



This is a repository copy of *Developing a preference-based utility scoring algorithm for the Psoriasis Area Severity Index (PASI)*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/148478/>

Version: Accepted Version

Article:

Matza, L.S., Brazier, J.E. orcid.org/0000-0001-8645-4780, Stewart, K.D. et al. (8 more authors) (2019) Developing a preference-based utility scoring algorithm for the Psoriasis Area Severity Index (PASI). *Journal of Medical Economics*. ISSN 1369-6998

<https://doi.org/10.1080/13696998.2019.1627362>

This is an Accepted Manuscript of an article published by Taylor & Francis in *Journal of Medical Economics (JME)* on 4th June 2019, available online:
<http://www.tandfonline.com/10.1080/13696998.2019.1627362>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



Developing a preference-based utility scoring algorithm for the Psoriasis Area Severity Index (PASI)

Louis S. Matza, John E. Brazier, Katie D. Stewart, Lionel Pinto, Randall H. Bender, Leon Kircik, Jessica Jordan, Katherine J. Kim, Alex Mutebi, Hema N. Viswanathan & Alan Menter

To cite this article: Louis S. Matza, John E. Brazier, Katie D. Stewart, Lionel Pinto, Randall H. Bender, Leon Kircik, Jessica Jordan, Katherine J. Kim, Alex Mutebi, Hema N. Viswanathan & Alan Menter (2019): Developing a preference-based utility scoring algorithm for the Psoriasis Area Severity Index (PASI), Journal of Medical Economics, DOI: [10.1080/13696998.2019.1627362](https://doi.org/10.1080/13696998.2019.1627362)

To link to this article: <https://doi.org/10.1080/13696998.2019.1627362>



Accepted author version posted online: 04 Jun 2019.



Submit your article to this journal [↗](#)



Article views: 34



View Crossmark data [↗](#)

Developing a preference-based utility scoring algorithm for the Psoriasis Area Severity Index (PASI)

Louis S. Matza¹, John E. Brazier², Katie D. Stewart¹, Lionel Pinto³, Randall H. Bender⁴, Leon Kircik⁵, Jessica Jordan¹, Katherine J. Kim⁴, Alex Mutebi⁶, Hema N. Viswanathan⁶, Alan Menter⁷

¹Evidera, Bethesda, MD, USA

²School of Health and Related Research, University of Sheffield, UK

³Amgen, Thousand Oaks, CA, USA

⁴Formerly at Evidera, Bethesda, MD, USA

⁵Derm Research, PLLC, Louisville, KY, USA

⁶Formerly at Amgen, Thousand Oaks, CA, USA

⁷Baylor University Medical Center, Dallas, TX, USA

Correspondence: Louis S. Matza, Evidera, 7101 Wisconsin Avenue, Suite 1400, Bethesda, MD 20814, email: louis.matza@evidera.com

Transparency

Declaration of funding

This study and the resulting manuscript was funded by Amgen (Thousand Oaks, CA).

Declaration of financial/other relationships

Three of the authors were employed by Amgen at the time this study was initiated (Mutebi, Pinto, Viswanathan). Five of the authors were employed by Evidera, an independent company that received funding from Amgen for time spent conducting this study (Bender, Jordan, Kim, Matza, Stewart). The other authors received funding for time spent on this research (Brazier, Kircik, Menter). All aspects of the study design, interpretation, and decision to submit for publication were determined by the authors. A peer reviewer on this manuscript has disclosed the receipt of grants/research supports for participation in clinical trials (paid to their institution) from Abbvie, Almirall, Amgen, Boehringer Ingelheim, Janssen, Leo-Pharma, Lilly, Novartis, Pfizer, Regeneron, Roche, Sanofi, UCB. The same reviewer discloses receipt of honoraria or consultation fees (paid to themselves) from Abbvie, Almirall, Amgen, Baxalta, Biogen, Boehringer Ingelheim, Celgene, Gebro, Janssen, Leo-Pharma, Lilly, Merck-Serono, MSD, Mylan, Novartis, Pfizer, Regeneron, Roche, Sandoz, Samsung-Bioepis, Sanofi and UCB. The same reviewer discloses participation in a company sponsored speaker's bureau for Celgene, Janssen, Lilly, MSD, Novartis and Pfizer. A separate reviewer on this manuscript has disclosed serving as an investigator for AbbVie, Boehringer Ingelheim, Eli Lilly, Janssen-Cilag, Merck, Novartis, Pfizer and UCB Pharma; participating in a clinical trial supported by Array Biopharma and MSD; being a member of advisory boards of Eli Lilly, LEO Pharma, MSD, Novartis, Pfizer and UCB Pharma; obtaining honoraria from ALK-Abello, AbbVie, Biotest, BMS, Janssen-Cilag, MSD, Novartis, Pfizer, Dr. Pflieger GmbH and Roche; and receiving support for conferences from AbbVie, Actelion, ALK-Abello, Alma Lasers, Almirall, ARC Lasers, Asclepion, BMS, Celgene, Dermapharm, Dermasence, Galderma, GSK, Janssen-Cilag, L'Oreal, La Roche Posay, LEO Pharma, Medac, MSD, Novartis, Pierre Fabre, P&M Cosmetics, Pfizer and Roche. The peer reviewers on this manuscript have no other relevant financial relationships or otherwise to disclose.

Author contributions

LM directed the study including protocol development, data collection, and data analysis. JB, KS, LP, JJ, KK, AM, and HV contributed to the study design, protocol development, data collection, and/or data interpretation. LK and AM provided clinical input to help shape the study design, methods, and interpretation. LM drafted the manuscript, and all other authors provided input and approval.

Acknowledgements

The authors would like to thank Zaneta Balantac, Evan Davies, Katie Gries, David Hengerer, Michael Iacobelli, Amanda Landrian, Amara Tiebout, and Erica Zaiser for assistance with data collection; Dennis Revicki for consulting on the study design; Christine Thompson for statistical programming; and Amara Tiebout for editorial assistance.

Ethics compliance

All authors followed the ICMJE requirements on privacy and informed consent from patients and study participants. Participants were required to provide written informed consent before completing study procedures, and all procedures and materials were approved by an independent Institutional Review Board (Ethical & Independent Review Services; Study Number 14145-01).

Accepted Manuscript

Abstract

Introduction

It is challenging to identify health state utilities associated with psoriasis because generic preference-based measures may not capture the impact of dermatological symptoms. The Psoriasis Area Severity Index (PASI) is one of the most commonly used psoriasis rating scales in clinical trials. The purpose of this study was to develop a utility scoring algorithm for the PASI.

Methods

Forty health states were developed based on PASI scores of 40 clinical trial patients. Health states were valued in time trade-off interviews with UK general population participants. Regression models were conducted to crosswalk from PASI scores to utilities (e.g., OLS linear, random effects, mean, robust, spline, quadratic).

