Validation of Patient Global Item for Quality of Life Impact on Hidradenitis Suppurativa

Joslyn S Kirby, MD, MS, MEd[1] *, Brittainy Hereford, BS[2] *, Linnea Thorlacius ,MD, PhD[3] , Bente Villumsen[4], John R Ingram[5], Amit Garg[6], Melissa Butt, MPH[1], Solveig Esmann[3], Tonya King, PhD[7], Jerry Tan, MD[8]**, Gregor BE Jemec, MD[3]**

*co-first authors
**co-senior authors

[1] Department of Dermatology, Penn State Health Milton S. Hershey Medical Center, Hershey, PA, USA
[2]Penn State College of Medicine, Hershey, Pennsylvania
[3] Department of Dermatology, Zealand University Hospital, Roskilde; Health Sciences Faculty, University of Copenhagen, Denmark
[4] The Patients' Association HS Denmark, Copenhagen, Denmark.
[5] Division of Infection & Immunity, Cardiff University, University Hospital of Wales, Heath Park, Cardiff, CF14 4XN, United Kingdom
[6] Department of Dermatology, Donald and Barbara Zucker School of Medicine at Hofstra Northwell, New Hyde Park, New York
[7]Department of Public Health Sciences, Penn State College of Medicine, Hershey, Pennsylvania
[8] Department of Medicine, University of Western Ontario, Windsor, Ontario, Canada

Corresponding author:
Joslyn S Kirby, MD, MS, MEd
Associate Professor
Penn State Milton S Hershey Medical Center
500 University Dr, HU14
Hershey PA 17033
Jkirby1@pennstatehealth.psu.edu
1-717-531-8307

Funding Sources: Dr Kirby received funding from the Agency for Healthcare Research and Quality for this research (K08HS024585).
Use of REDCap through Penn State is supported by NIH/NCATS Grant Number UL1 TR000127 and UL1 TR002014 through The Penn State Clinical & Translational Research Institute, Pennsylvania State University CTSA
Dr Ingram received funding from Health and Care Research Wales (Health Fellowship 14-08).

IRB approval status: IRB approved

Conflicts of Interest:
Kirby: AbbVie: Speaker, Advisory Board (Honoraria), Investigator; Incyte, ChemoCentryx: Consultant (Fees), Investigator; UCB: Investigator; InflaRx: Investigator
Thorlacius: Abbvie, Janssen: travel expenses. Regeneron: Investigator
Garg: Advisor for AbbVie, Pfizer, Janssen, Asana Biosciences, and UCB (honoraria)

Ingram: UCB Pharma, Novartis: Consultant; Abbvie: travel expenses.
Tan: UCB Advisory Board (Honoraria); Incyte: Investigator
Jemec: Advisory Board (honoraria): AbbVie, Chemocentryx, Coloplast, Incyte, Inflarx, Novartis, Pierre Fabre and UCB; Abbvie, Leo Pharma, Janssen-Cilag, Regeneron, Sanofi, Astra-Zeneca and Novartis: Investigator; AbbVie, Boehringer-Ingelheim, Galderma and MSD: speaker (honoraria); Abbvie, Leo Pharma and Novartis: unrestricted grants.
Hereford, Villumsen, Butt, Esmann, King: None

Keywords: hidradenitis suppurativa; acne inversa; measurement; Clinical Trial; Outcome Instrument; global assessment; PtGA, patient reported outcome; assessment

Key point word count: 93/140
Abstract word count: 228/250
Manuscript word count: 2,535/3,000
References: 5
Figures: 1
Tables: 4
Supplements: X

Bulleted Statements (104/140 words)

What's already known about this topic?

- HS can have a large negative impact on health-related quality of life.
- Patient global assessments are useful in eliciting the important aspects of the patient's perspective on an individual basis.
- There are few studies that developed and investigated the measurement properties of a patient global assessment item for HS.

What does this study add?

