Development, validation and clinical application of a patient-reported outcome measure in hyperhidrosis: The Hyperhidrosis Quality of Life Index (HidroQoL ©)

A thesis submitted in accordance with the conditions governing candidates for the degree of

DOCTOR OF PHILOSOPHY

At

CARDIFF UNIVERSITY

Presented by

Paul Kamudoni

School of Pharmacy and Pharmaceutical Sciences

Cardiff University

Cardiff

UK

February 2014

TABLE OF CONTENTS

Table of Contents	i
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
LIST OF ABBREVIATIONS	vi
GLOSSARY OF TERMS	vii
LIST OF TABLES	ix
LIST OF FIGURES	xv
CHAPTER 1: General Introduction	1
STUDY AIMS AND OBJECTIVES	60
CHAPTER 2: Rationale And Methodological Framework	61
CHAPTER 3: Development Of A Hyperhidrosis-Specific Patient-Reported Outcome Measure: Qualitative Study	93
CHAPTER 4: Development of a Hyperhidrosis-Specific Health-Related Quality of Life Instrument: Content Validation	133
CHAPTER 5: Development of a Hyperhidrosis-Specific Health-Related Quality of Life Instrument (Hidroqol): Factor Analysis	164
CHAPTER 6: Development of a Hyperhidrosis-specific quality of life instrument (HidroQoL): Rasch analysis	197
CHAPTER 7: Evaluation Of The Reliability Of The Hyperhidrosis Quality Of Life Index (HidroQoL)	243
CHAPTER 8: Evaluation Of The Validity Of The Hyperhidrosis Quality Of Life Index (HidroQoL)	264
CHAPTER 9: Responsiveness And Interpretability Of The Hyperhidrosis Quality Of Life Index (Hidroqol) Scores	321
CHAPTER 10: General Discussion	371

REFERENCES 400

PUBLICATIONS AND PRESENTATIONS 431

APPENDICES 432

ACKNOWLEDGEMENTS

I am indebted to Prof. Sam Salek, my supervisor, for his ongoing guidance, support and encouragement throughout the course of this research project. I wish to express my gratitude to Dr. Berno Müller, who was instrumental in conceptualising the financing of the project. I thank Riemser Pharma GmbH for providing the educational grant for this work. I also wish to thank Dr. Dagmar Braun for her support. I would like to thank Dr. Christian Müller for his input during the early phases of the project and for his helping me settle down in Greifswald (Germany).

I thank all the staff at Riemser Pharma GmbH, particularly, Mrs. Krist for her readiness to help with everything; and Mr. Scheider for being a great officemate and an excellent German language tutor. I am thankful for the hospitality and kindness of Dr. and Mrs. Garbe taking me out on their family exculsions and for providing me a place to stay in towards the end of my time in Greifswald.

I thank the staff at the Centre for Socioeconomic Research, particularly, Ms. Justine Jenkins, Ms. Helen Haron and Dr. Harun, for their support and contribution to this work.

There are many people who have stood by me along the way; unfortunately, I am unable to list them all. I am most grateful to my family, for being a source of inspiration and courage, ‘a haven in the stormy seas of life’. In particular, I thank Tante Penja, for her support, Ntwitwi, for her evening calls. Thank you Angela, for being there especially through the tipping points of the project.

This thesis is dedicated first to Jesus for saving me and giving me a purpose. I owe everything that I have attained to my mother and father for their careful investment in my life, spiritually, emotionally and intellectually.

ABSTRACT

Consideration of broader outcomes of disease, especially those exclusively experienced and reported by the patient, such as HRQOL, is not only consistent with the 'whole person' view of health contained in the 1948 WHO definition, but is also a prerequisite to building health-care systems that are responsive to the needs of the patients. For chronic skin diseases, such as hyperhidrosis, these provide a useful indicator of how a patient feels and functions disease for both practical and methodological reasons. The aims of this study therefore were to investigate the impact of hyperhidrosis on patients' HRQoL using a mix of qualitative and quantitative methods. In addition, a further aim was to develop and validate a disease-specific instrument for assessing HRQoL in hyperhidrosis. In pursuing the above aims, the feasibility of applying online social networking sites for outcomes research in dermatology was assessed.

Patients were recruited through online social networking communities related to hyperhidrosis for all stages of the study. Interviews, focus groups and surveys were used for collecting qualitative data from patients (n = 71) to understand quality of life issues of patients, and to provide the content of the new instrument. Dermatologists (n= 5) and patients (n=7) took part in the content validation of the HidroQoL©. Item reduction and the development of the scale's structure was carried out through several field-testing studies (n: USA, 559; UK, 115), using the item response theory (IRT) Rasch model and factor analyses. Further psychometric testing was performed in a separate study (n = 241). Distribution-based methods were applied in establishing minimum clinically important difference (MCID).

A thematic analysis of the qualitative data collected produced 29 quality of life themes and 102 sub-themes, forming the content for the initial 49-item HidroQoL©. The two expert panels judged the instrument as content valid, with a few suggestions. The Rasch analysis modelling led to the collapsing of response categories (from five to three) and the reduction in number of items (from 49 to 18), to ensure a perfect model fit. Factor analyses supported both a single- and a two-factor structure. In subsequent construct validation study the HidroQoL correlated with the DLQI (r_s = 0.572, $p < 0.01$) and the Skindex-17 (r_s = 0.551, $p < 0.01$). Reliability was high (Cronbach alpha = 0.9; test-retest ICC = 0.93). The scores were sensitive to change in patients' disease severity (standard response mean = 0.8, 95% C.I: 0.34-1.27). The scale banding proposed for the HidroQoL score is as follows: 0 – 1, no effect at all; 2 – 11, small effect; 12 – 22, moderate effect; 23 – 32, large effect; 33 – 36, very large effect. The MCID values were 1.94 – 3.07, for generalised

hyperhidrosis, 2.16 – 4.36, for axillary hyperhidrosis, 2.15 – 3.39, for palmo-plantar hyperhidrosis. An MCID of three is currently being proposed for all types of hyperhidrosis.

This study has provided the initial evidence supporting the appropriateness of the content of the HidroQoL and validity of inferences from its scores for assessing HRQoL in hyperhidrosis. In addition, the availability of MCID estimates for the HidroQoL will facilitate its clinical interpretation in both research and routine clinical practice. This study has also demonstrated how CTT and IRT can be integrated in the development and validation of a new generation of HRQoL instruments, using social network for patient recruitment.

LIST OF ABBREVIATIONS

CFA	Confirmatory factor analysis
CTT	Classical test theory
CVI	Content validity Index
DIF	Differential Item Functioning
DLQI	Dermatology Life Quality Index
EFA	Exploratory factor analysis
ETS	Endoscopic Thoracic Surgery
FLQA	Freiburg Life Quality Assessment
GQ	General Question on Global Impact of Hyperhidrosis
HRQOL	Health related quality of life
HDSS	Hyperhidrosis Disease Severity Scale
HHIQ	Hyperhidrosis Impact Questionnaire
HidroQoL	Hyperhidrosis Quality of Life Index
HS	Hyperhidrosis scale
IIRS	Illness Intrusiveness Rating Scale
MCID	Minimum Clinically Important Difference
NHP	Nottingham Health Profile
PBI	Patient Benefit Index
PGA	Patient's global assessment of change
PRO	Patient Reported Outcome
WHO	World Health Organisation
WTP	Willingness to Pay
SF-36	Short-Form 36
SF-12	Short-Form 12

GLOSSARY OF TERMS

Differential Item Functioning	This occurs where respondents with different characteristics have varying probabilities of getting a particular item response holding the underlying construct constant.
Applicability	This relates to the appropriateness of a measure's content and emphasis to the target population.
Content validity	this refers to the adequacy with which sampled items of an instrument reflect its aims as articulated in the conceptual framework.
Construct validity	The degree to which theoretical hypothesis relating to an underlying construct being assessed by an instrument are actually supported, providing evidence justifying particular or inferences or interpretation of scores.
Convergence validity	the relationship between a scale and other measures assessing a similar construct.
Divergent validity	the absence of a relationship between a scale and measures of dissimilar construct.
Factor Analysis	An analysis that identifies the least number of latent variables accounting for covariation among a set of items.
Primary hyperhidrosis	Excessive sweating beyond the physiological needs of the body that is without aetiology, often localised to the underarms, palms, feet or other body areas.
Secondary hyperhidrosis	Excessive sweating that has a known cause e.g. due to menopause, often generalised.
Internal Consistency	The inter-relationship between items in a multi-item scale, reflecting homogeneity within a scale.
Interpretability	Decoding of qualitative meaning from QOL scores

Kruskall-Wallis tests	A non-parametric alternative to the one way ANOVA typically applied for comparisons among two or more groups.
Rasch analysis	An analysis intended to assess how well data conforms to the unidimensional Rasch model, its assumptions and properties.
Responsiveness	The capability of a measure to capture true changes in the patient's conditions even if they are small.
Scale banding	Categorisation of scale scores to aid in their interpretation and application for decision especially in routine clinical practice.
Unidimensionality	Variation in the scores of an instrument can be attributed to a single dominant latent variable.

LIST OF TABLES

Table 1.1: Treatments for hyperhidrosis 20
Table 1.2: Criteria used in evaluating the measures 29
Table 1.3: Descriptive properties of instruments used in measuring QoL in hyperhidrosis 41
Table 1.4: Psychometric properties of instruments 42
Table 1.5: Studies using generic HRQoL instruments: SF-36 and NHP 44
Table 1.6: Effects of hyperhidrosis on lifestyle: IIRS 47
Table 1.7: Studies using a dermatology-specific QoL Instrument: Dermatology Quality of Life Index (DLQI) 50
Table 1.8: Studies using the Hyperhidrosis Impact Questionnaire (HHIQ) 54
Table 3.1: Sociodemographic characteristics of study participants 100
Table 3.2: Issues considered under each HRQoL themes 101
Table 3.3: List of issues forming the initial 75-item instrument proto-type 123
Table 4.1: Patient-panel ratings of language clarity, completeness, relevance, scaling of the HidroQoL 141
Table 4.2: Level of agreement and content validity index for the panel of patients 142
Table 4.3: Comments from patients 143
Table 4.4: Level of agreement and content validity index for the panel of dermatologists 146
Table 4.5: Dermatologists-panel ratings of language clarity, completeness, relevance, scaling of the HidroQoL 148
Table 4.6: Suggestions made by the panel of dermatologists 152
Table 4.7: Revision to the items of the HidroQoL 155
Table 5.1: Sociodemographic characteristics of study participants 170
Table 5.2: Access to and use of treatment 174
Table 5.3: Frequency of endorsement to the HidroQoL 174
Table 5.4: Polychoric correlations between items of the HidroQoL (part 1) 176
Table 5.5: Polychoric correlations between the items of the HidroQoL (part 2) 177
Table 5.6: Polychoric correlations between items of HidroQoL (part 3) 178
Table 5.7: Multicollinear items (correlations of at least 0.8) 179
Table 5.8: Item review based on correlation matrix (addressing multicollinearity) 181

Table 5.9: Eigen values for all 36 items 182
Table 5.10: Steps during EFA analysis: removal of poorly performing items 183
Table 5.11: Factor pattern and factor structure matrices for the 36 items of the HidroQoL 184
Table 5.12: Factor pattern matrix and residual variances for the 21 items of the HidroQoL 187
Table 5.13: Eigenvalues for all 21 items: observed and Horn's parallel analysis 187
Table 5.14: Correlation measures of internal consistency for the 21-item HidroQoL 189
Table 5.15: Distribution of responses to the HidroQoL in Sample 2 190
Table 5.16: Goodness of fit of the CFA models estimated 191
Table 5.17: Correlation between variables and construct (r-squared) and uniqueness (residual variance) 192
Table 6.1: Sociodemographic characteristics of study participants 202
Table 6.2: Access to and utilization of treatment 206
Table 6.3: Patients responses to the HidroQoL 207
Table 6.4: Overall model fit statistics for the 36 items HidroQoL and subsequent versions after rescoring 210
Table 6.5: Rasch model item parameters for the 36-items of the HidroQoL 211
Table 6.6: Impact of item reduction steps on overall fit of the RM and individual items. 218
Table 6.7: Parameter estimates for the final 18 items fitting the Rasch model following item reduction 219
Table 6.8: Pure set of items showing no DIF following the purification 221
Table 6.9: DIF in items according to patient characteristics 221
Table 6.10: Magnitude of DIF 224
Table 6.11: Impact of adjusting for response dependence on overall model fit. 227
Table 6.12: Eigen values of the HidroQoL 227
Table 6.13: Factor loadings and residual variances of the HidroQoL 229
Table 6.14: Sociodemographic characteristics of study participants 230
Table 6.15: Overall model fit statistics for the HidroQoL on participants from UK 231
Table 7.1: Sociodemographic characteristics of the study participants (during assessment 1) 248
Table 7.2: Disease-related characteristics of study participants 249
Table 7.3: Internal Consistency* of the HidroQoL 250
Table 7.4: Item-total scale correlations, for the HidroQoL, pooled/international, test 1 (n = 260) 251

Table 7.5: Item-total scale correlations, for the HidroQoL, U.S. Sample, test 1 (n = 142) 252
Table 7.6: Item-total scale correlations, for the HidroQoL, UK Sample, test 1 (n = 73) 253
Table 7.7: Correlations[§] among HidroQoL's items, test 1, pooled sample (n = 260) 256
Table 7.8: Correlations[§] among HidroQoL's items, test 1, USA Sample (n = 142) 257
Table 7.9: Correlations[§] among HidroQoL's items, test 1, UK (n = 72) 258
Table 7.10: Test-retest reliability for individual items of the HidroQoL, international Sample (n = 104) 259
Table 7.11: Test-retest reliability for individual items of the HidroQoL, USA Sample (n = 64) 260
Table 7.12: Test-retest reliability for individual items of the HidroQoL, UK Sample (N = 22) . 261
Table 8.1: Attributes of outcome measures used in data collection 268
Table 8.2: Sociodemographic characteristics of the patients 270
Table 8.3: Patients' treatment history 271
Table 8.4: Patient's level of disease burden: time[a] and money[b] spent in managing the condition and willingness to pay[c] 272
Table 8.5: Items of the HidroQoL receiving affirmation and missing responses 275
Table 8.6: Frequency of the HidroQoL Scores 276
Table 8.7: Comparison of HidroQoL scores by patient's gender (USA sample) 279
Table 8.8: Comparison of HidroQoL scores by patient's gender (UK sample) 280
Table 8.9: Comparison HidroQoL scores by patient's age (USA sample) 281
Table 8.10: Comparison of individual item and total HidroQoL scores by patient's age (UK Sample) 282
Table 8.11: Univariate regression analyses of HidroQoL score against patient's age 283
Table 8.12: Comparison of individual items and total HidroQoL scores by HDSS score (level of disease severity): USA sample 284
Table 8.13: Comparison of individual items and total HidroQoL scores by HDSS score (level of disease severity) (UK sample) 285
Table 8.14: Comparison of individual items and total HidroQoL scores by site of hyperhidrosis (USA sample) 287
Table 8.15: Comparison of individual items and total HidroQoL scores by site of hyperhidrosis (UK Sample) 288
Table 8.16: Comparison of individual items and total HidroQoL scores by GQ Score (global life impact): USA sample 289

Table 8.17: Comparison of individual items and total HidroQoL scores by GQ Score (overall HRQoL impact) (UK sample) 290
Table 8.18: Comparison of individual items and total HidroQoL scores by patient's WTP for complete cure for the sweating (USA sample) 292
Table 8.19: Comparison of individual items and total HidroQoL scores by patient's WTP for complete cure for the sweating (UK sample) 293
Table 8.20: Comparison of individual items and total HidroQoL scores by daily time spent in managing sweating (USA sample). 295
Table 8.21: Comparison of individual items and total HidroQoL scores by daily time spent in managing sweating (UK Sample). 296
Table 8.22: Univariate regression analyses of HidroQoL score against daily time spent with hyperhidrosis. 297
Table 8.23: Comparison of HidroQoL Scores by treatment history: US sample 298
Table 8.24: Comparison of HidroQoL Scores by treatment history: UK Sample 299
Table 8.25: Comparison of HidroQoL Scores by patient's co-morbidity (USA sample) 301
Table 8.26: Comparison of HidroQoL Scores by patient's co-morbidity (UK sample) 301
Table 8.27: Summary description of HidroQoL, DLQI, Skindex-17, EQ-5D scores. 303
Table 8.28: Multiple correlations between the HidroQoL scores and the Skindex, the DLQI, EQ-5D, general health, the HDSS, patients' WTP, and time spent in the daily management of the sweating (USA Sample) 305
Table 8.29: Multiple correlations between the HidroQoL scores and the Skindex, the DLQI, EQ-5D, general health, the HDSS, patients' WTP, and time spent in the daily management of the sweating (UK Sample) 306
Table 8.30: Results of univariate regression analyses of DLQI, Skindex, EQ-5D dimensions regressed on HidroQoL score. 307
Table 8.31: Predictors of HRQoL in hyperhidrosis (all variables included) based on the US sample (n = 127) 314
Table 8.32: Contribution of predictors included in 'all variables model' to explaining the variance in the HidroQoL Scores; with hierarchical inclusion of variables (USA sample) 315
Table 8.33: Predictors of hyperhidrosis-QoL based on stepwise backward regression analysis (USA sample, N = 127) 316

Table 8.34: Predictors of hyperhidrosis-QoL based on stepwise backward regression analysis (UK sample, N = 36) 316
Table 9.1: Sociodemographic characteristics of the patients 327
Table 9.2: Patients' treatment history and disease characteristics 328
Table 9.3: Distribution of the scale scores, in pooled sample 329
Table 9.4: Distribution of the HDSS score 330
Table 9.5: Distribution of the anchors: HDSS change score and PGA score 330
Table 9.6: Correlation of the HDSS and the PGA with the HidroQoL scores 330
Table 9.7: A comparison of the PGA against the HDSS change score in their comparison of patients. 332
Table 9.8: Sensitivity of the HidroQoL scores in patients experiencing 'no change', 'slight improvement' and 'slight deterioration' based on paired t-test, in the pooled sample.. 333
Table 9.9: A comparison of the HidroQoL's ability to detect change with that of the DLQI and Skindex in the pooled sample 334
Table 9.10: Estimating the responsiveness of the HidroQoL based on standardised response mean and effect size with patient groups (pooled sample) 336
Table 9.11: Sensitivity of the HidroQoL scores (USA sample) 337
Table 9.12: Relative efficiency Index of HidroQoL with DLQI and Skindex-17 in detecting change in the minimally improving group. 337
Table 9.13: Estimating the responsiveness of the HidroQoL based on standardised response mean and effect size (USA sample) 338
Table 9.14: Comparison of amount of change in patients between those with 'slight deterioration', 'no-change' and patients with 'slight improvement': Pooled sample 340
Table 9.15: Correlation of the HidroQoL score with the HDSS, DLQI and Skindex-17 340
Table 9.16: Distribution of the HidroQoL scores 342
Table 9.17: Frequency, mean, mode and median of GQ scores for each HidroQoL score (USA Sample). 345
Table 9.18: Alternative HidroQoL score banding for the USA sample 347
Table 9.19: Frequency, mean, mode and median of GQ scores for each HidroQoL score (UK patients) 348
Table 9.20: Alternative HidroQoL score banding based on the UK sample 350

Table 9.21: Frequency, mean, mode and median of GQ scores for each HidroQoL score (pooled sample) 351
Table 9.22: Alternative HidroQoL score banding based on the pooled sample 353
Table 9.23: Distribution of GQ scores for proposed scale banding (set 1) 354
Table 9.24: Distribution of GQ scores for proposed scale banding (set 2) 355
Table 9.25: Distribution of GQ scores for proposed scale banding (set 3) 355
Table 9.26: Distribution of GQ scores for proposed scale banding (set 4) 356
Table 9.27: Distribution of GQ scores for proposed scale banding (set 5) 356
Table 9.28: Distribution of GQ scores for proposed scale banding (set 6) 357
Table 9.29: Distribution of GQ scores for proposed scale banding (set 7) 357
Table 9.30: Distribution of GQ scores for proposed scale banding (set 8) 358
Table 9.31: Area under curve for the ROCs for each grouping 359
Table 9.32: Operating characteristics of the HidroQoL score cut-offs in classifying patients according to their GQ score. 361
Table 9.33: Mean HidroQoL score change in the ‘slightly improving’ patient group as an estimate of the MID 363
Table 9.34: Upper-bound of 1 tailed 95% CI for the mean HidroQoL-cs in the ‘no-change’ patient group as an estimate of the MCID 364
Table 9.35: Measures of precision of the HidroQoL, Standard deviation and Standard Error of Measurement (SEM) as MCID estimates. 365

LIST OF FIGURES

Figure 1.1: Health status domains proposed in Ware's model 10
Figure 1.2: Patrick and Chiang QOL and HRQOL model 11
Figure 2.1: The Hyperhidrosis disease severity scale 82
Figure 2.2: Global question on overall impact of hyperhidrosis 83
Figure 2.3: Question on Patients global assessment of change 84
Figure 2.4: Flow chart of study 92
Figure 3.1: Overview of data collection process 99
Figure 3.2: Conceptual framework for the new QoL impact questionnaire for hyperhidrosis ... 119
Figure 3.3: The 47-item developmental version of the new instrument 126
Figure 4.1: The developmental version of the HidroQoL with 49 items 158
Figure 5.1: Age distribution of the study participants 172
Figure 5.2: Duration of disease 172
Figure 5.3: Patient's self-reported disease severity 173
Figure 5.4: General impact of disease on patient's life 173
Figure 5.5 Scree-plot based on the Eigenvalues from the WLSMV for all 36 items 182
Figure 5.6: Scree-plot based on the Eigen-values from the WLSMV for all 21 items 188
Figure 5.7: Path-diagram of the CFA Model with two correlated factors. 193
Figure 5.8: Path-diagram for CFA model with 1 –factor 194
Figure 6.1: Age distribution of study participants 204
Figure 6.2: Duration of disease of the study participants 204
Figure 6.3: Patient reported disease severity based on the HDSS 205
Figure 6.4: Patient reported general impact of disease 205
Figure 6.5: Category Probability Curves of the original 5-category HidroQoL 213
Figure 6.6:Person-item distribution map of the 36-items of the HidroQoL showing an even spread of the items across the latent variable. 214
Figure 6.7: Item threshold map of the HidroQoL-36 after rescoring to 3 point response categories 215
Figure 6.8: An illustration of the impact of rescoring on category probability curves 216
Figure 6.9: Person-item distribution showing targeting of the HidroQoL following item reduction 220

Figure 6.10: An illustration of DIF as reflected in empirical group-specific Item characteristics curves 223
Figure 6.11: Test Characteristic Curves for the HidroQoL-18 225
Figure 6.12: Scree plots of the HidroQoL 228
Figure 6.13: Comparison of item hierarchical order between original calibration of the HidroQoL and from the UK sample 233
Figure 6.14: Scatter plot of item difficulty estimates of the HidroQoL, plotting estimates from the original calibration against those from the UK samples 234
Figure 6.15: The final version of the HidroQoL with 18-items 235
Figure 7.1: Patients included in the test-retest reliability study 255
Figure 8.1: Patients' age distribution 273
Figure 8.2: Patients' disease severity based on HDSS score 273
Figure 8.3: Overall health related quality of life impairment 274
Figure 8.4: Amount of money patients are willing to pay (WTP) for a permanent cure for their condition. 274
Figure 8.5: Distribution of the HidroQoL total Scores using box and whisker plot (USA, n = 127; UK, n = 36). 277
Figure 8.6: Mean scores for the HidroQOL's individual items 278
Figure 8.7: Scatter plot showing relationship between HidroQoL Score and age. 283
Figure 8.8: The relationship between the HidroQoL Score and daily time spent in managing the condition 297
Figure 8.9: Scatter plot illustrating the relationship between DLQI score and the HidroQoL score 307
Figure 8.10: Scatter plot illustrating the relationship between HidroQoL score and Skindex-17 score 309
Figure 9.1: Distribution of the Hyperhidrosis Disease Severity Score (HDSS) 341
Figure 9.2: Distribution of GQ Score 342
Figure 9.3: Distribution of the HidroQoL total scores using box and whisker plot (USA, N = 142; UK, N = 73) 343
Figure 9.4: Distribution of the HidroQoL total score using a Histogram (pooled sample, N = 260) 343

Figure 9.5: Relationship between the HidroQoL score and the mean, median and mode of the GQ score for USA patients 346
Figure 9.6: Relationship between the HidroQoL score and the mean, median and mode of the GQ score for UK patients 349
Figure 9.7: Relationship between the HidroQoL score and the mean, median and mode of the GQ score for pooled sample 352
Figure 9.8: Receiver operating characteristic curve for classifying patients between GQ 0-1 and GQ 2 – 4 using the HidroQoL 359
Figure 9.9: Receiver operating characteristic curve for classification of patients between GQ 0-2 and GQ 3 – 4 using the HidroQoL 362
Figure 9.10: Receiver operating characteristic curve for classification of patients between GQ score 0-3 and GQ 4 using the HidroQoL 362
Figure 9.11:Estimates for MCID for the HidroQoL across different patient sub-populations and reflecting multiple analytical approaches 366

CHAPTER 1

General Introduction

BACKGROUND

Outcomes reported by the patient, such as symptoms, physical function, well-being, HRQoL, perceptions of treatment effect, satisfaction with care received, have now gained recognition as a credible and key endpoint of therapy. For example, improvement in HRQoL is now recognised as the ultimate goal of health care (MacKeigan and Pathak 1992). This has paved the way for application of Patient Reported Outcome Measures (PROMS), for example in the assessment of efficacy of pharmaceutical therapies in clinical trials, as evidenced by the issuance of guidelines by drug regulatory agencies such as the US FDA and the European Medicines Agency (EMEA) on the use of PROMS for market authorisation application of new medicines. HRQoL and other PRO data appeared in the scientific discussions of 34% of products submitted between 1995 and 2005 (Bottomley et al. 2009). Gnanaskthy et al. (2013) observed that out of the 308 new molecular entities (NME) and biologic licence applications (BLA) granted approval by the US FDA between 2000 and 2012, Patient Reported Outcome (PRO) claims were approved in 70 (23%), with the PROs being the primary endpoint in the majority (81%) of cases. Furthermore, there has been a growing use of PROs in routine clinical practice in various therapeutic areas. For patient management, to assess the impacts of disease and its treatment; in screening for patients experiencing major effects; and in multidisciplinary teams discussions among others (Greenhalgh 2009). Recently, there have been efforts to use patient reported outcome measures (PROMS) for monitoring the performance of the health-care system, in the UK, but also in other countries such as Sweden and the USA. In Sweden, disease-specific clinical databases (quality registers) have been established under the watch of the medical profession and in the USA, this has covered spinal conditions in New England and for primary care in Pittsburgh (Black 2013). In the English NHS, since 2009, all health care providers treating NHS patients for hip or knee replacement, groin hernia repair and varicose vein surgery have been required to assess PROs before and after treatment involving 485 000 patients (Devlin and Appleby 2010). This highlights the pace and scope with the use of PROs.

Various developments in recent decades might explain the growing recognition of the need to capture patients' perspective of illness and health-care interventions. First, there has been a marked change in disease-epidemiology not only in the western world but also in the developing countries, with non-communicable and life-style diseases replacing communicable diseases. Conditions such

as cancer, diabetes and cardiovascular disease, which are long-term rather than acute, are increasingly a growing health challenge. Often complete cure from disease may not be realisable, which makes maintaining a comfortable, functional and satisfying life an important goal of therapy for such conditions (Salek and Luscombe 1992). Improvements in medical care and general living conditions have on the other hand meant that people are now living longer, leading to increase in conditions related to aging, such as Dementia and Parkinson's disease. While dealing with a condition such as cancer as a biological and pathological condition might be imperative, it is not necessarily of paramount concern to patients (Lohr 2002). Their worries extend to how the condition and its treatment will impact on their lives (Lohr 2002), for example: *Can I go out with friends without worrying that I may vomit due to the chemotherapy?* On the other hand, the management of conditions heavily relies on the experience of the patient in their everyday life to strive to conduct a "normal" life.

In chronic skin conditions, although decreasing the amount of sweating or the thickness, redness and number of lesions may be valid endpoints in assessing therapeutic effect, such endpoints may not necessarily be of most relevance from the patient's standpoint (Grob 2007). Moreover, such endpoints may not provide the full picture of the impacts of disease or therapy. In addition, there might be some aspects related to the disease-condition that only the patient might be aware of, and therefore able to report on. A broad perspective of the impacts of disease or therapy may offer a better framework for risk-benefit assessment of therapies (Finlay 1998). For example, where therapy is successful in eliminating the primary symptoms associated with a condition but results in other limitations in patient's life (e.g. Endoscopic thoracic Sympathectomy in hyperhidrosis may affect the nervous system or lead to compensatory sweating). The PROs such as HRQoL, provide a more comprehensive measure of the impact of skin disease apart from capturing what the patients care about most, for example, how their condition affects their daily life activities and their social lives. Thus, these offer a robust means for assessing disease activity in chronic skin disease (Grob 2007).

Although clinicians may want to forgo measuring QoL because of the associated challenges; a presumption that QoL is captured by observing biomedical outcomes; and that they (the clinicians) have the ability to judge the impact experienced by the patient, current evidence suggests otherwise. Schmitt and Ford (Schmitt and Ford 2007) showed the uniqueness of 'HRQoL' from

symptoms or 'disease severity', in spite of the two being linked, making it difficult to accurately infer one from the other. There is evidence of a poor relationship between disease severity and HRQoL, patients may experience great impairment even with low disease severity and the opposite can also be the case, they may have high disease severity and yet experience low HRQoL impact (Bowling 2001). Moreover, clinician's evaluation does not always agree with patients own assessment of their QoL (Jemec and Wulf 1996; Hermansen et al. 2002). Moreover, patients with a low severity but experiencing high levels of HRQoL impairment, often inaccurately assessed by clinicians, tend to be least satisfied with their care (Renzi et al. 2001).

HEALTH RELATED QUALITY OF LIFE: A THEORECTICAL FRAMEWORK

Definitions

The concept of QOL has been employed across multiple-disciplines and settings (academic research, policy making environment, everyday speech), with varying understanding of its components, determinants and interpretation used in each context, hindering the emergence of a unified definition. At the population level, the concept of QOL has been used to study national well-being by governmental agencies as well as international organisations such as the Organisation of Economic Cooperation and Development (OECD), the World Bank and the World Health Organisation (WHO), with a focus on social indicators of living conditions and how individuals evaluate their satisfaction. At this level, a high QoL may be indicated by high earnings; absence of poverty and unemployment, decent housing, health spending and life expectancy, an educated population, high levels of cultural participation and low rates of crime, equity in social opportunities, and the absence of political corruption in the broader context of responsible environmental management (Rapley 2003). A lack of attention to the priorities of communities and individuals in the population-level QoL due to the aggregation has provided the space for individual level QoL construct. Therefore, QoL fundamentally includes two elements, an objective and a subjective element. Although numerous definitions exist, simply put, QoL is how individuals consider their material situation.

- The WHOQOL group (1995) defined QOL as, *an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, level of*

independence, social relationships and their relationships to salient feature of their environment.

- Cummins (1997) considers QoL as *both objective and subjective, each axis being the aggregate of seven domains: material well-being, health, productivity, intimacy, safety, community and emotional well-being. Objective domains comprise culturally relevant measures of objective wellbeing. Subjective domains comprise domain satisfaction weighted by their importance to the individual.*

The list of objective elements considered important to well-being may be influenced by societal cultural and political values, which may differ across communities, cities or nations. On the other hand, how particular individuals relate to their material circumstances may be due to their personality, personal beliefs and values. In this regard a universally agreed understanding of good QoL may be challenging (Phillips 2006). Although socioeconomic factors such as sanitation, health, education, housing, employment and other factors are key to QOL more so as targets of social and economic policy, except for health, the others lie outside the remit of 'health policy' and may not necessarily be influenced by medical-care. The concept of HRQoL, therefore, has as its rationale the intention to distinguish those aspect of QoL that pertain to health, disease or its treatment from the rest (Coons and Kaplan 1992). Still, the extent to which this is feasible remains questionable, given that multidimensional HRQoL addresses most of QoL domains (Fitzpatrick 1996). Moreover, HRQOL may be influenced by factors other than those fitting the classification of ‘health-related’, linked to the socioeconomic factors such as level of education, income or family-relations. This redefines QoL in line with the intentions for measurement in health and medical care.