Results

A total of 245 participants completed utility interviews (51.4% female; mean age = 45.3y). Models predicting utility based on the four PASI location scores (head, upper limbs, trunk, lower limbs) had better fit/accuracy (e.g., R^2 , mean absolute error [MAE]) than models using the PASI total score. Head/upper limb scores were more strongly associated with utility than trunk/lower limb. The recommended model is the OLS linear model based on the four PASI location scores ($R^2 = 0.13$; MAE = 0.03). An alternative is recommended for situations when it is necessary to estimate utility based on the PASI total score.

Conclusions

The recommended scoring algorithm may be used to estimate utilities based on PASI scores of any treatment group with psoriasis. Because the PASI is commonly used in psoriasis clinical trials, this scoring algorithm greatly expands options for quantifying treatment outcomes in cost-effectiveness analyses of psoriasis therapies. Results indicate that psoriasis of the head/upper limbs could be more important than trunk/lower limbs, suggesting reconsideration of the standard PASI scoring approach.

Keywords: Health state utilities; Psoriasis Area Severity Index (PASI); psoriasis

JEL classification codes: I10; I19

Introduction

Decisions regarding allocation of healthcare resources are often informed by cost-effectiveness analyses in which treatments are assessed in terms of cost per quality-adjusted life year (QALY) gained [1]. QALYs are calculated using health state utilities, which are scores representing the strength of preference for a given health state [2, 3]. Although utilities for specific patient populations and treatment groups are often derived from generic preference-based measures such as the EQ-5D, these generic measures are not relevant or sensitive to every medical condition. Therefore, there is growing interest in development of preference-based scoring algorithms for condition-specific measures to estimate utilities in situations when generic instruments are not appropriate [4].

Dermatoses such as psoriasis may present a serious challenge for generic preference-based measures. For example, although the EQ-5D has demonstrated adequate measurement properties in some studies with psoriasis patients [5], this measure may be insensitive to psoriasis severity or treatment-related change [6-8]. The five items of the EQ-5D assess level of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. These concepts appear to have limited overlap with the impact of psoriasis [9]. To address this limitation of the EQ-5D, researchers have explored the possibility of adding a sixth item focusing specifically on psoriasis [7, 8]. In addition, there is a utility scoring algorithm associated with the Dermatology Life Quality Index (DLQI), a measure of dermatology-related quality of life impact [10, 11]. However, the DLQI is not psoriasis-specific, and it is not one of the most broadly used outcome instruments in psoriasis trials. Therefore, it has limited usefulness for deriving utilities specific to psoriasis.

One of the most commonly used psoriasis measures in clinical trials is the clinician-rated Psoriasis Area Severity Index (PASI) [12-16]. If utilities could be derived from PASI scores, it would be possible to obtain utilities from the wide range of clinical trials in which the PASI was administered. This would greatly expand the available options for estimating utilities that could be used in cost-effectiveness analyses of psoriasis treatments. Therefore, the purpose of this study was to develop an algorithm for deriving utilities from PASI scores.

Methods

Overview of study design

This study design followed three steps similar to other studies that have derived preference-based scoring algorithms for condition-specific measures [17-20]. First, health states corresponding to a range of PASI scores were developed. Second, utilities for each health state were obtained in a time trade-off (TTO) valuation study with general population respondents. Third, regression modeling was performed to identify a scoring algorithm for estimating utility based on any PASI score.

PASI

The PASI is a clinician-rated measure of psoriasis severity in four locations: head, upper limbs, trunk, and lower limbs [13]. For each location, a score is calculated based on psoriasis severity and surface area involvement. The severity score is comprised of three attributes (erythema [redness], induration [thickness], desquamation [scaling]) rated on a scale from 0 (none) to 4 (very severe) and summed for a severity score ranging from 0 to 12. The surface area score ranges from 0 (0%) to 6 (90 – 100%). Each of the four body location scores is calculated as the severity sum score multiplied by the surface area score (possible range = 0 to 72). Each location score is then multiplied by a specific correction score (head = 0.1, upper limbs = 0.2, trunk = 0.3, lower limbs = 0.4), with the adjusted scores summed to produce a total score ranging from 0 to 72.

Health state development

Studies developing preference-based scoring algorithms for condition-specific measures typically derive health states using psychometric techniques like Rasch analysis [4]. However, this

analytic approach was not appropriate for the PASI because it is not a conventional multi-item questionnaire. Instead, health states were developed for the subsequent utility valuation study based on PASI data from a clinical trial dataset [21]. The trial data for each patient included the PASI total score as well as the redness, thickness, scaling, and surface area scores for each of the four locations. Then, 40 health states were developed to correspond to the PASI scores of patients from this trial dataset. The 40 patients were selected to represent a wide range of PASI scores (listed in Table 1), while emphasizing the range most commonly observed in a clinical trial (i.e., roughly PASI 10 to 25). An effort was made to identify some health states with consistent severity across the four locations (e.g., health state 26) and other health states with unbalanced scores (e.g., health state 17, which has a head score that is more severe than the scores of the other three locations). This approach ensured that the 40 health states represented a wide range of psoriasis severity, and each health state corresponded to an actual PASI profile of a patient treated for psoriasis in a clinical trial.

With the input of four dermatologists, health state language was drafted based on the PASI scoring system. Each health state consisted of 16 statements describing the level of redness, thickness, scaling, and surface area coverage within the four locations (see Figure 1 for an example of a health state). An 11 x 17" page with photos (provided by the dermatologists) representing each level of redness, thickness, and scaling was presented along with the health states to ensure that respondents' TTO decisions were based on a reasonably accurate understanding of psoriasis severity levels.

Participants

Participants in the pilot and valuation studies were required to be at least 18 years old, reside in the United Kingdom (UK), and able to understand the interview procedures. Inclusion criteria did not specify clinical characteristics because interviews were intended to lead to a scoring algorithm for estimating utilities for cost-utility analyses in submissions to health technology assessment agencies, which often prefer that utilities represent general population values [22-24]. Participants were recruited via newspaper and online advertisements.

Pilot study

The health states were tested in a pilot study in December 2014 with 19 general population participants in London (29.4% female; mean [SD] age = 36.1 [12.7] years; age range = 19 to 67). The time horizon used in the TTO method and the number of health states per participant were varied. Participants generally reported that the health states were clear and easy to understand. Some participants suggested minor revisions in formatting, and the health states were edited accordingly. All TTO methods yielded utility scores in a reasonable range with logical discrimination among health states.

Valuation study

The valuation study was conducted in Edinburgh and London in January and February, 2015. Participants were required to provide written informed consent before completing study procedures, and all procedures and materials were approved by an independent Institutional Review Board (Ethical & Independent Review Services; Study Number 14145-01).