- This study describes the development, validation, and psychometric properties of a patient global assessment item for HS.
- This patient global assessment item for health-related quality of life shows promise with its validity and reliability, allowing for its further development as a clinical research tool.

Abstract (247/250 words)

Background

Hidradenitis suppurativa (HS) is a chronic inflammatory disease that is not well understood. The HS core outcome set calls for a patient global assessment (PtGA), yet an outcome measure for this domain has not been established.

Objectives

Our aim is to assess the reliability and validity of a candidate single-item PtGA for HS-specific health-related quality of life (HRQOL).

Methods

Qualitative concept elicitation interviews followed by cognitive debriefing interviews were conducted with HS patients in Denmark (DK) and the United States (US),. A cross-sectional observational study was conducted with adults with HS in the US and DK. The candidate PtGA item, demographic items, and multiple patient-reported scales including the Hidradenitis Suppurativa Quality of Life (HiSQOL), Dermatology Life Quality Index (DLQI), numerical rating scale (NRS) for pain, and others were concurrently administered.

Results

Convergent validity of the PtGA was supported with large correlations with the DLQI (r = 0.78, [95CI: 0.74-0.82]) and HiSQOL score r = 0.82 [95CI: 0.78-0.85]. The PtGA displayed known-groups validity with the DLQI score bands based on significance of an analysis of variance ($p<.0001$). Good test-retest reliability was supported by the intraclass correlation coefficient (ICC value = 0.82, [95CI: 0.78-0.85]) for those who reported stable HS. Responsiveness was assessed by differences in PtGA score against a patient reported assessment of change, and significant differences were found.

Conclusions

The single-item PtGA has exhibited reliability and validity in assessing HS-specific HRQOL in HS patients, making it a good provisional tool for HS clinical research.

Introduction

Hidradenitis suppurativa (HS), also known as acne inversa, is a chronic, recurrent inflammatory disease known to cause painful nodules in the axillae, groin, breast, and buttocks regions. There is a frequent delay in diagnosing patients with HS, as it is commonly mistaken for a skin infection.[1] HS has an estimated prevalence of 0.1-4% worldwide.[2] Those ranging from puberty to age 50 are often affected and there is a predilection for women.[1] Treatments such as topical antibiotics, biologics, and systemic antibiotics are routinely prescribed, however, there is a need to develop and investigate additional treatments to reduce the negative impact of this condition.[2]

The HIdradenitis SuppuraTiva cORe outcomes set International Collaboration (HISTORIC) was formed to construct a Core Outcome set (COS) to guide assessments in future clinical trials.[2,3,4] Six domains were recommended in the core domain set: pain, physical signs, HS-specific quality of life, progression of course, symptoms, global assessment, both patient- and physician-rated. Single-item assessments such as a patient global assessment (PtGA) assess a construct comprehensively, such as health-related quality of life, and allows the patient to consider all contributory factors (such as multiple symptoms, treatments, risks, and benefits) related to the condition or disease. While a single item PtGA is needed for clinical trials, this type of broadly encompassing item will likely provide an effective and agile method for physicians to assess patients in the outpatient setting. To our knowledge, there have not been any studies describing the development or measurement properties of a PtGA specifically for HS. Thus, the objective of this study was to develop and investigate the measurement properties of a single-item PtGA, to serve as a valid and reliable method of assessing HS-specific HRQOL.

Methods

The single item PtGA item was concurrently developed along with the HiSQOL, a multi-item HS-specific HRQOL measure.[5] In short, a mixed methods study design was utilized and included four phases aligned with guidance from the US Food and Drug Administration.[6] Conduct of the study was overseen by the international group of investigators, which included patient research partners, clinicians with expertise in HS, and researchers with expertise in instrument

[1]Jemec, G. B., MD. (2012). Hidradenitis Suppurativa. *New England Journal of Medicine*, 366, 158-164. doi:10.1056/NEJMcp1014163