Numerous definitions of HRQoL have been proposed, reflecting differing perspectives on the content and scope of what should be relegated as 'health-relevant' aspects of QoL as well as differences in theoretical underpinnings of the concept. It is difficulty to deny the influence of WHOs monumental whole-person definition of health on the current understanding of HRQoL, where “health is not merely the absence of disease but a state of complete physical mental, social well-being” (Group 1948).

- Bowling (2001) defines HRQoL as “optimum levels of physical role (e.g. work, carer, parent etc.) and social functioning, including relationships and perceptions of health,

fitness, life satisfaction and well-being. It should also include some assessment of the patient's level of satisfaction with treatment outcome and health status and with future prospects". Short of including material circumstances such as education, physical environment and level of income, this definition encompasses both positive and negative aspects of health. Some aspects included may not easily be affected by medical care. Nonetheless, this definition strongly reflects WHOs 'whole person' definition of health (Phillips 2006).

- Schipper et al. (1990) defines HRQoL as "the functional effect of an illness and its consequent therapy upon a patients, as perceived by the patients". They further highlight physical and occupational function; psychological state; social interaction and somatic sensation as the being the domains influencing overall effect. Intrinsic in this definition is a presumed level of functionality, which is considered as the norm, thus this perspective focuses on assessing "dysfunction" or negative aspects of disease. This definition, therefore, focuses on those aspects that can be influenced by clinicians and medical care.
- Ebrahim (1995) has offered an alternative definition of HRQoL, where it is defined as self-perceived well-being related to or affected by the presence of disease or treatment. With 'well-being' understood as an individual's inner-personal state (Doward and McKenna 2004), this definition opens up to other impacts of disease, including those that may not be directly influenced by clinical therapy, for instance, personal relationships, self-image and future health concerns. Nonetheless, there are some reservations over the inter-changeable use of the notion of well-being with QoL (Doward and McKenna 2004).
- Padilla et al (1996) defines HRQoL as a personal, evaluative statement summarizing the positivity or negativity of attributes that characterize one's psychological, physical, social, and spiritual well-being at a point in time when health, illness, and treatment conditions are relevant. This definition takes a broad view, encompassing all aspects of life of an individual with a disease condition, including negative, while acknowledging potentially positive influences of a medical condition.

The consequences of disease on the patient's life have also been captured using other concepts. For example the ICIDH described the concepts of 'impairment', disability and handicap. Impairment refers to the loss of normal physiological or psychological function due to disease for example, symptoms and adverse events from treatment (Doward et al 2004). Disability refers to

the restriction on normal performance of activities in various aspects of a patient's life as a result of impairment from disease (Finlay and Kelly 1987; Doward et al. 2004). The disadvantage in the fulfilment of roles in society due to impairments and disability associated with a disease is referred to as 'handicap' (Finlay, 1998). The concepts of impairment and disability directly relate to QoL as they represent clinical and functional status perspective of the impacts of disease, respectively. On the other hand, handicap is not directly related to an individual's QoL as it is measured from a societal point of view and reflects impacts on society (Doward et al. 2004).

Theoretical Foundation

A number of theories have been presented as genesis or foundation to the measurement of QoL, including the *needs-satisfaction-approach*; the *being, belonging and becoming model* and *Calman's gap theory.* The *needs-satisfaction approach* is rooted in Maslow's human motivation theory. Maslow (1973) cited in Bowling (2001) argued that people are motivated by a desire to fulfil their needs, which are grouped as physiological, safety, love and belonging, esteem and self-actualisation, presented hierarchically. Once a given set of needs are met, they cease to be a source of motivation, the focus is on the next level of needs which then become a source of motivation. As such, QoL reflects the ability and capacity of the individual to satisfy certain needs, where needs are fulfilled QoL is good, otherwise it is poor (Doward and McKenna 2004). QoL, in this perspective, is seen as quite distinct from health status and function, for example whereas 'walking' is a function, it may affect 'doing shopping' or 'recreational activities' which may be seen as 'needs' to an individual.

In the *Raphael's being, belonging and becoming* model (Raphael et al. 1996), QoL is considered as the extent to which an individual enjoys the important possibilities in their life (Phillips 2006). Key to this theory is the individual who is crucially at the centre of determining what a possibility is and how it is important to them. QoL is seen as a multidimensional concept comprising of three domains: *being*, representing who one is, with physical health, psychological and spiritual elements; *belonging*, reflecting a person's relations to their surrounding environment, social and community; and *becoming*, reflecting deliberate activities pursued to express personal goals, hopes and aspirations (Raphael et al. 1996). While this approach concentrates on functioning and role performance, the importance of the interaction of the individual with others and the environment in influencing QoL is highlighted.

In *Calman's gap theory* QoL is determined by the discrepancy between a patient's expectations and achievements (Calman 1984). This may also refer to the gap between patients' potential achievement and their actual achievements (Powell and Powell 1987). Where such discrepancy is small, QoL would be high, while a large discrepancy would lead to a low QoL. Other factors may influence the size of such a discrepancy, which would in turn affect QoL. This theory highlights the role of interventions such as communication/discussions about treatment options and their side effects; patient education supporting patients in coping with their condition in influencing QoL. The connection between previous experiences, current situation and aspirations for the future with QoL is clarified (Bowling 2001). *Calman's gap theory* seems to share much in common with Raphael's theory presented above, in both cases lowering expectations has a positive influence on QoL (Phillips 2006). On the other hand, this theory can be seen as an adapted version of the *needs-satisfaction theory,* replacing the notion of 'need' with 'relative deprivation'. The Patient Generated Index (PGI) is an example of a measure underpinned by this theory.

The measurement of HRQoL must take into account its unique properties and the associated idiosyncrasies. First, the multidimensional nature of HRQoL as a concept is in keeping with the broad and holistic definition of 'health' presented by the WHO (Schipper et al. 1990) presented above. In reality however, medical therapy, relates only to a narrow aspect of 'health' and consequently HRQoL (Bowling 2001), questioning the relevance of the concept relative to, for example, 'health status'. HRQoL measures often lack reference to therapeutic goals, in their development, in comparison to 'Clinimetric' or 'Health Status' measures (Testa 2000). Furthermore, the multi-dimensionality of HRQoL introduces challenges with weighting of the importance of each domain in producing summary measures. While this may be done statistically, individual patients may value the domains differently.

In its broadest sense health entails absence of infirmity, which reflects a negative aspect, but even more importantly it encompasses 'well-being', the ability of individuals live their life to its fullest potential, which reflects the positive elements of health. Both elements indeed belong to the concept of HRQoL, even though the focus within clinical research or routine clinical practice has tended to be on assessing the negative aspects of HRQOL. Cummins et al. (1998) critiques such an approach as leaning more towards a biomedical model of medicine. This also explains the

interchangeable usage of HRQOL with 'health status', due to the relevance of the latter in most clinical situations. Although a more comprehensive assessment of HRQOL is considered ideal, these may come at the cost of applicability and practicality of measurement instruments, explaining why the typical approach is to concentrate on those aspects of HRQOL that may be relevant for a particular patient group or which might be affected by therapy.

Components Of HRQoL

The plurality of views on HRQoL is reflected in varying scope of the definitions, for example in the inclusion of negative and/or positive aspects of health; inclusion of aspects of HRQoL amenable to the influence by medical care and lying in the ambit of health-care (Ferrans 2005). It is also important to consider the variation in the components of HRQoL. Nevertheless, instrument developers seem more comfortable with focusing on negative elements of health, as these seem more linked to aims of medical therapy. The different views on the components of HRQoL seem to focus on the following core domains: emotional status; physical functioning; social functioning; and medical symptoms (Fitzpatrick 1996). Various conceptual models have proposed for illustrating the components of HRQOL and how they are linked to each other and other outcomes. These are illustrated below.

Ware's model

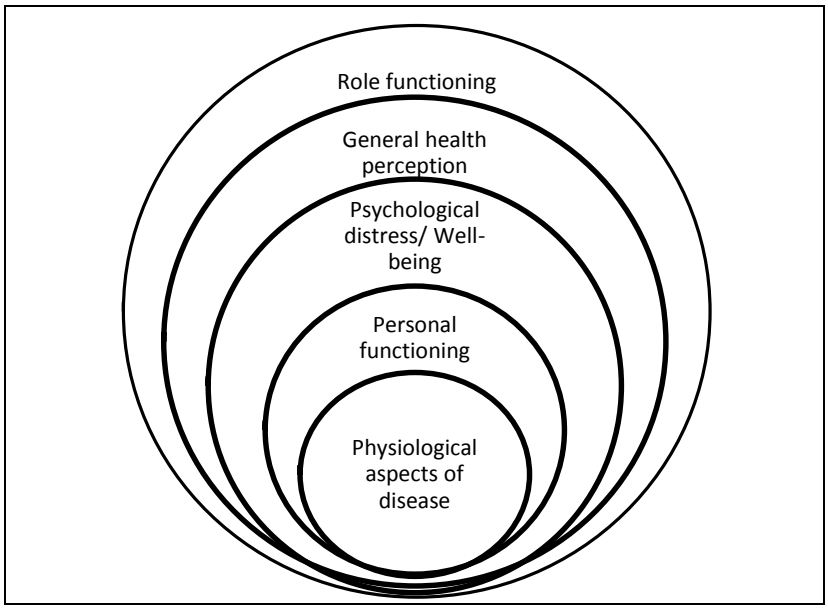
In one of the earliest models on HRQoL, Ware (1984) presented a framework for discussing disease and its impact on the patient's life. Health status or well-being was argued to have the following components (domains):

- *physiological status*, measurable physiological parameters of the disease such as symptoms, lab-values;
- *personal functioning*, performance of daily tasks such as self-care and other physical activities;
- *Psychological health and wellbeing*; this includes psychological effects of disease such as anxiety or frustration, but might also include positive affect, better mental health.
- *General health perception,* how the individual looked at their overall health considering their physical functioning, personal functioning, psychological distress and wellbeing.
- *Social well-being*, performance of usual roles whether in the community, school/work

or home.

The model is presented in Figure 1.1 where each domain, presented in a layer was hypothesised to influence the concentric outer layer, while allowing some feedback influence of outer layers to the domains inside. For example, while it is clear, problems in performance of tasks (personal functioning) may result in frustration or anger (psychological distress); similarly, anxiety (psychological distress) may not only result in limitations in some activities but may also affect individual's immune system.

Figure 1.1: Health status domains proposed in Ware's model



Source: adapted from Ware et al. (Ware 1984).

The lack of personal interpretation and evaluation from the patient, assessing how important their perceived status is, means that this framework may not reflect full impact of disease (Bloom 1984). On the other, individual's overall consideration of their health status is likely to encompass their ability to function socially, thus general perception and social role ought to be swapped.

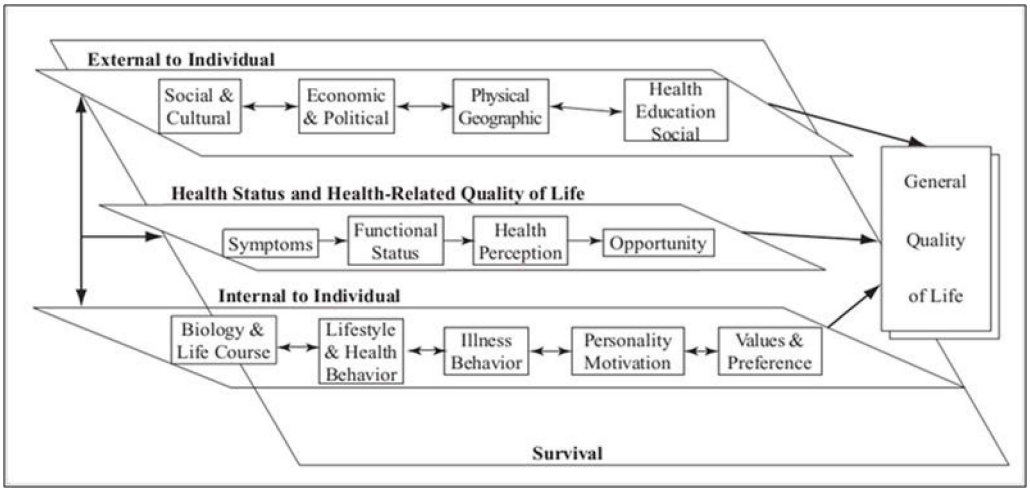
Recently, Ware (2003)suggested a re- conceptualisation of health status, into a model of three inter-related multi-layered health onions, comprised of a physical component, a mental component and a participation component. He cites a huge volume of evidence for a two components model which includes empirical studies on generic health status measures such as the SF-36, SIP, NHP and HIE. The rationale for a third separate component for measuring participation is that: this is consistent with the view reflected in the new ICF, disability and health to measure and interpretation role

participation separately. In addition, from health economics perspective, the utility implication of reduced role participation is the same regardless of its causes. Considering that social functioning is already seen to be a part of HRQOL, the practical implications of the newly discovered-and-proposed importance of social functioning, is yet to be understood.

Patrick and Chiang model

Patrick and Chiang (2000) proposed a comprehensive model illustrating the relationships between health/disease and the concepts of HRQoL and QOL (Figure 1.2). Their framework builds on a more basic model proposed earlier by (Wilson and Cleary 1995), where the causal relationship among 5 levels of health outcomes, from *biologic/physiological outcomes*; through *symptom status*; *functional status; general health perceptions* to *overall QOL* was proposed and illustrated. The new model has included greater clarification on the external environment as well as aspects internal to the individual; and illustrated their integration with health status and HRQOL. This is not only valuable in interpreting results when evaluating treatments, but may also reveal effect modifiers, which may be targeted by interventions (Ferrans 2005). Ultimately, therefore, general QOL is determined in a three-way interaction between external environment, health related quality of life, and the internal individual characteristics.

Figure 1.2: Patrick and Chiang QOL and HRQOL model



Source: adapted from Patrick and Chiang (2000).

Still, delineating determinants or inputs, and components of health status or HRQOL is essential in the conceptual purity of HRQOL, indispensable for the future development of the concept.

While this model seems to have successfully distinguished the "inputs" from the "components" of the health outcomes, the demarcation between "health status", QOL and HRQOL still seems fuzzy.

Practical issues

Assessing QoL is like assessing the beauty of a rose, irrespective of the aspects considered, colour, smell and height, capturing its full beauty is almost impossible and ultimately lies in the eyes of the individual (Mount and Scott (1983) cited in Carr et al. (2002)). Subjectivity is one of the core attributes of HRQoL concept. The importance attached to the different elements and domains of the HRQoL vary across individuals apart from changing over time for the same individual. How individuals experience, value and perceive their illness/disease condition, may reflect internal values, preferences, priorities, personality, their ability to adapt to their condition, support network, beliefs and other socio-demographic factors. This suggests the same level of disease severity may show different HRQoL profiles. The majority of patients with chronic conditions may still experience favourable QoL; which may not be worse than that seen in healthy individuals (Fitzpatrick 1996). Moreover, disease severity has shown to be weakly related to HRQoL (Fitzpatrick 1996; Bowling 2001). The subjective nature of HRQOL has been further highlighted and crystallised in the 'disability paradox', whereby people disabled from birth have shown higher than expected QoL; while an increase in self-reported QoL was seen in individuals with recently acquired disabilities, or post-diagnosis of terminal illnesses (Rapley 2003; Phillips 2006).

Furthermore, an often cited issue with applying HRQoL in clinical trials and routine practice relates to perceived lack of objective often assumed from the subjective nature of HRQoL outcomes (Grob 2007). While such perception is not necessarily obvious, the issue of objectivity of measurement is quite central to validity and reliability of any empirical measurement, therefore, it has a relevance to both subjective "soft" and objective "hard" endpoints. Unlike endpoints such as systolic BP or blood glucose levels, QoL is a theoretical construct, whose existence hinges upon a set of agreed upon indicators. Moreover, the lack of conceptual agreement implies that QoL construct is not only defined differently in different groups, but also the choice of its indicators will show dissimilarity, suggesting absence of a common unit of measurement. As a hypothetical construct, the link between what is being measured and the process of measurement may not necessarily be directly observable (Testa 2000).

The subjectivity and individualistic nature of QOL information may cause a moral-dilemma, with ethical implications. To be ethically acceptable, the design and implementation of research need

to pay attention to the interests of research subjects, for example respecting the expertise of patients in living with their condition (Rapley 2003). This means involving patients in the development or selection of QOL instruments upon which they will be assessed. The measured QOL therefore would be a reflection of individual values, preferences and socio-cultural characteristics of the patients.

Selecting HRQOL measures

Choice of the most appropriate instrument for assessing HRQoL depends on the objectives for collecting data, the environment of the application, methodological and practical considerations (Patrick and Deyo 1989). Researchers may need to choose between different types of instruments including individualised measures; disease specific; therapeutic-area specific; and generic instruments. *Individualised instruments* allow patients to choose what items are included in an instrument, as well as indicating how important they are to them (Luckett et al. 2009). These measures have failed to be generally accepted due to their large respondent and administrative burden (Luckett et al. 2009).

Disease-specific measures are intended for assessing QoL in specific patient populations, including content that clinicians and patient consider important for a given condition. Thus, they may have greater sensitivity to the clinical conditions allowing better discrimination between patients with different levels of disease severity, or for detecting change over time in the patient's condition (Bowling 2001). These measures are also important for their high practicality as irrelevant content is excluded. These measures are most useful in routine clinical practice and in clinical research and more sensitive to change over time.

Therapeutic-area specific instruments are a hybrid between disease-specific and generic instruments. They have a broader scope than disease-specific instruments, to allow application in more than one disease while on the other hand they maintain content that is relevant to the group of diseases beyond generic measures (Salek 1998). This makes these instruments relevant for various applications, both clinical research and routine clinical practice. Examples include the Dermatology Life Quality Index (DLQI) and the Skindex, used in dermatology.

Generic instruments capture impacts of the disease and its treatment on the general health-related quality of life. A major rationale for their use is their comprehensive outlook, going beyond impacts associated with particular symptoms or problems, to consider overall HRQoL. Therefore, they are broadly applicable across disease, health interventions, demographic and cultural

subgroups (Patrick and Deyo 1989). This makes them useful for addressing broad policy questions, addressing epidemiological questions, making comparisons across disease and patient groups (Bowling 2001). Nonetheless, these measures have limited usefulness in routine clinical practice which is related to their poor practicality and applicability as well as a relatively lower responsiveness (Greenhalgh et al. 2005).

Attributes of HRQOL instruments

Choice of instrument must also consider a number of key psychometric attributes in order to ensure the reliability, validity and interpretability of the data collected. These include:

- *Content validity*: The content of an instrument needs to be comprehensive, possess the right focus and emphasis, relevant to the target patient population and should reflect the stated aims of the instrument (Nunnally and Bernstein 1994; Salek 1998).
- *Construct validity*: Empirical evidence from several validation procedures supporting particular interpretation of scale scores (Frost et al. 2007a).
- *Reliability:* the instrument produces consistent and reproducible results (Streiner and Norman 2008).
- *Responsiveness:* The instrument must be capable of detecting important changes in the patient's condition over time even when such changes are small (Guyatt et al. 2002).
- *Interpretability*: Information supporting the qualitative meaning of the scale scores e.g. cut-offs for minimal clinical important difference (MCID) (Lohr 2002).
- *Practicality:* completing the instrument should impose minimal burden on respondents; and minimal administrative effort should be required in implementing the measure (Salek and Luscombe 1992).

HYPERHIDROSIS: DISEASE BACKGROUND

Hyperhidrosis is a pathologic skin condition characterised by sweat secretion in excess of the physiologic needs of the body necessary for thermo-regulation purposes (Atkins and Butler 2002). The condition has been classified in the WHO ICD-10 under code R-61 as *localised and generalised excessive sweating symptoms*. A distinction can be made between primary and

secondary hyperhidrosis. The former is related to excessive function of the sudo-motor sweat control system, otherwise the exact underlying cause is not fully understood (Vorkamp et al. 2010). The latter represents situations where the excessive sweating is caused by an underlying stimuli, for example infection, cancer, diabetes, medication, social phobia/anxiety or other disorders (Vorkamp et al. 2010).

The human skin contains 1.4 to 1.6 million *apocrine* and *eccrine* sweat glands (Leung et al. 1999). Hyperhidrosis is mainly associated with the latter, which are mainly concentrated on the palm, soles, face, axillae and to a lesser degree the chest and the back, while the former are confined to the axillary, the areola of the nipple, the anogenital area, and the excessive auditory areas (Leung et al. 1999). Although the exact cause of the dysfunction is not fully understood, size or number of sweat glands have been ruled out (Wörle et al. 2007). It is believed that the excessive sweating results from non-thermal sympathetic over stimulation, which explains why *primary excessive sweating* does not occur during sleep (Vorkamp et al. 2010). On the other, emotional stimuli seems to have a role in sweating as demonstrated in epidemiological studies reporting stress, emotions and social relationship as more important aggravating factors than physical exertion or heat (Park et al. 2010). The genetic transmission of the autonomic dysfunction has been hypothesised and explored (Ro et al. 2002). For instance, epidemiological studies have reported 30 – 65% positive family history in primary hyperhidrosis patients (Haider and Solish 2005; Lear et al. 2007). The limited understanding of the disease process and the fact that patients are not able to predict and know exactly when a sweating episode will break out, is not only a hindrance to the development of a cure, but is also a major source of anxiety and fear associated with the condition (Hornberger et al. 2004).

Diagnosis

The diagnosis of hyperhidrosis involves a thorough evaluation of the patient's history and a physical examination (Solish et al. 2007).The following clinical criteria have been included in various treatment guidelines: sweating that is bilateral and relatively symmetric, positive family history, more than one episode of sweating a week, onset of excessive sweating condition below age of 25, the sweat must cease during sleep, any secondary causes must have been ruled (Hornberger et al. 2004; Solish et al. 2007; Wörle et al. 2007). Assessing the severity of impacts in daily life activities experienced by the patient is key to confirming the diagnosis (Solish et al. 2007).

The clinical management of the hyperhidrosis depends on its severity. Clinical based severity assessment utilises measures including gravimetry, minor's iodine test and evaporimetry. The assessment of impairment in daily life activities, on the other hand, uses quality of life instruments and self-reported disease severity scales such as the HDSS or the DLQI (Vorkamp et al. 2010). Only the clinical-based measures are detailed in this section, a presentation of the QOL and disease severity measures will be presented in a separate section of this chapter.

Gravimetric assessment quantifies the amount of sweat produced over a particular skin area, within a given time period, by use of a filter-paper and a microbalance (Hund et al. 2002). The paper is weighed before and after its application to a thoroughly cleaned affected skin area, over a given time. Weight per unit of time can then be calculated from the before-and-after weight measurements. Kalkan (Kalkan et al. 1998) applied modification of this method in palmar hyperhidrosis. A pad glove made from gauze material was used in place of the filter paper and; surgical gloves were worn on top to prevent moisture from escaping. The weighing then used an electronic scale with sensitivity of 0.0001 g. The minor's iodine test is used in demarcating the area affected by hyperhidrosis (Glogau 2001). The affected area is wiped by an iodine tincture, then a starch is applied after thorough drying. New sweat secreted leads to a colour change demarcating the area affected, following the reaction between the iodine molecule and the starch. Assessment may be facilitated by taking a digital photo. The Ventilation capsule method assesses sweat production based on moisture evaporating from the skin measured using an electronic device (skin moisture meter) (Keller et al. 2009) A cup of 1 cm diameter connected to the device is used to capture moisture leaving the skin; with the amount of sweating over time (e.g. mg/cm/minute) read off a digital sweat meter reading (Ohhashi et al. 1998).

Clinical measures may have, however, limited usefulness in the management of hyperhidrosis for a variety of reasons. First, their anticipated objectivity is questionable. Cut-off quantity of sweat between patients and non-patients is unclear, an artefact of intra-individual variation of sweat production at different times and situations (Hund et al. 2002). Currently suggested cut-off values of 50 mg/5 min for females and 100 mg/5 min for males are arbitrary, their specificity or sensitivity has not been established (Hornberger et al. 2004). The other issue relates to the practicality of these measures. The cumbersome nature of these clinical tests makes them challenging to apply in

routine clinical practice, limiting their usefulness to a few cases and research settings (Chang et al. 2011). The assessment of quality of life impairment resulting from the sweating, therefore, is important to the diagnosis and management of hyperhidrosis in routine practice.

Treatment

Treatment guidelines for hyperhidrosis have been proposed and published in Canada, the U.S. and Germany (Hornberger et al. 2004; Solish et al. 2007; Wörle et al. 2007) but not in the UK. The management of hyperhidrosis largely depends on its severity, impact on quality of life and the body area affected (Table 1.1). The available therapies include topical creams containing aluminium chloride, tap water iontopheresis, interdermal injection with botulinum toxin, systemic treatment with pharmaceutical intervention and surgery.

Topical treatment with salts:Aluminium chloride (10% - 35%) based topical creams are regarded as first line treatment for mild focal hyperhidrosis (Hill and Glade 2012). These have shown to be most efficacious in axillary hyperhidrosis (Goh 2007; Streker et al. 2012). Lower concentration creams are available over the counter.

Tap water Iontopheresis: Tap water Iontopheresis, used for treating palmo-plantar hyperhidrosis, involves using a device to apply an electrical current to introduce ions into the affected skin, leading to an obstruction of the sweat glands (Hill and Glade 2012).

Interdermal botulinum toxin-a injection*:* Interdermal injection with botulinum toxin-a results in flaccid paralysis and autonomic dysfunction, providing a treatment effect lasting 4 – 17 months (Heckmann et al. 2001). This therapy is used in palmar, plantar and axillary hyperhidrosis.

Oral systemic treatments*:* Systemic treatment uses agents with an anti-cholinergic or anti-depressant effect such as methanthelium bromide or glycopyrrolate (Wörle et al. 2007). Though generally not recommended this therapy is relevant for generalised hyperhidrosis or as an adjunct therapy in focal hyperhidrosis (Hornberger et al. 2004; Solish et al. 2007).

Surgical treatment*:* In axillary hyperhidrosis sweat glands may be locally removed using retrodermal curettage or liposuction (Henteleff and Kalavrouziotis 2008). An alternative procedure, endoscopic thoracic sympathectomy, which targets disrupting the sympathetic nerves is used in axillar, plantar and palmar hyperhidrosis.

Epidemiology

Quantifying the number of people with hyperhidrosis within the general population at a given point in time, incidence and understanding the epidemiologic characteristics of the suffers is not only useful for understanding the full burden and impacts of the condition, but is also a prerequisite to determining and meeting the specific health care needs of this patient population. Therefore, articles publishing epidemiological information of hyperhidrosis were reviewed. Six studies have published figures on prevalence and an additional four provide useful epidemiological details about the condition.

Prevalence

Brown et al. (Brown et al. 2005) using a retrospective analysis of medical and prescription codes obtained from the General Practice Research Database (GPRD) which contains medical records of up to 3 million patients, estimated the prevalence of hyperhidrosis in the UK to be at 1.6%. Strutton et al. (2004) in a mail survey of 150, 000 households in the U.S., found prevalence to be at 2.8%. Prevalence was higher in the 25 – 34 age group (4.5%) and did not show significant differences between male and females. Besides national and environmental differences which might give rise to differences, use of medical records in the UK study assumes that patients were correctly diagnosed and classified in records, which may not always be guaranteed. Moreover, prevalence may be different in the patient population not seeking for treatment. Three more studies estimated prevalence in different sub-populations. Schäfer et al. (Schäfer et al. 2012) estimated prevalence in a sample comprised of 14,336 employees of 52 companies in Germany. Overall prevalence was estimated to be 16.3%; of those affected, 6.1% had frequent or continuous sweating. Higher rate of generalised sweating (68%) reported in this study, suggests that the prevalence rate might have included those with secondary hyperhidrosis.

A retrospective analysis of data from routine medical examination carried out on military recruits, from 94,806 Israeli adolescents (16 – 22 yrs) reported a prevalence of 0.2% for males and 0.1% for females (Wohl et al. 2007). While hyperhidrosis was the most prevalent skin problem observed, these figures seem relatively lower than prevalence rates observed in Europe and North America. The differences in geography and culture, and their influence on the actual amount of sweating and general views on what sweating is considered normal, cannot be ignored. Tu et. al. (Tu et al. 2007) carried out a survey of 13,000 adolescents (15 – 22yrs) from 10 high schools and 3 colleges

in Fuzhou People's Republic of China based on cluster-sampled from 42 high schools and 12 colleges. Prevalence was estimated at 4.59% for the full sample, but was significantly slightly higher in the high-school group; with no significant differences between male and female. Setting aside the differences in how data used in estimating prevalence was collected and analysed, these figures indicate a wide variation in hyperhidrosis according to age, geography & culture. Prevalence was highest in those of the active working group. There was no consensus on gender differences in prevalence rates.

Disease characteristics

The studies by Schäfer et al. (Germany) and Strutton et al. (U.S.) reported *axillary* hyperhidrosis as the commonest body site affected. Tu et al. (China), on the other hand, reported *palmar-plantar* hyperhidrosis as the most prevalent. Mean age at onset has been reported to range from 15 – 25 yrs, with palmer-plantar hyperhidrosis showing childhood onset and axillary starting during or post-puberty (Strutton et al. 2004; Lear et al. 2007). This might explain the observed higher prevalence for palmar-plantar in the study with adolescents, while the studies including post-adolescent participants showed the axillary as the commonest site. Conflicting conclusions were made on gender differences in self-reported disease severity in patients receiving treatment. Kirimian-Teherani et al. (2009) reported a higher level of self-assessed severity in women, while Lear et al. (2007) and Strutton et al. (2004) found no differences. Individuals with hyperhidrosis showed more *dermatologic co-morbidity*. *Type IV allergies* were more common in females with hyperhidrosis than in general female population (Karimian-Teherani et al. 2009); Risk for *Psoriasis* and *tinea pedum* was also higher in persons with hyperhidrosis than those without (Schäfer et al. 2012) which might be a consequence of the constant wetness or topical treatments being used.

Hyperhidrosis remains largely untreated: 47% of women and 28.6% of men were reported to have discussed their sweating problem with their health care provider in the US study (Strutton et al. 2004). As much as 66%-88% of hyperhidrosis patients had tried various forms of self-treatment to manage their condition (Hamm et al. 2006; Lear et al. 2007). On the other hand, patients who seek for treatment tend to utilise health services more, for instance, 85% of participants in a study using a clinic sample had visited their physician at least once in the previous year (Hamm et al. 2006).