The 40 health states (listed in Table 1) were categorized into five severity groups ranging from mild (health states 1 to 8) to severe (health states 33 to 40). Each participant rated five health states, including one selected randomly from each of the five severity groups. As an introductory task, participants first ranked the health states in order of preference. Then, utilities for each health state were obtained in a TTO task. A 20-year TTO time horizon with 5% trading (1-year) increments was used because it was easy for participants to understand in the pilot study, and this relatively long time horizon is consistent with the chronic nature of psoriasis. Following standard utility assessment methodology [1-3], participants were offered a choice between living 20 years in the health state being

rated versus living shorter amounts of time in full health. For each health state, choices were presented in an order that alternated between longer and shorter durations in full health (e.g., 20 years, 0, 19, 1, 18...). The utility score was calculated based on the choice in which the respondent was indifferent between y years/months in the health state being evaluated and x years/months in full health (followed by dead).

The resulting utility estimate (u) is calculated as $u = x / y$. Each health state rated as better than dead received a utility score on a scale with the anchors of dead (0) and full health (1). When participants indicated that a health state was worse than dead, respondents were offered a choice between immediate death (alternative 1) and a 20-year life span (alternative 2) beginning with varying amounts of time in the health state being rated, followed by full health for the remainder of the time horizon. For these health states, the current study used a bounded scoring approach ($u = -x / t$, where x is the time in full health, and t is the total life span of alternative 2), which is commonly used to avoid highly skewed distributions for negative utilities [25, 26].

Regression analyses for modeling health state values

Regression models estimated the relationship between PASI scores and TTO utilities, following an established approach [4, 18-20, 27]. First, eight models were run using the PASI total score as the independent variable, followed by a parallel set of models with the four PASI location scores as independent variables.

The eight models began with the (1) ordinary least squares (OLS) linear regression model estimating the best-fitting linear relationship between PASI scores and utility: $E(y_i) = \beta_0 + \beta_1(x_i)$ where β_0 is the intercept, β_1 is the coefficient for PASI score x_i , and y_i is the utility. Subsequent models included: (2) random effects mixed model (accounting for the fact that each respondent provided multiple utilities, which may not be independent of each other); (3) mean model (providing an aggregate level prediction model, equally weighted across the 40 health states); (4) robust linear regression model (reducing the influence of outliers by down-weighting such data points in the estimation) [28]; (5) linear spline model (allowing for a bend or “knot” in the data; essentially two connected linear models); (6) robust regression spline model; (7) OLS model with a quadratic term (allowing for a non-linear relationship between PASI and utility); and (8) robust model with a quadratic term. Models were evaluated and compared based on indices of model fit (R^2 , adjusted R^2), accuracy of estimation (mean absolute error [MAE], frequency of large absolute prediction errors), plots of the mean residuals for each model, and line graphs depicting the relationship between PASI and utility.

Results

Valuation survey: respondent characteristics

A total of 285 potential participants (157 London; 128 Edinburgh) were scheduled for interviews, and 257 (139 London; 118 Edinburgh) attended and completed their interviews. Nine of the 257 participants (five London; four Edinburgh) were either unable to complete the utility interview procedures (seven who had difficulty understanding the TTO task and/or the health state content) or were found to be ineligible during the interview process (one with cognitive impairment and one with a psychiatric disorder that interfered with their ability to complete the study procedures). Thus, a total of 248 valid utility interviews were conducted (134 London; 114 Edinburgh). Three participants who provided negative (i.e., worse than dead) utilities for all five health states were excluded from analyses. Therefore, the analysis sample includes 245 respondents (131 London; 114 Edinburgh). The sample was 51.4% female ($n = 126$), with a mean age of 45.3 years. The most commonly reported ethnic/racial backgrounds were White (74.3%), Black (9.4%), and Asian (9.0%). Self-reported marital status included single (51.0%), married/living with a partner (36.7%), divorced (8.6%), and widowed (2.4%). Most participants reported being employed (35.1% full-time and 30.6% part-time). A little under half of the

sample had completed a university degree (33.9%) or a postgraduate degree (13.1%). The most commonly reported health conditions were depression (19.2%), anxiety (14.7%), hypertension (9.0%), arthritis (8.2%), and diabetes (7.8%).

TTO utility values

Each health state received at least 29 valuations. As expected, health states representing higher total PASI scores (i.e., greater psoriasis severity) tended to have lower mean utilities (Table 1). Some health state utilities deviated from the expected order based on total PASI score, likely due to the influence of the PASI location (e.g., head vs. upper limb) on respondent preferences. For example, the utility score for health state 17 (mean utility = 0.77; total PASI = 14.4; head = 32) was lower than the utility of health state 22 (mean utility = 0.80; total PASI = 18.4; head = 0).

A series of t-tests compared mean utilities across subgroups, finding minimal differences. There were no significant differences between men and women. Of the 40 health states, there was a statistically significant age difference (median split: older vs. younger) for only one health state (health state 13; $p = 0.01$). Comparisons between the London and Edinburgh subgroups found significant differences for only three of the 40 health states (health states 3, 17, and 27; $p < 0.05$). Comparisons between participants who worked full-time or part-time ($n = 161$) versus those who did not ($n = 84$) also found significant differences for only three of the 40 health states (health states 6, 16, and 26; $p < 0.05$). Similarly, comparisons between participants who were married or living with a partner ($n = 90$) versus the other participants ($n = 155$) found significant differences for only four of the 40 health states (health states 10, 26, 27, and 39; $p < 0.05$). Comparisons between participants with a university degree or higher ($n = 115$) versus those who did not have a university degree ($n = 130$) found significant differences for four of the 40 health states (health states 3, 7, 23, and 34; $p < 0.05$). When differences were observed, they were not in a consistent direction (i.e., no group had consistently higher or lower scores than the comparison group). Furthermore, with an adjustment for multiple comparisons (e.g., a Bonferroni correction), none of these differences would be considered statistically significant. Overall, these comparisons suggest that demographic variables did not have a meaningful or consistent impact on utility scores.

Regression modeling based on the PASI total score

Generally, model fit and accuracy are similar across the regression models based on the PASI total score, although some differences emerged (Table 2, Figure 2). There is minimal difference among the parameter estimates of the first three models (OLS linear model, the random effects model, and the mean model) among all truly comparable statistics. Although R^2 for the mean model (0.7847) is substantially higher, this cannot be compared to any other model because this model does not assess the same data points.

The robust models, which minimize the influence of outliers and high leverage points, yield results that are notably different from their OLS model counterparts. The robust models have higher intercepts and less steep slopes, resulting in higher utility scores, particularly in the more severe part of the PASI score range. Fit/accuracy statistics are not directly comparable between robust and OLS models (e.g., R^2 is not assessing the same data variance).