[2]Thorlacius, L., Ingram, J. R., Villumsen, B., Esmann, S., Kirby, J. S., … Gottlieb, A. B. (2018). A core domain set for hidradenitis suppurativa trial outcomes: an international Delphi process. *British Journal of Dermatology*, 179(3), 642–650. https://doi.org/10.1111/bjd.16672

[3]Thorlacius, L., Garg, A., Ingram, J. R., Villumsen, B., Theut Riis, P., Gottlieb, A. B., Merola, J. F., Dellavalle, R., Ardon, C., Baba, R., Bechara, F. G., Cohen, A. D., Daham, N., Davis, M., Emtestam, L., FernÃ¡ndez-PeÃ±as, P., Filippelli, M., Gibbons, A., Grant, T., Guilbault, S., Gulliver, S., Harris, C., Harvent, C., Houston, K., Kirby, J. S., Matusiak, L., Mehdizadeh, A., Mojica, T., Okun, M., Orgill, D., Pallack, L., Parks-Miller, A., Prens, E. P., Randell, S., Rogers, C., Rosen, C. F., Choon, S. E., van der Zee, H. H., Christensen, R., & Jemec, G. B. E. (2018). Towards global consensus on core outcomes for hidradenitis suppurativa research: an update from the HISTORIC consensus meetings I and II. *The British Journal of Dermatology, 3,* 715–721

[4]Thorlacius, L., Ingram, J. R., Garg, A., Villumsen, B., Esmann, S., Kirby, J. S., Gottlieb, A. B., Merola, J. F., Dellavalle, R., Christensen, R., & Jemec, G. B. (2017). Protocol for the development of a core domain set for hidradenitis suppurativa trial outcomes. *BMJ open, 2,* e014733.

[5]Joslyn S Kirby, MD, MS, MEd, Linnea Thorlacius ,MD, PhD, Bente Villumsen, John R Ingram, Amit Garg, Karl Bang Christensen, Melissa Butt, MPH, Solveig Esmann, Jerry Tan, MD, Gregor BE Jemec, MD The Hidradenitis Suppurativa Quality of Life (HiSQOL) score: Development and Validation. British Journal of Dermatology [submitted]

[6]Food and Drug Administration (FDA). *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.* Rockville, MD.2009.

development. People with HS who were 18 years or older were identified based on diagnostic code for HS (International Classification of Diseases, Ninth Revision (ICD-9) code 705.83) in the medical record at academic institutions in the United States and Demark. People who gave informed consent, had a confirmed diagnosis of HS, and were fluent in English or Danish were recruited by phone and in clinic. Semi-structured concept elicitation interviews were conducted with people with HS by two investigators with experience (JS, ES). The PtGA item was developed using the qualitative data and the expertise of the investigators and four patient research partners. The initial item was translated into Danish based on recognized methods for cross-cultural adaptation.[7] Cognitive debriefing interviews and focus groups were conducted with people with HS to evaluate the item. All sessions were audio recorded, transcribed, and the qualitative data was analyzed with Nvivo 11 software (QSR International, Burlington, MA).

Field testing and psychometric assessment of the item was conducted with an observational non-interventional study conducted in the US and DK. The candidate PtGA item, Hidradenitis Suppurativa Quality of Life (HiSQOL) measure[7], Dermatology Life Quality Index (DLQI)[8,9], the Hospital Anxiety and Depression Scale (HADS) and numerical rating scale (NRS) for pain[10] were administered concurrently.[5] A web version of all instruments and items was developed in REDCap (Research Electronic Data Capture), a secure, web-based application designed to support data capture for research studies.[11] All measures and a patient-rated perception of change in HS item were administered a second time (24-72 hours later) to evaluate test-retest reliability. This timeframe was chosen due to the unpredictable, intermittent, and rapid onset of HS worsening. This study was approved by the ethics committee of each institution and the Danish Data Protection Agency. All participants gave informed consent prior to participation in the study.