Table 1.1: Treatments for hyperhidrosis

Region	First Line	Second Line	Third Line
Axillary	*Mild* Aluminium Chloride 10 – 12 % to 35%; applied at bedtime; up to twice daily if necessary and if well tolerated *Severe* i. 2nd Line treatment iii. Oral systemic medications (e.g. anti-cholinergics) alone or as adjuvant treatment.	Intra-dermal injection with Botulinum Toxin 50 - 100 U/Axilla	i. Combination of 1st and 2nd line treatment. ii. Oral systemic medications such as anti-cholinergics, either alone or as adjuvant treatment. iii. Surgical intervention - Liposuction iv. Surgical intervention – ETS.
Palmar	*Mild* AC 10 – 12% up to 50% *Severe* Aluminium Chloride or BTX-A or Combining both.	*Mild* i. BTX-A ii. Tap-water iontopheresis therapy *Severe* i. Oral systemic medications (anti-cholinergics) alone or as adjuvant therapy ii. Use anti-cholinergics (glycopyrrolate solution) instead of water during iontophoresis	Surgical intervention – Liposuction or ETS.
Plantar	*Mild* AC in absolute ethanol or Salicylic acid gel, 20 – 50% concentration *Severe* i. AC ii. BTX-A iii. Iontophoresis	*Mild* i. Intra-dermal injection of BTX-A. ii. Tap water Iontophoresis *Severe* i. Oral systemic medication (anti-cholinergic)	*Mild* Add topical AC to line 2 treatment *Severe* i. Surgical intervention - ETS ii. Iontophoresis therapy using Glycopyrrolate solution

Psychological co-morbidities

The literature review also uncovered a strand of literature on psychiatric morbidity in patients with hyperhidrosis. While this is not part of the construct of QOL, its possible implications for QOL justified the inclusion of this set of evidence. Ak et al. (2013) investigated Alexithymia[1] in Turkish patients with hyperhidrosis (N = 50) attending a dermatology clinic, in comparison with non-patients (N = 44). The structured clinical interview (SCID-I) for Diagnostic and Statistical Manual of Disorders fourth edition (DSM-IV Disorders) and the Toronto Alexithymia Scale (TAS-20) instruments were employed for psychiatric diagnosis. Among the patients with PFH, 45.6% were alexithymic, comparison to 2.23% among the non-hyperhidrotic subject. A personality trait such as alexythymia might be a mental reflection of the abnormal central nervous system characteristic of hyperhidrosis (associated with abnormal response to emotional stresses) providing further indication of the genetic basis for hyperhidrosis (Ak et al. 2013). Whether such personality trait abnormalities are exclusive to HH, apart from other dermatological illnesses is unclear. Furthermore, the involvement of a third key parameter such as depression or anxiety, may not be ruled out.

Another study on psychopathology in hyperhidrosis was carried out by Ruchinskas et al. (2002) in patients from the USA awaiting treatment with ETS. The Spielberg State-Trait Anxiety Inventory (STAI) was used for assessing anxiety while other psychological pathologies were identified by Minnesota Multiphasic Personality Inventory 2 (MMPI-2). Eighty-eight percent of the patients had normal MMPI-2 scores, with patients showing normal values for psychologic conditions including hypochondrias, depression, schizophrenia, hypothermia and social introversion. STAI scores for 86% of the patients was within ranges for normal population. Nonetheless, it is noteworthy that the patient population used does not represent the full continuum of patients with hyperhidrosis. The results obtained, therefore, may not apply across the board to other patients.

Weber et al. (2005) investigated the possible association between PFH and psychopathology (anxiety, depression and social phobia) in 70 patients, with palmar, plantar and axillary hyperhidrosis across multiple centres in Germany. The STAI, Hospital Anxiety and Depression

1 Personality disorder involving insufficiency in identification and expression of emotions.

Scale – Depression (HADS-D), Symptom Checklist-90-Revised (SCL-90-R), Social Phobia Scale (SPS) were used in patient assessment. Values of STAI (55.6±10.3), HADS-D (55.5±11.3), SCL-90 (51.6±10.6 - 56.3±13.3) and SPS (15.2 ± 13.5) were within normal ranges. No significant differences were observed across patients suffering from HH affecting different sites. Still, the subgroup of patients which was subsequently treated with Botulinum Toxin – A (BTX-A) (31 patients) showed elevated scores for SPS reflecting social phobia. Eligibility for BTX-A includes having failed to benefit from first line treatment e.g. Aluminium chloride, which may reflect greater disease severity.

The absence of anxiety disorders in patients with HH has been further supported by Ramos et al. (2006) in Spanish patients awaiting ETS surgery (N = 158) and Schneier et al. (2012) in two samples: HH patients awaiting ETS (N = 40) and patients suffering from Social Anxiety Disorder (N = 40) from the USA. Although Ramos et al. found the levels of social anxiety to be well within those for the normal population, based on the STAI, the anxiety experienced was considered to still have a severely incapacitating effect on the patient's life. The patients included in either studies represent highly selected patient populations.

Cost of illness

Evidence related to the costs associated with hyperhidrosis or its treatment was scarce (Kowalski et al. 2005) reported on the budget impact of including BTX-A in the treatment pathway for severe primary axillary HH inadequately managed with topical agents in US managed care populations. Their results are from the perspective of a 1 million member US managed care plan over a 1 year period and consideration of costs is limited to treatment and medical costs. The inclusion of BTX-A in the treatment plan results in an incremental cost of 1,400 US$ per successfully treated patient. Annual costs per severe primary axillary hyperhidrosis patient were reported at $ 578 where BTX-A is included in the treatment pathway and $ 312 for a treatment pathway without such costs. Considering that costs of office visit may be unique for a particular insurance organisation, these figures may be markedly different for members of a different insurance organisation. Ambrogi et al. (2009) has reported costs associated with treatment primary palmar hyperhidrosis, in a prospective comparative open-label study comparing BTX-A and ETS surgery, in Italian patients (N = 154). They considered medical and treatment costs, including the costs of hospital stay. They report € 2654 ± €145 per case of ETS surgery, and €655 ± 23 per case of BTX-A therapy. The role of the study protocol in the costs incurred on the patients in their study may not be ruled out. On

the other hand, as the reported figures are based on the actual costs at the hospital where the study was carried out, figures reported may also reflect organisational and operational features of the institution. Either issues may have a bearing on the generalizability of the findings

Assessing HRQoL in hyperhidrosis

The nature and extent of the handicap and impairment in the patient's life resulting from skin disease is well understood (Jowett and Ryan 1985; Finlay and Ryan 1996). Its impact extends across various areas of life (such as emotional distress, impact on social life such as in relationships, professional-life, physical discomfort from itching or wet-skin, and the burden associated with managing the condition). This also has to be seen in the light of skin's high visibility as well as its particular role in self-image (Beltraminelli and Itin 2008). On the other hand, for conditions such as hyperhidrosis, the laboratory-clinical- measures of sweat are difficult to interpret apart reliability and practicality issues (Hund et al. 2002), leaving self-reported impacts on the patient's life as 'vital sign' of disease activity (Chren 2005) especially in routine clinical practice.

This means that the assessment of the impacts of hyperhidrosis on QoL is crucial in studies evaluating effectiveness of treatments or in clinical management of hyperhidrosis patients. As a long term condition, treatment therapies (non-surgical treatments) in hyperhidrosis are largely concerned with enhancing the patients' quality of life: their ability to manage everyday routine such as performing housework; interacting with others; participating and contributing to social activity; and performance at work/school. On the other hand, treatment therapies in hyperhidrosis are often associated with unbearable side effects such as compensatory sweating (ETS surgery, Inter-dermal Botox Injection), mouth dryness (anti-cholinergics) or transient hand weakness (Inter-dermal Botox Injection), raising the question whether benefits of treatment outweigh the burden associated with side effects. Assessment of a patient's quality of life, in such contexts/situations, therefore, may offer a comprehensive framework for a more holistic evaluation of benefits and risk. The assessment of patient reported outcomes, therefore, is just as important in hyperhidrosis as in other long term skin and non-skin conditions. Facilitating the understanding of the burden associated with the disease at the society level, for instance by looking impairment in quality of life and the lost of productivity. Moreover, understanding QoL impairment is key not only in determining the needs of patients for various health care services, but plays a vital role in

determining the severity of condition, useful for the effective diagnosis and management of the condition.

Impacts on QoL

The humanistic consequences of disease are not only an important element of the overall burden of disease but reflect issues that are of most relevance to patients, their families and society at large. As the patient is the expert in their experience with a condition, their voice should matter most when considering such outcomes. Qualitative research methods are, therefore/as such quite useful for delving into phenomenon, particularly obtaining insights into the beliefs, values and perceptions of informants capture in their own words (Pope and Mays 2008). This makes them the first choice in understanding the disease experience of patients. Thus, a literature review was carried out to uncover the impacts of HH on patients' QoL. This was key to understanding the areas of HRQoL of importance in HH and how patients perceive and describe such effects. Apart from providing a rationale for assessing HRQoL, this provides an important foundation for developing a new measure for assessing QoL.

A structured process was following in sourcing and selecting studies for inclusion in the review. The literature searches were carried in multiple bibliographic databases including Pubmed; Google Scholar, Ovid/Embase and Scopus. A combination of 3 blocks of terms was applied to the title, abstract and keywords of the databases: block 1: *hyperhidrosis*; block 2: *effects, effects on patients, impact, impact on patients*, block 3: *health related quality of life, quality of life, patient's life, daily life, everyday life, lifestyle*. The initial eligibility criteria was that studies should be investigating QOL in patients with primary HH using qualitative research methods. When only one relevant study was found, eligibility criteria was changed to include studies that had employed quantitative methods. In this case the instrument used in the investigation should be validated and should be readily interpretable, in addition to reporting baseline results. Only the qualitative study is reported in this section. The quantitative studies are reported in the next section as clinical application of quality of life instruments.

Thomas et al. (2006) investigated lifestyle impact, compensating behaviours and treatment experiences of female hyperhidrosis patients, through three focus group discussions with 21 female patients with HH from the US. Patients were recruited through the database of the

international hyperhidrosis society (IHHS). Patients reported effects on their relationships with family and friends, in their professional interactions. Additionally, effects were reported on patient's self-confidence and self-esteem, besides the psychological distress. Patients mentioned feeling their life was taken over by the hyperhidrosis all the time. They worried about their clothes getting soiled which led embarrassment when it happened. One participant was quoted as follows:

> "we *were running around...I had to put my shirt around my waist because I had a spot on the back of my pants from the waist down to the knees. It looked like I wet myself and I didn't want people to make fun of me on the last day of school*"

Patients reported on the inconvenience, effort and cost associated with strategies employed to deal with the sweating and its symptoms, for example choosing clothing that hid the sweat, using tissues and pads, using a fan when getting dressed. While this study provides valuable insights, exclusion of males means that gender-specific experiences of males were not reflected in the results. On the other hand, there is no indication whether the issues of relevance to patients had been exhaustively explored.

Critical appraisal of HRQOL measures

In view of the importance of HRQOL in hyperhidrosis, its accurate evaluation depends on the availability of the availability of robust and suitable instruments. Previously, critical reviews of HRQOL instruments used in hyperhidrosis have been carried out. Panhofer et al. (2005) focused on measures used in hyperhidrosis patients treated with sympathetic surgery. Other reviews by Cetindag et al. (2008) and Solish et al. (Solish et al. 2008) covered measures used across all types of hyperhidrosis and associated with all forms of treatment, although they did not include a critical evaluation of the psychometric properties of the measures identified. Therefore, an important aim of this chapter was to review the instruments used in HRQoL measurement in hyperhidrosis particularly including an appraisal of their psychometric properties. As a comprehensive overview of the field, such a review would be a useful resource to all stakeholders involved in measuring HRQoL in hyperhidrosis, including patients, healthcare practitioners, and health care decision makers.

To identify instruments that have been used in HRQoL assessment in hyperhidrosis a literature search was carried out in PubMed, PsycINFO and EMBASE. The initial search was based on the following terms: "hyperhidrosis and quality of life"; "hyperhidrosis and daily life"; "hyperhidrosis

and clinical trial"; "hyperhidrosis and impact". References of the papers initially extracted were also searched to identify more material for our review. An additional search strategy was based on the identified instrument e.g. "SF-36 and hyperhidrosis", "DLQI and hyperhidrosis" to identify all studies using the instruments in hyperhidrosis patients. A study was included if it reported the measurement of HRQoL in hyperhidrosis patients using a HRQoL instrument; or if it reported the psychometric properties of such an instrument. An instrument was included if it was developed for the measurement of HRQoL and if it had been used in hyperhidrosis patients. Such instruments could be disease specific, dermatology -specific or generic. We limited ourselves to HRQoL self-assessed by patients, either self-completed questionnaire or interviewer administered. Study-specific instruments were excluded.

Information related to the instruments was extracted following standard quality criteria for HRQoL instruments (Lohr 2002; Both et al. 2007). Information extracted included key psychometric properties, descriptive information and additionally details of studies applying the instrument. The criteria used in the evaluation of psychometric properties are presented in Table 1.2. Thirteen instruments have been applied in assessing HRQoL in hyperhidrosis; this includes four generic instruments, four dermatology-specific instruments and five hyperhidrosis-specific instruments (Table 1.3, Table 1.4) .

Generic HRQoL Instruments

Short form 36 (SF-36)

The SF-36 is a generic measure of health status developed for use in population surveys and studies supporting health policy development (McDowell 2006). Its application, however, has extended beyond its initial purpose to include the evaluation of effectiveness of therapies. The content of the instrument was obtained from the most frequently measured concepts in widely used surveys and those mostly affected by disease and treatment (Stewart and Ware 1992; Quality-Metric 2013). A 5 point Likert scale is used for the 36 items of the second version of the instrument, SF36v2, translating into scores of 0 to 100. These use a recall period of 4 weeks. An additional item assesses change in general health over last year (i.e. 1 year recall period).

The 36 items form 8 domains (physical functioning, emotional role functioning, physical role functioning, social role functioning, mental health, vitality, bodily pain, and general health

perception) and are further aggregated into two higher order scales (summary measures), the physical component scale and the mental component scale (Kini and DeLong 2012). This structure is supported by results of factor analysis and correlation analysis (Ware et al. 1994; Quality-Metric 2013). Numerous construct validation studies have performed various tests of validity for the SF-36, including content, concurrent, criterion, predictive validity (Bowling 2005; McDowell 2006; Quality-Metric 2013). Adequate internal consistency, test-retest reliability (2 weeks – 6 months) and responsiveness have, likewise, been established in diverse patient populations (Ware et al. 1993; McDowell 2006). Only some psychometric attributes of the SF-36 in patients with skin disease are known and nearly no attributes are known in patients with hyperhidrosis. For example, the internal consistency in patients with skin diseases has not yet been reported (Both et al. 2007). Shikiar et al. (Shikiar et al. 2006), based on a phase II RCT of Adalimumab in psoriasis patients, assessed the construct validity, responsiveness and MID of the SF-36. The bodily pain and social functioning scales of the SF-36, correlated well with the DLQI and clinical endpoints. All SF-36 sub-scales were sensitive to changes in clinical anchors. Another study, by Chren et al. (Chren et al. 1997) has also reported acceptable correlation between the SF-36 and the Skindex, although the former was more sensitive to physical symptoms or social effects.

The SF-36 may be administered in various ways, as a self-completion paper and pencil version, self-completed electronic delivery or using interview delivery. Other versions of the instrument are available, the acute version, using a 1-week recall period, the SF-12, with 12 items, which is discussed in the next section, and the SF-6D, where it is possible to calculate health utilities for use in economic evaluation studies. Furthermore, the practicality and acceptability of the SF-36 has been established, with completion taking less than 10 minutes (Brazier et al. 1992; Bowling 2005).

A number of limitations have been identified on the SF-36. For example with respect to the content, concepts including sleep adequacy, sexual function, health distress, eating, recreation and hobbies, communication have not been included. Hunt and McKenna (1993) have observed an over-reliance on psychometric techniques particularly in the early development of the instrument. The instrument has been used in at least four studies in hyperhidrosis, all of which were focused on assessing the efficacy of surgical therapy, alone or in comparison with other therapies. Sayeed et al. (1998) concluded that the SF-36 was not suitable for hyperhidrosis due to its irrelevance in hyperhidrosis patients.

Short form 12 (SF-12)

The SF-12 was designed as a short-version of the SF-36, to be brief enough to fit on a single A4 page with completion taking no more than two minutes (McDowell 2006). Selection of items included was based on their psychometric properties with the intention of explaining at least 90% of variance in (MCS and PCS) scales of the SF-36 (Ware et al. 1996). Studies on the factor structure of the SF-12 have reported varied results. Jenkinson (1997) based on a diverse UK community based patient population (N = 9332) supported the two factor structure. On the other hand, Jakobsson et al. (2011) found the contrary evidence. Adequate test-retest reliability correlation has been established (Both et al. 2007; Jakobsson et al.).

Capability of the SF-12 to discriminate among diagnostic groups of patients was considered comparable to that of SF-36 (McDowell 2006). Significant but weak correlations have been observed between DLQI and the MCS and PCS of the SF-12 (Grozdev et al. 2012). The PASI, a measure of disease severity in psoriasis showed significant association with PCS, and only showed negligible association with MCS.

Table 1.2: Criteria used in evaluating the measures

Aspect	Definition	Criteria & Code
Content Validity	Evidence that the domain of an instrument is appropriate relative to its intended use (Lohr, 2002). The conceptual and empirical basis for the items of the instrument. The involvement of the target population	++ Target patients and experts were involved + No patient involvement, other form of content validation given. - Inadequate content validity 0 No information reported.
Construct validity	Evidence that supports a proposed interpretation of scores based on theoretical implications associated with the constructs being measured (Ibid, 2002). Does the tool confirm hypothesised differences (Both et al, 2007)?	++ At least 75% of results in accordance with hypothesis, based on robust design and method. + Under 75% of results in accordance with specific hypothesis, adequate methods used. - Hypothesis not confirmed or inadequate methods used 0 No information reported
Convergent Validity	Does the tool relate to other tools assessing the same construct (Both et al., 2007)?	++ Correlation > 0.70 + Correlation < 0.70 - Correlations not statistically significant. 0 No information reported
Internal Consistency	The precision of the scale based on the homogeneity of the scale's items at one point in time (Lohr, 2002).	++ Cronbach alpha 0.70 – 0.95 + Cronbach alpha below 0.70 or above 0.95 - Very low Cronbach alpha, inconsistencies observed. 0 No information reported
Test – retest reliability	Does a repeated administration of the tool within a reasonable period of time results in similar results (Both et al. 2007)?	++ ICC above 0.70 + ICC below 0.70 - No correlation observed 0 No information reported

Table 1.2 (continued)

Aspect	Definition	Criteria & Code
Responsiveness	The ability of the instrument to detect changes over time OR differences between patients, due to therapy or impact of disease	++ 75% of results showed confirmation of hypothesis, based on an adequate measure. + Less than 75% results confirm hypothesis/conflicting evidence. - Poor or solely based on statistical evidence. 0 No information reported.
Floor and Ceiling effects	Does the tool capture the detail and breadth of real differences among persons?	++ Less than 20% in extremities. - More than 20% in extremities. 0 no information reported.
Interpretability	Can qualitative meaning be assigned to the scores (Veenhof et al. 2007)?	++ Thresholds provided based on anchor or banding techniques + Distribution based techniques used 0 no information reported.
MCID	Has the minimal change relevant to patients been reported?	++ MCID reported. 0 MCID not known.
Respondent Burden	Is length and content acceptable to patients?	++ Less than 10 minutes - More than 10 minutes or problems with acceptability 0 no information reported
Structure	Evidence in support of the proposed structure or scaling of the instrument	++ Item Response theory confirms proposed structure + Factor analysis/regression analysis confirms proposed structure - Factors analysis and item response theory does not confirm proposed structure 0 No information provided

Nottingham health profile (NHP)

The NHP was developed as a measure of perceived social, physical and emotional health problems in primary health-care setting (Hunt et al. 1985). Its content development involved patients with acute and chronic ailments (McDowell 2006). The 38 items of the NHP, scored on a binary scale (yes/no), form 6 domains. A frame of reference of 'at the moment' is used. Scores can be presented as a profile for each domain by summing all affirmed items, alternatively a weighting can be employed to give scores from 0 to 100 (2007). The proposed structure is not supported by empirical evidence, a study based on factor analysis, supports two higher order domains (Prieto et al. 1998). Convergence validity has been demonstrated, for instance moderate-strong correlation with SF-36 has been reported. Favourable internal consistency and test-retest reliability has been reported in varied patient populations for instance in patients with Asthma, Migraine and leg-ulcers although no such evidence exists for hyperhidrosis. Ceiling effects have been reported, indicating poor sensitivity of the NHP in minor levels of disability; or in distinguishing between levels of good health (McDowell 2006). This raises further concerns regarding its suitability for hyperhidrosis. The application of this instrument in hyperhidrosis has been limited, two studies using the NHP were found.

Illness intrusiveness ratings scale (IIRS)

The Illness Intrusiveness Rating Scale (IIRS) was developed as a measure of the degree of disturbance to a patient's wellbeing, including lifestyle, activities and interests, resulting from a particular health problem, disease or its treatment (Devins 2010). The IIRS was originally developed for assessing the declining control over various life aspects in end-stage renal disease patients (Devins et al. 1983). The items of the instrument are based on a major social research project by Flanagan et al. (Flanagan 1978) undertaken in the U.S. to identify factors relevant to quality of life for Americans in the 1970's (Devins 2010).

Responses to the 13 items of the instrument are given on a 7-point Likert scale, with a total score calculated as a simple summation of item scores, ranging from 91 (highest intrusiveness) to 13 (lowest intrusiveness). The IIRS has three domains established on various language versions of the instrument (Korean and English) and in diverse disease conditions (chronic illnesses and anxiety disorders) (Bieling et al. 2001; Devins et al. 2001; Kim et al. 2005). However, the factor analysis supports exclusion of 1 item (related to diet) from the subscale scores (Devins 2010).

The validity, reliability and sensitivity of the IIRS in hyperhidrosis patients has been evaluated by (Cinà and Clase 1999). Moderate correlation was reported between total IIRS scores and new items assessing the global severity of hyperhidrosis, including limitations in the number of clothing-changes and choice of clothing (Table 1). Strong reliability was established based on strong internal consistency and test-retest reliability test.
The IIRS has been translated into, and validated in, other languages including French, Hungarian, and Korean although no formal cultural equivalence studies have been undertaken (Devins 2010). The IIRS is can be administered in various means, self-completion paper and pencil, web-based version or interview (Devins et al. 1983; Cinà and Clase 1999; Ritter et al. 2004). The IIRS takes no more than 15 minutes to complete and on the other hand, has an easy scoring system.

A few concerns regarding the IIRS are as follows. First, the use of a 7 point scale has an unclear rationale. While presenting a cognitive burden, it may unnecessarily introduce noise in measurement. The interpretation of the IIRS poses a unique challenge, especially in deciphering its clinical meaningfulness. This is because it is not possible to "experience illness intrusive in the absence of a health problem ...except those experience lifestyle disruptions vicariously" (Devins 2010). Conceptually, illness intrusiveness may not be assessed in a non-diseased population.

Skin-Specific Quality of Life Instruments

Skindex

The Skindex is a 30-item questionnaire that measures the effects of skin disease on the quality of life of patients, based on a 4-week recall period. Items assess the frequency of bother from skin condition with responses scored on a 5 point scale, from 0 -'never' reflecting no effect, to 100 – 'all the time' for maximum effect. Excluding one item, the remaining 29 items form three subscales (physical functioning, symptoms and emotional-wellbeing). This structure has been confirmed by factor analysis (Abeni et al. 2002; Augustin et al. 2004b).
Early development involved patients attending private dermatology clinics and a VA hospital (dermatology clinic) in the U.S. (Chren et al. 1996). Construct validation has been carried out. For example, patients with inflammatory dermatoses showed significantly higher scores than isolated skin lesions ((Chren et al. ; Abeni et al.)). Moderate correlations were obtained between: the Skindex scales and comparable scales of the SF-36; 'symptom scale' and the 'physical functioning

scale' of the FLQA-d; 'symptom scale' and MHF's circle of itching and scatching scales (Chren et al. 1997; Augustin et al. 2004b). Additionally, the Skindex showed sensitivity to change in the patients condition, reflecting disease severity measures (PASI and EASI) ((Chren et al. ; Augustin et al.)).

To facilitate interpretation of the Skindex scores, scale banding systems applying both anchor and distribution based approaches have been developed (Nijsten et al. 2009; Prinsen et al. 2010; Prinsen et al. 2011). The Skindex has been culturally adapted into German, Italian, Spanish (Both et al.). Apart from a paper and pencil version, an electronic version is available in the Netherlands. Two short-versions of the Skindex have been developed and validated, Skindex-16 and Skindex-17, following two alternative psychometric theories, classical test-theory and item response theory (Rasch model) (Chren et al. 2001; Nijsten et al. 2006b). Skindex – 16 assesses extent of bother and also resolved the substantial floor effects some of the items. On the other hand, in the Rasch-based Skindex 17, response categories were reduced to three levels with items regrouped into two scales: *emotional* and *social functioning* scale (with 17 items) and symptoms scale (with 5 items). This version can explain up to at least 85% of the variance in the Skindex 29 scores. A single study has applied the Skindex in hyperhidrosis (Weber et al. 2005).

Dermatology life quality index (DLQI)

The DLQI is a 10-item questionnaire developed as an easy to use assessment tool, for measuring dermatology specific health related quality of life in routine clinical practice (Finlay and Khan 1994). The measure was developed for use in routine clinical practice and it fits on a single A4 paper. Items assess the intensity of the effects of skin condition on patient's QoL on a 4 point scale, from not at all (0), for minimal effect to very much, (3) for maximal effect. These cover six domains including symptoms, daily activities, leisure, work/school, personal relationships and treatment. A summary score calculated as the summation of the item scores is used. However, neither the proposed six-domain structure nor the total score subsuming unidimensionality are unequivocatnly supported by evidence. Factor analytic studies have reported from one to four factors (Basra et al. 2008). Rasch analysis on the DLQI has confirmed unidimensionality in patients with Atopic Dermatitis, but not those with Psoriasis (Twiss et al. 2011).

The development of the DLQI involved patients with various skin diseases attending an outpatient dermatologic clinic in the UK (Finlay, 1994). In the initial validation study which compared

patients with a normal population, the DLQI scores showed high specificity, repeatability and internal consistency (Finlay and Khan 1994). The validity, reliability and responsiveness has been extensively demonstrated in at least 115 studies (Basra et al. 2008). The DLQI has been used widely including: in 33 different skin conditions; formal translations for 55 languages, although only 9 cultural adaptation studies have been reported; use in parallel with at 30 generic and disease specific questionnaires. Data supporting interpretation of results for instance or banding system or MID is readily available (Hongbo et al. 2005; Basra et al. 2008). Specific MCID cut-off values for axillary and palmar hyperhidrosis have been proposed (Kowalski 2007). Nevertheless, a few shortfalls have been identified. High floor and ceiling effects have been noted (Both et al.). Furthermore, differential item functioning has been noted in some items with respect to culture, age and gender (Nijsten et al. 2007). Furthermore, some components of QoL such as emotional well-being have not been adequately covered (Basra et al.).

Patient benefit index (PBI)

The Patient Benefit Index (PBI) was developed for assessing the therapeutic benefit from dermatologic therapy, based on patient's own therapeutic goals (Augustin et al. 2009). Its content was based on interviews with patients, regarding the burden resulting from their disease and relevant benefit from treatment; and an expert panel (including clinicians) (Augustin et al. 2008). The instrument is comprised of a set of two questionnaires, the Patient Needs Questionnaire (PNQ), which assesses the patient's therapeutic needs (goals), and the Patients Benefit Questionnaire (PBQ), which measures the magnitude to which the needs (goals) captured by the first questionnaire are met. Each questionnaire has 23 items, which are paired between the two instruments. Importance of the treatment goals (captured by the 23 items) is rated on a 5 point Likert scale, from not at all important (0), for an issue/area considered unimportant by the patient, to very important (5) for an issue/are considered very important. The final score, representing the PBI, is calculated as the PBQ score of the individual weighed by the respective importance given to a specific need (item score on the PNQ) expressed as a proportion of total score on the PNQ (Augustin et al. 2008).

Construct validation of the instrument was carried out on a group of German dermatologic patients, with various diagnoses including hyperhidrosis. The PNQ was sensitive to differences in the treatment needs of patients with different skin conditions, Herpes Zoster, Chronic hand and foot

Eczema. On the other hand, the PBI demonstrated sensitivity to therapeutic benefit from treatment (t-test between PBQ scores from different time periods showed significant differences – indicating the effect of treatment). The PBI has shown adequate internal consistency. Disease-specific versions of the PBI have been validated in Acne, Vitiligo and Pruritus, Chronic hand eczema (Augustin et al. 2008; Augustin et al. 2009; Blome et al. 2009a; Blome et al. 2009b). One study has used the PBI in an observational study in hyperhidrosis (Muller and Augustin 2013).

The PBI has some floor effects - a disproportionately high frequency of zeros, which the authors attribute to "the proportion of patients in an early stage of treatment" in the validation sample (Augustin et al.). Similar problems have persisted in the other modifications of the PBI. Differential Item Functioning with regard to culture and demographic factors has not yet been assessed. Furthermore, language translation or cultural adaption studies were not found.

Disease-Specific Quality of Life Instruments

Hyperhidrosis impact questionnaire (HHIQ)

The Hyperhidrosis Impact Questionnaire (HHIQ) was designed for evaluating the effects of focal hyperhidrosis and its treatment on the daily life of patients (Teale et al. 2002). Its content development was based on the input of patients and physician-specialists from the UK and Germany, ensuring the cultural equivalence of concepts (Teale et al. 2002). The HHIQ consists of 41 items administered at baseline and 10 items for subsequent follow-up visits (applicable to longitudinal studies). The baseline module includes items on disease and treatment background and the impact of the disease on medical and non-medical resource utilisation. Although Teale et al. (Teale et al. 2002) state that convergent and discriminant validity of the HHIQ was established based on SF-12 and DLQI in hyperhidrosis patients seeking treatment (n = 345) and a matched control group (n= 154), the magnitude of the correlations were not reported. This was also noted for reliability and responsiveness.

Although a number of clinical trials (Naumann et al. 2002; Naumann et al. 2003; Solish et al. 2005) have shown change in HHIQ scores after treatment, the appropriate methods for assessing responsiveness have not been applied. Score change was not linked to criterion to ensure that the patient's condition had indeed changed. Although the studies collected other information on the

patients, such as gravimetric sweating measures, scores on DLQI and SF-12, correlations or cross-tabulation was not calculated, needed for establishing convergence validity.

No alternative versions of the instrument e.g. an electronic or interview delivered format have been published. Based on the parallel development of the instrument in Germany and in the UK it is assumed that a German and English version are available. Lack of brevity in this instrument may restrict its use in routine clinical practice. No data facilitating the interpretation of scores of the HHIQ has been published. No evidence on how to summarise results for example into subscale totals is available. The published studies using HHIQ have interpreted their findings at the item level.

Hyperhidrosis disease severity scale (HDSS)

The *Hyperhidrosis Disease Severity Scale* (HDSS) is a patient centred scale for assessing the disease severity and daily life impairment in daily activities, caused by hyperhidrosis. Its development involved patients with hyperhidrosis attending Canadian and U.S. clinics as well as a population of U.S. households. The HDSS has a single item, asking the patient to rate the severity of their hyperhidrosis on a 4-point scale, from 1, "*My sweating is never noticeable and never interferes with my daily activities*", to 4, "*My sweating is intolerable and always interferes with my daily activities*" (Solish et al. 2005). This makes it very practical in clinical practice, with the Canadian treatment guidelines recommending its application as a diagnostic tool for determining disease severity, for hyperhidrosis in a number of studies, including the Canadian clinical guidelines for hyperhidrosis (Solish et al. 2007).

Adequate validity, reliability and responsiveness has been demonstrated for this measure (Kowalski et al.). HDSS score corresponded to the levels of limitations in daily life activities in the individuals surveyed by Strutton et al. (Strutton et al. 2004). Changes in the HDSS score corresponded to: changes in gravimetric measurement (amount of sweat produced) (Lowe et al. 2004); changes in scores for the HHIQ and the DLQI (Solish et al. 2005). The changes observed before and after treatment on HDSS score compared with that observed on the other instruments. Translation and cultural equivalence studies of the HDSS were not found, although a study published Spanish (Baez et al. 2007) was found. Additionally, although Kowalski et al. (Kowalski et al. 2004) mention that the instrument can be administered either via self-completion or interview, evidence comparing the two methods was not found. The HDSS has been used in a variety of settings as a primary efficacy outcome measure in clinical trials (Lowe et al. 2004; Solish

et al. 2005; Flanagan et al. 2008) as a patient selection criteria in clinical research or trials (Kowalski et al. 2004) and in epidemiological studies (Strutton et al. 2004; Connor et al. 2006).