The two types of nonlinear models (spline and quadratic) offer some improvement in model fit/accuracy over corresponding linear models. Model differences are highlighted by the predicted plots (Figure 2). The two nonlinear models are similar in terms of utility estimates, fit, and accuracy. The spline model is preferred over the quadratic for two reasons. The spline performed slightly better than the quadratic (e.g., R^2 , MAE, frequency of absolute errors), and perhaps more importantly, its linear trends are either descending or flat. In contrast, the quadratic curves upward for PASI scores over 51.9

for the OLS quadratic and 43.0 for the robust quadratic (Figure 2), which could lead to nonsensical utility estimates among patients with severe psoriasis.

Regression modeling based on the four PASI location scores

Results of models with the four PASI location scores followed the same general patterns as the models described above using the PASI total score. However, running the models using the four PASI location scores resulted in small but consistent improvements in model fit/accuracy over the PASI total score, including increased R^2 as well as reduced MAE and frequency of absolute errors. Parameter estimates indicate that head and upper limb scores were more strongly associated with utility than trunk and lower limb (Table 3).

For the spline and quadratic model, however, the same problem emerged as described above for the total score quadratic model. These models yielded utilities that begin to increase, rather than continuing to decrease, in the far upper range of PASI scores that represent greatest psoriasis severity. This pattern, which results in illogical utility scores for the most severe health states with the trunk and lower limbs scores in the quadratic, as well as the trunk in the spline model (Figure 3).

Discussion

Given that the four body locations clearly differed in terms of impact on preference, it is recommended that a model with the four PASI location scores as independent variables be used to estimate utilities based on the PASI. The models predicting utility based on the four PASI location scores generally had better fit/accuracy than models using the PASI total score. Among the models using the four location scores, none offered notable improvement over the simple linear model. Therefore, the OLS linear model based on the four PASI location scores is recommended as the primary algorithm for deriving utilities from PASI scores (see scoring algorithm in Table 4).

In situations where the four PASI location scores are not available, it would be reasonable to use a scoring algorithm based on the PASI total score. Among the models based on the total score, the spline model with a knot at 29 offers the best model fit and accuracy. Therefore, this model is recommended for use when it is necessary to derive utilities from the PASI total score (see scoring algorithm in Table 4).

It was concluded that the robust models would not be recommended for future use. When evaluating these models, one must determine the extent to which outliers can be disregarded. Utility valuation studies, including the current study, typically elicit a wide range of values for each health state, with substantial heterogeneity and outliers. Because outliers such as negative scores and scores at the ceiling are a common and expected part of the utility valuation process, it is difficult to determine cut points at which these data points should be disregarded. Therefore, the robust models that largely disregard these values were ultimately not considered optimal for this scoring algorithm. Models that could lead to illogical relationships among utilities for individuals with severe psoriasis are also not recommended for future use (i.e., all quadratic models and the spline model using the four PASI location scores).

When deciding whether to derive utilities with a condition-specific preference-based (CSPB) approach, such as the scoring algorithm provided in the current study, advantages and disadvantages should be considered. CSPB measures are generally thought to be more sensitive to disease severity and treatment-related change than generic utility measures such as the EQ-5D [4]. However, CSPB measures sacrifice some comparability across diseases, studies, and cost-utility models, whereas utilities derived from generic measures have the advantage of maximizing comparability [4, 29]. Despite this limitation, there are situations when CSPB measures are the best available option for deriving utilities, such as when generic instruments do not assess content relevant to a particular disease. Psoriasis may present such a situation because the commonly used generic instruments do not assess dermatological

symptoms or their impact. Other recent research has addressed this challenge by adding psoriasis-specific dimensions to the EQ-5D to improve sensitivity to skin clarity, particularly at the highest PASI levels [7]. Future research could examine the relationship between utilities based on the EQ-5D with added dimensions and utilities derived directly from the PASI using the algorithm derived in the current study.

Findings from the models using the four PASI location scores could have important implications for scoring and interpretation of PASI results in clinical trials. In the regression model results, the parameter estimates and statistical significance levels indicate that psoriasis of the head had the strongest impact on health state preference, followed by the upper limbs, while the trunk and lower limbs had a much weaker association with utility. In addition to the statistical results, this difference among the four body locations is apparent from visual inspection of the lines/curves in Figure 3. The importance of head/upper limb psoriasis was also apparent during the interviews. Many respondents explained their preferences by reporting that skin patches of the head and upper limbs would be most troubling because these locations are most visible and least likely to be covered by clothes.

In contrast to current results highlighting importance of the head and upper limbs, the standard PASI scoring algorithm is more heavily weighted toward psoriasis of the lower limbs and trunk. Therefore, the standard PASI scoring may not optimally represent the impact of psoriasis. This finding should be interpreted with caution because the current preference values were derived from a general population sample. It is not known whether patients would agree that psoriasis of the head and upper limbs is most important. Therefore, it would be useful to replicate this preference task with a patient sample. If results from patients were similarly weighted toward the head and upper limbs, this finding would suggest re-evaluating the standard scoring approach of the PASI so that scores would better represent the impact of psoriasis.

A limitation of the proposed scoring algorithm is that utility is derived only from the four visible aspects of psoriasis assessed by the PASI, rather than the patient's broader experience of psoriasis. Symptoms such as pain and itching are not represented. While these symptoms likely co-occur with the visible signs captured by the PASI to some extent, it is possible that utilities derived from the PASI may not be entirely consistent with the patients' experience of psoriasis. To address this potential limitation, future research could examine validity of the proposed utility scoring algorithm by comparing the resulting utility scores to scores in instruments designed to assess patients' experience of psoriasis. For example, with a dataset that includes both the PASI and DLQI, it would be possible to run correlations between PASI-based utilities and the DLQI scores, which represent the quality of life impact associated with dermatological conditions. This analysis would provide an indication of the extent to which the proposed utility scoring algorithm is consistent with the patient's overall experience of psoriasis and its impact.

In sum, the current study provides a new method for obtaining utilities corresponding to a wide range of psoriasis severity. With the scoring algorithm provided in this study, utilities can be derived from any treatment group or patient sample whose psoriasis was rated with the PASI. Because the PASI is commonly used in clinical trials, the recommended scoring algorithm greatly expands options for quantifying outcomes in cost-effectiveness analyses, facilitating more accurate modeling to inform resource allocation decisions.