Analysis

Patient demographics and characteristics were reported as means with standard deviations and sum totals with percentages. Psychometric properties of the PtGA item including test-retest reliability, convergent validity, known-groups validity, and responsiveness were assessed. Test-retest reliability was used to evaluate score reproducibility over time when the condition is stable. Test-retest reliability was assessed with participants who reported stable HS and had two instances of data. Intraclass correlation coefficients (ICCs) were used to appraise test-retest reliability, where an ICC $\geq$ 0.9 is an indication of excellent reliability, 0.75 to 0.9 is good reliability, 0.5 to 0.75 is moderate reliability, and ICC < 0.5 is poor reliability.[12] It was hypothesized the item would have at least good test-retest reliability. Convergent validity was assessed with the Spearman correlation using Fisher's z Transformation of the PtGA and other measures using baseline responses (Day 0). Correlations were interpreted as 0.00-0.19 poor,

[7]Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000;25(24):3186-3191.
[8]Basra MK, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. *Br J Dermatol.* 2008;159(5):997-1035.
[9]Lewis V, Finlay AY. 10 years experience of the Dermatology Life Quality Index (DLQI). *The journal of investigative dermatology Symposium proceedings.* 2004;9(2):169-180.
[10]Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *Journal of pain and symptom management.* 2011;41(6):1073-1093.
[11]Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;421(2):377-381.
[12]Portney LG, Watkins MP. Prentice Hall; New Jersey: 2000. Foundations of clinical research: applications to practice.

0.20-0.39 average, 0.40-0.59 moderate, 0.60-0.79 significant, and 0.80-1.0 strong.[13] It was hypothesized the item would have at least strong correlation with the DLQI and HiSQOL. Known-groups validity was based on an analysis of covariance (ANCOVA) for DLQI known groups with the PtGA responses as the dependent variable, and adjustment for country. It was hypothesized there would be significant differences of the PtGA item scores for the DLQI groups. Responsiveness was assessed using ANCOVA for the mean differences in the PtGA score based on the patient global impression of change in HS as improved, stable, or worsened, and adjusted for country. It was hypothesized there would be significant differences in the PtGA across these groups. Data analysis and statistical tests were computed using SAS software, version 9.4 (SAS Institute, Inc., Cary NC, US).

Results

Development of the Patient Global Assessment Item (PtGA)

After concept elicitation and cognitive debriefing, the candidate PtGA for HS-specific HRQOL was finalized as: “In the past 7 days, how much has HS influenced your quality of life?” Five response levels were provided and included: not at all, slightly, moderately, very much, and extremely. The score applied to each response was 0, 1, 2, 3, and 4 respectively.

Field Testing and Psychometric Assessment

Overall, 441 patients with a diagnosis of HS participated in this study; most were female and the majority reported a moderate or very large impact on HRQOL based on DLQI score (Table 1). The novel PtGA item responses included the full range of responses (Figure 1) and the most frequent response categories were ‘slightly’ or ‘moderately’.

The PtGA item demonstrated convergent validity through correlation with existing measures (Table 2). The strongest overall correlation of the PtGA, a very strong correlation of 0.82 (0.78-0.85), was with the HiSQOL total score. The PtGA item had overall correlations of 0.78 (0.74-0.82) with the DLQI and 0.65 (0.60-0.71) with the NRS for pain. The PtGA item had less strong overall correlations, 0.53 (0.46-0.59) and 0.55 (0.48-0.61) with the HADS depression and anxiety subscale scores, respectively. Correlations for the US sample were lower than the Danish sample, but as described here, the HiSQOL had the highest correlation for both groups, followed by the DLQI, and NRS for pain. In addition, the PtGA demonstrated known groups validity in a comparison of its scores across DLQI score groups (Figure 2).

