VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

Generating EQ-5D-3L Utility Scores from the Dermatology Life Quality Index: A Mapping Study in Patients with Psoriasis

Niall J. Davison, MSc[1,], Alexander J. Thompson, MSc[1], Alexander J. Turner, PhD[1], Louise Longworth, PhD[2], Kathleen McElhone, PhD[3], Christopher E.M. Griffiths, MD, FRCP, FMedSci[3], Katherine Payne, PhD[1], The BADBIR Study Group[†]*

[1]*Manchester Centre for Health Economics, The University of Manchester, Manchester, UK;* [2]*PHMR, Berkeley Works, London, UK;* [3]*The Dermatology Centre, Salford Royal NHS Foundation Trust, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK*

ABSTRACT

Objectives: To develop an algorithm to predict the three-level EuroQol five-dimensional questionnaire (EQ-5D-3L) utility scores from the Dermatology Life Quality Index (DLQI) in psoriasis. **Methods:** This mapping study used data from the British Association of Dermatologists Biologic Interventions Register—a pharmacovigilance register comprising patients with moderate to severe psoriasis on systemic therapies. Conceptual overlap between the EQ-5D-3L and DLQI was assessed using Spearman rank correlation coefficients and exploratory factor analysis. Six regression methods to predict the EQ-5D-3L index (direct mapping) and two regression methods to predict EQ-5D-3L domain responses (response mapping) were tested. Random effects models were explored to account for repeated observations from the same individual. Estimated and actual EQ-5D-3L utility scores were compared using 10-fold cross-validation (in-sample) to evaluate predictive performance. Final models were selected using root mean squared error, mean absolute error, and mean error. **Results:** The data set comprised 22,085 observations for which DLQI and EQ-5D-3L were recorded on the same day. A moderate correlation was found between the measures ($r = -0.47$). Exploratory factor analysis showed that two EQ-5D-3L domains (pain/discomfort and depression/anxiety) were associated with all six DLQI domains. The best-performing model used ordinary least squares with DLQI items, age, and sex as explanatory variables (with squared, cubic, and interaction terms). A tool was produced to allow users to map their data to the EQ-5D-3L, and includes algorithms that require fewer variables (e.g., total DLQI scores). **Conclusions:** This study produced mapping algorithms that can generate EQ-5D-3L utility scores from DLQI data for economic evaluations of health interventions for patients with psoriasis.

Keywords: dermatology, DLQI, EQ-5D, mapping, psoriasis, utility.

Introduction

Psoriasis is a commonly occurring immune-mediated condition affecting about 2% of the UK population [1]. In addition to affecting the skin, psoriasis is associated with significant mental and physical comorbidity [2]. Treatments for psoriasis, such as biologic therapies, have been shown to improve health-related quality of life (HRQOL) as measured by the dermatology-specific Dermatology Life Quality Index (DLQI) and the generic EuroQol five-dimensional questionnaire (EQ-5D) [3].

The DLQI is the most widely used measure to assess HRQOL related to skin disease in psoriasis studies [4]. The reliability, construct validity, and responsiveness of the DLQI have all been demonstrated in patients with psoriasis [4,5]. Nevertheless, DLQI scores are not preference-weighted and therefore do not account for the views of the public on the relative importance of one-unit changes on the DLQI scale. The measure is also applicable only in the context of skin conditions. These limitations in using the DLQI are problematic for decision makers who need to compare the cost-effectiveness of health interventions across multiple disease areas.

Decision-making bodies such as the National Institute for Health and Care Excellence (NICE) recommend the quality-adjusted life-year (QALY) as the standard metric of benefit for economic evaluations of health care interventions [6]. The EQ-5D is the recommended preference-based measure of health status used to calculate QALYs in the context of NICE appraisals [6], and the value set of the original three-level version (EQ-5D-3L) is currently favored by NICE [7]. The construct validity and responsiveness of the EQ-5D-3L have been confirmed for people with psoriasis [8].

To estimate the impact of a treatment in terms of QALYs in psoriasis studies that do not administer the EQ-5D-3L, a "mapping" algorithm can be used to convert DLQI data to EQ-5D-3L

* *Address correspondence to:* Niall J. Davison, Manchester Centre for Health Economics, The University of Manchester, 4th Floor Jean McFarlane Building, Oxford Road, Manchester M13 9PL, UK.

E-mail: niall.davison@manchester.ac.uk.

†List of members: Jonathan Barker, MD FRCP; Marilyn Benham; David Burden, MD FRCP; Ian Evans; Chris Griffiths, MD FRCP FMedSci ; Sagair Hussain, PhD; Brian Kirby, MRCP; Linda Lawson; Kathy McElhone, PhD; Ruth Murphy, MB ChB PhD FRCP; Anthony Ormerod, MB MD FRCP; Caroline Owen, MBChB FRCP; Nick Reynolds, MD FRCP; Catherine Smith, MD FRCP.

1098-3015$36.00 – see front matter Copyright © 2017, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

https://doi.org/10.1016/j.jval.2017.10.024

index scores (direct mapping) or EQ-5D-3L domain responses (response mapping). Mapping algorithms are developed using a data set containing both the source measure (DLQI) and the target measure (EQ-5D-3L). Regression methods are used to estimate a relationship between the measures in an "estimation sample" and the predictive ability of the resulting algorithm is assessed in a "validation sample." The algorithm can be applied to DLQI data from trials, observational studies, or economic models to calculate QALYs when EQ-5D-3L data do not exist.