References

1. Brazier JR, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and Valuing Health benefits for Economic Evaluation. New York: Oxford University Press, 2007.
2. Feeny D. Preference-based measures: utility and quality-adjusted life years. In: Fayers P, Hays R, eds. Assessing quality of life in clinical trials 2nd ed. New York: Oxford University Press, 2005:405-31.
3. Torrance GW. Measurement of health state utilities for economic appraisal. *Journal of health economics* 1986;**5**(1):1-30
4. Brazier JE, Rowen D, Mavranouzouli I, et al. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health technology assessment* 2012;**16**(32):1-114 doi: 10.3310/hta16320[published Online First: Epub Date]].
5. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *Eur J Health Econ* 2014 doi: 10.1007/s10198-014-0638-9[published Online First: Epub Date]].
6. Pereira FR, Basra MK, Finlay AY, Salek MS. The role of the EQ-5D in the economic evaluation of dermatological conditions and therapies. *Dermatology* 2012;**225**(1):45-53 doi: 10.1159/000339865[published Online First: Epub Date]].
7. Pickard AS, Gooderham M, Hartz S, Nicolay C. EQ-5D health utilities: exploring ways to improve upon responsiveness in psoriasis. *J Med Econ* 2017;**20**(1):19-27 doi: 10.1080/13696998.2016.1219359[published Online First: Epub Date]].
8. Swinburn P, Lloyd A, Boye KS, Edson-Heredia E, Bowman L, Janssen B. Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value Health* 2013;**16**(8):1156-62 doi: 10.1016/j.jval.2013.10.003[published Online First: Epub Date]].
9. Lebwohl M. Psoriasis. *Lancet* 2003;**361**(9364):1197-204 doi: 10.1016/S0140-6736(03)12954-6[published Online First: Epub Date]].
10. Blome C, Beikert FC, Rustenbach SJ, Augustin M. Mapping DLQI on EQ-5D in psoriasis: transformation of skin-specific health-related quality of life into utilities. *Arch Dermatol Res* 2013;**305**(3):197-204 doi: 10.1007/s00403-012-1309-2[published Online First: Epub Date]].
11. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994;**19**(3):210-6
12. Feldman SR, Krueger GG. Psoriasis assessment tools in clinical trials. *Ann Rheum Dis* 2005;**64** **Suppl 2**:ii65-8; discussion ii69-73 doi: 10.1136/ard.2004.031237[published Online First: Epub Date]].
13. Fredriksson T, Pettersson U. Severe psoriasis--oral therapy with a new retinoid. *Dermatologica* 1978;**157**(4):238-44

14. Gaspari AA, Tying SK. Commonly Applied Clinical Trial Endpoints for Psoriasis and PsA. *Clinical and Basic Immunodermatology*. London, UK: Springer, 2008:526.
15. Navarini AA, Poulin Y, Menter A, Gu Y, Teixeira HD. Analysis of body regions and components of PASI scores during adalimumab or methotrexate treatment for patients with moderate-to-severe psoriasis. *J Drugs Dermatol* 2014;**13**(5):554-62
16. Spuls PI, Lecluse LL, Poulsen ML, Bos JD, Stern RS, Nijsten T. How good are clinical severity and outcome measures for psoriasis?: quantitative evaluation in a systematic review. *J Invest Dermatol* 2010;**130**(4):933-43 doi: 10.1038/jid.2009.391[published Online First: Epub Date] |.
17. Mavranezouli I, Brazier JE, Rowen D, Barkham M. Estimating a preference-based index from the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM): valuation of CORE-6D. *Med Decis Making* 2013;**33**(3):381-95 doi: 10.1177/0272989X12464431[published Online First: Epub Date] |.
18. Rentz AM, Kowalski JW, Walt JG, et al. Development of a preference-based index from the National Eye Institute Visual Function Questionnaire-25. *JAMA Ophthalmol* 2014;**132**(3):310-8 doi: 10.1001/jamaophthalmol.2013.7639[published Online First: Epub Date] |.
19. Yang Y, Brazier J, Tsuchiya A, Coyne K. Estimating a preference-based single index from the Overactive Bladder Questionnaire. *Value Health* 2009;**12**(1):159-66 doi: 10.1111/j.1524-4733.2008.00413.x[published Online First: Epub Date] |.
20. Yang Y, Brazier JE, Tsuchiya A, Young TA. Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the asthma quality of life questionnaire. *Med Decis Making* 2011;**31**(2):281-91 doi: 10.1177/0272989X10379646[published Online First: Epub Date] |.
21. Papp KA, Leonardi C, Menter A, et al. Brodalumab, an anti-interleukin-17-receptor antibody for psoriasis. *N Engl J Med* 2012;**366**(13):1181-9 doi: 10.1056/NEJMoa1109017[published Online First: Epub Date] |.
22. CADTH (Canadian Agency for Drugs and Technologies in Health). Guidelines for the economic evaluation of health technologies: Canada [3rd] ed. Ottawa: Canadian Agency for Drugs and Technologies in Health, 2006.
23. NICE (National Institute for Health and Care Excellence). Process and methods guides: Guide to the methods of technology appraisal 2013. In: NICE (National Institute for Health and Clinical Excellence), ed. London, UK: NICE, 2013.
24. PBAC (Pharmaceutical Benefits Advisory Committee). Guidelines for preparing submissions to PBAC, Version 4.4. Canberra, Australia: Australian Government Department of Health, PBAC, June 2013:344.
25. Matza LS, Sapa SJ, Dillon JF, et al. Health state utilities associated with attributes of treatments for hepatitis C. *Eur J Health Econ* 2015;**16**(9):1005-18 doi: 10.1007/s10198-014-0649-6[published Online First: Epub Date] |.
26. Rowen D, Brazier J. Health Utility Measurement. In: Glied S, Smith P, eds. *The Oxford Handbook of Health Economics*. New York: Oxford University Press, 2011:788-813.

27. Brazier J, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. *Med Decis Making* 2008;**28**(1):113-26 doi: 10.1177/0272989X07301820[published Online First: Epub Date]].
28. Ryan TP. *Modern Regression Methods*. 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2009.
29. Versteegh MM, Leunis A, Uyl-de Groot CA, Stolk EA. Condition-specific preference-based measures: benefit or burden? *Value Health* 2012;**15**(3):504-13 doi: 10.1016/j.jval.2011.12.003[published Online First: Epub Date]].