The PtGA demonstrated good test-retest reliability (ICC 0.84 (0.80, 0.87)) among patients with stable HS (Table 2). The test-retest reliability was similar for respondents from DK and the US, as well as across age groups and race. Responsiveness of the PtGA item was measured by assessing the mean score change from the first to the second response for three groups of respondents that reported HS worsening, stable disease, or improvement. For those reporting improved HS, the mean PtGA score was significantly lower, with mean (95%CI) decrease of -0.54 points ($p<.0001$) indicating decreased impact on HRQOL. For those with stable HS, there was a mean difference of 0.28 point ($p<.0001$). For those with worsening of HS, there was no

[13]Landis J. R., Koch G. G. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi: 10.2307/2529310.

significant mean change (0.03 point, p=.78). In regression analyses, there was a significant difference in mean PtGA score as defined by disease status (better, stable, worse) (p=0.009) and country (p=.68), age (p=.68), and sex (p=.87) did not significantly contribute to scores.

Discussion

In this study, the single-item PtGA has shown suitable reliability and validity to appraise HS-specific HRQOL in adults with HS. Convergent validity was exhibited in comparing the PtGA item with the DLQI, HADS, HiSQOL, and NRS scales. The known score bands of the DLQI were used to demonstrate differences and support known-groups validity for the PtGA. A valid PtGA serves a great clinical tool to integrate and emphasize the patient perspective into clinical care. Its essence lies in allowance of variability from patient-to-patient to voice key issues and needs during both the disease and treatment course. Patient needs being met leads to greater adherence in treatment, thus increasing patient and physician satisfaction and overall outcomes. The use of a physician global assessment (PGA) in clinical trials has been effective in assessing and tracking dermatologic disease severity over time. Simplicity in its use in a clinical setting can result in a high physician compliance rate[14], thus allowing for the PGA to dually serve as a measurement of physician performance as well as a tool for justifying changes in care, both clinically and for cost. Comprehensive data along with follow-up gives a quantifiable metric of disease and patient improvement. PGA use does have its limitations; notably, accounting for intra-rater versus inter-rater variability in disease assessment. Analyzing disease severity at various time points is also a limitation. The use of photographs at each time point can be helpful, but there is a need for precision and consistency in assessment at each time point to accurate assess true disease state improvement.

Test-retest analyses demonstrated good reliability among patients that reported stable HS. Power calculations are highly dependent on sample size and measurement error for reliability in validating clinical hypotheses[15]. A small sample size along with high measurement error would not indicate good test reliability. Thus, a conservatively large sample size gives sufficient room to measure for generalized reliability. It is difficult to have faultless measurements with human subjects and survey administrators. Measurement error can be derived from the timeliness of survey administration and time allotted to complete the survey. The second survey was administered 24-72 hours after initial survey completion. Attempting to administer both surveys at the same time of day reduces variability in disease state that might be dependent on time of day. Additionally, ensuring all patients have adequate time to complete the survey can decrease the incidence of incomplete survey responses as dependent on time. The survey tool may also need improvement to minimize measurement error. Recommendations include revisions to the wording of the questions to enhance broader understanding as well as providing an alternative numerical (0 to 10) or visual analog (10 cm) rating scale.

Item responsiveness, however, needs to be improved since item showed differences in those with reportedly stable disease and those with reported changes in disease. The five-item

[14] Pascoe VL, Enamandram M, Corey KC, et al. Using the Physician Global Assessment in a Clinical Setting to Measure and Track Patient Outcomes. *JAMA Dermatol.* 2015;151(4):375–381. doi:10.1001/jamadermatol.2014.3513

[15] Kanyongo, Gibbs Y.; Brook, Gordon P.; Kyei-Blankson, Lydia; and Gocmen, Gulsah (2007) "Reliability and Statistical Power: How Measurement Fallibility Affects Power and Required Sample Sizes for Several Parametric and Nonparametric Statistics," Journal of Modern Applied Statistical Methods: Vol. 6 : Iss. 1 , Article 9. DOI: 10.22237/jmasm/1177992480