Hyperhidrosis scale

The Hyperhidrosis Scale was developed for assessing the severity and quality of life impacts of palmar and plantar hyperhidrosis in particular physical symptoms and social impairment in evaluating the effectiveness of surgical treatment (Neumayer et al. 2005). Its development involved patients being considered for surgical treatment (Keller et al. 2009). The instrument's items are scored on a 10 point Likert scale, from no distress (0) to most distress (10), and are organised under three domains. The first two domains cover distress related to palmar and plantar hyperhidrosis; the third domain covers hyperhidrosis of the axillar and the rest of the body. The authors of the measure proposed calculating an average of the item scores as a means of summarising measurement. In practice, item scores have been simply summed up to create a total score (Neumayer et al. 2004; Neumayer et al. 2005). No evidence has been published supporting the structure of the instrument.

Further validation showed strong association with gravimetric of sweating (Keller et al.). A normalised summary score generated by dividing the sum of the item scores with the number of completed items has been suggested. A cut-off score has been determined based on ROC, supporting interpretation of scores (Keller et al. 2009). The suggested structure of the instrument has not been tested. Furthermore, no study has reported on DIF across different socioeconomic and demographic. The current response scaling for the items may affect practicality of the measure, 10 response categories may unnecessarily overburden respondents.

Hyperhidrosis questionnaire

The Hyperhidrosis questionnaire (HQ) was developed not only as a tool for evaluating QoL in hyperhidrosis but also as an educational tool in routine clinic (Kuo et al. 2004). Its development involved hyperhidrosis outpatients (n=85) awaiting thoracic surgery at a regional teaching hospital in Taiwan. Content validation was done by a group of experts (CVI = 0.70) following the generation of the original items. The instrument has a total of 34 items covering five domains (functional, psychological, social, affective and physical function). This structure has been

confirmed by a factor analysis. Adequate internal consistency was established based on optimal Cronbach alpha values for the instrument's five domains. The authors indicate that completion takes between 8 – 10 minutes. Construct validity has been assessed using factor analysis and inter-item/item-subscale correlations (Kuo et al.). The extracted factors explained 68.9% of variance in total scores. Convergence validity, test-retest reliability and responsiveness have not been assessed. No study applying the instrument was found.

'Amir' quality of life questionnaire

Amir et al. (Amir et al. 2000) published findings on the early development of a questionnaire for evaluating the impacts of hyperhidrosis on patient's quality of life. The instrument contains 35 items covering five domains (functional, social, Inter-personal, emotional, condition) scored on a 7-point Likert scale from strongly agree to strongly disagree. No frame of reference has been provided. The content of the measure is based on in depth interviews with 10 patients, subsequent construct validation is based on 48 patients. Construct validity was demonstrated by evaluating the relationship between subjective suffering and QoL using stepwise regression analysis; and testing for expected relationship between gender and disease onset and the QoL impairment.

This instrument remains of limited usefulness. First, as development involved patients awaiting surgery at a dermatology clinic in Israel, the attributes established may be applicable to patients with a severe form of hyperhidrosis. Second, lack of a specified timeframe to reference responses means that this instrument can be applied in evaluative the impact of treatment or monitoring change over time. Moreover, further validation studies are needed to demonstrate the test-retest reliability and responsiveness of the instrument.

Hyperhidrosis quality of life questionnaire

The Hyperhidrosis Quality of life Questionnaire (HHQLQ) was developed for assessing quality of life following surgical therapy (de Campos, 2003a). The HHQLQ consists of 20 items covering 4 domains (functional/social, personal, emotional self/others, under special circumstances) with responses based on a 5-point Likert scale, from highest quality of life (1) to lowest quality of life (5). Patients rank their QoL when undertaking a number of activities (or in various contexts) that may be influenced by or that may lead to sweating (Ambrogi et al.). A total score is obtained by adding the item scores, thus the worst/lowest score is 20, while the best quality of life is reflected

by a score of 100. Two additional general items seek for a general evaluation of quality of life before and 30 days following surgery.

Construct validity is supported by a number of studies. In one prospective open label comparative study, (Ambrogi et al. 2009) found significantly greater improvement in disease severity and in QoL among palmar hyperhidrosis patients receiving surgery (n=86) from the sixth month following surgery ($p = 0.007$) and on subsequent follow-up visits ($p < 0.001$) relative to the improvement in BTX-A group (N = 68). The HHQLQ was used in parallel with clinical (Minor Iodine starch test and Pad glove test) and quality of life measures (DLQI, SF-36 and the Nottingham Health Profile). Reliability has not been reported. Although observational studies using the HHQLQ have reported statistically significant change in the measure's scores, this may seem insufficient as an assessment of the responsiveness.

Studies applying the instrument published in Portuguese and Italian were found, although the formal translations of the measure into these languages were not found. This instrument may have minimal usefulness in patients with mild hyperhidrosis, as the purpose of its development, evaluating outcomes after surgical treatment, reflect patients with severe hyperhidrosis.

Freiburg life quality assessment (FLQA)

The Freiburg Life Quality Assessment is a measure of dermatologic quality of life. It is comprised of a core set of items applicable to all dermatologic diseases and additional items specific for various dermatologic diseases, making up the various disease specific modules of the instrument e.g. for allergies, chronic skin diseases, leg ulcer and wounds (Augustin et al. 2000; Augustin et al. 2004a; Augustin et al. 2010). The hyperhidrosis module has 46 items under six domains and an additional 4 VAS, with each item assessing the frequency with which patients experienced various restrictions within the previous week. Responses are given on a 5-point scale. The FLQA-hyperhidrosis has an additional 4 generic items on general health, overall QOL and disease severity scored as VAS. The hyperhidrosis module can be scored in similar way as the wound module (Augustin, 2010, personal communication). A summary score can be obtained both at the scale and subscale levels, by calculating the item mean scores for subscales and for total scale, respectively (Augustin et al. 2010). Nonetheless, the proposed structure is not supported by evidence.

The validity, reliability and responsiveness of the FLQA-hyperhidrosis module is yet to be established. Such information, on the other hand has been published for the core module (Augustin et al. 2004a). Subscales of the module show moderate correlation with comparable scales of DLQI and ALLTAG questionnaires. All sub-scales of the core module except the 'treatment subscale' showed optimal internal consistency ($\alpha > 0.75$). Test-retest reliability has been reported for the chronic disease module (Augustin et al. 2000). The core module was able to capture an improvement in all domains of quality of life, in psoriasis patients following treatment, reflecting responsiveness. A scale-banding system has been developed using distribution based techniques to facilitate interpretation of the core module scores (Augustin et al. 2004a). The modular design of the FLQA in general and more specifically the number of items in the FLQA hints limited applicability in routine clinical practice.

Clinical Application of Measures: HRQoL Impacts Of Hyperhidrosis

Six studies investigating generic HRQOL in patients with HH using the SF-36 were found, only four reported pre-treatment (baseline) scores (Table 1.5). There was consensus among the studies on the effects of hyperhidrosis on SF-36's vitality and mental health domains as well as on the absence of any impacts on physical functioning domain (Sayeed et al. 1998; Young et al. 2003; Schmidt et al. 2006). There was a lack of agreement on impacts associated with Bodily Pain and Social Function domains. The differences in the results may be partially attributable to a relatively small size of the sample of the study by Sayeed et al. (1998). Naumann et al. (Naumann et al. 2002) and Hamm et al. employed the SF-12 in evaluating QOL in HH patients. Naumann (Naumann et al. 2002) is a multi-centre RCT of the impact of treatment with BTX-A on QOL in patients with bilateral primary axillary hyperhidrosis (N = 320). The pre-treatment were 52.5, for the PCS, 47.8 for MCS scores. On the other hand, Hamm et al. (Hamm et al. 2006) used a cross-sectional design to compare hyperhidrosis patients with controls without hyperhidrosis

Table 1.3: Descriptive properties of instruments used in measuring QoL in hyperhidrosis

Questionnaire	Target Population	Concept assessed	No. of Scales	No. of Items	Response options	Range of Scores
HHIQ	Primary focal HH	daily life impairment	-	41- baseline; 10- follow up	Varies	-
HDSS	Primary axillary HH	subjective disease severity; daily life impairment	1	1	Likert type; 4	1-4
HS	ETS* treated palmar- plantar HH patients	Physical Symptoms & Social impairment	1	15	Likert type; 10	0 -150
HQ	Surgically treated HH patients	disease specific HRQoL	5	34	Likert type; 5	34 -170
HQLQ	Outpatients awaiting for surgery	disease Specific HRQoL – daily life impairment	-	20	Likert type; 5	20 - 100
FLQA	Dermatology patients	Dermatology - HRQoL	6	46 plus 3 VAS	5 plus	NA
DLQI	Dermatology patients	Dermatology- HRQoL	-	10	Likert type; 4	0 -30
Skindex-29	Dermatology patients	Dermatology- HRQoL	3	30	Likert type; 5	0 - 100
PBI	Dermatologic patients incl. HH	Therapeutic benefit	5	23	Likert type; 5	0-4
SF-36	General population	HR-QoL	8	36	Varies	0 - 100
SF-12	General population	HR-QoL	8	12	Varies	0 -100
NHP	General population	HR-QoL	6	38	2	0-100
IIRS	Chronic illness patients	disease severity (Illness intrusiveness)	3	13	7	13 -91

* ETS – Endoscopic thoracic sympathectomy

Table 1.4: Psychometric properties of instruments

Questionnaire	Content Validity	Construct Validity	Convergent Validity	Internal Consistency	Test-Retest Reliability	Responsive-ness	Floor & Ceiling effects	Inter - pretability	MCID	Respondent Burden	Structure
Disease Specific HRQol instruments											
HHIQ[1,2,3]	++	++	++	++	0	0	++	0	0	0	0
HDSS[4,5,6,7]	++	++	++	na	++	++	0	+	+	++	na
HS[8,9]	++	++	+	++	0	++	0	+	0	0	na
HQ[10]	++	+	0	++	0	0	0	0	0	++	+
HQLQ[11,12,13]	-	-	0	0	0	-	0	0	0	0	0
AMIR[14]	++	+	0	++	0	0	-	0	0	0	+
Dermatology HRQol Specific Questionnaire											
FLQA[15,16]	++	++	+	+	++	++	+	++	0	0	0
DLQI[17,18,19]	++	++	++	++	+	++	+	++	++	++	-
Skindex[20,21,22,23,24,25]	++	++	++	++	++	++	++	++	0	+	+
PBI[26,27,28]	++	++	+	++	+	+	+	+	0	0	+
Generic HRQol Questionnaire											
SF-36 [29, 30]	++	++	++	++	++	++	+	+	++	+	+
SF-12[30]	++	++	++	++	++	+	++	+		++	+
NHP[24, 30]	++	++	++	+	++	+	+	0	0	++	+
IIRS[31,32,33]	+	++	++	++	++	++	0	0	0	0	+

Note: Although Teale, Roberts et al. 2002 claim that the HHIQ has favourable internal consistency, test-retest reliability, construct and convergent validity, the relevant correlations were not reported.

References: (1) Teale, Roberts et al. (2002); (2) Naumann, Hamm et al. (2002); (3) Jonathan, Nina et al. (2004); (4) Lowe, Campanati et al. (2004); (5) David, Jonathan et al. (2004); (6) Solish, Benohanian et al. (2005); (7) Solish, Bertucci et al. (2007); (8) Keller, Bello et al. (2009); (9) Keller, Sekons et al. (2001); (10) Kuo et al. (2004); (11) Panhofer, Zacherl et al. (2006); (12) de Campos, Kauffman et al. (2003); (13) Ambrogi, Campione et al. (2009); (14) Amir et al. (2000); (15) Augustin, Zschocke et al. (2000); (16) Augustin, Lange et al. (2004); (17) Finlay and Khan (1994); (18) Basra, Fenech et al. (2008); (19) Kowalski (2007); (20) Chren, Lasek et al. (1996);(21) Augustin, Wenninger et al. (2004); (22) Abeni, Picardi et al. (2002); (23) Nijsten, Sampogna et al. (2006); (24) Both, Essink-Bot et al. (2007); (25) Prinsen, Lindeboom et al. (2010); (26) Augustin, Radtke et al. (2009); (27) Augustin, Reich et al. (2008); (28) Blome, Augustin et al. (2011); (29) Both, Essink-Bot et al. (2007); (30) McDowell (2006); (31) Devins, Binik et al. (1983); (32) Cinà and Clase (1999); (33) Bieling, Rowa et al. (2001).

Patients achieved a significantly lower mean score than non-patients (patients vs. controls: MCS, 44.4 vs. 50.8, $p < 0.01$; PCS, 52.9 vs. 54.9, $p < 0.01$). Such differences would be considered to be of clinical significance considering MCID value of 1.3 and 2.3 for the PCS and MCS scores, respectively (Bennett et al.2003). The two studies reflect a higher QoL impact in aspects related to patient's mental health relative to the mental component of HRQoL, echoing results from studies using the SF-36. The MCS in patients with HH reported by Hamm et al. is lower than that seen in patients with psoriasis [45.6±11.4], dermatitis [47.1 ±10.7], Acne [45.1± 11.3], implying that the impairment resulting from HH is much greater or comparable to these conditions. On the other hand, the level of impairment in physical aspects of QoL (based on PCS) seems less than in psoriasis [48.1± 9.2], dermatitis [49.6±7.9] but slightly worse than that in Acne [55 ± 5.4] (Tabolli et al. 2011).

Cina and Clase (Cinà and Clase 1999) investigated illness intrusiveness in a population of hyperhidrosis patients from an email discussion board (N = 84) following a cross-sectional design (Table 1.6). The total score obtained (45±18) was less than IIRS scores obtained by Cina et al. (Cina et al. 2006) in a prospective observational study in HH patients awaiting ETS surgery (N = 30) (IIRS total score = 57±14). Cina et al. (Cina et al. 2007) compared their earlier findings from the sample of patient awaiting ETS surgery with a control group (N = 13). Scores on the control group indicated the absence of intrusiveness (13.5±17). HH patients took more showers and baths, changed their clothes more often, were limited in the type of wardrobe that they could use and sweated more with the consumption of alcohol, spicy foods and caffeine containing foods (Cina et al. 2007). The greatest intrusion was reported in relation to work, social relations, relations with spouse and in recreational activities (Cina et al. 2006). The level of disruption in patient's life in hyperhidrosis seems worse than in other chronic conditions such as ulcerative colitis [IIRS total score = 27.6 ±16.62], renal transplant [38.7±18.42], multiple sclerosis [44.8±18.59], schizophrenia [50.5±16.68]. The lower level of disruption observed in the study by Cina and Clase (1999) might be alluded to differences in their respective patient population, the former included patients with mild forms of disruption as well as those not seeking for treatment. On the other hand, the later represents a highly selected group.

Table 1.5: Studies using generic HRQoL instruments: SF-36 and NHP

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After	Remarks
Lee et al. 2012	Korea		patients with HH involving hands, feet, axillar, head and neck, perineum, and other	Retrospective, case-series	36	M = 41.6%	27± 14.9	NR	NR	Highly selective patients, treated with glycopyrrolate
Schmidt et al. 2006	Germany	Hospital	Patients with HH (palmar, axillary, facial) treated with ETS	Open-label, observational	178	M = 28.7 %	32.9 +/- 9.7	NR	NR	Sample is highly selective, includes patients previously receiving surgical intervention
Elia et al. 2005	Italy	University hospital	Patients with severe palmar HH	case-series	45	Male = 42%	28.76± 5.25	General, 54.6; PF,38.7; RP,70.0; SF,64.4; RE,55.1; MH,57.4; BP,35.5; GH,48.9; VT,68.9; MCS,46.5; PCS,38.3	MCS: 55.2 PCS: 43.3	Severe cases of HH, where previous treatments had failed Follow-up after 6 months
Elia et al. 2005* [same study as above]								GE, 23.04; MO, 12.09; EN, 30.83; SL, 10.98; PA, 47.84; IS, 1.58; EM, 26.7.	GE,14.12; MO,8.52; EN,16.02; SL,5.71; PA,29.98; IS,0; EM,10.15	

Notes: 1. *Study used Nottingham Health Profile (NHP);

2. Domains of the SF-36: General, general health perception; PF, physical functioning; RP, Role physical; SF, 64.4; RE, role emotional; MH, mental health; BP, Bodily Pain; GH, general health; VT, vitality; MCS, mental component summary score; PCS, physical component summary score.

4. Domains of the NHP: MO, mobility; PA, pain; EN, energy; SL, sleep; EM, emotional reaction; IS, social isolation; GE, general score.

Table 1.5 (continued)

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After	Remarks
Kumagai et al.	Japan	University hospital	3 patients (7%) palmar only; 14 (33%), of craniofacial HH; 35 (83%), axillar; 28 (67%), planter; and 17 (40%), blushing	Case series; follow-up at 1, 3 and 6 months.	40	M=42%	*Mean (range)* 29. 7 (18 - 55)	NR	NR	Involved treatment with ETS, reflecting severe cases of HH. Patients experienced reduced impairment except for bodily pain and physical functioning
Sayeed et al.	UK	University hospital	Patient with palmar and axillary HH, electing ETS surgery	Retrospective, case-series; Patient follow-up at 6.2 months (range: 5.1 - 9.9)	16	M = 45%	Median (range) 26 (18 - 48)	PF, RF, SF, RE, BP: 100 MH: 78, VT: 70, GH: 82 **MCS: 52.7** **PCS: 59.4**	PF, RF, SF, RE, BP: 100 MH: 82, VT: 75, GH: 84 **MCS: 54.4** **PCS: 59.3**	SF-36 not sensitive to impairment in patients QoL, most patients achieved high/maximum scores in 4 dimensions of the SF-36.
Young et al. 2003	Ireland	Hospital	Patients with palmar HH receiving ETS	Retrospective, case-series	62	M = 34%	Mean (range) 29 (17 - 64)	PF: 950, RP: 300, RE: 300, MH/E: 320, VT/EN: 200, SF: 220, BP: 100, GH: 350, Overall: 2635	PF: 950, RP: 400, RE: 300, MH/E: 380, VT/EN: 200, SF: 200, BP: 100, GH: 325, Overall: 2835	Highly selective patient population; severe HH. Mean follow-up was 38.46 months

Table 1.5 (continued)

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After
Hamm et al. 2006	Germany	Hospital	patients with palmar/axillary HH and non-patients	cross-sectional study	Controls: N =154 axillary: N = 165 palmar: N = 116	Control: M = 43% axillary: M = 42% palmar: M = 421%	Control: 27+/-7 Axillar 32 +/-12 Palmar 30+/-9	Mean score MCS: 44.4, patients, 50.8, controls PCS: 52.9, patients, 54.9, controls	
Naumman et al. 2002	Multiple European centres, plus DE and UK	hospital/clinic	Patients with persistent HH	Multi-centre RCT	320	M = 46%	Mean age(range) 31.5 (17 - 74)	PCS = 52.6 (No Diff. Between placebo and treatment group) MCS = 49.1, BTX-A; 46.4, placebo	PCS Change BTX-A: - 0.9, Placebo: - 1.2 MCS Change: BTX-A: - 1.7, Placebo: 0.5

Table 1.6: Effects of hyperhidrosis on lifestyle: IIRS

Author	Country	Setting	Patient population	Study design	n	Gender	Age (years)	Before	After	Remarks
Cina et al. 2007	Canada	Hospital	primary HH patients electing for surgical treatment	Case-series/open label	30 patients 11 non-patients	M = 50% (patients) M= 36% (controls)	*Mean* 39± 13 (patients) 26±10 (non-patients)	57 ± 14 (patients) 13.5±0.7 (non-patients)	19.3 ± 15 (patients)	Number of controls small. Concept becomes hypothetical if presented to controls.
Cina et al. 1999		Community based; online HH discussion group	Community patient population; online	Cross-sectional; Observational, exploratory	80	M = 65 %	*Mean* 32± 9 (patients)	45 ± 18		Reflects patients in community Might have included those not suffering from HH
Cina et al. 2006	Canada		primary HH patients electing for surgical treatment	Prospective multi-centre cohort design	22			*Mean Scores* 57 ± 14		Follow up took place at 2 and at 4 months.

The authors also admit that caution is required in using their results, especially considering the risk of including patients who may not necessarily have had hyperhidrosis, based on their sampling strategy. On the other hand, exaggeration of the level of reported pre-treatment impairment may not be ruled out in the Cina et al. (2006) study

In the fifteen studies using the DLQI, scores ranged from 10 – 14 for axillary hyperhidrosis, 8.8 – 15, for palmar hyperhidrosis, 13, for cranio-facial hyperhidrosis, 9.4 for sweating of the trunk (Table 1.7). Hamm et al. (2006) compared dermatology-specific-QoL in hyperhidrosis patients to that in controls. The DLQI total score was lower in non-patients than patients (DLQI scores: axillary HH, 10 ±5.6; palmar HH, 8.8±5.9; controls, 0.7±2). In patients with axillary HH, HRQoL effects were mainly linked to limitations on daily life activities (influence on choice of clothes), symptoms and feelings (embarrassment and self-consciousness) and effects on leisure and social activities activities (Swartling et al. 2001; Hamm et al. 2006) In palmar HH these were related to symptoms and feelings, daily life activities, leisure and social activities and personal relationships (Swartling et al. 2001; Hamm et al. 2006). Amini et al. Amini et al. (2008) in a retrospective, exploratory study including patients (N = 94) receiving treatment at a dermatology clinic in the Netherlands, obtained the highest baseline DLQI scores in patients with hh involving the 'axillae and face' (15±5.62). The lowest scores were in patients with palmar and/or plantar hyperhidrosis (9.24±5.08). The baseline DLQI score reported by Müller et al.(Muller et al. 2012) is notably higher than that seen in rest of the studies (DLQI Score = 16.5). This may reflect the inclusion/exclusion criterion of the study, being an RCT.

Four studies reported using the Hyperhidrosis Impact Questionnaire (HHIQ) including Naumann et al. (2002), Solish et al. (2005) and Hamm et al. (2006) and Strutton et al. (2004) (Table 1.8). Patients (34% - 79%) experienced limitations at work, resulting in reduced effectiveness, thus patients changed how they worked. The majority of the patients (64.7 % - 86%) were moderately/severely affected emotionally. For example, all four studies reported a majority of patients (70% - 94%) having less confidence than they would like to have; Thirty six to seventy-one percent reported being depressed or unhappy. Additional challenges were experienced in personal relationships and social situations. For example, 59% - 70% of patients reported difficulties with meeting new people for the first time; 25 % - 79% reported an inability to participate in family events or to spend to with friend. Various social situations presented

challenges to patients: meeting people for the first time was a concern in 47% - 90% of the patients; being in public places and shaking hands with people also reported to be challenging.

The condition also affected patients' daily life activities, with 30.4% - 61% of the patients getting frustrated with every day chores. To mitigate such effects, patients purchased additional items or accessories to help in completing routine tasks or by seeking help from family and friends. Solish et al. and Naumann et al. represent highly selective patient populations owing to their inclusion/exclusion despite including patients from multiple centres. A similar critique holds for the Hamm et al. study albeit to a smaller degree, in this case, while the patient population represents HH patients seeking care at a dermatology clinic, those with more severe condition may have been overrepresented. While the study by Strutton et al. is by and large free of the issues noted in the above studies, still some limitations can be noted. The reliance on one member to report on an entire household poses a risk on accuracy of reported information, as second hand experience may differ from first hand experiences. In particular, QoL impacts tend to differ depending on the source of such information, whether family member or patient is the one reporting.

DISCUSSION

The quest to understand impacts of disease and its treatment as experienced and reported by patients has become ubiquitous in health-care, with Gill and Feinstein (1994) describing the change in the field as *growth from a small cottage industry to a large academic enterprise*. The drive towards a highly transparent drug regulatory regime and stricter risk-benefit assessment for new pharmaceutical products has provided a platform for consideration of broader and more comprehensive set of outcomes beyond clinical and pharmacologic endpoints in clinical research. Furthermore, the fact that patients may be more interested in how treatment of their condition affects their daily life or what sacrifices they may have to make in terms of their QOL, than just the alleviation of symptoms (Lohr and Zebrack 2009) has enhanced the relevance and importance of PROs such as HRQOL in routine clinical practice. Given that patients are experts in their experience with a condition, they are the best source of information on their HRQOL (Salek 1998). This, the subjective nature of patient experiences and idiosyncratic perception on the same, lies at the heart of the conceptual issues related to HRQOL.

Table 1.7: Studies using a dermatology-specific QoL Instrument: Dermatology Quality of Life Index (DLQI)

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After	Remarks
Rosell et al. 2013	Sweden	specialist clinic	patients attending a specialist clinic	Open-label, observational study; comparison of BTX-A and BTX-B	58, axillary 26, palmar	NR	*Mean+/-SD* 32.3 +/-10.4, axillary 26 +/-10.2, palmar	*mean +/- SD* 12+/- 5.5, axillae: 10.3 +/7.3, palmar	axillae: 1.7+/-2.6 palmar: 1.2+/-1.5	atients represent severe cases of HH
Amini et al. 2008	Netherlands	Dermatology clinic	HH patients attending clinic	Retrospective, observational	94	M = 34 %	32.6	9.24 ±5.08, hands and/or feet 10.98±4.51, axillar 12.91±2.95, axillar and/or feet 15.75±5.62, axillar and face 12.27±6.76, generalised		Patients and range of treatment represent typical for dermatology clinic
Tupker et al. 2006	USA	Dermatology clinic	patients with generalised HH treated with oxybutynin	Prospective, observational	13	NR	NR	15.9±6.9	3.7±5.2	Patient characteristics were not reported in study
Solish et al. 2005	Canada	Dermatology clinic	Patient with axillary HH treated with BTX-A	Multicentre (N = 30), prospective, open label	146	M = 33%	35 (range: 18 - 73)	10.6	1.7 DLQI decrease > 5, 76% of patients DLQI = 0, 53% of patients	Very stringent inclusion criterion; patients reflect severe cases (HDSS = 4, for 64% of patients)

Table 1.7 (continued)

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After	Comment
Tan et al. 2002	Canada	Hospital	patients with focal HH awating treatment with BTX-A	open-label, retrospective observational	22, axillary 2, forehead 10, palms	M = 38%	mean (range): 29.6 (17 - 56)	**Mean score** 18, axillae 13, forehead 18.5, palmar	**Mean score** Axillae: 4 Forehead: 2.5 Palms: 9	The DLQI was modified for greater relevance for sweating. Highly selective patient population, attending one practice. Small number of patients with forehead HH
Müller et al. 2012.	Germany	Specialist practices & Hospital	Patients with axillary or palmar and axillary HH	multicentre RCT	339, randomised 267, analysed *sites* 267, axillar 217, palmar	M = 42%	range: 18 - 66	16.6 (methantheline) 16.4 (placebo)	*day 14* 11+/-6.4, treatment group 13.2, +/- 6.6, placebo group *day 28* 9.7 +/-6.8, Treatment group 12.2 +/-6.8, placebo	At least three quarters of the patients were severely affected by their condition. RCT involved treatment with Methantheline (Vagantin) over 28 days.
Swartling et al. 2001	Sweden	Hospital	Patients attending Neurology and dermatology depts. of Uni hospitals treated with BTX-A	Prospective open-label study	58 Palmar: 46 Plantar: 31 Axillary: 30	M = 45%	Range: 15-49	10.3 (2 - 22), all patients 10.6, relapse-patients 9.9, relapse-free 9.1, palmar, relapse free 11.6, axillary, relapse free	4.3, all patients 8.8, relapse-patients; 2.4, relapse-free 1.8, palmar relapse free 2.4, axillary, relapse free	2 -15 months follow-up Patient population reflects severe HH,

Table 1.7 (continued)

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After	Comment
Campanati et al. 2003	Italy	Hospital	patients with HH treated with BTX-A	open label, observational study	41 Axillae: 14 Palms: 16 Axillae and Palms: 11	M (%) 29.3%, palmar 36.8 %, axillary	*median (IQR)* 32.5 (24 -43), axillae 26 (22-5 - 41), palms 25 (20 - 35), palms&axillae	*median (IQR)* 13 (11 - 15), axillae 13 (12 - 17), palms 14 (13 - 17), palms and axillae	0.5 (0 -1), axillae 1 (1 - 3), palms 1 (1 - 4), palms and axillae	patients reflect severe cases of HH, not responding to other treatments
Campanati et al. 2011	Italy	Hospital	plantar hyperhidrosis not responsive to other forms of treatment	Open-label, observational study	79 Palmar: 41 Axillary: 38	M = 32 %	median (IQR) 27 (25 - 34), axillae 29 (27 - 40), palmar	14 (11 - 17), axillae 15 (12 - 18), palmar	1 (0 - 1), axillae and palmar	patients reflect severe cases of HH, not responding to other treatments
Harper et al. 2010	UK	Hospital	patients with HH treated with BTX-A	Open label, service audit	37	NR	NR	Mean 12.9	2.5	includes severe cases of HH
Kim, Kil et al. 2010	S. Korea	Hospital	patients with CS in the trunk	Retrospective design	17	M = 55%	*Mean +/- SD* 26.3 +/- 4.9	*Mean +/- SD* 9.4+/-2.0	2.8+/-1.0	R highly selective patient group: previous surgery/ severe CH
Bechara et al. 2007	Germany	Specialist dermatology (hyperhidrosis) clinic	patients with axillary HH treated with suction curettage	Open-label, observational study	51	M = 37.3%	*mean+/-SD (range)* 28.6+/-10.6 (19 -48)	*Median score(range)* 12 (9 - 18)	4 (2 - 8)	Patient were severe cases of hyperhidrosis 9 months follow-up period
Innocenzi et al. 2005				Open-label, exploratory	20			NR	NR	
Lupin et al. 2012	Canada	hospital/clinic	patients with axillary HH treated with microwave-based device	Multi-centre, open-label study	31	M = 26%	*Mean (range)* 33 (18 - 65)	*Mean* 11.8	1.6	highly selective patient population

Table 1.7 (continued)

Author	Country	Setting	Patient population	Study design	N	Gender	Age (years)	Before	After	Comment
Hamm et al. 2006	Germany	Hospital	patients with palmar/axillary HH and non-patients	cross-sectional study	controls: N =154 axillary: N = 165 palmar: N = 116	Control: M = 43% axillary: M = 42% palmar: M = 421%	Control: 27+/-7 Axillar 32 +/-12 Palmar 30+/-9	Control: 0.7 +/-2 Axillary: 10 +/-5.6 Palmar 8.8 +/-5.9		Patients seeking for dermatology treatment at a University clinic

Table 1.8: Studies using the Hyperhidrosis Impact Questionnaire (HHIQ)

Author	Hamm et al. 2006	Strutton et al. 2004	Naumann et al. 2002	Solish et al. 2005
Number of patients (N)	345, patients 154, non-patients	150,000 households	320	146
Country	Germany	USA	Germany, UK & other	Canada
Patient population	Single centre, dermatology clinic (Uni. Hospital)	community based patient population		Multicentre, dermatology clinics
Setting, study design	Single centre, dermatology clinic	National survey	International Multicenter RCT	Multicentre, open label, dermatology clinic
Impact on career choices and work habits				
Moderate/extreme limitation at work	axillary: 65%, palmar: 62%	33.6%		79%
Moderate/extreme impact on effectiveness at work	axillary: 46%, palmar: 38%			
Accomplished less at work	axillary: 33%, palmar: 24%			34%
Made changes on how they worked	axillary: 41%, palmar: 41%			43%
Patients worked less carefull/accurately	axillary: 27%, palmar: 24%			
Overall satisfaction with ability to perform work activities			Btx-A: 20% Placebo: 15%	8%
Time and effort spent treating HH				
Changed clothes at least twice a day	axillary: 71%, palmar: 31%		BTX-A: 76% Placebo: 75%	77%
Shower or bath at least once a day	axillary: 27%, palmar: 10%			
Spend at least 15 min a day symptoms	axillary: 38%, palmar: 22%			30%
spent 15 min or less			BTX-A:62.8% Placebo: 52%	

Currently, a unified standard conceptual definition for HRQOL is not available, as reflected in the multiple definitions provided in this chapter, reflecting the outstanding this. Complicating the situation further, how individuals understand health varies across cultures, age, gender,

socioeconomic status, and may even vary depending on whether one is sick or healthy (Bowling 1995 cited in Haas 1999).