Figure and table legends

Table 1. Time Trade-Off Utility Scores for 40 Health States Representing PASI Severity

Table 2. Regression Models Estimating Utility Based on the Total PASI Score: Model Coefficients and Performance Statistics

Table 3. Regression Models Estimating Utility Based on the Four PASI Location Scores: Model Coefficients and Performance Statistics

Table 4. Recommended Models for Deriving Utilities from PASI Scores

Figure 1. Example of a Health State

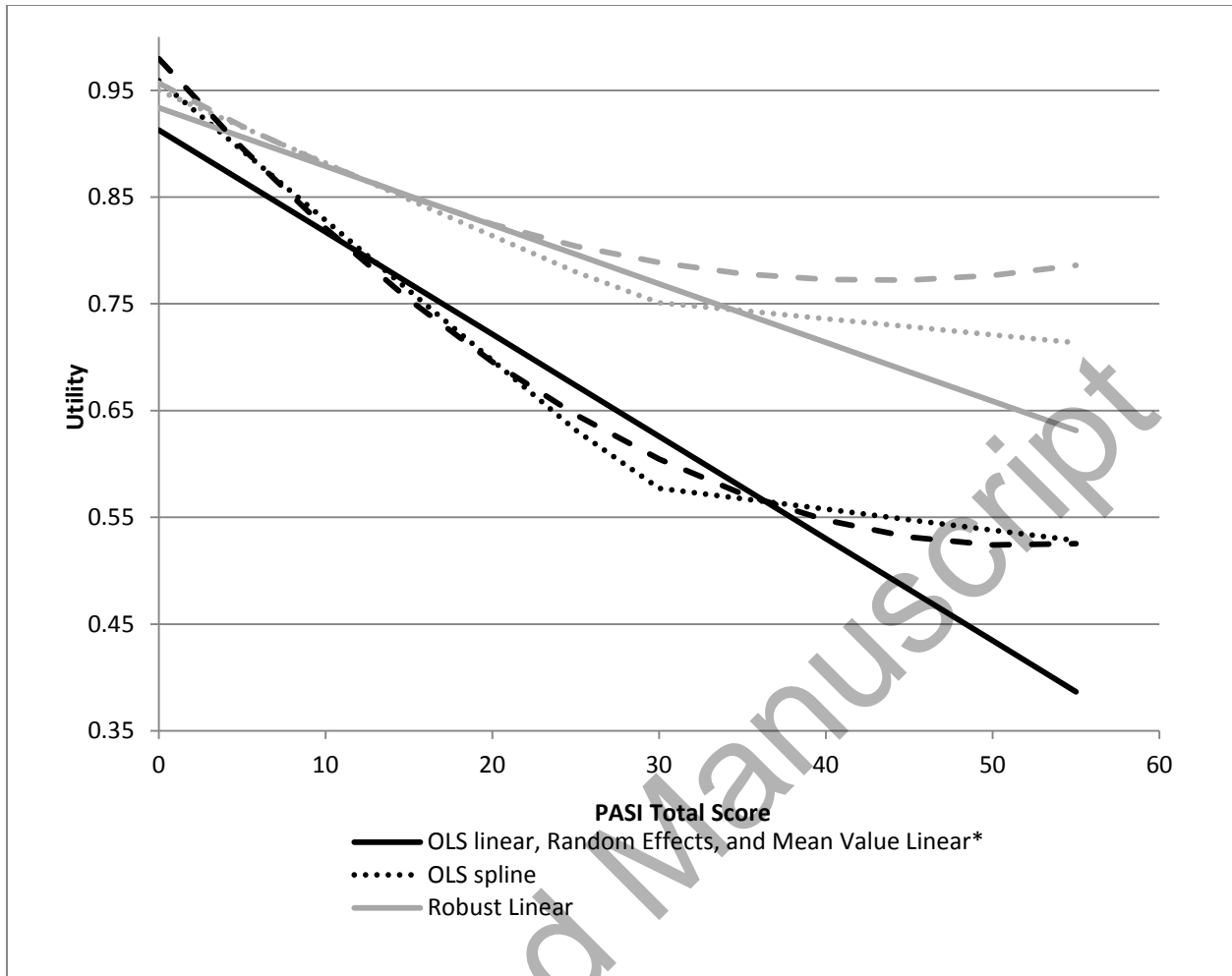
Figure 2. Utility Curves Predicted by Regression Models Estimating Utility Based on the Total PASI Score

Figure 3. Marginal Utility Curves Predicted by Regression Models Estimating Utility Based on the Four PASI Location Scores

Figure 1. Example of a Health State^a

<p>You have been diagnosed with a chronic skin condition that produces thickened, red, scaly patches on your body. These patches are frequently itchy and dry.</p>			
<p>Head (including face, scalp, and neck)</p>	40%	Surface Area	Covering 40% of this body surface area
	4	Redness	Fiery red
	3	Thickness	Very raised and thickened
	3	Scaling	Most of your patches are covered with thick, rough flakes
<p>Upper Limbs (including arms, shoulders, and hands)</p>	20%	Surface Area	Covering 20% of this body surface area
	1	Redness	Light red
	1	Thickness	Almost flat, but slightly raised and thickened
	1	Scaling	Some of your patches are covered with fine flakes
<p>Trunk (including anything above the hips, such as the stomach, back, and chest)</p>	20%	Surface Area	Covering 20% of this body surface area
	1	Redness	Light red
	1	Thickness	Almost flat, but slightly raised and thickened
	1	Scaling	Some of your patches are covered with fine flakes
<p>Lower Limbs (including all areas below the waist such as hips, buttocks, legs, and feet)</p>	40%	Surface Area	Covering 40% of this body surface area
	2	Redness	Moderately red
	2	Thickness	Moderately raised and thickened
	1	Scaling	Some of your patches are covered with fine flakes