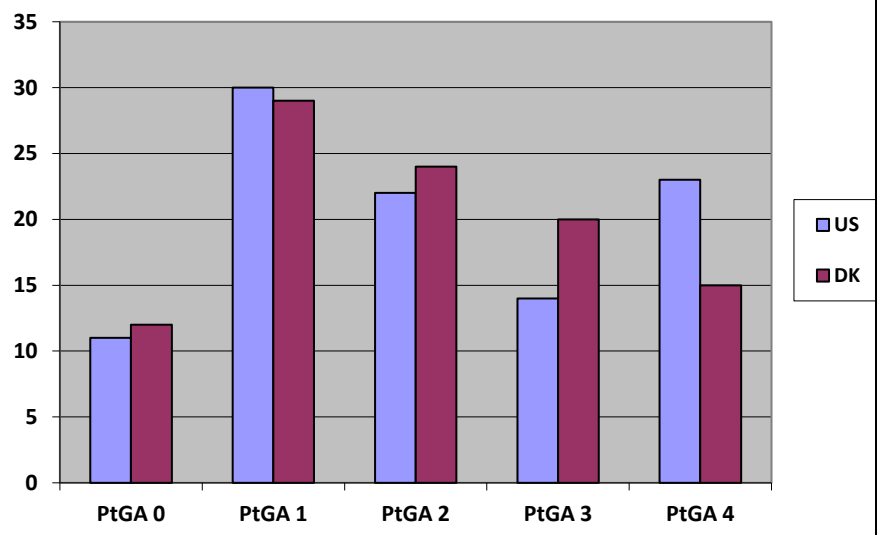
response options may constrain the ability of the item to discriminate between changes in disease impact. Work is underway to investigate the psychometric properties of a revised version of the PtGA with a revised response option, which may show improved responsiveness. There were instances of missing data though the highest rate of missing data was 49.2%, thus the observed responses are likely sufficient to produce significant and robust results for interpretation. Notably, participants of this study were selected from academic institutions, where the population tends to have more severe disease and the astute physicians are known for treating HS. This serves as a limitation in that the level of impact may be inflated and thus, the PtGA may not be as generalizable in non-academic settings. Additional clinical trials in private practice and community clinic settings would be needed to show reproducibility of PtGA effectiveness through responsiveness, validity, and reliability of patients with a lesser dermatologic disease severity.

HS greatly impacts one's well-being in many aspects of life, whether it be physically, mentally, socially, or emotionally. A reliable and valid measurement tool in assessing the state of a patient's HS is key in physician assessment and treatment plan. Working in a clinical setting does not always allow for the issuance of a lengthy patient survey due to time constraints. A single-item PtGA can be used quickly in a clinic visit, allowing for the physician to review the patient's responses and engage in patient-physician dialogue. The foundation of future HS clinical research as it relates to patient-reported outcomes and HRQOL has been established by providing a candidate PtGA item.

Table 1: Demographic characteristics and baseline assessments of the field testing participants				
Characteristics	**US**		**DK**	
Total participants, n	224		217	
Age in years, mean (range)	39.6 (19-77)		42.9 (19-72)	
Sex, n (%) Female Male	 195 (87%) 29 (13%)		 194 (91%) 20 (9%) *3 missing*	
Race, n (%) White Asian Black North American Indian Hispanic/Latino Mixed Other	 159 (71%) 3 (1%) 49 (22%) 0 (0%) 6 (3%) 1 (0%) 6 (3%)		 NC	
Education <12 years of school High School/GED Occupational, technical, or vocational Some College/No Degree Associate's Degree Bachelor's Degree Master's Degree Doctoral Degree	 6 (3%) 45 (20%) 20 (9%) 58 (26%) 25 (11%) 49 (22%) 19 (8%) 2 (1%)		 28 (13%) 11 (5%) 50 (23%) 0 (0%) 46 (21%) 70 (33%) 10 (5%) 0 (0%) *2 missing*	
Instrument	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
PtGA	2.09 (1.34)	2	1.96 (1.26)	2
HiSQOL, total	28.98 (17.53)	27	26.92 (17.86)	27
DLQI	12.97 (8.33)	12	11.17 (8.02)	9
NRS for pain	3.29 (2.83)	3	3.15 (2.61)	3
HADS, Anxiety subscale	8.14 (4.91)	8	7.05 (5.09)	7
HADS, Depression subscale	6.43 (4.67)	6	5.13 (2.96)	4
US: United States sample, DK: Denmark sample, NC: not collected				