Table 1.8 (continued)

Author	Hamm et al. 2006	Strutton et al. 2004	Naumann et al. 2002	Solish et al. 2005
Emotional				
Feel Unhappy		54.8%		
Feel Depressed		35.7%		
Unhappy or depressed	axillary: 71%, palmar: 54%		48.70%	46%
Less confident than they would like	74%	69.8%	71.80%	94%
Moderately or significantly emotionally injured		64.7%	Btx-A: 86% Placebo: 85%	67%
Personal relationships and social situations				
Moderately or extremely limited in establishing personal relationships	59%			70%
Moderate or extreme limitation with participation in family events or spending time with friends	54%		25%	79%
Moderate or extreme limitation with sexual activities	34%	46%	31.80%	52%
Moderate or extreme limitation in social situations e.g.	severe axillary, 50%	45.8%		
meeting people for the first time	71%	46.7%	BtX-A: 80.5% Placebo: 75%	90%
being in public places	56%	35.1%	65%	84%
shaking hands	58%	31%		50%
changed type of leisure activities	53%	41.6%	44.60%	
decreased leisure time	42%	34.6%	19.30%	
Moderate or extreme limitations in sports		17.2%		
Purchase additional items or accessories to help complete routine daily activities		31.4%		
Become frustrated with some daily life activities		58. 2%	30.40%	61%
Require help (eg from family and friends) to perform tasks that would otherwise do on own		17.8%		

Notes: 1. Results from Strutton et al. 2004 are for patients with severe axillary hyperhidrosis

Moons (Moons et al. 2006) discussed six major conceptual issues in HRQOL. For example, does HRQOL include 'subjective' or 'objective' elements or both? Is HRQOL separate from other concepts such as health-status and functioning? Can its determinants be separated from QOL and its indicators? Can HRQOL be clearly defined and distinguished from overall QOL? These issues

have implications on the measurement of HRQOL, beyond their relevance for conceptual debate. Not only do they complicate the design of appropriate measurement tools, but also present a challenge to the interpretation, reporting and application of study results. Perhaps rather than seeking one theocratic concept to supersede or unite current definitions, the pluralistic nature of QOL should be accepted as *the nature of the beast*. Gill and Feinstein (1994) have suggested a practical way forward to ensure rigorous HRQOL measurement in the current conceptual conundrum. Researchers should be transparent in their use of QOL, by clarifying a definition. The domains being assessed should be made clear. The choice of measurement instruments used in data collection should be justified. But even more importantly, patients ought to be involved, in defining relevant HRQOL issues of study, and how important they are. Undertaking these steps would ensure that HRQOL is appropriately measured, and that such reference should only be done where it is warranted.

Besides the above concerns, choice of the most appropriate instrument for measuring QOL involves multiple-decisions: whether to use a generic, therapeutic area-specific or disease-specific instrument; whether the developmental purpose and patient population for the measure relate to the research question and target population; psychometric properties of the measure (Streiner and Norman 2008). In routine clinical practice additional considerations include, whether the instrument is easy to complete and easy to administer; applicable – consider as relevant by the patients, responsive to individual change, with interpretable scores (Higginson and Carr 2001).

One therapeutic area where QOL considerations have a particular role is in skin disease, in view of the profound impact on the QOL of patients, often exceeding that in various chronic diseases conditions (Finlay 1998; Basra and Shahrukh 2009) Simultaneously, typical for most skin diseases, and hyperhidrosis in particular, easily measurable laboratory values is either scarce or difficult to interpret (Rani et al. 2005) as well as making QOL measurement particularly crucial. A core aim of this chapter was to review the QOL impacts of hyperhidrosis as well as the instruments that have been used in assessing them. Impairment in generic HRQOL was alluded to social functioning and emotional role limitation, whereas dermatology-specific impacts were related to daily activities, personal relationships and symptoms and feelings. The necessity for more showers, sweating after consumption of alcohol, spicy foods or coffee, and facing a limited choice of clothing representing disruptions in patients' lives.

Still, these results do not fully explain why dermatology-specific QOL impairment is worse in hyperhidrosis in comparison to other conditions such as psoriasis as noted by Hamm and colleagues (Hamm et al. 2006). Several explanations are plausible. First, patients with hyperhidrosis report feeling their life as being taken over by hyperhidrosis all the time. This reflects the persistence (and frequency) of sweating episodes and their related impacts. Sweating episodes are accompanied by feelings of anxiety (besides other negative emotions), in a chicken-egg circle. Patients get anxious that they may sweat and in turn more anxiety leads to more sweating. This is consistent with earlier views on the disease, where psychiatric underpinnings were suspected (chicken-egg) (Ruchinskas 2007). This is also consistent with the view of hyperhidrosis as a multi-factorial condition (Beltraminelli and Itin 2008). The greater impact on QOL in this case being alluded to the strong feedback between the psychiatric impacts and sweating.

Unlike other dermatological conditions, hyperhidrosis is poorly treated, with only 35% visiting the doctor (Strutton et al. 2004). Nonetheless, even those seeking treatment face a hard choice between expensive treatments such as Botox, high-risk surgical interventions such as ETS surgery or other much cheaper but less effective treatment. Consequently, the majority of patients survive with an unmanaged condition. In the long term this poses a real risk of patients developing psychological sequelae owning to persistent impact. Clinical trials in hyperhidrosis have tended to use all three types of measures, generic, therapeutic-area specific, and disease specific, and more often than not in combination. On the other hand, clinicians have tended to use more disease specific measures, (Solish et al. 2007). Among the generic measures, only the IIRS had been validated in hyperhidrosis patients (Cinà & Clase 1999). This only adds on their natural limitation, namely, their inclusion of content which may be irrelevant to patients with particular disease (Halioua et al. 2000) threatening content-validity or ‘applicability’. Generic measures may still be of use in hyperhidrosis when making comparisons with the other non-dermatologic conditions, still this would have to be preceded by an evaluation of their psychometric attributes most especially content validity and responsiveness.

Among the skin-specific measures, only the DLQI and the PBI had been validated in hyperhidrosis patients. These measures are overall more relevant for hyperhidrosis patients in comparison to

generic measures. Given the DLQI's brevity, simpler scoring system and availability of cut-offs for clinical significance (in the form of MID) for hyperhidrosis, it would be preferred of the two. Of the disease-specific measures found, the HDSS and the HHIQ were the most validated. Although the HDSS has be included in this review and has also been used in other trials as a measure of QoL, its developers intended it as a measure of disease severity and interference of hyperhidrosis in everyday life. Moreover, as a single item instrument it does provide a detailed picture of QoL On the other hand, while the HHIQ is indeed promising both in terms of coverage issues relevant to patients, it was not designed for use in routine clinical practice; its internal structure has not been tested, it has no scoring system, and no information has been provided for interpretation of scores.

The other disease-specific measures found, had incomplete details regarding key psychometric information. Even more worrisome, is the unusually high number of studies in hyperhidrosis assessing QoL using adhoc QoL questionnaires reflecting that considerations of rigor of measurement were not at all made. Practical recommendation on developing instruments in dermatology from the EADV task force on QoL is to apply both classical test theory and modern test theory (Prinsen et al. 2013). Particular emphasis was placed in testing that measures remain invariant in different patient populations and measuring situations. None of the current disease-specific measures in hyperhidrosis has been tested using modern test theory or demonstrated invariance

This review has revealed a deficit in the current measurement of HRQOL in hyperhidrosis. There is need for a new measure which would assess HRQOL specific to patients with hyperhidrosis, with its content underpinned by patient experiences and quality of life issues they face, with demonstrated optimal psychometric attributes of construct validity, inter-temporal stability, internally consistent, tested internal structure and unidimensional scales. In order to ensure clinical feasibility of such a measure, adequate attention would have to be given to ensuring its practicality, for example having a small number of questions as much as possible to allow all questions to fit on one side of an A4 page, and using a simple scoring procedure.

SUMMARY

- There is a growing recognition of the need to incorporate the patient's perspective in understanding the impacts of disease and its treatment largely driven by a revolution in societal values as well as changes in disease epidemiology and the organisation and delivery of health care services.
- Increasingly, HRQoL is being acknowledged as the ultimate goal of health-care, leading to its growing application as an endpoint in clinical trials, in routine clinical practice, and much recently as a tool for driving system-innovation through its system wide application.
- Controversy surrounds the definition of HRQOL, its constituents and how to measure them and their linkage with other clinical endpoints, presenting challenges on its measurement. HRQOL is subjective, varies over time and is a multidimensional concept
- The choice of the most appropriate instrument has to consider purpose for measuring QoL against those of available instruments. Psychometric attributes and appropriate conceptual coverage should also be considered.
- Hyperhidrosis results in profound HRQoL impairment, covering occupational related impacts, emotional distress, physical discomfort, limitations in social life, the extra activities involved in managing the condition. This is comparable or worse than other common skin conditions such as psoriasis or eczema.
- The measurement of HRQOL in hyperhidrosis has made use of generic, skin-specific and disease-specific measures. While generic measures may not be recommended, the DLQI seems viable. For understanding disease-specific issues, none of the current measures can be considered optimal, although HHIQ seems the best among the lot.
- There is need for a new HRQOL instrument for measuring disease-specific QoL, based on patient's input, with well tested psychometric properties, applicable to all forms of hyperhidrosis, having paid particular attention to ensuring its feasibility in the routine clinical

STUDY AIMS AND OBJECTIVES

AIMS

- To investigate the impact of hyperhidrosis on patients' HRQoL.
- To conceptualise, develop and validate a disease-specific instrument for assessing HRQoL in hyperhidrosis that would be applicable in clinical research as well as routine clinical practice.

OBJECTIVES

- To explore the experiences of patients with hyperhidrosis in order to obtain an in-depth understanding of the extent and nature of QoL impacts
- To create a conceptual framework for HRQoL in hyperhidrosis.
- To develop a disease-specific instrument for evaluating QoL impacts in hyperhidrosis based on the experiences of patients.
- To assess whether the content of the new disease-specific instrument was relevant to patients with hyperhidrosis; adequate and appropriate for measuring the concept of quality of life.
- To establish the dimensional structure of the new instrument and to perform item reduction.
- To assess the reliability and the construct validity of the new instrument.
- To establish the minimum important clinical difference (MCID) value of the new instrument.
- To develop and propose scale banding for the interpretation of scores from the new instrument.
- To assess the feasibility of applying online social networking sites for outcomes research in dermatology (using hyperhidrosis as an example).

CHAPTER 2

Rationale and Methodological Framework

PART I: RATIONALE

The articulation of humanistic aspects in the 1948 constitution of the WHO signalled international consensus around the final outcome of healthcare ought to be: complete physical, mental and social well-being. This new paradigm meant that treatment of disease was not only curative but rather aimed at improving the outcomes that patient cared for most, related to their everyday functioning and quality of life. A broad view of treatment goals is also in keeping with the epidemiologic developments where non-communicable diseases such as diabetes, cardiovascular diseases and cancer are increasingly accounting for a greater portion of burden of disease globally. As most of such conditions tend to be chronic, maintaining or improving the functionality and quality of life of patients is considered a key endpoint of treatment. Furthermore, quality of life facilitates the understanding of the full impacts of disease on the patients, particularly for those elements of the impact which can only be known by them (Patrick et al. 2007). This therefore, elevates the importance of patient reported outcomes and QoL from merely a consequence of treatment to the ultimate goal that should be measured and attained by each treatment.

QoL considerations are of particular importance in skin diseases. First, as most skin conditions are not life-threatening, their major burden is associated with morbidity (Basra and Shahrukh 2009), more so considering that the degree of severity tends to be influenced by psychosocial effects (e.g. effects on self-image and social life) and physical discomfort more than symptoms (Grob 2007). As doctors are not always able to predict the QoL of their patients (Basra and Shahrukh 2009), the measurement of QoL in patients with skin disease becomes even more imperative. In hyperhidrosis, further practical considerations may make the assessment of QoL quintessential. First, the amount of sweating defining the condition is unclear (Wörle et al. 2007). Moreover, the reliability of current methods for quantifying sweat, such as gravimetric measures, has not been assessed. Additionally, these measures are neither feasible nor practical in the busy set up of a routine clinical practice. This places patient reported disease impacts such as effects on QoL as reported by the patient as a particularly useful piece of information in the clinical management of hyperhidrosis.

Addressing the need for QoL measurement in hyperhidrosis requires instruments that are appropriate and fit for purpose. Current measures were reviewed in this regard and the full results have been presented in the general introduction chapter. The following observations were made:

- Generic HRQoL measures such as the SF-36 were seen to include items irrelevant for hyperhidrosis while omitting key issues relevant in this patient population. This also

applies to dermatology specific measures though to a lesser degree. Moreover, except for the Patient Benefit Index (PBI), the other measures have not been validated in this patient population.

- For the current disease-specific measures temporal stability is not known. Evidence of construct validity and responsiveness is not based on robust approaches. For example, only factor analysis exploring factorial structure has been undertaken in one measure only (Hyperhidrosis questionnaire).

Significance of a new instrument in hyperhidrosis

In view of the above, an adequately validated instrument for evaluating hyperhidrosis-QoL that would be feasible and practical in both routine practice and clinical research would enable the efficient measurement of the impacts of hyperhidrosis on the life of patients. This is not only vital in quantifying the burden associated with the condition but would also have implications on the diagnosis and management of the condition for instance leading to better decision making regarding use of systemic treatment in patients with great HRQoL impairment (Grob 2007). Moreover, the availability of an instrument with better acceptability and applicability in routine clinic practice would facilitate more discussions of patients HRQoL in hyperhidrosis during consultation, which constitutes an aspect of care important to patients (Salek et al. 2007). This would also support the understanding of the patients' healthcare needs, leading to more appropriate care being given, which might include counselling or other forms of care as necessary. Furthermore, the improved precision and ability to detect important changes, in the measurement of HRQoL would have practical implications on clinical trials (Streiner and Norman 2008), particularly by reducing the sample sizes needed to achieve desired statistical power, reducing a major hurdle for research in the field.

A major obstacle to the use and acceptance of QoL information emanates from the subjective nature of such information and the lack of credibility of the approaches used in quantifying these (Grob 2007). This is further complicated by the fact that QoL is an unobservable hypothetical construct. These concerns are in part addressed in measurement theories such as classical test theory widely applied in psychology and education, to develop and validate instruments for assessing psychological constructs such as depression and intelligence. The efficiency and level of objectivity of measurement of such constructs has been further enhanced by the development

of modern test theory such as the Rasch model, which permits the translation of raw scores into an interval scale variable (Fisher Jr 2000). Both measurement theories provide a set of attributes to be reflected in an instrument for optimal measurement of constructs. This has been of even greater interest in the context of patient reported outcomes and several groups have made recommendations of the key attributes (Patrick and Chiang 2000; Lohr 2002; Frost et al. 2007b; Terwee et al. 2007).

KEY ATTRIBUTES OF HRQOL INSTRUMENTS

Conceptual Framework

A clearly articulated conceptual framework based on literature and the relevant qualitative data (from patients or therapeutic clinical experts), outlines the concepts being measured by an instrument; the rationale for assessment; domains and their inter-relationships (Lohr 2002; US-F.D.A. 2009). This in turn facilitates the appropriate organisation of the instrument, in line with the intentions of measurement including the scaling structure and the related measurement model (Rothman et al. 2007). The process of developing a conceptual framework not only ensures that the purported constructs are measured but also facilitates the interpretation of the data.

Acceptability And Practicality

Ultimately, the acceptability of an instrument by the final users, the patients, has a bearing on not only whether the respondents are motivated to complete the questionnaire but also the integrity of the data obtained. The instrument must be easy to complete, imposing the least burden on the respondents (Both et al. 2007). For instance, the measure should not be unnecessarily lengthy and must be well organised such that navigating through and completing the questionnaire should be easy. Loss of spontaneity to the response process as the respondents become fatigued may lead to avoidable errors or undesirable response behaviours, for example ‘satisficing’. On the other hand the effort required to administer the instrument, collect and process data must be minimal to make the application of the instrument in a clinical situation feasible (Salek 1998). Again if data collection and processing are excessively burdensome avoidable errors may creep in during the process as the administrators become less careful.

Validity

As HRQoL is unobservable, evidence that an instrument assesses what is intended, demonstrating validity of the measurements, is a key property (Fayers and Machin 2007). Where there is lack of evidence supporting this, inferences may be misleading as there is no certainty regarding what is actually being measured (Haynes et al. 1995). Thus, the validity attribute relates to particular use of the scale and is not an inherent trait of the instrument (Messick 1988). This then means that the validation exercise is a continuing process providing evidence supporting various inferences based on the instrument(Streiner and Norman 2008). Validity is typified in three forms: content validity, construct validity and criterion validity.

Content validity refers to the adequacy with which sampled items of an instrument reflect its aims as articulated in the conceptual framework (Salek 1998). First, this reflects the appropriateness of the items, domains and other elements of the instrument to the underlying construct being measured. This also captures how comprehensively the underlying construct has been covered by the instrument (Patrick et al. 2011b). Determination of content validity is both a quantitative and qualitative process based on judgement by experts (Streiner and Norman 2008). This form of validity is ensured by having a clear conceptual framework, as a basis for the instrument, and following an organised and structured process in the development of the instrument, which should be carefully documented (Terwee et al. 2007). By capturing the connection between intended measurement concept and the way patients from the target population understand and discuss that concept (Patrick et al. 2011b), this form of validity serves as vital proof of concept for later development of the instrument. Quite often content validity is confused with face validity. The latter relates to the acceptance conferred by lay persons that the instrument appears to be sound and relevant (Lynn 1986). The conclusion that the instrument is indeed measuring what it is supposed to be measuring is by perception and thus assumption, there is no rigorous quantification or measurement (Lynn 1986; Frost et al. 2007b).

Construct validity assesses the degree to which theoretical hypothesis relating to an underlying construct being assessed by an instrument are actually supported, providing evidence justifying particular or inferences interpretation of scores (Terwee et al. 2007). Studies for demonstrating construct validity vary in design, although they all have common footing on testing hypothesis

related to the underlying construct. Common study designs for construct validity are presented below. A unified view of validity considers all forms of validity as being subsumed under construct validity based on the argument that construct interpretation undergirds all score based inferences, thus the various forms validity are indeed only supporting that the construct is valid (Messick 1988)

- ***Known groups validity*** involves testing hypothesis relating to group differences in scores of patient, anticipated to differ. Such groups are usually based on some important clinical variable for instance level of disease severity or the localisation of the sweating, in the case of hyperhidrosis.
- ***Convergence and divergent validity*** is based on expected relationships between a scale and other measures assessing a similar construct. A scale is expected to show high correlation (convergence) with other scales assessing similar constructs and; conversely a low correlation (divergence) would be expected with other measures (scales) assessing unrelated constructs (Streiner and Norman 2008).

Criterion validity assesses the extent to which a measure agrees with an external gold-standard measure, how well the new measure is consistent with and captures the essence of the gold-standard (Frost et al. 2007b). This includes situations where the gold-standard is measured at the same time as the new measure, reflecting *concurrent criterion validity*; as well as where the gold standard is only observed at a later date, *predictive criterion validity.* Nonetheless, criterion validity may not be applicable in the context of QoL measurement due to lack of measures of proven validity suitable to be gold-standards (Salek and Luscombe 1992). Otherwise, existence of such gold-standards would obviate any needs to develop a new measure.

Reliability

Reliability refers to the degree to which scores of an instrument reflect the true score (representing the underlying condition): the proportion of total variance in measurement accounted for by true score (underlying latent variable) after measurement error is accounted for (Streiner and Norman 2008). This is also viewed as the signal-to-noise ratio associated with measurement (Guyatt et al. 1993). Reliability relates to both the consistency and reproducibility of scores from an instrument, as seen in the different forms of reliability. This suggest that reliability has an impact on other psychometric attributes such as construct validity and responsiveness

Internal consistency looks at 'homogeneity' among items belonging to a single scale or domain, whether the items are tapping into the underlying construct equally strongly (Fayers and Machin 2007). The assumption made here is that as items in a single scale are meant to be assessing different aspects of the same underlying construct, the items are inter-related through their relationship with this construct. Internal consistency, therefore, captures the proportion of scale score total variance attributable to a common source among the items (DeVellis 2011).
Split-half reliability looks at a scale's consistency by evaluating the inter-correlation between two halves of a scale (Streiner and Norman 2008). The many possibilities for creating the two halves can be challenging, the following common approaches are used: first half-second half split; odd-even items; balanced based on external criteria to divide the scale in different ways may create challenge. Correlating one half of the scale to the other implies that the correlation represents reliability of a single half of the scale, to obtain the reliability of the entire scale the Spearman's - Brown formula has to be applied (DeVellis 2011).

Inter-rater reliability assesses the level of agreement in ratings on subjects made by different judges, following two alternative approaches: 'consistency' or 'absolute agreement'. Inter-rater consistency relates to the amount of proportion that deviates from means as different experts rate an item while absolute agreement constitutes the exact agreement in the ratings made by different judges (Wynd et al. 2003). This form of reliability is of particular importance in observer- or interviewer-administered instruments (Salek 1998).

Temporal stability (test-retest reliability) assesses the reproducibility of scores. It is expected that if the instrument is used in patients whose condition has not changed, the scores obtained on the two assessments should be similar. As the underlying construct will not have changed, the correlation of the two assessments gives the degree to which the measured concept actually determines the observed scores (DeVellis 2011, p.51).

Interpretation of results needs to be qualified by the factors that influence reliability. As total variance is the denominator in the reliability equation, assuming measurement error is held constant, it is possible to increase reliability of a measure simply by increasing total variability. This may be a consequence of increasing the number of items, sample heterogeneity and the

number of response options as suggested by the *Spearman-Brown's prophecy* (Nunnally and Bernstein 1994).

Responsiveness To Change

An important attribute in evaluative instruments, is the capability to capture important changes in the construct (patient's condition) where such changes have taken place (Epstein 2000). As the purpose of such longitudinal use of an instrument is in identifying true change over and above inter-temporal variability in the scores, reliability becomes a prerequisite to responsiveness (Streiner and Norman 2008). Establishing responsiveness requires assessing an instrument in study following a longitudinal design. Testing for responsiveness then involves testing hypothesis relating to amount of change relative to the classification based on the anchor variable. On this basis, two major schools of thought have emerged: the first treats responsiveness as a separate attribute unique from validity or reliability (Guyatt et al. 1987), the other views responsiveness as part and parcel of an instrument's validity (Hays and Hadorn 1992; Liang 2000). This demarcation seems to be of little practical importance.

Interpretability

Availability of data supporting the decoding of qualitative meaning from QoL scores is quintessential to the feasibility and usefulness of instruments evaluating HRQoL, thus from that standpoint it is one of the key attributes of such measures. Observing an effect beyond what chance can explain (i.e. statistical significance) alone is not informative as to what a given magnitude of effect means in practical terms, this in turn requires the definition/identification of cut-off scores for clinical significance, the amount of change in the score that is large enough to require a change in treatment (Wyrwich et al. 2005). Banding systems, normative-values, or minimal clinically important difference (MCID), utility conversion algorithm have been utilised to support interpretability. This reflects two major approaches to interpretability. The first, based on distribution of sample scores and the second, based on scores of an anchor. The latter uses an external variable to define amount of change taking place in the patient's QoL, with the change score in the smallest change patient group giving MCID (Wyrwich et al. 2005).In the distribution-based method, MICD is given as half-standard deviation of the scores or standard error of measurement (Norman et al. 2003) (Revicki et al. 2008).

PRACTICAL ISSUES

Response Scaling

The scaling employed for capturing item responses plays a key role in how well the concept being addressed by an instrument's item are assessed. The appropriateness of the choice of scale (e.g. *visual analogue scales (VAS)*, *rating scales, Likert*, *adjectival scale*) may depend on the study aims, the nature of the disease condition and its treatment, the concept being measured, the mode of administration and target population (Patrick et al. 2011b). VAS and rating scales offer a continuous continuum hence may be more suitable for symptoms such as pain. On the other hand, adjectival and Likert scales are ordinal and might be more appropriate for assessing variables like frequency or intensity daily life impacts. While VAS seems more sensitive, the available finesse offered by the continuous nature of the scale is beyond the human capability to detect or distinguish small changes, such that this may introduce noise in the process (Streiner and Norman 2008).

For Likert and adjectival scales the number of response categories and the adjectives used as labels influence their meaning and usage (Streiner and Norman 2008, p.45) nonetheless these seem to easier to complete and have less administrative burden. Offering less categories than people can discriminate unnecessarily leads to loss of information and on the other hand offering many categories beyond the recommended five (+/-2) also leads to unnecessary cognitive burden and more noise in the measurement (Streiner and Norman 2008, p.49). The choice of response may be influenced by the scaling used. Among fully labelled scales, polar-point labelled scales and number based ranking, labelled scales had the highest number of extreme positives (Dillman 2006, p.462).

Frame of reference

A frame of reference needs to be specified for an instrument reflecting the period of time respondents are to consider in providing their responses. As longer periods, for instance exceeding one month, are associated with greater recall bias (Frost et al. 2007b) the shortest recall period feasible is always preferred. (Norquist et al. 2011) propose criteria for judging the appropriateness of a recall period, where the construct being measured, its time course; the purpose for measurement, for instance assessing treatment benefit in clinical trials; and the target patient population, and burden on respondents. For example acute symptoms that show rapid fluctuation such as pain may best be assessed with a shorter time frame, such as 'at present' while concepts

related to psychosocial functioning or activities of daily living may not show much fluctuation on day to day basis, thus may optimally be assessed with a weekly to monthly recall period (Frost et al. 2007b). Choice of recall, therefore, needs to appropriate for the condition and the timing of assessment, while imposing minimal burden on the patient (Kerr et al. 2010).

Mode of administration

The decisions related to the means to use for collecting the data and how the instrument is employed in data collection not only affects the psychometric attributes of the instrument but the final data collected to the extent that validity needs to be uniquely demonstrated for each data collection mode (Dalal et al. 2011). Because of this, pooling together QoL data collected using different modes is not recommended during clinical trials, as the observed effect size may be attenuated, unless equivalence of the modes is demonstrated (Coons et al. 2009). The appropriateness of a mode depends on purpose of HRQoL assessment, target population, their reading and writing abilities, the aims of the study, the characteristics of the disease condition and its treatment, the particular construct being assessed and the recall period (Patrick et al. 2011b).

Observation is perhaps one of the most common means used to collect QoL information. Although the doctor's conclusions regarding the patients quality of life may not always be accurate (Basra and Shahrukh 2009), this constitutes an important aspect of the patient-doctor interaction during consultations. This approach may also see further use in clinical trials involving populations not capable of judging their own quality of life, for instance young children (Salek 1998), the elderly and the terminally ill.

Alternatively, QoL data can be collected through **interviews,** where an interviewer reads out the items of an instrument to a study participant and subsequently notes down the provided responses. The interaction between respondents and the interview offers an opportunity to verify who is actually responding to the instrument while simultaneously giving the respondents an opportunity to ask questions where they do not understand, resulting in better compliance (Salek and Luscombe 1992). This method can be resource intensive in terms of both finances and time and tends to be more prone to biases related to social desirability or faking good, as people intentionally or un-intentionally provide more desirable expected responses to the interviewer (Streiner and

Norman 2008). Moreover, the interviewer-respondent interaction is prone to influences of the interviewer on the process and the information provided.

An alternative approach to collecting data is to present an instrument to patients for **self-completion**. This approach is attractive for its practicality, as it imposes the least administrative burden relative to other approaches. Moreover, there are several ways for implementing this approach: interview-delivered, mail-delivered or web-delivered. On the other hand, the instrument can use paper-and-pencil or electronic device as a medium depending on the method of delivery.

- **Interview-delivery** involves an interviewer being present during the completion process to provide instructions and respond to any questions that may come up (Salek and Luscombe 1992). This approach is practical for collecting QoL information in both routine clinical practice and research situations. For instance patients can complete the instrument in the waiting room prior to consultation, with a nurse being available to provide instructions and to respond to any questions and issues that may arise during the process. This approach requires that the instrument is of the reading level of the respondents, and that they are able to write, making it challenge in groups with reading and writing problems.
- QoL information can also be collected through **mail-delivery.** The questionnaire is mailed to the respondents; with a stamped return envelope for the completed questionnaire. This approach is favoured for its relative low cost in comparison to the other approaches in addition to the minimal human resource investment needed for collecting the data (Bowling 2009). Moreover, the respondents complete the instrument in their own environment, which might further enhance quality of data. This, on the other hand, entails that it may not possible to verify who is actually completing the questionnaire. A number of measures contained in total design method have been proposed to resolve low response rates associated with this approach (Dillman 2006).
- The **internet** is increasingly being utilised to deliver HRQoL instruments. The phenomenal growth in internet usage means that the relevance of the approach will continue to grow and making the issue of representativeness obsolete. Data collection is not restricted by geography on the other hand web-administration facilitates data collection from hard to reach populations(Tweet et al. 2011). Furthermore, respondents may feel anonymous and more comfortable to disclose sensitive information due to the perceived impersonal nature of the computer, leading to minimal influences of biases such as social desirability, faking

good, in the responses. Nevertheless, in spite of the minimal variable costs of adding one additional questionnaire, the initial costs of setting up a web-system might be substantial (Dalal et al. 2011). Furthermore, the required computer literacy and internet connectivity may limit the use of this method in certain populations.

Electronic data collection

Electronic devices, including tablet computer, personal digital assistants and smart phones, are increasingly replacing paper-and-pencil questionnaires in interview delivered self-completion questionnaires. Also, interview-administration has also seen the encroachment of electronic devices leading to computer assisted interviewing, for instance in the sequencing of the questions and recording of the responses by interviewees (Streiner and Norman 2008). In self-completed delivery, reduced error rates and respondent burden have been reported, reflecting previously observed preference of patients for electronic devices (Dillman 2006). Complex item skip patterns and additional information aiding the data completion process are easily implementable (Dillman 2006). Administrative burden is reduced as some secondary data processing and management tasks can be automated which also helps to mitigate some data entry errors.

STUDY POPULATION

The rise in social media has provided a new platform and channel through which patient interact with other patients with their condition, sharing information, finding support and advocating for greater public awareness of their condition. Moreover, online data collection may have a number of advantages such as overcoming geographical limitations and providing a degree of anonymity to study subjects (Idriss, 2009). This suggests that social media could be an important source of subjects in outcomes research in hyperhidrosis. Participants in all stages of this study will be recruited from online social networking communities on hyperhidrosis including the International Hyperhidrosis Society, and the UK hyperhidrosis society, Very Sweaty Betty Forum

In the qualitative stage, two Facebook pages will be created for the study. Facebook advertising campaigns, targeting users aged 16 and above, resident in Germany or the UK will be carried out. Searches of existing patient groups dealing with hyperhidrosis on Facebook will be carried out. Identified pages will be reviewed for their focus and relevance to the study, based on wall postings,

introductory information provided about the group. Depending on the nature of the identified relevant groups (whether open or closed) a posting about the study will be made on the page or alternatively a request to post study related information on such pages will be sent to the administrators of such groups.

To reach patients on other social networking sites platforms outside Facebook a search for discussion forums and patient groups will be carried out on Google, using the search terms [‘hyperhidrosis’, ‘excessive sweating’, ‘sweating’] AND [‘forum’, ‘support’]. Similarly, identified groups will be contacted and provided with information about the study for their membership.

In the quantitative phases of the study patients will be recruited through the two largest patient support groups, the *International Hyperhidrosis Society (IHHS)* and the *UK Hyperhidrosis support group*. Both groups maintain a website, a Facebook page and circulate a periodic email-based newsletter among their members. In December of 2012, their *Facebook* accounts had at least 800 and 200 followers (‘fans’), respectively. Their newsletters have a circulation of 50,000 and 2000 subscribers, respectively. The normal sharing of information among friends and group-members and the typical cross-group membership (Abram 2012), suggests that the number of persons receiving an original posting of information is exponential as it includes second and third parties and more.