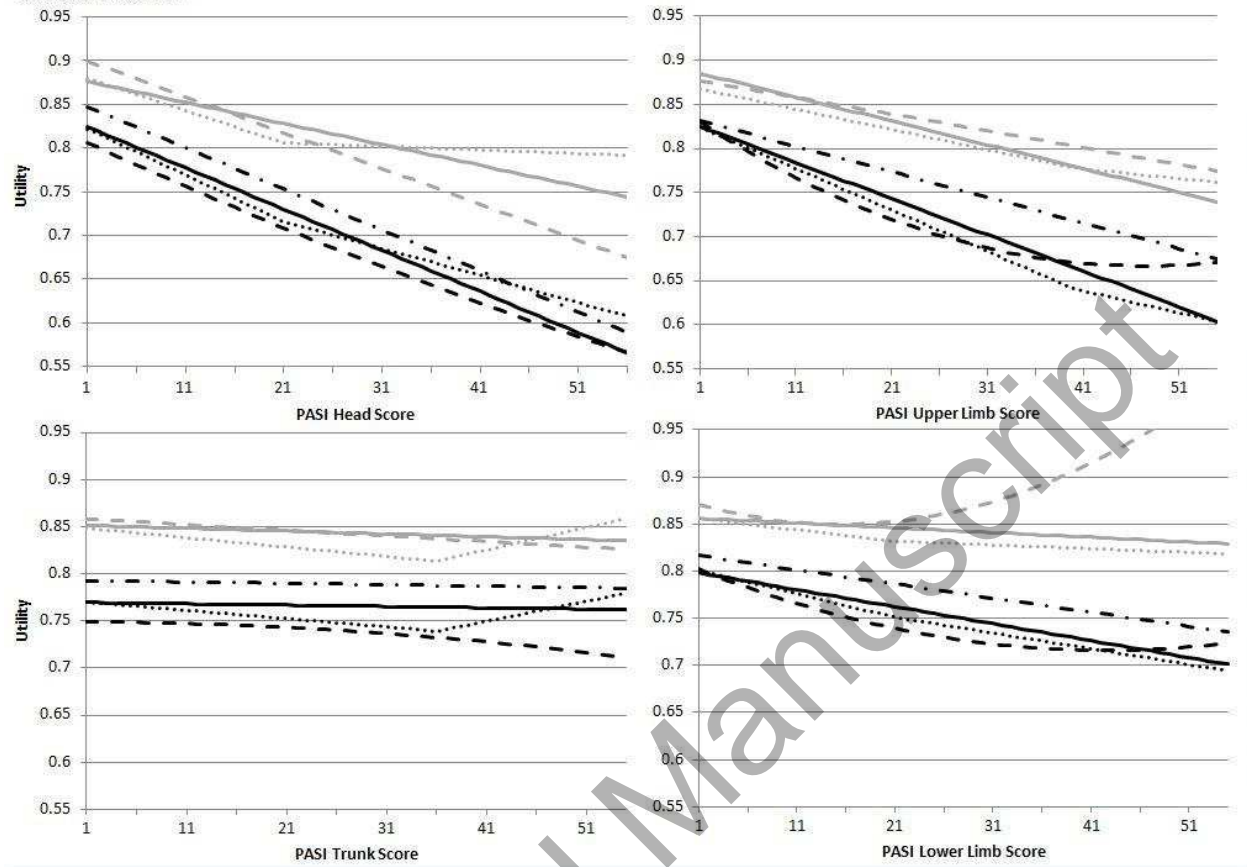
^a Health State 13 representing a PASI total score of 12.



Footnote:

* The OLS Linear, Random Effects, and Mean Value Linear models are represented by a single line because their utility predictions are indistinguishable from each other.

Figure 2. Marginal* Utility Curves Predicted by Regression Models Estimating Utility Based on the Four PASI Location Scores



Caption: Figure 2. Marginal* Utility Curves Predicted by Regression Models Estimating Utility Based on the Four PASI Location Scores

Footnotes:

* Each graph presents the model predicting utility based on scores for the specified PASI location score when the other three location scores are held constant at their median across the 40 health states valued in this study (medians: head = 12, upper limbs = 15, trunk = 14, lower limbs = 18).

** The OLS Linear and Mean Value Linear models are represented by a single line because their utility predictions are indistinguishable from each other.

Figure 2 Legend:

- The OLS Linear Regression and Mean Models**
- Linear Spline Model
- - - Random Effects Mixed Regression Model
- - - OLS Model with a Quadratic Term
- Robust Regression Model
- Robust Regression Spline Model
- - - Robust Model with a Quadratic Term

Table 1. Time Trade-Off Utility Scores for 40 Health States Representing PASI Severity

Severity Group	HS #	Health State Symptom Severity: PASI Scores					Valuation Study Results		
		Total Score	Head	Upper Limbs	Trunk	Lower Limbs	N	Mean Utility ^a	SD
1	1	0.2	0	1	0	0	31	0.96	0.03
1	2	2.6	2	3	2	3	31	0.90	0.16
1	3	3.4	2	0	0	8	31	0.89	0.17
1	4	4.1	10	3	3	4	29	0.89	0.11
1	5	5.9	3	4	0	12	30	0.85	0.17
1	6	6.2	4	4	10	5	31	0.90	0.08
1	7	6.8	6	6	6	8	31	0.87	0.16
1	8	7.8	4	14	6	7	31	0.85	0.20
2	9	8.4	10	12	6	8	31	0.81	0.33
2	10	8.8	0	3	6	16	31	0.92	0.08
2	11	10.9	1	4	12	16	31	0.89	0.13
2	12	11.4	10	10	12	12	31	0.83	0.20
2	13	12	30	6	6	15	31	0.73	0.22
2	14	13.2	12	14	12	14	30	0.82	0.22
2	15	13.5	3	12	12	18	30	0.77	0.29
2	16	14.2	6	16	24	8	30	0.78	0.30
3	17	14.4	32	0	16	16	31	0.77	0.36
3	18	14.6	12	10	14	18	31	0.76	0.29
3	19	16.5	12	24	3	24	29	0.64	0.49
3	20	17.6	24	16	12	21	31	0.71	0.31
3	21	18	0	18	12	27	31	0.81	0.24
3	22	18.4	0	7	14	32	31	0.80	0.29
3	23	18.6	40	8	30	10	31	0.64	0.33
3	24	19.3	15	14	14	27	30	0.74	0.30
4	25	19.4	8	21	32	12	30	0.73	0.33
4	26	20	20	20	20	20	31	0.70	0.35
4	27	20.3	36	22	9	24	31	0.62	0.40
4	28	21.6	21	24	21	21	30	0.71	0.28
4	29	22.2	6	24	24	24	31	0.80	0.28
4	30	23.4	18	24	24	24	30	0.69	0.35

Severity Group	HS #	Health State Symptom Severity: PASI Scores					Valuation Study Results		
		Total Score	Head	Upper Limbs	Trunk	Lower Limbs	N	Mean Utility ^a	SD
4	31	25.2	28	28	28	21	31	0.57	0.43
4	32	26.9	14	27	27	30	31	0.58	0.44
5	33	29.1	28	36	45	14	31	0.58	0.44
5	34	31.2	24	24	20	45	31	0.58	0.44
5	35	34.3	35	28	28	42	31	0.57	0.48
5	36	35.1	27	36	36	36	30	0.53	0.45
5	37	36.4	40	48	36	30	29	0.51	0.44
5	38	41.4	50	40	36	44	31	0.47	0.54
5	39	44.3	20	33	55	48	31	0.61	0.44
5	40	52.2	27	55	55	55	31	0.56	0.37

^a Time trade-off utility scores are on a scale anchored with 0 representing dead and 1 representing full health.

