the formation of health states, then the chosen statistical methods could be applied to ensure that the developed health states were entirely data driven.

Selection of HAQ Items That Best Describe Disability

Rasch analysis. Rasch analysis verifies that the scale of the instrument is unidimensional, a fundamental requirement of