Selection of participants

The generalizability of results of a study critically depends on how participants are sampled, particularly whether such process is independent of characteristics of patients and whether each member of the target population has an equal probability of being selected into the sample (Bland 1995). This implies random sampling. However, given the potential costs involved in implementing this, the exploratory nature of the current study and the interest in patients with specific characteristics, purposive sampling was employed. In addition, the target study population (online patient social networking communities) entails snowballing sampling. This sampling approach involves a researcher asking an initial group of study participants to recruit other potential study participants they may know (Bowling 2009). Thus, patients recruited will be asked to invite other patients to participate in the study through their connections in the online social

communities. This is consistent with normal activities within an online patient social networking community, where members interact and share information, most especially studies going on.

Data Collection

A new website will be created for the web-version of the new instrument. The landing page will contain background information to the study, with additional patient related information (e.g. downloadable full patient information sheet) placed in another location of the website. Access to the questionnaire-area will require a valid email address and a password. Apart from being a security/validity measure, the latter will facilitate completion of the questionnaire on multiple occasions. Before accessing the study instrument patients will be asked for their informed consent, which will be collected electronically by entering their name and email address. The completion of the study instruments was logically established to allow a logical flow will lead patients logically through the process from the landing page through the security check, receiving the patient information sheet and giving informed consent, to the screening process, and then completing the questionnaire.

PART II: METHODOLOGICAL FRAMEWORK

The new instrument will be developed with the following objectives in mind: i) to realise a tool that is practical and feasible in clinical practice as well as research setting; ii) with content and emphasis that is appropriate for all forms of hyperhidrosis, all level of disease severity, and relevant in evaluating the benefit from all forms of treatments, iii) to have robust scaling properties supported by both classical and modern test theories, iv) to have optimal reliability, validity and responsiveness, sufficient for evaluative use in individual patients.

The development of new instrument will involve the following steps (Figure 2.1):

Step 1: a review of the literature and existing hyperhidrosis-QoL instruments.

Step 2: investigation of the QoL issues relevant to hyperhidrosis patients using qualitative research methods

Step 3: development of a conceptual framework and drafting of new instrument based on results from the previous phases

Step 4: content validation by expert panel

Step 5: the new measure would be assessed for practicality and acceptability in target population.
Step 6: Initial construct validation and item reduction.
Step 7: Validation of the final version of the new instrument.

Step 1: Review Of Literature And HRQoL Instruments

The development of a new QoL instrument requires a clear rationale which contributes to the definition and measurement of the construct under assessment. Moreover, a strong theoretical basis for an instrument is essential to construct validity (Bond 2004). This involves a comprehensive review of the literature of the disease condition, its impacts and existing HRQoL measures. Such work has already been carried out for the current research and the results are reported in the first chapter. This established the need for a new QoL measure in hyperhidrosis overcoming the inadequacies of the existing measures.

Step 2: Qualitative Study Of HRQoL Issues In Hyperhidrosis

The content of a patient reported outcomes instrument has to be relevant to the target population to ensure content validity (Lasch et al. 2010). The items should reflect the way in which the patients view and describe their experiences with the disease (Patrick et al. 2011a). This entails patient involvement in the instrument development process. Qualitative research, useful for gaining insights into beliefs, views, and conceptual understanding held by subjects on an issue (Pope and Mays 2008) might provide a means for doing this. Therefore, a qualitative study will be carried out in patients with hyperhidrosis with the aim of understanding the various impacts experienced by patients. Semi structured interviews, focus group discussions and a survey containing open questions will be used for data collection. A topic guide will be developed for interviews and focus group discussion. Its use will be limited to probing on issues omitted by the patient which are known to be important in hyperhidrosis based on previous studies. Patients will be encouraged to elaborate more on their answers by probing them for reasons why or asking them for specific examples in their narratives. Subsequently the questions in the open survey will be generated based on the results from the focus groups and interviews. Combining multiple qualitative methods, for instance focus groups and interview is important in ensuring the validity of findings in qualitative research (Whittemore et al. 2001). In this way the data collected is enriched by the strengths of each method: interviews enable in-depth insights, focus groups offer unique data through the

interaction among subjects (Brod et al. 2009), survey is capable of reaching a wider larger of participants at a relatively lower cost.

During the qualitative phase of the study, a stand-alone PhP-based online discussion board will be developed. Discussion boards are typically used as platforms for internet forums, allowing text-based discussions among any number of members and guests, and are managed by an administrator. In order to include only patients recruited to the study in discussions, participants will be given a username and password, for accessing the discussion board. Data will also be collected using instant messaging platforms (e.g. Skype and Windows Live Messenger). Data collected from the interviews and focus groups will be tape recorded and transcribed verbatim. Thematic analysis, an atheoretical approach, will be used for data analysis. This means analysis will commence without a preconceived theory, rather a framework will be developed from the data as analysis proceeds, driving further data analysis and data collection (Braun and Clarke 2006). Issues emerging from the qualitative study analysis of the data transcripts will be organised as themes and common themes will form domains.

Step 3: Development Of The Conceptual Framework And The New Questionnaire

An instrument's **conceptual framework** depicts the relationship among the observable items, their domains and the underlying latent variable (Rothman et al. 2007) reflecting how a particular construct is understood. It is critical that the development of the conceptual framework precedes the actual drafting and development of the instrument because of its influence on later phases of instrument development and validation. Therefore before the drafting of the new instrument its conceptual framework will be developed first, based on results of prior steps (the literature review and the qualitative study). The new instrument will be drafted based on the conceptual framework and the qualitative issues collected from patients. Ensuring a structured process at this stage is an important aspect of ensuring content validity (Lynn, 1986). A team comprised of experts in clinical research and patient outcomes measurement will be created to undertake the drafting of the new questionnaire. Criteria relating to the inclusion of content, wording of the actual questions, and other elements of the questionnaire (layout, formatting, response options, instructions), will be set beforehand to guide the process. A transparent decision making process will be followed by the team.

Step 4: Expert Panel for Content Validation

As critical proof of the link between the content of an instrument and the underlying concepts, content validation of the new instrument will be carried out by asking therapeutic experts to make judgements on the appropriateness of the content for the intended concepts i.e. whether the items included are relevant, whether all important issues are included, whether language is clear enough for someone with a reading ability of a 12 year old to understand; whether the response categorisation is appropriate for each item. In order to provide adequate guidance to the experts, as well as to enhance the credibility of the process a validation questionnaire will be used for this purpose. Inter-rater agreement will be used in assessing the reliability of the ratings and the content validity index will be used for summarising ratings.

Step 5: Acceptability And Practicality Of The New Questionnaire

The usefulness of an instrument depends on whether it addresses issues relevant to patients, it is simple, easy to complete and does not required long time to do so (Thornicroft and Slade 2000). On the other hand, a measure should not impose undue burden on those administering in terms of data collection or analysis. The latter is of particular importance in routine clinical practice, where there might be particular constraints on time and monetary resources (Higginson and Carr 2001). A pilot study will be carried out following a cross-section design. Study participants will complete the new instrument and a supplementary questionnaire collecting information on: relevance of the items, ease of completion, time to completion for the new instrument. Suggestions on possible issues to be added will also be sought. Furthermore, problems encountered in completing the new instrument reflected in missing item responses or errors in completion will be noted. Items highlighted as unclear or causing any difficulties will be reviewed.

Step 6: Item Reduction and Construct Validation of the new instrument

The summation of item scores at the scale level is underpinned by strong assumptions relating to the nature of the underlying construct i.e. that there is a single latent variable (Fayers and Machin 2007). Such assumptions have implications on any inferential use of the instrument, as they touch upon the definition of the underlying construct. Therefore, to assess the latent structure of the new instrument a prospective cross-sectional study will be undertaken. Patients will complete the new instrument on a single assessment occasion. Exploratory factor analysis (EFA) will be carried out

to explore the most optimal dimensional structure for the new instrument. This will subsequently be tested in confirmatory factor analyses.

As the development of a new instrument tends to start with a large number of items undertaking item reduction is useful step in the development process (Terwee et al. 2007). Considering that the initial item development of the new instrument will aim to be inclusive, item reduction will be necessary subsequently. Classical test theory's correlation analysis and exploratory factor analysis and modern test theory's Rasch model will be used to carry this out. Nonetheless, apart from results from statistical models, qualitative considerations will also be made (e.g. importance of issues to patients and overlap with existing items) (Guyatt et al. 1993). The intention is to realise a set of items contributing to the measurement of the latent variable. On the other hand, this step will elucidate on the internal structure of the new instrument.

Step 7: Validation of the final version of the new instrument

Validity

As there is no single 'ultimate test' for construct validity (Streiner and Norman 2008), its assessment involves testing for various hypotheses relating to the relationship between the underlying variable and the items of the instrument in different situations. Therefore, assessing the validity of the final version of the instrument will involve testing a number of hypotheses.

Known-groups validity: Patients with more severe hyperhidrosis will be expected to show greater quality of life impairment. Patients where hyperhidrosis affects multiple areas (for instance, axilla, feet, and palms) are expected to experience greater quality of life impairment than those with a single area, controlling for everything else.

Convergence validity: It was hypothesised that patients' hyperhidrosis-specific QoL has a positive relationship with their skin-specific QoL and their generic QoL. Therefore, a prospective cross sectional study of patients with hyperhidrosis will be carried out. Participants will complete the final version of the new questionnaire and additional questionnaires assessing dermatology-QoL (the DLQI and the Skindex-17) and generic HRQoL (EQ-5D). Using both the DLQI and the

Skindex would allow the new instrument to be compared against the two key instruments for evaluating dermatology-QoL.

Reliability

The degree of measurement error in an instrument has practical implications on the practical use of an instrument (Streiner and Norman 2008). For example, in clinical trials a less reliable measure may require a larger sample to show a particular effect size relative to a more reliable measure. Thus, internal consistency and test-retest reliability of the score for the new instrument will be tested.

A study following a cross-sectional design will be carried out to assess internal consistency. Participants will complete the new questionnaire on a single assessment. Inter-item correlations, item-partial total correlations and Cronbach's alpha coefficient will be estimated based on the item scores.

A longitudinal study, where participants complete the new instrument on two assessment occasions, 7 days apart, will be carried out to assess the test-retest reliability. A period of 7 days has been recommended in test-retest assessment studies, to prevent practice effect, yet on the other hand, the condition should have remained stable (Salek and Luscombe 1992). The latter was ensured by collecting data on an additional variable for instance self-rated disease severity (using HDSS). Test – retest reliability will be assessed by measuring the level of agreement in the baseline and follow-up scores, assuming the patient's condition should have remained the same.

Responsiveness

Establishing responsiveness of an instrument requires not only showing that an instrument can capture statistically significant changes (changes beyond chance), but more importantly that it can capture minimal changes considered important by the patient (Revicki et al. 2008). A longitudinal study with three assessments (at baseline, on 8^{th} day and on 21^{st} day) will be carried out to establish this attribute. During each assessment patients will complete the new instrument and an additional questionnaires for determining magnitude of the experienced by the patients and its importance to them.

In particular, the following hypotheses will be tested:

- the new instrument can capture changes in the group of patients experiencing minimal but important changes in their condition,
- the magnitude of change in patients with minimal improvement in their condition was greater than those with no change in their condition.
- change would be greater over a longer period (between baseline and 21st day in comparison to baseline and day 8) in those patients receiving active treatment.

Interpretability

The qualitative meaning of scores from HRQoL scales is not intuitively apparent (De Vet et al. 2006) despite the importance this has on the credibility and usefulness of HRQoL information especially in clinical practice (Higginson and Carr 2001; Grob 2007). Thus, to facilitate interpretability of scores of the new instrument a banding system and MCID estimates will be established. A banding system establishes the score ranges of measure reflecting qualitative categorisations corresponding to a mild, moderate or severe level of impact of a condition (Prinsen et al. 2010). For the new instrument, this will be established based on data collected from the cross-sectional validation study, on the new instrument and an anchor variable. The MCID, on the other hand, reflects the smallest change considered important to patients (Revicki et al. 2008). This will be estimated based on data collected for the cross-sectional validation study and the data from the longitudinal responsiveness assessment study. As a triangulating of multiple anchors is recommended in establishing MCID (Guyatt et al. 2002), two instruments, the HDSS and the PGA will be used as anchors for assessing change in the patient's condition. The standard deviation and standard error of measurement from validation study will provide the distribution based estimates of the MCID.

Sample size

Sample size considerations differ between qualitative and quantitative research. In the former, it is not possible to determine the needed sample size prior to data collection; rather sample adequacy is determined in the course of data collection. Data collection continues until 'saturation' has been reached, which reflects a situation where further data collection (e.g. interviews) is not yielding new data (Kerr et al. 2010). On the other hand in quantitative research sample size is dependent on the particular statistical analysis performed. Required sample size will reflect the intended

power of analysis, the magnitude of effect size to be observed and chosen level of significance and reliability of measurement (Lipsey 1990). Exploratory studies, where magnitude of effect size and reliability are unknown apriori may present some challenges in this regard. A useful recommendation is to use a sample matrix based on key disease or treatment characteristics for a particular disease, where each sub-category (each cell) should have at least 15 subjects (Johnson et al. 2011). For initial estimates of reliability and validity at least 200 subjects are recommended (Frost et al. 2007a). If a test-retest correlation of 0.85 is observed with a sample size of 100, the 95% confidence interval is 0.78 – 0.90, while a sample size of 150 would narrow this to 0.8 – 0.89 (Johnson et al. 2011).

Rules of thumb on sample size requirements for correlation analysis and factor analysis vary in their guidance, ranging from 5 to 20 observations per variable with more suggestion above and below this ratio (Costello and Osborne 2005). However, the minimum sample size required for accurate recovery of population factor pattern matrix is influenced by many factors including the distribution and reliability of the variables, and degree of association among variables, communalities, degree to which factors are over identified (Reise et al. 2000; Schmitt 2011). Thus power and precision ought to be core consideration in parametric estimation based factor methods (Schmitt 2011), while in non-parametric approaches when communalities are high, sample size of 100 may be adequate (Reise et al. 2000).

Assessment of adequacy of sample size for given statistical test should be made along with other key considerations relating to the sample for instance ensuring that the target population is adequately represented along with all important disease characteristics. Otherwise, appropriate tools will be applied to indicate uncertainty surrounding estimates e.g. using confidence intervals in presenting results.

Data Collection Instruments

Apart from the new hyperhidrosis-HRQoL measure, other instruments/questions will be used in collecting data from patients regarding their HRQoL, level of disease severity, overall impact of disease and change in their condition over time. Disease severity was assessed using the Hyperhidrosis Disease Severity Index (HDSS) (Figure 2.1). The HDSS is a validated single item scale which measures the severity of hyperhidrosis and the related daily life interference on a four point scale (Kowalski et al. 2004). A 1 point decline represents up 50% reduction in sweating; while a 2-points reduction reflects a decrease of 80% (Solish et al. 2007). Data on generic HRQoL

Figure 2.1: The Hyperhidrosis disease severity scale

How would you rate the severity of your hyperhidrosis?

☐1. My sweating is never noticeable and never interferes with my daily activities

☐2. My sweating is tolerable but sometimes interferes with my daily activities

☐3. My sweating is barely tolerable and frequently interferes with my daily activities

☐4. My sweating is intolerable and always interferes with my daily activities

was collected using the EQ-5D, an instrument designed for use in both clinical and economic evaluation research and intended to be highly practical and useful in international-studies (The-EuroQoL-Group 1990; Brooks 1996). The EQ-5D consists of a descriptive part, containing 5 items reflecting unique domains, with each rated on 3 levels; and a VAS scale assessing health. Responses to the descriptive component can be combined into a 5 digit number, which can in turn be used to identify a patient's health status. A utility value, for calculating QALY's can be read off from country-specific reference preference values generated from the general public. A modification of the instrument offering 5 levels for each domain/item and minor revisions to the item-descriptors, EQ-5D-5L, has been developed and has shown content validity (Herdman et al. 2011). Although studies to generate reference values are currently underway algorithms for mapping the value sets for the original 3 level EQ-5D into the EQ-5D-5L have been published (Rabin et al. 2011).

Dermatology-specific QoL was assessed using the DLQI and the Skindex-17. The DLQI was developed as a practical measure of the impairment in patients QoL resulting from skin disease (Finlay and Khan 1994). The instrument consists of 10 items assessing the intensity with which patients experienced various impacts in preceding week. Items are scored from 0, not at all, to 3, very much and can be summed up to give an overall scale score (0, minimum impairment; 30, maximum impairment).The Skindex is a validated instrument developed as a measure of effects of skin disease on patients HRQoL (Chren et al., 1996). Its 30 items assess frequency with which patients experience various effects using a 4 weeks recall period. A brief version of the measure with 17-item, the Skindex-17, has been developed based on the Rasch model scaling (Nijsten et

al. 2006b). This has demonstrated strong correlation to the original measure as well as optimal psychometric properties consistent with modern test theory (Sampogna et al., 2013). A full review of these measures is available in chapter 1.

In addition, two general questions were administered, an overall-impact global question (Figure 2.2) and a global assessment of change by the patient (Figure 2.3). This is the first time that these questions are administered in hyperhidrosis, although a similar questions have been applied in dermatology to establish the scale banding for the DLQI (Hongbo et al. 2005) and in renal replacement therapy to establish the scale banding for the Renal Quality of Life Profile (Aawar 2011).

DATA PROCESSING AND ANALYSIS

The analysis of data will be carried out using STATA 11, SPSS and more specialised software including M-PLUS for CFA, and RUMM 2030 for Rasch analysis. This study will collect a variety of variables of different types including continuous, discrete, ordinal, categorical and binary scale: item scores from the QoL questionnaires; patient characteristics e.g. age and gender; disease characteristics and treatment such as location of hyperhidrosis and duration of disease; resource utilisation including time in minutes and amount of money in currency

Figure 2.2: Global question on overall impact of hyperhidrosis

Over the last seven days including today, how much has your sweating condition affected your life?	
Extremely large effect	☐
Large effect	☐
Moderate effect	☐
A small effect	☐
No effect at all	☐

Figure 2.3: Question on Patients global assessment of change

How would you describe your condition today, in comparison to your last assessment?	
Better	☐
Slightly better	☐
No change	☐
Slightly worse	☐
Worse	☐

The data will initially be explored through descriptive analysis of each variable, calculating measures of central tendency (mean, median), variability (SD), and interquartile range for continuous variable; Frequency counts for ordinal and categorical variables. Further analyses will involve making inferences based on various hypotheses tests. In order to reject a null hypothesis observed probability of a false positive, type I error, as reflected in P-value needs to be less than, level of significance (α) (Altman et al. 2000). This study will use a level of significance (α) of 5% . Where several hypotheses will be simultaneously tested Bonferroni adjustment will be applied to the level of significance, as (α / k), where K is the number of tests (Fayers and Machin 2007).

- Testing for differences between two means will use independent or paired t-test, depending on whether the two means are mutually exclusive or are related. The Mann-Whitney and Wilcoxons tests are the non-parametric alternatives, respectively, for situations where assumptions of the t-tests are not met.
- Hypothesis tests involving differences among more than two groups will be carried out using ANOVA test. Where the core assumptions of this test are not met, particularly, the assumption of homogenous variances across group, the Kruskall-Wallis test will be used alternatively.
- Testing of hypothesis relating to associations between means of variables will be carried out based on Pearson's correlations. Where the data is not continuous Spearman's rank correlation will be used.
- Polychorric correlations will be estimated in order to assess multicollinearity among the items. This type of correlation produces consistent and robust results in ordinal data. They are based on the assumption that the variable is linear and continuous but divided up in a series of categories (Holgado–Tello et al. 2010). Multicollinearity is identified when correlations coefficient is 0.8 or greater.

Possible influences on the magnitude of observed inter item correlations including range of score values, homogeneity of items, distribution of the data (particularly departures from normality) and existence of outliers in the data (Fayers and Machin 2007, p.33) will be explored. Normality assumption implies skewness not exceeding |3|, while Kurtosis must not be greater than |7| (Ozer et al. 2009; Byrne 2011). While the former impacts on means, covariance tends to be vulnerable to kurtosis values (Byrne 2011).

Further statistical analyses carried out during construct validation will use various forms of regression methods, modelling latent variable including exploratory factor analysis, confirmatory factor analysis and the Rasch model:

Exploratory factor analysis

Exploratory factor analysis provides a way for explaining variability in a large set of indicators using a few latent variables (Kline 1994) which is quite handy for investigating the latent structure of new instruments. The aim is to identify the smallest number of interpretable factors explaining the covariation among items (Muthén and Muthén 1998 - 2010)**.** This involves first generating the variance-covariance matrix, followed by the estimation of the factors which entails putting together those items sharing the highest co-variation. Considering that a factor solution is not unique the initial estimated solution needs to be rotated in order to achieve a simple structure that is more interpretable (DeVellis 2011).

To perform an EFA on the instrument, first a polychorric correlation matrix will be generated. These more appropriately take into account the ordinality of the data and remain robust when data are skewed, in comparison to the conventional Pearson's correlation coefficients (Byrne 2011). The initial factor estimation will be carried out using robust diagonally least squares estimator (WLSMV) which yields robust test statistics, parameter estimates and standard errors when indicator variables are categorical and where normality assumptions are violated (Byrne 2011, p.132). Rotation will be performed using the Geomin routine available in M-PLUS software, which allows correlation among factors. This rotation is particularly suitable for psychosocial domains known to be highly related (Lackey et al. 2003). Where the factors are not related, Geomin still performs well yielding results comparable to orthogonal rotation routines. Choice of appropriate number of factors to be extracted will be based on the parallel analysis and will be confirmed by statistical goodness of fit measures (Schmitt 2011). Kaiser's rule, based on size of eigenvalues; scree-plot, which is a graph of number of factors against eigenvalues and parallel

analysis, comparing actual against ones randomly generated, will also be reported. The following criteria will be applied:

- Kaiser's rule: factors with eigenvalues greater than are included (Kaiser (1960) in DeVellis 2011, p.148).
- Scree-plot: all factors to the left of the 'ankle' are extracted, where there is a change in the slope.
- Parallel analysis: the last factor to be retained must have an eigenvalue greater than the one that would be produced randomly (Williams et al. 2010).

Goodness of fit indices go alongside factor estimation based on Likelihood methods, in assessing how well the hypothesised model fits the data. These can be classified into three groups: i) chi-square based indices, based on the null hypothesis that compared with a single factor model the chosen number of factors (k) are adequate; ii) practical fit indices, evaluate proportionate improvement in model by comparing a hypothesized model against a less restricted baseline model (Byrne 2011, p.70); and ii) absolute fit indices, which are based on analysis of residuals after fitting the model to the data (Brown 2006). The following indices will be used for the study:

- The 'chi-goodness of fit test': a non-significant chi-statistic represents good fit (Lackey et al. 2003, p.121).
- Practical fit indices. Comparative Fit Index and Tucker-Lewis Index, where values of below 0.9 and 0.95 have been suggested for acceptable and adequate fit , respectively (Schmitt 2011).
- Absolute fit indices. Root Mean Square Error of Approximation (RMSEA), where value below .05 show good fit, .08 to .1 mediocre fit and; above 0.1 poor fit (Browne and Cudeck 1992; MacCallum et al. 1996); The Standardized Root Mean Square Residual (SRMR), is seen as reflecting 'adequate fit' when less than .05 and acceptable fit when less than 0.8; The Weighted Root Mean Square Residual uses a cut off value of 0.95 for good fit (Byrne 2011, p.76, p.140).

Confirmatory factor analysis

The goal in confirmatory factor analysis (CFA) is to evaluate whether a hypothesised factor structure fits a given dataset (Brown 2006). Thus unlike exploratory factor analysis CFA is a hypothesis testing tool and fits a regression model of the hypothesised latent variables and the indicator variables as specified by the researcher. Inferences are based on overall model fit, the

significance of the individual item parameters (loadings) and magnitude of the residuals. The goodness of fit indices presented above are also applied. Residuals of 0.05 are indicative of good fit (Byrne 2011). CFA implemented in this study will test the hypotheses of a single factor solution and; the factor structure obtained from the EFA

Rasch model analysis

The item response theory, in particular the Rasch model (RM), offers a framework for scaling unidimensional instruments. The model expresses the probability of choosing a particular response to an item as a function of the relative difference between the severity level assessed by an item and that of the respondent, respectively. As both are measured on a common linear scale, this represents the distance between the item location and respondents location on the single linear scale of the latent variable (Tennant and Conaghan 2007). The relationship between the latent variable and the item responses follows a monotonic logistic ogiv function, reflected in the item characteristic curve (ICC) (Masters 1982). This is similar to the curve representing a typical binary logistic function. The RM is based on core assumptions of unidimensionality and local independence, such that once the single latent variable (ϑ) is accounted for no further relationship should exist between any two items (Reeve and Mâsse 2004). This gives rise to a probabilistic Guttman pattern whereby for any given item, persons with greater severity (ability) should have a higher probability of choosing a higher category on an item in comparison to persons with less severity; the opposite also applies that for a given person, the probability of choosing a 'higher category' should be higher for items at lower severity level than those at a higher severity level for any person (Tennant et al. 2004).

Appropriate fit to the RM ensures that an instrument is sufficiently unidimensional and that it complies with conjoint measurement principles, a precondition for converting the data from the instrument into interval scales (Bond 2004). The intention of Rasch analysis, therefore, is to evaluate whether data have sufficient fit to the model to warrant such claims.

Assessing conformity to the Rasch model, its assumptions and properties involves the following:

1. Assessing whether response categories are functioning optimally. *Average latent measure* across observations in a response category and *category thresholds* should monotonically increase with the category; each response category should have a distinct peak on the *category probability curve* graph reflecting the space along the latent variable where it is most probable (Linacre 1999). Category characteristic curves define the most likely response

category for a specific person location value on the latent variable. The category threshold indicates a location on the latent variable where probability of selecting adjacent categories is equivalent (Linacre 1999).

2. Testing item and person fit to the model. This uses residuals obtained after fitting data to the model, calculating a fit residual statistic and the item trait interaction chi-statistic.

   The residual statistic for items is calculated as the squared summation of the standardized residuals of the responses of all persons to an item (Andrich et al. 2012b). Fit residuals exceeding |± 2.5| indicating poor fit (Andrich et al. 2012b). As the RM does not distinguish between items and persons (Bond, 2007), the residual fit statistic for persons is calculated and interpreted in a similar way.

3. The item trait interaction test of fit, assesses the discrepancy between actual and model scores of class intervals (which group patients according to ability), visually reflected by discrepancy between the ICC and empirical counterpart. An item chi-value is generated by adding all standardized differences for class intervals (Andrich et al. 2012b, p.21) (Andrich et al.).

4. Testing of overall model fit. Mean fit residual value of 0 and standard deviation of 1 reflect overall model fit (Shea et al. 2009). The item-trait interaction statistics for all items are summed up into total item-trait interaction statistic. Optimal fit is reflected in a non-significant statistic (Chi-squared statistic, p-value > 0.05). Good fit to the RM implies that the hierarchical ordering of the items remains invariant across the different levels of severity assessed by the construct.

5. How well the instrument can differentiate persons according to severity should be assessed. This is reflected in the Personal Separation Index (PSI) which reflects the proportion of variance explained by the model out of the total person variability (Wright and Masters 1982; Bond and Fox 2007). A PSI of 0.8 reflects capability to reliably distinguish patients into at least 2 groups of severity e.g. high and low severity.

6. Assessing targeting of items. The item-person map is visually examined for adequacy in spread of the items along the breadth of the latent variable, ideally there should not be be large gaps in between items (Wright and Masters 1982, p.90); mean location of persons should be close to 0 to match the item mean location centred at 0 logits. (Gorecki et al. 2011).

7. Assessing unidimensionality. First, a principal component analysis is carried out on residuals after fitting the RM. Unidimensionality is supported if the first component accounts for no

more than 30% of the variance in the data and has eigenvalue of 3 or less (Linacre 1998). A more stringent assessment of unidimensionality has been suggested by (Smith 2002). Items are grouped according to their loading on the first residual factor, comprised of high positive and high negative loading items, respectively. Pairs of person estimates generated from the two item sets are compared using a series of t-tests. If the proportion of significant tests (or the lower bound of its confidence interval) exceeds 5% unidimensionality is ruled out (Tennant and Pallant 2006).

8. The assumption of local independence will be assessed by examining the correlation matrix of the item residuals. Residual correlation exceeding 0.2 – 0.3 reflect a violation of this assumption. The magnitude of the response dependence is calculated as the shift in the latent variable range representing a given response choice on the dependent item, induced by a particular response choice on the independent item (Andrich et al. 2012a).
9. Assessing for invariance across demographic factors. DIF can be assessed for key demographic factors using a two way ANOVA test. A significant main effect (demographic variable) at 0.05 level of significance, with Bonferroni adjustment, indicates presence of uniform DIF. On the other hand a significant interaction effect (demographic variable X class interval representing ability groups along the latent trait), after Bonferroni adjustment, indicates non-uniform DIF (Andrich et al. 2012b). Identification of DIF requires a pure set of items, upon which the scale is anchored (Teresi and Fleishman 2007).

Any action on DIF requires an understanding of its magnitude and impact. Magnitude indicates the difference between item difficulty estimates based on all patients and comparable estimates specific for each demographic group (Linacre 2009). The impact of the DIF on estimation of person estimates is assessed by comparing person estimates generated from the DIF-free items against estimates based on all items including those with DIF (Tennant and Pallant 2007). Using a t-test, significant results, at 0.05 level of significance, indicate that DIF has an impact. The Item Characteristic Curve (ICC) of the two series may also be useful in assessing whether the pairs of person ability estimates agree. Impact of DIF can also be explored by assessing whether the Test Characteristic Curves (TCCs) from different demographic groups are comparable i.e. whether the relationship between the raw score and the underlying latent variable varies across the demographic groups. Identical TCCs indicate the absence of impact of DIF on the total score (Edelen et

al. 2006). The criterion for magnitude of DIF is also relevant for differential scale functioning

Missing Data

Situations where a question or an entire questionnaire has not been completed are common during data collection in QoL research. The reason behind the missing data has an influence on choice of tools for dealing with the consequent problems in data analysis, for example, whether an item is skipped by mistake or due to its irrelevant. There are three main classifications of patterns of missing data, missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) (Fayers and Machin 2007). MCAR arises where the probability of having a missing item (questionnaire) is independent of previous or unobserved current and future scores. MAR occurs where missingness is dependent on known covariates and scores of previous items, but not on the unobserved scores. The third case, relates to where the unobserved HRQoL influences the missingness. The presence of MAR and MCAR is not worrisome, as their impact on accurate measurement of HRQoL is minimal (Leidy et al. 1999). MNAR causes the greatest concern as its presence may lead to an over or underestimation of HRQoL, highlighting the need for transparent approaches in addressing its presence.

There are no clear guidelines on the number of missing items to warrant the exclusion of an entire respondent's questionnaire although Streiner and Norman (2008) have mentioned a ceiling of 5% of items. However, it's worth noting that where Rasch scoring is applied, a higher number of missing items may be tolerated without much bias in measurement(Fayers and Machin 2007). On the other hand, in some situations (e.g. during instrument development work) data imputation to replace the missing data offers a viable alternative. This is done in various ways including using the last observed value carried forward; by calculating a simple mean; or using regression methods (Fairclough 2010). Other more sophisticated imputation approaches such as hot-deck and markov-chain are capable of preserving variability in the data.