HS = Health State; PASI = Psoriasis Area Severity Index

Table 2. Regression Models Estimating Utility Based on the Total PASI Score: Model Coefficients and Performance Statistics

	OLS Linear	Mean Value Linear	Random Effects Linear	Robust Linear	OLS Spline ^a	Robust Spline	OLS Quadratic ^c	Robust Quadratic ^c
Intercept	0.91***	0.91***	0.91***	0.93***	0.96***	0.95***	0.98***	0.96***
PASI	-	-	-	-	-	-	-	-
	0.0096** *	0.0096** *	0.0097** *	0.0055** *	0.013** *	0.0068** *	0.018***	0.0086** *
Add-on PASI > 29	--	--	--	--	0.011** *	0.0053** *	--	--
Quadratic Term	--	--	--	--	--	--	0.00017* *	0.0001**
Random Intercept Variance	--	--	0.063***	--	--	--	--	--
N	1225	40	1225	919	1225	919	1225	919
R ²	0.1103	0.7847	--	0.0684	0.1197	0.0721	0.1181	0.0712
Adjusted R ²	0.1095	0.7791	--	--	0.1182	--	0.1167	--
MAE-I ^b	0.2094	0.0457	0.1101	0.1946	0.2084	0.1935	0.2085	0.1938

weighted MAE	--	--	--	0.1061	--	0.1050	--	0.1053
MAE-M ^b	0.0457	0.0457	0.0382	0.1020	0.0381	0.1008	0.0410	0.1009
Count >.05 ^c	15	15	12	26	13	26	13	26
Count >.10 ^c	4	4	2	16	2	17	1	17
Count >.05 ^c weighted ^d	--	--	--	4	--	3	--	3
Count >.10 ^c weighted ^d	--	--	--	0	--	0	--	0

^a This is the alternate model recommended for deriving utilities from the PASI total score in situations when the four PASI location scores are not available.

^b MAE-I (mean absolute error at the individual level) is the mean of the absolute value of the difference between each observation (i.e., every utility score for every health state) and the predicted value based on the regression model. For MAE-M (mean absolute error of utility means), the mean utility score of each health state was calculated first. Then, the MAE-M was computed as the absolute value of the difference between this mean (i.e., the mean utility score for each health state) and the predicted value based on the regression model.

^c The number of mean residual errors greater than 0.05 or greater than 0.10 (i.e., the number of health states for which the difference between predicted and observed mean utility values exceed the specified threshold of 0.05 or 0.10)

^d The weighted observed mean utility scores used to calculate the mean residuals are obtained from the robust regression, which down-weights outliers and high leverage points.

*p < 0.05; **p < 0.01; *** p < 0.001

Table 3. Regression Models Estimating Utility Based on the Four PASI Location Scores: Model Coefficients and Performance Statistics

	OLS Linear ^a	Mean Value Linear	Random Effects Linear	Robust Linear	OLS Spline	Robust Spline	OLS Quadratic ^c	Robust Quadratic ^c
Intercept	0.92***	0.92***	0.92***	0.93***	0.95** *	0.95***	0.96***	0.96***
Head	- 0.0047** *	- 0.0047** *	- 0.0047** *	- 0.0024** *	- 0.0053 *	- 0.0037** *	-0.0052	-0.0041*
Upper	-0.0041*	- 0.0041** *	- 0.0029**	- 0.0027** *	- 0.0047 *	- 0.0023**	-0.0072*	-0.0019
Trunk	-0.00014	-0.00011	-0.0015	-0.0003	- 0.0008	-0.0010	-0.000029	-0.0006

	OLS Linear ^a	Mean Value Linear	Random Effects Linear	Robust Linear	OLS Spline 8	Robust Spline	OLS Quadrati c	Robust Quadrati c
Lower	-0.0018	-0.0018*	-0.0015*	-0.0005	- 0.0025	-0.0012	-0.0042	-0.0031
Add-on Head > 20	--	--	--	--	0.0022	0.0033*	--	--
Add-on Upper > 40	--	--	--	--	0.0068	-0.0012	--	--
Add-on Trunk > 36	--	--	--	--	0.0030	0.0034	--	--
Add-on Lower > 21	--	--	--	--	0.0008 3	0.0008	--	--
Head- squared	--	--	--	--	--	--	0.000014	0
Upper- squared	--	--	--	--	--	--	0.000076	0
Trunk- squared	--	--	--	--	--	--	-0.000012	0
Lower- squared	--	--	--	--	--	--	0.000049	0.0001
Random Intercept Variance	--	--	0.063***	--	--	--	--	--
N	1225	40	1225 (245)	1225 (920)	1225	1225 (921)	1225	1225 (921)
R ²	0.1276	0.9073	--	0.0789	0.1322	0.0824	0.1320	0.0821
Adjusted R ²	0.1248	0.8967	--	--	0.1265	--	0.1263	--
MAE-I ^b weighted	0.2070	0.0308	0.1085	0.1919	0.2062	0.1913	0.2063	0.1913
MAE MAE- M ^b	--	--	--	0.1032	--	0.1033	--	0.1026
Count >.05 ^c	5	5	6	28	4	28	4	28
Count >.10 ^c	1	1	1	21	0	15	0	16
Count >.05 ^c weighted a	--	--	--	1	--	0	--	0
Count >.10 ^c weighted	--	--	--	0	--	0	--	0

	OLS Linear^a	Mean Value Linear	Random Effects Linear	Robust Linear	OLS Spline	Robust Spline	OLS Quadrati c	Robust Quadrati c
--	-----------------------------------	----------------------------------	--------------------------------------	--------------------------	-----------------------	--------------------------	-------------------------------	----------------------------------

^a This is the primary model recommended for deriving utilities from the PASI.

^b MAE-I (mean absolute error at the individual level) is the mean of the absolute value of the difference between each observation (i.e., every utility score for every health state) and the predicted value based on the regression model. For MAE-M (mean absolute error of utility means), the mean utility score of each health state was calculated first. Then, the MAE-M was computed as the absolute value of the difference between this mean (i.e., the mean utility score for each health state) and the predicted value based on the regression model.

^c The number of mean residual errors greater than 0.05 or greater than 0.10 (i.e., the number of health states for which the difference between predicted and observed mean utility values exceed the specified threshold of 0.05 or 0.10)

^d The weighted observed mean utility scores used to calculate the mean residuals are obtained from the robust regression, which down-weights outliers and high leverage points.

*p < 0.05; **p < 0.01; *** p < 0.001

Table 4. Recommended Models for Deriving Utilities from PASI Scores

Name of Model	Scoring Algorithm with Coefficients
Primary Model: OLS Linear Model based on the Four PASI Location Scores	Utility = 0.92 + (-0.0047 * PASI head) + (-0.0041 * PASI upper limbs) + (-0.00014 * PASI trunk) + (-0.0018 * PASI lower limbs)
Alternate Model ^a : Spline Model based on the PASI Total Score	Utility = 0.96 + (-0.013 * PASI total) + (0.011 * Max[0, PASI total - 29]) ^b

^a Recommended for use when the four PASI location scores are not available.

^b The term “Max[0, PASI total - 29]” is a derived variable which takes the value 0 when the PASI total score is less than or equal to 29, which is called the spline change point or “knot.” When the PASI total score is greater than 29, the value of this variable is PASI total score - 29.