HiSQOL: The Hidradenitis suppurativa quality of life (HiSQOL) instrument is a recently developed HS-specific health-related quality of life instrument. A higher score (maximum of 68) indicates a higher negative impact of HS. The three sub-scale are Symptoms, Psychosocial Effects, and Activities & Adaptations.
DLQI: The Dermatology Life Quality Index (DLQI) is a skin-specific health-related quality of life index. This 10-item scale has a score range from 0 (better) to 30 (worse) indicating more negative impact on quality of life.
NRS for pain: Numerical rating scale is a single item 11-point scale was used to assess severity of HS-related pain.
HADS: The Hospital Anxiety and Depression Scale is used to measure symptoms of anxiety and depression. The scale was administered to the patients in the form of 14 questions, 7 questions were dedicated to depression and anxiety respectively with scoring ranges from 0 (better) to 21 (worse) for each sub-scale.

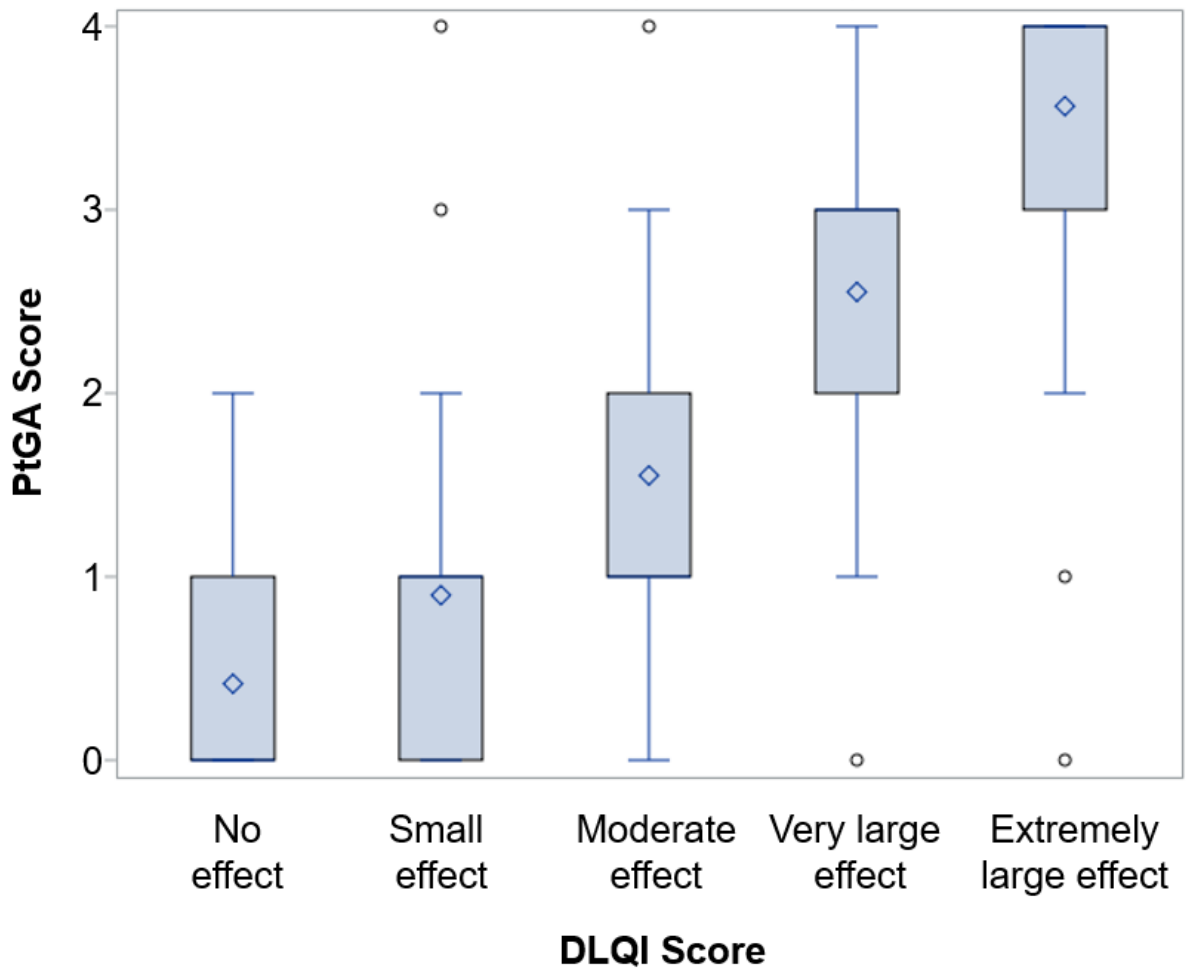
Figure 1: Participant Responses for PtGA Item