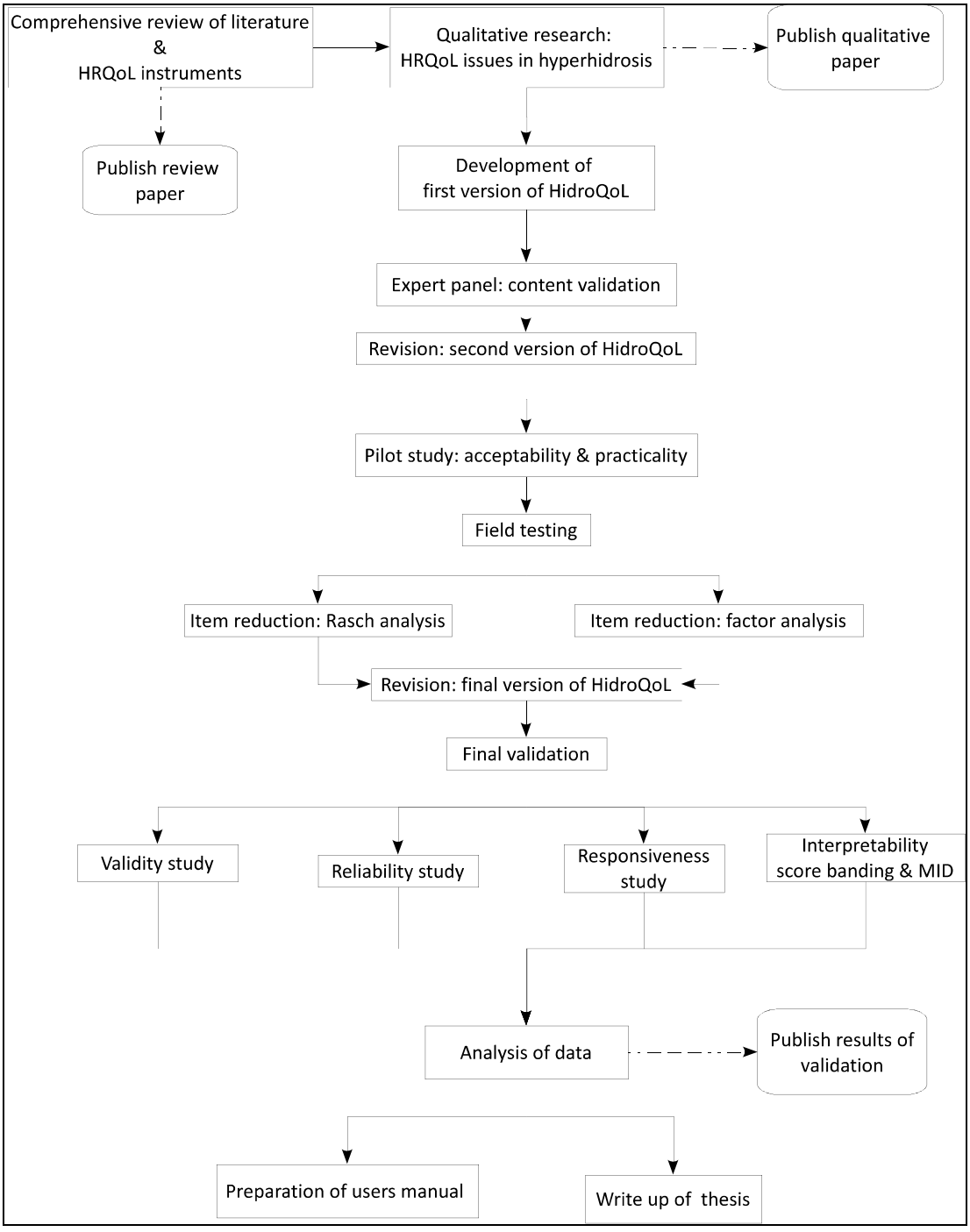
To prevent the problem from arising in the first place, respondents will be instructed to cross-check their questionnaires to make sure they have completed all items for the paper-and-pencil version of the instrument in this study.

SUMMARY

- This chapter has provided the justification for a new instrument for evaluating QoL in hyperhidrosis, by elucidating on the role and need for HRQoL assessment in hyperhidrosis and the inadequacies in the existing instruments. The attributes of a measurement instrument necessary for the credibility of HRQoL information were highlighted. On that basis, the steps necessary to produce such evidence were outlined reflecting the design of this study, to develop and validate a new instrument for measuring HRQoL in hyperhidrosis, for application both in the clinic and in research settings

- An overview of the study is presented in Figure 2.4

Figure 2.4: Flow chart of study



CHAPTER 3

Development of a Hyperhidrosis-Specific Patient-Reported Outcome Measure: Qualitative Study

INTRODUCTION

Skin disorders are unique in that the subjective experiences of patients tend to be a more powerful determinant of the patient's quality of life (QoL) as well as the overall burden of the disease, in comparison to the degree of objective severity for instance (Russo et al. 2004; Jobling and Naldi 2006). Patients care less about specific symptoms of skin for instance redness; flacking; or being wet; rather of more importance to them is the actual physical discomfort experienced; change in the patients self-image and; the wider implications for the patient's psychological functioning and social life (Grob 2007). On a practical level, evaluating symptoms may not be as straightforward, in spite of their alluring objectivity. For example in hyperhidrosis, the amount of sweat considered pathologic is unclear; on the other hand, and the tools for quantifying amount of sweating are impractical for the routine clinic (Wörle et al. 2007).

Patients with hyperhidrosis have reported various impacts on their QoL, for instance physical discomfort associated with the continuous dampness; feelings of embarrassment and anxiety; and difficulties in meeting strangers; limitations in everyday life activities and in occupational activities (Strutton et al. 2004; Hamm et al. 2006; Solish et al. 2008). This current understanding is based on survey instruments and quantitative methods; a fuller picture of the extent and nature of the impacts for instance the complex interrelations among disease severity, individual adaptation and public response may require methods that can unveil patient thoughts, beliefs and interpretations of their experiences (Jobling and Naldi 2006). Further, other issues with possible implications on the QoL of patients such as patient's information need; self-management strategies; experiences in obtaining care are not well understood. Thus a study rooted in qualitative research methods would be very useful in exploring and elucidating on these issues.

This, therefore, means that QoL impacts ought to be a key consideration in the diagnosis and management of hyperhidrosis in routine clinical practice and the evaluation of treatments; a point also recognised in existing treatment guidelines (Tan and Solish 2002; Solish et al. 2008). This would require a credible way of capturing and evaluating such QoL impacts. Such a measure needs to be appropriate and relevant for assessing QoL impacts in hyperhidrosis and should at least have demonstrated evidence of its precision, reliability, validity, sensitivity to change and practicality.

OBJECTIVES

The objectives of this study were to:

- To explore the experiences of patients with hyperhidrosis in order to obtain an in-depth understanding of the extent and nature of QoL impacts
- To develop a disease-specific instrument for evaluating QoL impacts in hyperhidrosis based on the experiences of patients.

METHODS

Ethics

Ethics-approval for this study was obtained from the University Hospital of Greifswald Ethics Committee in Germany on 31st July 2011, where the data collection was based. Prior to this, guidance had been obtained from the South Wales NHS Research Ethic committee considered the study to lie outside their remits as data collection was not based within the UK.

Written informed consent was obtained from participants before their participation in the study.

Recruitment

Materials, including the background to the study, information sheet and an invitation to patients were placed on the study's Facebook page and other online social networking communities for patients with hyperhidrosis. Patients interested in the study contacted the research team by E-mail to participate in FGD or interviews.

Study participants

The study included participants fulfilling the following inclusion criteria:

- With self-reported sweating problems.
- Seeking for treatment;
- 18 years or older;
- Able to speak, understand and write in English.
- 18 years or older;

- Sweating linked to other underlying health problems or treatments

Sample size

In qualitative research, unlike in quantitative data collection, it is not feasible to objectively predict optimal sample size prior to data collection. Thus, we continued with data collection up until no new themes were emerging i.e. a point where 'saturation' of the content had been reached.

Data collection

This study utilised qualitative research methods including semi-structured interviews, focus group discussions (FGD) and a survey with open ended questions for data collection. Interviews and focus groups are especially useful in facilitating detailed and deep understanding of social phenomena (Gill et al. 2008). Both allow the framework of understanding, thoughts, feelings, perceptions and emotions of study participants to be aptly explored (Bowling 2005). Focus groups bring together a group of respondents to meet and discuss a particular topic or issue and; they involve a moderator, who is usually also the researcher, facilitating the sessions (Fayers and Machin 2007). The interaction among respondents adds to the richness of the data collected(Bowling 2005). Interviews on the other hand provide an appropriate setting for exploring sensitive topics (Gill et al. 2008).

Procedures

- Two FGD sessions were conducted in the form of online text-based discussions, each over a period of two weeks and were moderated by a member of the research team (P.K.) who posted topics, probes and prompts on the board. Participants were encouraged to read postings from the moderator and the responses of the others as well as make their own contribution to ongoing discussions. Access to the online discussion board (*based on php*) required a password and username which were given to each participant.

- Interviews were carried out by telephone and internet instant chat facility e.g. Skype. Each interview began with inviting the patient to share their experience of the effects of the disease condition in general, through the following question: "In what ways does hyperhidrosis affects your life?" Each of the areas mentioned by the patient was probed

further. The interviewer also raised questions in relation to specific areas of life previously known to be heavily impacted by hyperhidrosis overlooked by the patient.

- Based on the topic guide, results from the FGD and interviews an online survey with open ended questions was developed according to *Zoomerang survey platform*. This was then posted on various online social networking communities for patients with hyperhidrosis.

Data Processing And Analysis

Interviews were tape recorded then transcribed verbatim. The focus group discussions and open surveys were already in text format. Content analysis of transcripts was carried out using NVIVO 9. This form of qualitative analysis is focused on both the context and content of source material, with the aim of identifying major themes, their frequency and their relationship to external factors such as demographic characteristics of study subjects (Robson 2011). As an atheoretical approach, the analytical framework i.e. the thematic structure is data driven and not imposed from known theories or previous studies (Braun and Clarke 2006). Analysis of data started as soon as data collection started, and was continued during the data collection process. Transcripts were thoroughly studied in order to gain an understanding of the data and to build an initial overview of topics. This process was repeated with further data collection. The transcripts were indexed and sorted by assigning common labels to chunks of the transcript considered to be about the same topic, a process referred to as topic (Saldana 2009). During the early stages of analysis, the coding was aimed at indexing all data. In later stages, the initial coding was revised, not only to reflect subsequent data collected, but also to combine a number of codes addressing a similar topic, to enhance meaning and understanding of data. The higher level coding produced major themes from the data. Grouping themes addressing a common topic identified major domains in the data, representing areas important to patients. Further analysis aimed at identifying the number of subject contributing to a particular theme. The inter-relations between themes was also analysed by exploring the themes coding common material i.e. overlap in reference material. Throughout the study a clear audit trail of decisions taken was kept including a codebook and a saturation matrix. In order to provide clear examples of what each identified theme covered, quotations based on transcripts of what the patient actually said were included.

RESULTS

Sociodemographic Characteristics Of The Study Participants

Out of 13 potential participants recruited for the FGD, 9 (69%) patients participated; out of 41 recruited for the interviews, 32 took part; and out of 46 who started filling the survey 30 completed (65%) (Figure 3.1). In total 100 patients were recruited and 71 took part in the study. The mean age of the study participants (males = 21, female = 50) was 34.9 years (range 16 - 67) and the mean duration of the condition was 23 years (3 – 60 years). The study participants experienced sweating in different areas: axillary plus other (n = 24), generalised (n = 19), palms and feet (n = 13) (Table 3.1)

Data saturation and Qualitative themes

Saturation of data was achieved at the 33rd study participant. Because of the novelty of the patient recruitment strategy employed, further interviews were conducted as well as data collection using another mode of data collection. The data analysis identified 103 HRQoL issues, grouped under 26 themes. These reflected seven main areas of QoL impact including daily life (mentioned by 95.8% of patients), psychological life (91.5%), social life (90.1%), professional life (74.6%), dealing with the condition (74.6%), unmet health care needs (64.8%) and physical impact (53.5%) (Table 3.2). Mean number of themes reported per person was 11 (2-23 themes).

No statistically significant differences (based on χ-test) were observed in number of themes reported by females and males, or among participants reporting different affected body areas.

For the purpose of clarity, the study findings will be reported under two sections, Part 1, concentrating on the experiences of hyperhidrosis patients, covering the issues that impact on their QoL. Part 2, reports on the development of the Hyperhidrosis Quality of Life Index (HidroQoL) based on the data collected from the qualitative data collected from patients.

Figure 3.1: Overview of data collection process

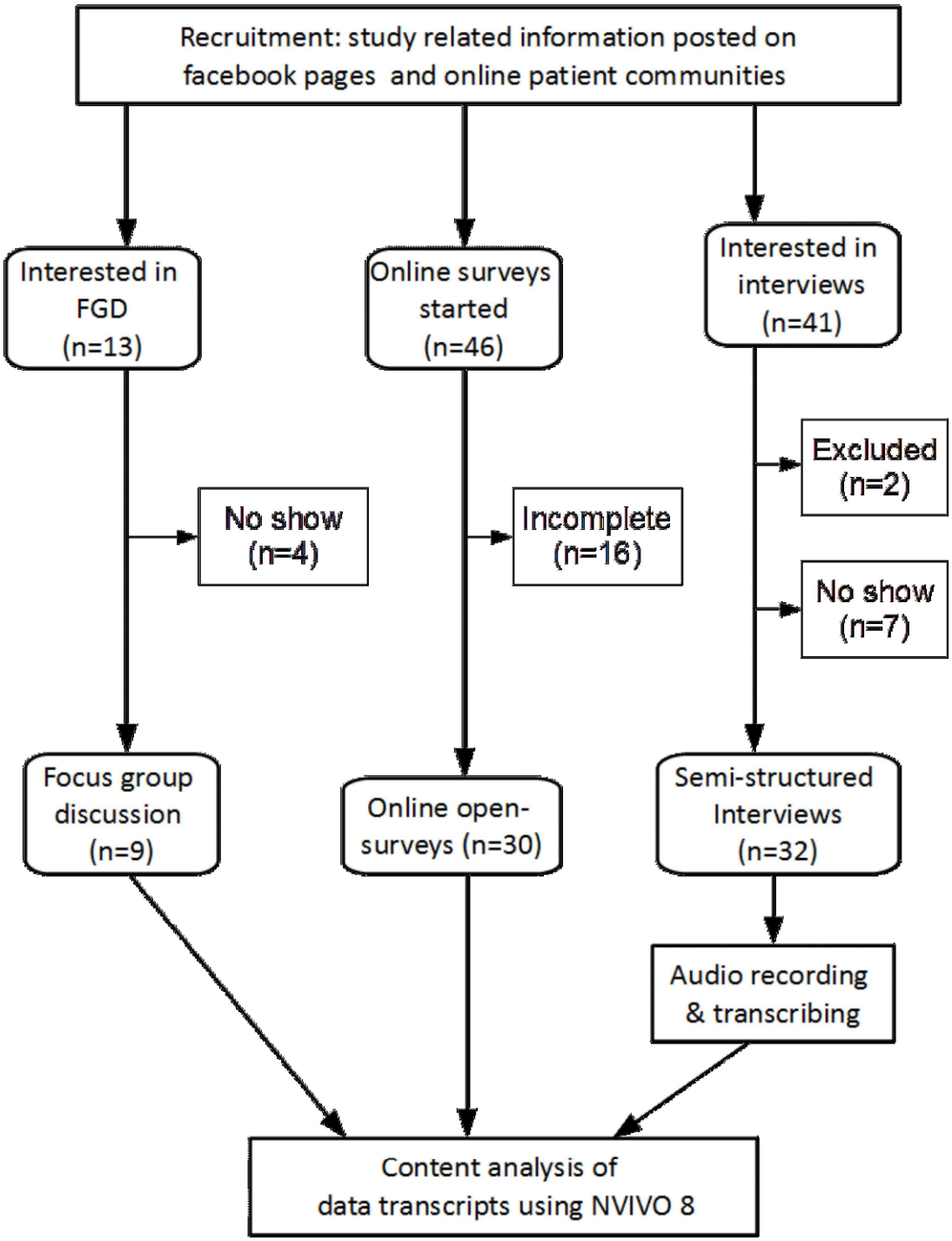


Table 3.1: Sociodemographic characteristics of study participants

Gender, n (%)	
Male	21 (30%)
Female	50 (70%)
Age (years)	
Mean (SD)	34.9 (13.2)
Median	31
Range	16 - 67
Duration of disease (years)	
Mean (SD)	23.3 (13.6)
Median	21
Range	3 - 60
Body site affected (n)*	
Armpits only	3
Palms only	2
Feet only	3
Armpits plus other sites	24
Palms and feet	13
General (whole body)	19
Face	3
Trunk & lower body	7
Country	
UK	41
U.S.A	14
Canada	2
Netherlands	2
Other	12

Part I: HRQoL Issues In Hyperhidrosis

Qualitative analysis of data collected from patients through interviews, focus groups and the open-ended survey questions identified seven main areas of HRQoL affected by hyperhidrosis: *daily life activities*, *psychological functioning*, *social life, managing the condition*, *professional life, physical discomfort*, and *unmet medical needs*. The thematic analysis identified a total of 104

issues (sub-themes) which were subsequently organised under 29 themes grouping together similar issues.

Daily Life Activities

Life Style

As a long term condition that patients live with on a daily basis immediate impacts on the patient's way of life were reported. Nearly three-quarters of study participants reported lifestyle impacts making this the most prevalent theme.

Table 3.2: Issues considered under each HRQoL themes

HRQoL area	Theme	%	Issues (subthemes)
Daily life (96%)	Touch technologies	21.1	Touch technologies
	Hobbies	40.8	Sports
			Recreational activities
	Limitations – general activities	39.4	Transacting with money (notes & coins)
			Sitting for long periods
			Light movements (e.g. dressing up)
			Holding objects
			Driving
			Doing physical activities
			Activities involving barefoot
	Lifestyle	76.1	Travel & holidays
			Choice of clothing
			Choice of footwear
			Appearance e.g. hairstyle & make-up
			Food choices
	Limitations – home	32.4	Daily household chores (cooking, cleaning)
			Shopping
			Caring for children
			Handling pets
	Summer-activities	32.3	Activities typical for summer months

Table 3.2 (continued)

HRQoL area	Theme	%	Issues (subthemes)
Psychological-functioning (92%)	Self-image	49.3	Think about sweating all time
			Low self-esteem
			Low self-confidence
			Looking untidy
			Feeling less attractive
			Self-conscious
	People's judgement	64.8	Feel stigmatised
			Fear reactions of others to my sweating
			Worry about leaving sweat marks on things
			Fear that sweating will be noticeable (e.g. avoid raising arms)
	negative emotions	69	Feeling nervous
			Feeling embarrassed
			Feeling depressed
			Feeling sad
			Feeling lonely
			Feeling hopeless
			Feeling frustrated
			Feeling angry
	health concerns	15.5	Fear that condition is worsening
			Worried that something is medically wrong
			Worried of passing condition to offspring
	Restricted life	50.7	Sweating greatly hinders my life
			Inconvenience
			Taken over my life
			Negatively affects my satisfaction with life
			Controls my thoughts
			Influences all my decisions
			Fear of doing new things
			More passive – laid back

Table 3.2 (continued)

HRQoL area	Theme	%	Issues (subthemes)
Social functioning (90%)	Relationships	39.4	Am a virtual recluse
			Stops one from having friends, partner
			Negatively affects personal relationships
			Interferes with communication
	Physical contact	66.2	Avoid holding or shaking hands
			being near others e.g. sitting, queuing, dancing
			physically expressing affection e.g. hugging and cuddling
	Interacting with people	57.7	meeting new people
			Not able to socialise with others.
	Being in public	29.6	Being in public e.g. going out
			Travel on public transport
Professional Life (75%)	Career	32.4	major career decisions e.g. choosing a career or retiring
	Work-tasks	63.4	Reduced performance at work
Dealing with the condition (75%)	Special chores	60.6	Wearing additional layers of clothing
			Controlling the sweating or keeping dry
			Aides to assist in carrying out daily activities
			Carrying spare clothing and towel (e.g. handkerchief, kitchen towels)
			Actively disguise sweat
	Personal hygiene	16.9	Shower several times a day
			Change clothes several times a day
	Financial burden	12.7	Financial burden
	Time concerns	29.6	Take things at a slower pace
			Spend more time in daily body hygiene
			Can't do things spontaneously (i.e. need to plan in advance for everything)

Table 3.2 (continued)

HRQoL area	Theme	%	Issues (subthemes)
Unmet health care needs (65%)	Poor management of condition	57.7	Treatments not working
			Unhappy with how doctors treat condition
			Disappointed with poor access to treatment
	Information needs	32.4	Knowledge of health care professionals about condition
			Patient information not adequate
			Lack of public awareness
Physical impact (54%)	Skin problems	16.9	Sore and cracked skin
			Painful skin due to soreness
			Skin infections e.g. athletes feet, dermatitis.
	physical discomfort	36.6	Cold sweating
			Constantly sweating
			Constantly feeling hot
			Sweat dripping into eyes
			body and clothes wet
			Slide in and out of shoes
	Body odour	14.1	Feet
			Body & clothes

Patients (61%) frequently mentioned effect on their choices of fabric, colour and design of their clothing. Many reported avoiding colours as 'red' or 'blue', favouring 'black' or 'white'. Others would stick to cottons, staying away from fabrics like polyester, which do not allow good aeration.

- *"You decide your whole wardrobe of clothes around that one thing...rather than thinking oh that's a nice fashionable type of thing you know...so you end up wearing black a lot"*
- *"...I have to wear black ...if I wear white it's going to show up and it's the most embarrassing thing ever"*
- *"I never wear a skirt. I would wear either short trousers or linen trousers coz the top of my legs and the middle of my legs rub as well...and I get blisters"*

- *Take care not to wear clothes that are too thick, try to wear thinner layers so that they be removed easily.*

Several explanations were offered for these effects including: physical discomfort and; because of concerns over whether other people would notice the sweating. Thus, choice of clothing was among the various strategies employed in managing the sweating.

- *"You have to watch what clothes you wear...obviously to try and hide the perspiration coz no matter what you did...you clothes were marked...it was really restricting as far as where going out socially was concerned...and you are always aware...it could be a smell.."*

Choice of shoes was also affected: particularly in patients with sweating of the feet. One patient said: *"every time I wear flip-flops it is impossible for me to walk without slipping"*. A small number of respondents (12.7%) reported effects on their food choices, they avoided spicy or hot foods, drinks containing alcohol or caffeine. Other less prevalent issues, included interferences with holiday decisions (choice of activities or destination); effects on appearance, a number of female participants reported 'make-up melting away' or 'hair getting messy'.

Hobbies

Forty-one percent of participants reported effects on their hobbies. Frequently mentioned activities included sports and recreational activities like playing musical instruments. One participant talked of their 'pastime reading' being affected because they can't hold a book. In as much as these primarily relate to the physical challenges of sweating for instance wet hands, there was still an element of concern about what others thought of the sweating.

- *... I don't like exercising on the street or anything like that...like going for a walk people tend to look at me if am really sweaty ...and that makes me really nervous...*
- *"..Cycling I used to enjoy...walking I will do...but I'll only do short distances...if I get too hot and bothered I sort of give up...coz it pushes my boundaries...I used to like ice skating...can't do that anymore...swimming is the only one probably am comfortable with...because its wet...I can't go in Saunas...any Gymn activities....unfortunately I put on weight in the last 6 years... I want to lose it...I don't want to go to the Gymn for people to comment...so I find it a struggle...to do those sort of things"* (Female, 27yrs)

General activities

Forty percent of the subjects reported difficulties with everyday life activities. The most commonly mentioned included trouble with holding objects, turning door knobs, opening jars, working with hand tools and driving, which were linked to sweating of the palms. Doing manual work, dressing up, and activities involving being barefoot were reported as being uncomfortable. A few participants mentioned avoiding making payments using cash or notes as they tend to stick to the palms which resulted in embarrassment.

- *"...when the temperature rises above about 22 degrees [celcius] I cannot do anything except sit still indoors with a fan running, this helps but does not stop it completely. I can't go outside, walking or do gardening, in fact anything that involves movement of my body, even moving my arms like using a whisk when cooking, or sewing affects the complaint, even in the winter if I am active, like walking. This all means that I cannot participate in anything physical as perspiration just pours from my head and neck and runs down my face and soaks my clothes, my hair looks like I have just washed it as it is wet all over"* (Female, 67yrs)
- *" I don't drive because of hyperhidrosis....with the steering wheel,...that has been a huge effect...not so much living in London...where...I doubt I would drive anyway...but living in Australia, New Zealand, you need to drive I'd make any excuses not to drive...because it just seems so difficult"*
- *"Simple everyday tasks become nonstop worries – holding an object, opening a jar or writing with a pen. I even found driving a car a test, my hands would leave the steering wheel soaking wet, making it very slippery, and subsequently dangerous"* (Male, 24)

Activities at home

Nearly a third of the study participants reported challenges in their life at home. Tasks such as cleaning, cooking, ironing and other household chores were affected. Two mothers mentioned challenges with caring for young children.

- *"The minute I start to do anything the least strenuous I stream with sweat, so housework is a nightmare. I have to change all my soaked clothes if I hoover one room. Have to tie strips of towel around my forehead and neck when I do anything that involves movement, and have to keep changing them as they get soaked'.*

Summer activities

The respondents reported that their sweating worsens during the summer months (in the US & Western Europe), activities that tend to be done during this season were affected. Nearly a third of the study participants experienced trouble with such activities, for instance, outdoor activities tending their gardening and mowing; or going to open air concerts. For some the mere approach of the summer season make them worried.

- *" I have a son who I'd love to take out in the summer, but sometimes we don't because I've had so many years of fearing summer"* (Male, 32 yrs)

Touch technologies

Constantly having hands that are wet, was reported to cause difficulties in using technologies that rely on touch. This was a problem in nearly a fifth of the subjects. Mostly, difficulties were experienced in the 'use of computer keyboards', 'laptop mouse', texting and sending short messages or general use of mobile phone, and using touch-screen interfaces. Two patients mentioned damaging their cell phones and keyboards because of the humidity from their hands.

Psychological Functioning

Emotions

Study subjects experienced various emotions as a result of their condition. This was mentioned by 69% of the participants and was the most prevalent psychological impact of hyperhidrosis. The sweating made the majority of subjects feel embarrassed; while a slightly lesser number reported feeling anxious about their sweating, and that others would notice it. Other participants said they were frustrated with life because of the condition. A handful reported feeling depressed. Less frequent emotions included sadness, anger, and hopelessness

A female participant shared her experiences as follows:

- *"...little things... when you are wearing a ring on your finger and people want to have a look so they grab your hand and you feel all embarrassed cause they are sweaty. Having your nails done, they are constantly working with your half and once again you feel embarrassed because you are sweating"* (Female)
- *"...the biggest problem is that it is horrendously embarrassing particularly if its problems with the hands, am quite an easy going type of person so if I meet someone for the first*

time or even if I meet someone whom I know I would like to shake their hand but I am very reluctant to do so, the first thing that goes through my head is...are my hands sweating..” (Male, 37 yrs)

Although negative emotions resulted from the sweating, the reverse was also seen, certain negative emotions lead to sweating.

- *“I'm on an antidepressant. I feel like I sweat because I'm nervous and I'm nervous because I sweat. It's a vicious cycle. I feel like an antidepressant helps to relieve that nervousness and helps to relieve a tiny percentage of my sweating”*
- *“It's like a catch 22 or what came first, The chicken or the egg?"Kind of like what comes first with us, the sweat or the anxiety? They both go hand in hand”*

People’s judgement

Concerns over how other people would react to the sweating was reported by nearly 64% of the study participants, making it the third most prevalent theme. A feeling that others misunderstood led to fears over how people would react. As such, participants often worried over how noticeable their sweating looked. Many feared leaving behind sweat marks on objects they came into contact with e.g. chairs, door knobs. On the most extreme, participants felt stigmatized, although this was not commonly reported.

Concerns over people’s reactions emerged as a key underlying theme behind other impacts experienced. For instance, impacts on hobbies; summer activities; feeling anxious and; embarrassed; impacts on social life were all linked to the perception of the judgment of others.

- *I can’t raise my hand all the way without showing my huge puddle of sweat...I can’t tell you how many times I heard “your back’s wet” from the person behind me. Just pure embarrassment”*
- *“if you are giving something to someone and you have sweat marks all over it... its nasty...and it’s not something that you can control...”*
- *“I remember at school when I would win a certificate in a subject, I knew I would have to go up on stage and collect it, just the thought of standing there with everyone looking, thinking that they might see me sweating, it was so upsetting!”*

- *"Sitting on a plastic chair or a leather one is a no-no in hot weather, or I would leave a sweat mark (sometimes it doesn't even have to be a hot day, it can be any day). And you can't see if your back sweat is showing through, so leads you constantly thinking that it is and everyone is laughing at you behind you, when you hear a laugh".*

Self-image

Hyperhidrosis also had an impact on the way participants viewed themselves, with nearly half of them reporting issues related to this theme. Many mentioned being preoccupied by their sweating condition. A heightened self-consciousness was commonly experienced; while some thought of themselves as being less attractive; or as dirty. On the extreme the participant's sense of self-worth was diminished leading to low self-esteem; and reduced feeling of self confidence.

- *"I am disgusted in myself for it and so it massively eats away at my self confidence, it makes me feel awful and dirty and gives me low self esteem - this has certainly been the route cause of my severe lack of confidence in everything I do or I am"*
- *"it makes me feel quite dirty even though I've had quite a few showers per day, even though I've changed my clothes three times a day....it makes me feel very depressed...and very alone"*
- *.well...it takes over what you have to think about before you leave home...whether you take extra clothing...whether you are wearing a black shirt...or t-shirt.....it constantly doesn't leave you...it affects your life all the time....it makes you feel very uncomfortable...and very aware...* (Male, 42 yrs)
- *Socially it made me very insecure...its embarrassing shaking hands and sweating through your clothes, when you're a girl its even worse.*

Restricted life

Half the study participants expressed feeling hindered by their sweating. Patients extensively described feeling their whole life is being held back, one patient said '*there has been countless things I haven't done because you first sort of thought ohh God that's going to place me in an uncomfortable position*'. Some patients talked of how they avoided any form of new challenges in their life. A handful of patients considered their life taken over by their sweating condition. One participant said '*everything you do in life you have to think the sweating will become a problem*.

Even for a handful who said they considered their sweating as something they can get on with, they nonetheless called it an inconvenience. This suggests that patients with hyperhidrosis perceive some degree of loss of control over their life.

- *"My life revolves around how sweaty I am that day. If I'm really sweaty, I stay inside and to myself. If I'm not too sweaty, I will more than likely go do something.*
- *I have to make sacrifices over this disorder* (Female, 18yrs)
- *I can't take subway after April, because its too hot on the platforms, so it takes twice as long to get anywhere. I don't own a car in NYC its more of a hindrance than a help .*
- *"I just don't want to go out...I wan be in the background I don't want anyone coming near me...its quite inhibiting"* (Female, 26 yrs)

Health concerns

The sweating resulted in concerns over their health for sixteen percent of the participants. Participants were worried that their condition was worsening.

Social Functioning

Participants mentioned experiencing disruption in their social life. To a large extent this reflected behavioural avoidance of activities or circumstances where embarrassment, anxiety feelings, self consciousness or other negative emotions might be experienced. This was also related to a diminished self-image in other cases.

Physical contact

Driven by concerns over how others would react to the sweat and fears of rejection, the majority of participants (57.7%) mentioned feeling uncomfortable with being in close proximity to others. For instance, touching others; holding or shaking hands; was avoided by many. Some participants avoided being close to others for instance when seated, queuing and; dancing. A smaller number had trouble with hugging, cuddling or any other forms of expressing affection.

- *"you distance yourself away from people. So they probably think I am a cold person... you cant expose yourself to the rest of them knowing that I've got this condition"* (Female, 36 yrs)
- *"was always too scared to dance with anybody because my back was always soaking wet...and the hotter I got the worse I got"* (Female, 58 yrs)
- *I am unable to touch my husband, daughters and grandchildren, without first thinking about how to do it without them actually having contact with my skin*
- *"You were very aware of it...when people would get close to you...it was really embarrassing...I didn't really like that...you were afraid that they would notice...that you were sweating*..(Female, 55 yrs).
- *"I suppose you don't really wanna get too close to someone if you are constantly sweating...they might feel very uncomfortable...you are so aware of your sweating and its not nice to be that close to someone*"...(Male, 42 yrs)

Interacting with people

The participants described how the condition inhibits their social interaction with other people, as expressed by 57% of the participants. A high self-consciousness; reduced self confidence and the fear of being judged by others made interacting with people in various social situations for instance weddings, seminars or when visiting friends, a challenging task. This also presented when meeting people for the first time, with a third struggling with this.