A search of the Health Economics Research Centre database of mapping studies [9] identified four published algorithms from the DLQI to the EQ-5D-3L [10–13]. None of these studies explored regression methods outside of linear regression, and none of them used validation to assess the predictive ability of their algorithm (one study reported cross-validation for an algorithm to the EQ-5D visual analogue scale). In 2017, a study estimated ordinal logistic regressions (response mapping) in patients with any skin disease [14]. The study concluded that response mapping, estimated on 3452 patients with age and sex variables, was superior to linear regression using coefficients from a previous algorithm estimated on 86 patients without item-level data, age, or sex [13]. Consistency in model specifications and estimation data sets would allow for a fair comparison between regression methods. This study aimed to produce mapping algorithms from the DLQI to the EQ-5D-3L in a large sample of patients with psoriasis as well as to produce a user-friendly tool to aid implementation.

Methods

This mapping study explored a wide range of regression models to develop algorithms to predict EQ-5D-3L utility scores from the DLQI. Methodological guidance produced by the NICE Decision Support Unit and reporting standards guidance outlined in the 2017 International Society for Pharmacoeconomics and Outcomes Research Task Force report were followed [15,16].

Data

This study used data from the British Association of Dermatologists Biologic Interventions Register (BADBIR). BADBIR is a UK and Eire observational, prospective cohort involving 154 dermatology centers, set up in 2007 to assess the long-term safety of biologic and conventional systemic therapies for patients with moderate to severe psoriasis [17]. NICE has recommended that all patients in the United Kingdom receiving biologics for psoriasis should be registered to BADBIR [18]. Because of the longitudinal nature of the data set, some patients in the sample received treatments that reduced the severity of their psoriasis. The range of severities within the sample means that mapping algorithms are likely to be valid for all severities of the disease. Ethical approval for data collection for the registry was obtained in March 2007 (National Health Service Research Ethics Committee North West England; reference 07/MRE08/9). All patients gave written consent for their participation in the registry.

The DLQI and EQ-5D-3L are two patient-reported outcome measures collected within BADBIR. All observations with a DLQI and EQ-5D-3L record on the same date were included in the sample. Data were classified as missing when the measures were administered on the same date but at least one measure was incomplete. Missing data were not imputed to avoid the introduction of an additional source of error associated with the imputation method [19]. Data were available up to January 1, 2016.

Outcome Measures

The EQ-5D-3L asks respondents to describe their health using five domains (mobility, self-care, usual activities, pain/discomfort, and depression/anxiety). In each domain, respondents indicate whether they have 1) no problems, 2) some problems, or 3) extreme problems on the day of EQ-5D-3L completion. The UK tariff was used in this study to produce values between −0.594 and 1, where 1 represents "full health," 0 represents death, and negative values represent health states valued as worse than death [20].

The DLQI consists of 10 items and covers six domains (symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment) [21]. The questionnaire asks respondents how they have been affected by their skin problem over the last week. Each item is scored from 0 to 3, and items are summed to produce DLQI scores between 0 and 30, where a lower score indicates better HRQOL.

Conceptual Overlap

The estimation of a mapping algorithm relies on there being conceptual overlap between the source (DLQI) and the target (EQ-5D-3L) measure domains. If there is no overlap in content, then mapping is unlikely to appropriately capture the relationship between the measures to estimate health-related utility [15,22].

Some overlap was expected because the psychometric properties of both measures have been validated in patients with psoriasis [5,8]. Spearman rank correlation coefficients between both the total scores and the domains were generated to investigate the degree of conceptual overlap. The strength of correlation was based on absolute values and described by categories defined a priori, which were very weak (0–0.19), weak (0.20–0.39), moderate (0.40–0.59), strong (0.60–0.79), and very strong (0.80–1) [23].

Exploratory factor analysis (EFA) was also used to understand whether the domains of the measures could be explained by the same factors (latent constructs). In EFA, "eigenvalues" are the variances of the factors and "factor loadings" are the correlations of a domain with the respective factor. Factors were varimax-rotated to enable interpretation and retained if they had an eigenvalue greater than 1 [24]. Factor loadings greater than 0.3 were considered to be "meaningful" [25,26]. If a factor had meaningful loadings on EQ-5D-3L and DLQI domains, there would be reason to suspect that these domains captured the same latent construct.

Model Development

A priori, the following aspects were not known and therefore investigated: the optimal type of regression (functional form), the optimal set of independent variables (model specifications), whether direct mapping would outperform response mapping, and whether adjusting for multiple observations from the same individual would improve predictive ability.

The typical EQ-5D-3L distribution is trimodal with one peak at full health, one peak for moderate states, and another peak for more severe states [15]. Linear regression is used widely to derive mapping functions [27], but has theoretical drawbacks when modeling outcomes of this type [28]. Therefore, in mapping studies, a range of functional forms have been used in attempts to account for the properties of the EQ-5D-3L distribution [15,29]. Nevertheless, comparisons of model performance between linear regressions and more complex models have resulted in mixed conclusions regarding the best approach [29]. Established mapping guidelines do not recommend a specific set of methods or functional forms [16,30], and therefore multiple functional forms were tested in this study.

Six types of direct mapping models were used: ordinary least squares (OLS), Tobit, generalized linear models (GLMs), two-part models, and adjusted limited dependent variable mixed models with two latent classes (ALDVMM-2) and with three latent classes (ALDVMM-3). OLS was used because it is the most common