	US Sample n (%)	DK Sample n (%)
PtGA		
0	24 (11%)	27 (12%)
1	67 (30%)	62 (29%)
2	48 (22%)	52 (24%)
3	32 (14%)	43 (20%)
4	52 (23%)	32 (15%)

Table 2. Reliability and convergent validity of the PtGA item			
Test-retest reliability	**ICC (95%CI)**		
Overall	0.84 (0.80, 0.87)		
Country			
Denmark	0.88 (0.83, 0.91)		
US	0.81 (0.79, 0.87)		
Sex			
Female	0.84 (0.79, 0.87)		
Male	0.84 (0.73, 0.92)		
Age Group			
≤20	1.00 (1.00, 1.00)		
21-30	0.81 (0.71, 0.88)		
31-40	0.87 (0.80, 0.92)		
41-50	0.85 (0.77, 0.90)		
51-60	0.82 (0.70, 0.89)		
61-70	0.84 (0.64, 0.94)		
71-80	0.93 (0.68, 0.99)		
Race			
White	0.82 (0.76, 0.87)		
Asian	NC		
Black	0.77 (0.61, 0.88)		
North American Indian	NC		
Hispanic/Latino	NC		
Mixed	NC		
Other			
Convergent validity	**Spearman correlation (95%CI)**		
	Overall	**US**	**DK**
DLQI	0.78 (0.74, 0.82)	0.72 (0.64,0.77)	0.85 (0.81,0.89)
HiSQOL	0.82 (0.78,0.85)	0.76 (0.69,0.81)	0.88 (0.84,0.90)
HADS			
Anxiety	0.55 (0.48, 0.61)	0.52 (0.41,0.68)	0.58 (0.48,0.66)
Depression	0.53 (0.48, 0.59)	0.57 (0.48,0.66)	0.48 (0.37,0.58)
NRS for pain	0.66 (0.60, 0.71)	0.60 (0.51,0.68)	0.72 (0.65,0.78)

Figure 2. Known groups validity of PtGA item for established DLQI score groups



		n (%)	PtGA		p-value*
			Mean	SD	
DLQI score bands	**No effect**	24 (5.5%)	0.42	0.65	<0.001
	Small effect	89 (20.6%)	0.90	0.74	
	Moderate effect	107 (24.8%)	1.55	0.77	
	Very large effect	127 (29.4%)	2.55	1.00	
	Extremely large effect	85 (19.7%)	3.56	0.75	

*p-value for comparison of means across groups; 9 respondents with missing data
Based on complete case analysis, n=432 for combined US and DK responses