- *"[the sweating] affects my social competence because it is difficult focusing on a conversation when trying to hide sweating or thinking about how disgusting it feels against the body and having wet clothes".*
- *"when meeting new people I have a constant worry and fear of shaking hands. Do I shake their hand? Do I pretend I didn't see them offering the hand? Do I tell them I have sweaty hands? endless excuses echo around your head"*
- *"I sweat more if in social gatherings. Visiting, depending on whom, if my family then I'm comfortable and ask to put the fan on or drink my tea or coffee cold or iced. If with friends or people who do not know about my sweating disease then I will sweat more and flush badly"*

- *"The thought of meeting new people and having to shake their hand is terrifying! You just know it will be stone cold and sweating, then they look at you with a funny look!"*

Being in public

Furthermore, a third of the study participants described various situations in public for instance going out to a party, restaurant, cinema which made them uncomfortable. Some subjects had trouble using public transport. As with the other themes related to social life, self-image impacts and concerns over the reactions of other people was the underlying issue.

- *'there have been times where I've been very nervous about going to a party or work gathering fearing the sweating'.*
- *"travelling to and from work is mortifying. Sitting or standing in sodden clothes for 8 hours and travelling on public transport is horrendous"* (Male, 32 yrs)
- *"I can't go out... to like parties....or anything coz when I do...like... when I fix my hair or anything it will ... I would literary be blow drying my hair and I would be sweating all over"* (Female, 18yrs)

Relationships

The study participants mentioned finding support through friends and family who were understanding of their condition. Nonetheless, a much larger number (40%) reported that their condition had impacted their personal relationships. Avoiding going out and being in public meant that they would not be able to mix and interact with friends which is an ingredient to sustaining such relationships, loosing touch in the end. Low self-esteem also resulted in communication problems.

- *"I have not been in a relationship, as I feel too embarrassed to explain, I lost touch with most of my friends after school, because by this time, the sweating had got worse and they were wanting to go out, I would feel too anxious about it and would make an excuse".* (Female, 24 yrs).

Dealing with Hyperhidrosis

Special chores

The majority of participants (63%) employed various strategies in managing their condition. The burden of mitigating the symptoms and impacts of sweating presents an additional burden on the patients in their daily life. These were thought of as 'little rituals' by one patient, and included carrying a towel, tissue or handkerchief or; a pair of extra clothes for changing. Some patients said they needed to have a fan or air-conditioning running when at home or at the office; carry around a small hand-fan. Two participants mentioned drying up in the rest rooms when in public facilities.

- One woman said '*we become masters of disguise... I used to hide under layers of clothing even in the hot summers. Just to hide the sweat*!' while others reported putting on gloves in summer to hide the sweat.

Personal hygiene

Staying clean and maintaining body hygiene required extra effort, as indicated by some participants (17%). To stay fresh participants took several showers, changed clothes or shoes several times in a day.

- *"it makes you feel very unclean some times...you are constantly bathing three, four or five times a day".*

For those working or studying; this can be particularly challenging .as they have be out for the whole day.

- *'I have to wash my uniform each night and sometimes take spare set to work to change during the day'* (Female Nurse, 23 yrs)
- *'the fear that the sweat will start to smell...I tried to avoid long days at school as much as possible or at least have two shirts with me to school so I could change'* (Female Student, 17 yrs)

Time concerns

Dealing with the condition was time-consuming for the study participants. They required additional time mostly for various activities in managing the sweating including personal hygiene.

Moreover performing daily life activities (walking; dressing up) would usually require more time, as they took things at a slower pace to avoid the sweating from breaking out. Moreover, managing the sweating in and out of the home involved advanced planning, on things like clothing, which meant some loss of spontaneity.

- *One patient said 'I try not to be short of time to get to any appointment as rushing will cause a problem.* (Female, 50 yrs)
- *'You know like when I do take a shower I have to wait... for an hour to completely dry ...or relax...not to do anything... It's not nice at all'.* (Female, 18 yrs)
- *"it makes me feel quite dirty even though I've had quite a few showers per day, even though I've changed my clothes three times a day....it makes me feel very very depressed...and very alone* (Male, 28 yrs)"

Physical Impact

Physical discomfort

Some level of physical discomfort was associated with the condition; as reported by 40% of the study participants. Being in wet clothes day in day out was a concern. Further discomfort was associated with having wet feet, particularly in those with plantar hyperhidrosis. Those with facial sweating were annoyed with sweat dripping into their eyes.

- One subject said putting on shoes '*feels like paddling*'.
- *'my legs sweat, every inch of my body...it is so uncomfortable living in wet clothing. I have to some times change several times a day as it ends up smelling like a vinegar sweaty smell*'
-

Skin problems

Constant dampness made the skin vulnerable to other problems, as reported by 17% of participants. The commonly reported skin problems were soreness and cracked skin. Skin conditions such as hand eczema, Athletes feet, were also reported. A handful of patients experienced excessive sweating concomitantly with facial blushing.

- *I moved 150 miles away to a cooler part of the country. My skin often becomes so sore that it cracks and bleeds. I have constant chafing.*

Body odour

Fourteen percent of the subjects voiced concerns over body odour. This theme was directly related to being in public view and how the participants perceived other people's reactions to their condition; or body odour. Thus, they were more worried about the smell in situations where they were in close proximity to others or were in a enclosed public space for instance in a bus.

- "*The worst is the effect of unpleasant odour of my feet. I remember taking a bus ride and everybody noticed the offending smell. I try to avoid enclosed places like elevators, conference room, airport lounge*". (Male, 41 years)

- *"I find, perhaps because I am Asian, spicy food affect the sweating, that there is a horrible smell. Even when I cut back on problem foods my sweat has very distinct smell, acidic even. Its both the wet patch and the smell I worry about"* (Female, 27 yrs)

Unmet Medical Needs

Clinical management

How the study participants viewed the care they received was also an important theme for them, with 58% feeling that their condition had been poorly managed. Patients raised concerns related to their relationship with their doctor; the effectiveness of the treatments they received and; the side-effects associated with them. For many, obtaining a correct diagnosis of the condition was not easy; others felt their doctors neither gave them adequate attention nor showed any understanding. Further concerns were mentioned in relation to access to treatments: subjects frequently mentioned not findings a treatment that worked for them, this was particularly a problem in post-surgery treated patients. Respondents reported experiencing dry-mouth after oral treatment; compensatory sweating following surgery, with secondary impacts on their feelings and heart functioning; a bitter taste following iontophoresis.

- *'my GP [primary care doctor] didn't diagnose that I had hyperhidrosis or identify that it was a condition...even though I took a magazine article along...he sort of acted like it was rubbish... told me to apply more anti-perspirants...told me to get a stronger anti-perspirants...it was quite humiliating...'.*(Male, 32 yrs)
- *"what I would like to see over time are doctors who do understand, ..., who show compassion and do understand, and give us the time of day. this alone would take away*

from us that awful feeling of being alone and that no one understands. we all have come out feeling worse than walking in as they do not understand the whole of how we feel and how this affects us in our daily lives."

- *"getting treatment can be as bad as pulling a tooth with no Novocaine! Topicals burn and irritate your skin making it red raw...oral meds dry you out to the point you can't spit or pee and leave you in a dozy daze all day. Botox is like liquid gold, the one thing that does give us a break is the most hardest to get because of the greed of the Doctors costs and each vial. The iontophoresis machines are so expensive if not covered through heath care or insurance. ETS has left people worse than before with horrific side effects, the Dr's who did not tell the truth about how the surgery is really preformed!!!...Yet all the organizations claim how easy it is if you go to see a Dermatologist...It has taken me 12 years to finally get where I am and the frustration of not being understood, heard or fobbed off with another medicine because the side effect is less sweating but can cause you to overheat and be hospitalized is just in my opinion not bloody good enough!"* (Female, 42 yrs)

Information needs

Lack of information was also highlighted as an aspect of care that was reported as a concern to patients. Nearly a third of the study participants found currently available information inadequate. Patients pointed out their need for more information about their condition. During the data collection several participants wanted to find out what causes it. Worry was expressed over the level of knowledge of healthcare practitioners about hyperhidrosis. Others considered the lack of awareness about the condition in the general public was responsible for the lack of public sympathy for the condition.

- *'The lack of knowledge about hyperhidrosis among the medical community is also frustrating... I feel that it is difficult to find a provider that is knowledgeable about this condition because it is not a 'sexy' diagnosis, there is so little funding that goes into research for hyperhidrosis. I am very thankful for research such as this that allows for any insight into living with hyperhidrosis'.*
- *Even though the family doctor knew from an early age...likes 13/15...when I had gone to see him... he said I would grow out of it. ...felt really let down by the family doctor...ignorance probably* (Male, 42 yrs)

- *"[it is] a very difficult thing to explain to someone that...you're not just sweating because you are nervous or sweating because you have bad personal hygiene...or sweating because you're weird...but you've actually got a condition...you've actually got something that causes you to do that because its not a very publicly known condition...its not like you're saying to someone you've got cancer coz everybody knows what ...that is or you've got MS which everyone knows its sort of seen as something weird."*

Professional Life

Work-tasks

The study participants reported difficulties in their occupational life as a result of the excessive sweating. Performance of tasks at work or school was affected in 63.4% of the subjects, the majority of whom regarded this as the most important impact of the condition. Having wet palms presented a challenge in the performance of certain tasks: for instance any manual work requiring the hands; operating machinery; writing; and using a computer. For example, a participant working as a nurse found it hard to put on latex gloves or administer an IV. Some limitations were also attributable to concerns over how clients or co-workers would react. For example a pharmacist found it difficult to dispense medication which had sweat marks all over, due to her sweating.

- "*Jobs are difficult as you can imagine handing someone's change back wet, as they look at the beads of sweat on your hands reflecting under the light. or when your writing the paper sticks to your hand and smudges the ink".*
- *" I am a nurse, but I no longer perform patient care. When I was performing patient care, my hyperhidrosis interfered with numerous work tasks- simple things such as putting on sterile gloves, connecting IV tubing, examining a patient, etc. I was always able to work though the hand sweat, but it did cause me stress. Also, the constant hand sweat inside of gloves caused irritation and dermatitis on my hands. My hands were always red and raw. Since leaving patient care, I still have to be conscious of the hand sweat when filling out forms on paper, typing on the computer".*
- *"In recent months, in between my two Botox treatments, my arms were red raw for 4 days, the first two I couldn't lift my arms without being in excruciating pain. I was miserable at work and barely did any lifting e.t.c. and just said I had pulled a muscle. This prompted me to pay for the 2nd botox as i just could not continue like that. I smothered creams on for a*

few days and barely went out and then went for my Botox treatment. I had it 3 weeks ago now and i am back to being very confident and not having to worry about sweat marks, smells, i can put a little bit of deo on for a nice smell but not worry about it irritating my arms either

Career

Participants reported being influenced by their condition in making major career decisions. As mentioned by a third of the participants, they had chosen their career, then, just to accommodate their sweating. Some participants believed that they had not progressed in their careers due to the condition, while others had opted for early retirement because of their sweating condition. One participant said she had let go of an opportunity to become a policeman, which was her career of choice, settling for a 'really boring office job', because of her condition. One patient had opted for an early retirement due to the condition.

- *if I am to be absolutely honest...when I was at school I would have applied to do medicine...but because I knew I would have to examine people with my hands which I could never do...I opted to do pharmacy because I knew I would be able to do that'* (Female, 40 yrs).

Part II: Development of the Preliminary HidroQoL

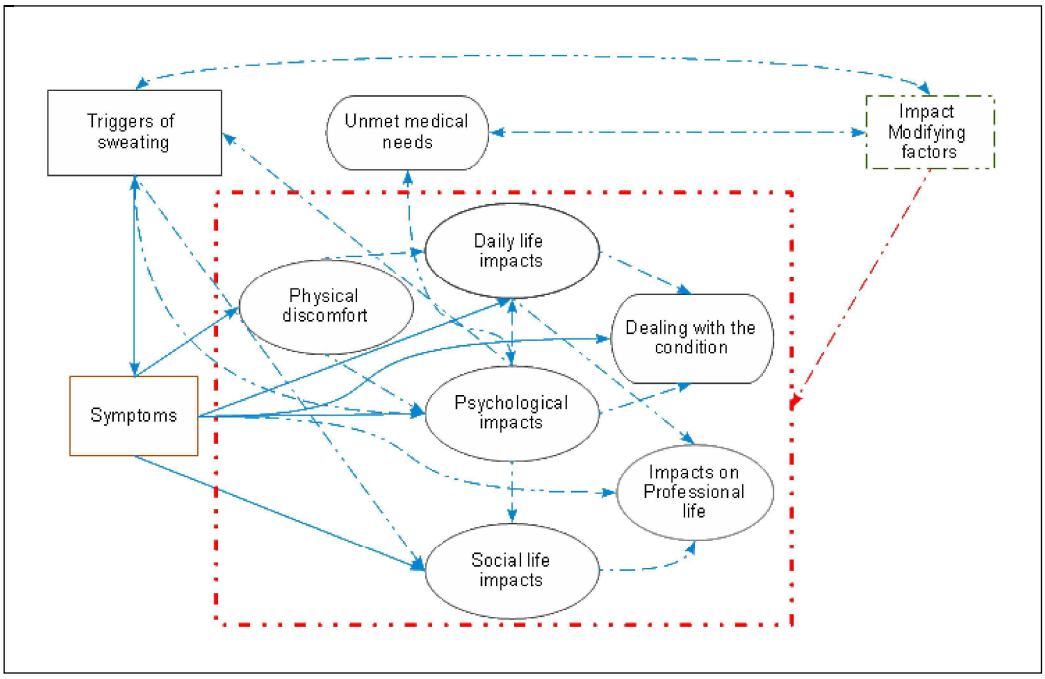
Conceptual framework

An important first step in the development of HRQoL instruments involves the development of a conceptual framework. This clearly defines the concept being measured; the rationale for undertaking the measurement; the target population; as well as the context of measurement (DeVellis 2011, pp.73-75). Skipping this critical step, has consequences for the clarity of the dimensional structure; may complicate subsequent data analysis and; may ultimately obscure interpretation of scale scores (Rothman et al. 2007). The conceptual framework of the new instrument presented in Figure 3.2 was based on results from the qualitative study undertaken in hyperhidrosis patients, reported in the previous section:

- The new instrument was developed for assessing hyperhidrosis-specific QoL. The main impacts of hyperhidrosis on QoL included physical discomfort, daily life impacts,

psychological impacts, social life impacts, impacts on professional life, and the burden associated with managing the condition.

Figure 3.2: Conceptual framework for the new QoL impact questionnaire for hyperhidrosis



Note: The core impacts of hyperhidrosis on QoL are enclosed in the dashed red box. Unmet needs considered outside the core QoL impacts is a results of the interaction between impacts of hyperhidrosis and available care given to patients.

- The target patient population for the new measure included persons with primary hyperhidrosis, including patients with different body areas affected.
- The new measure was to be used for evaluative purposes: evaluating the HRQoL of individual patients with hyperhidrosis, both in routine clinical practice and in research settings.

Initial item pool

Following content analysis of the data collected through interviews, focus groups and open surveys with hyperhidrosis patients, major QoL issues were identified. These were subsequently used to create the preliminary version of the instrument following the criteria listed below:

- Issues with a prevalence of 5% or more in the qualitative sample were included.
- Age or gender specific issues, regardless of being mentioned by less than 5% of the sample, were also included.

- Language considerations were made in crafting each item:
  - The wording of the item was meant to follow and reflect the language used by patients, technical jargons were to be avoided;
  - Readability was to be such that a 12-year old would understand;
  - The items were to be written in concise and simple sentences, aiming for six words or less;
  - Each item was to represent a single concept; not be double barrelled; and not use ambiguous words;
  - Items were to correspond to response formats (DeVellis 2011; Patrick et al. 2011b).

Response format

Once the items were drafted an appropriate response categorisation fitting the item stem, the concept under measurement, recall period and the mode of administration was chosen. This decision would involved deciding on type of scaling to use for responses, the corresponding number of categories, and their labels. While visual analogue scales may seem attractive for the range over which respondents can distinguish their condition; this very advantage may make the response task burdensome; and may also result in variation unrelated to the underlying condition of the patients. Seven, plus or minus two, has been recommended as maximum number of categories that people are capable of distinguishing (Streiner and Norman 2008, pp.48-49). On the other hand, caution ought to be exercised as offering people less choices than their capability to discriminate may lead to a loss of information (Streiner and Norman 2008, pp.48-49). Initially a 7-point Likert scale was used for the first prototype questionnaire with descriptors unique for each item. This was revised during the developmental process, to a 5-point Likert scale, with an option for 'not relevant'; with common descriptors used for all items. This was seen to strike a balance between applicability as well as need to offer sufficient choices and precision.

Frame of reference

Further, the period of time respondents needed to consider in producing answers, the recall period, was set. In turn, the wording of the items and responses and instructions would reflect this decision. The suitability of the recall period depends on measurement goals, for instance long-term impacts versus efficacy of intervention; the nature of the construct, symptoms or HRQoL impacts; frequency of assessments and; ultimately the target population (Norquist et al. 2011). The shortest

recall period feasible is recommended, a recall period that is too short may unnecessarily overburden the respondent; on the other hand exceeding one month may be associated with increased recall bias (Frost et al. 2007b). 'At present' was chosen as recall period for the new instrument. Responses based on the condition of the respondents during the time of assessment would be subject to minimal recall bias as the respondents would produce answers spontaneously, minimising noise in the measurement process.

Mode of Administration

The choice of how the instrument will be applied during data collection, whether in-person interview; telephone; paper and pencil; electronic; web-based tends to have an influence on data obtained (Frost et al. 2007b). Suitability of the mode depends on a number of factors: the preferences of the target participants and the construct under assessment; the content of the instrument for instance recall period; number and frequency of assessments among others (DeVellis 2011, p.189). Paper and pencil administration was chosen for the new instrument because of its ease of administration, making it easy and practical for routine clinical practice while avoiding 'social desirability' issues salient in modes such as in-person interview. Furthermore, to reach patient populations outside the clinic, it would also be administered via the internet, which may allow coverage of patients outside the clinic; besides other advantages for instance a stronger sense of anonymity for respondents.

Layout and structure

The structure of the instrument including its formatting is an important element of the instrument, with impacts on the accuracy and reliability of data collected (Haynes et al. 1995). For example formatting has potential implications, on navigational errors (such as item non-response and misinterpretation) and; respondent and administrative burden (Mullin et al. 2000). In order to ensure a simple, clear, consistent and natural design, the following decisions were taken:

- Items containing similar content were grouped together
- A light grey shading of 0.4 cm thickness was used to separate items, in order to reflect the responses that related to a particular item; grid lines were avoided.
- Tick boxes were provided for giving responses
- Responses categories followed a natural ordering from 'no, not at all' on the extreme left to 'very much' on the extreme right.

- Instructions were provided on what was being measured and the relevant recall period for participants to use in recalling their answers and how to choose responses. Instructions presented on the first page were circumscribed in a border to enhance their visibility. Instructions were also included on each page throughout the instrument.

The earliest proto-type of the new questionnaire contained a total of 75 QoL issues (Table 3.3). Subsequently this was re-organised to form the 47 item HidroQoL mostly by combining similar issues (Table 3.3) Its items were scored on a 5 point Likert scale with an additional 'not applicable' option.

DISCUSSION

Recognising patients as the experts in their personal experiences with their disease (Rothman et al. 2009), this study utilised a mix of qualitative methods to dig into the thoughts, perceptions and beliefs of patients with hyperhidrosis in relation to their experience of living with the condition, particularly to understand the extent and nature of QoL impacts. An additional aim of the study was to use the information collected from patients to develop a new instrument for assessing QoL impacts in hyperhidrosis. The findings obtained in this study are in accordance with previous published studies on the impacts of hyperhidrosis on patients HRQoL (Tan et al; Hamm et al; Solish et al.; Neumann et al.). The QoL impacts of hyperhidrosis were cross-cutting. Patients have previously also reported feeling that their life is taken over by hyperhidrosis (Thomas et al. 2006). For example, aspects of daily living including choice of clothing and relationships with family and friends have been reported to be affected (Thomas et al., 2006). In a study by (Solish 2006), respondents reported limitations when in public places (74%); meeting people for the first time (70.2%); developing personal relationships (58.5%). Patients mentioned feeling less confident than they would like (69.8%); frustration with some daily activities (58.2%); changing (41.6%) or reducing time spent (34.6%) on leisure and reducing time spent working. Patients reported being emotionally impaired (74%), having less confidence (74%), reduced work performance (63%), influences on career choice (42%); while a comparative control group registered no impairment in a study based at German University clinic (Hamm et al. 2006).

The HRQoL impacts of hyperhidrosis are comparable to those experienced in other chronic conditions. For instance, the condition had an influence on major life changing decisions (e.g. career choice) and location, which have been previously observed in psoriasis, cystic fibrosis or

diabetes (Bhatti et al. 2011). Impairment in dermatology-QoL was comparable to other skin conditions: the DLQI scores from patients with axillary (17 – 11.6) or palmar (18 – 9.1) hyperhidrosis were comparable to, or worse than those from patients with dermatitis (inpatient) (16.2) or psoriasis (13.9). Cina and Clase (Cinà and Clase 1999) found the lifestyle intrusiveness associated with hyperhidrosis to be worse than in other known chronic conditions, such as end-stage renal disease, rheumatoid arthritis or multiple sclerosis.

The use of qualitative methods in this study provided further insights beyond merely identifying QoL effects, but also the main factors influencing those issues were explored. This was highly useful for two main reasons, first in supporting the understanding of the interconnection between various impacts. For example, the fear of others noticing and judging the person lead to feelings of 'embarrassment', which in turn result in avoiding situations in which others would notice the sweating or they would feel the negative emotions, which essentially reflects the daily life and social life impacts. The same can be said about feelings of anxiety, experienced whenever the person sweats in the presence of others or; whenever they think their sweating is noticeable. Consequently, such situations are avoided. Furthermore, as in psoriasis (Magin et al. 2009),

Table 3.3: List of issues forming the initial 75-item instrument proto-type

1	Sweating influences my choice of clothing (e.g. design, colour or material)
2	I avoid exposing soaked clothing around the armpits area sweating (e.g. I avoid raising my arms)
3	I do activities at a slower pace due to the sweating (e.g. physical activities such as walking)
4	Sweating influences my choice of footwear
5	Holidays are less enjoyable because of sweating
6	I have trouble handling money with my hands because of the sweating
7	I have trouble giving care to children because of my sweating
8	I avoid certain foods e.g. spicy foods because they make me sweat (gustatory sweating)
9	I find it difficult to do hobbies that involve physical activities (e.g. walking, cycling, exercising, playing musical instruments)
10	Doing work-related activities is difficult (e.g. dealing with clients, caring for patients, working with tools)
11	Sweating restricts my life (e.g. stops me from travelling)
12	Sweating influences my career decisions (e.g. choice of work)
13	Handling paper documents and writing is difficult because of my sweating
14	I avoid outdoor activities (sun-basking or gardening)
15	I have trouble using hand operated electronics due to my sweating (e.g. computer keyboards, cell-phone, touch-screens)
16	My sweating makes shopping difficult
17	Activities involving walking barefoot are difficult because of my sweating

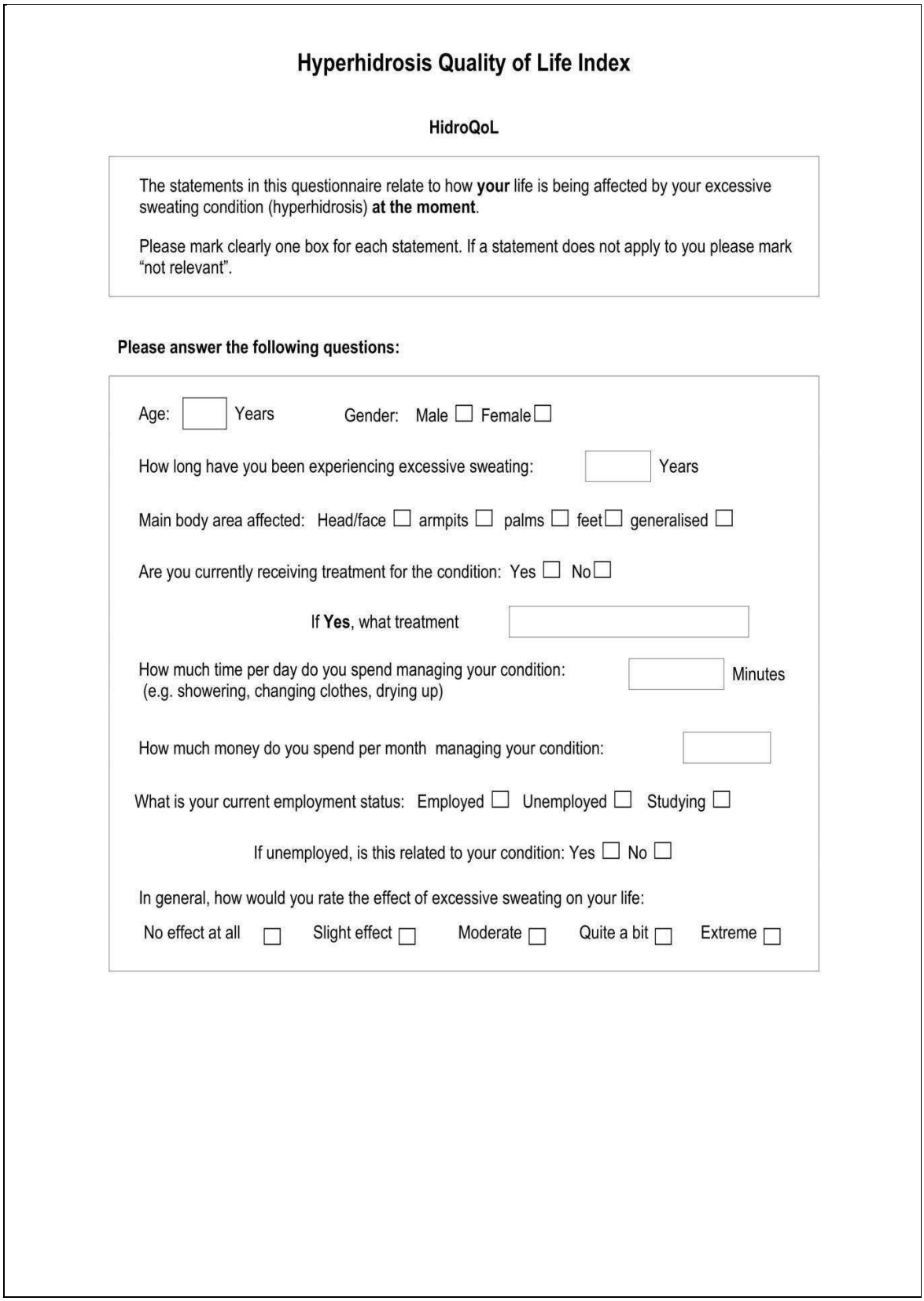
Table 3.3 (continued)

18	I dread holding or shaking hands with others
19	Sweating interferes with my personal relationships (e.g. with friends or partner)
20	I feel embarrassed because of the sweating
21	I can't socialize as much as I would like to
22	I am afraid of meeting new people
23	I fear speaking to groups of people because of my sweating (e.g. doing presentations, meetings, interviews)
24	I avoid going out (e.g. to parties, eating in restaurants)
25	I am a virtual recluse because of my sweating
26	I can't find a treatment that works for me
27	My doctor does not understand my condition
28	Adapting to the sweating is difficult (e.g. maintaining body hygiene, need to keep fan or air condition on)
29	I disguise my sweating (e.g. wear gloves, jacket, socks)
30	I carry spare clothes or towel with me because of my sweating
31	I fear that my sweating will be noticed by others
32	I look untidy
33	I change clothes...
34	I shower…
35	I feel less attractive
36	I can't wear a hairstyle or make-up of my choice
37	Sweating makes me feel nervous
38	Sweating has taken over my life
39	I fear doing new things because of my sweating
40	I feel hopeless
41	My sweating makes me feel sad
42	I feel miserable because of my sweating
43	I dread summers because of the sweating
44	I fear that my sweating is worsening
45	I think about sweating ...
46	My self-esteem is low because of my sweating
47	I feel less confident because of my sweating
48	I am emotionally drained because of the sweating
49	I feel more self-conscious because of my sweating
50	Sweating makes my sexual life less enjoyable
51	I fear leaving sweat marks on objects
52	I have trouble being in crowded spaces because of my sweating (e.g. in bus or train)

Table 3.3 (continued)

53	I am drenched in sweat (e.g. my clothes are wet)
54	My sweating is physically uncomfortable
55	Light movements make me sweat (e.g. getting dressed)
56	I slide in and out of my shoes
57	Sweat gets into my eyes
58	My feet give an unpleasant odour
59	It is difficult to grip objects in my hands because of my sweating (e.g. tools, door knobs)
60	I am afraid to physically express affection because of my sweating (hugging and cuddling)
61	I avoid getting close to people (when sitting, queing, dancing)
62	I feel that others judge me because of my sweating
63	I feel depressed because of my sweating
64	I fear rejection from others because of my sweating
65	My sweating makes housework difficult (e.g. cleaning, cooking)
66	My sweating exerts a financial burden on my life
67	Casual walking makes me sweat
68	My skin is sore and cracked because of my sweating
69	I can't do things spontaneously
70	Sweating makes driving difficult
71	Doing physical activities is difficult because of my sweating (e.g. manual work)
72	My body (or clothes) gives a bad odour because of the sweating
73	I sweat even in winter
74	I get other skin problems as a result of my sweating
75	I feel hot even in winter

Figure 3.3: The 47-item developmental version of the new instrument



Hyperhidrosis Quality of Life Index

HidroQoL

The statements in this questionnaire relate to how **your** life is being affected by your excessive sweating condition (hyperhidrosis) **at the moment**.

Please mark clearly one box for each statement. If a statement does not apply to you please mark "not relevant".

Please answer the following questions:

Age: ☐ Years Gender: Male ☐ Female ☐

How long have you been experiencing excessive sweating: ☐ Years

Main body area affected: Head/face ☐ armpits ☐ palms ☐ feet ☐ generalised ☐

Are you currently receiving treatment for the condition: Yes ☐ No ☐

If **Yes**, what treatment ☐

How much time per day do you spend managing your condition: (e.g. showering, changing clothes, drying up) ☐ Minutes

How much money do you spend per month managing your condition: ☐

What is your current employment status: Employed ☐ Unemployed ☐ Studying ☐

If unemployed, is this related to your condition: Yes ☐ No ☐

In general, how would you rate the effect of excessive sweating on your life:

No effect at all ☐ Slight effect ☐ Moderate ☐ Quite a bit ☐ Extreme ☐

Hyperhidrosis Quality of Life Index

The statements in this questionnaire relate to how **your** life is being affected by your excessive sweating condition **at the moment**.

Please mark clearly one box for each statement. If a statement does not apply to you please mark "not relevant".

		Not at all	A little	Some what	Quite a bit	Very much	Not relevant
1.	My choice of clothing is affected	☐	☐	☐	☐	☐	☐
2.	My choice of footwear is affected	☐	☐	☐	☐	☐	☐
3.	My holiday is affected	☐	☐	☐	☐	☐	☐
4.	I have difficulties gripping objects	☐	☐	☐	☐	☐	☐
5.	I have difficulties handling money	☐	☐	☐	☐	☐	☐
6.	I have difficulties with physical contact with	☐	☐	☐	☐	☐	☐
7.	My hobbies are affected	☐	☐	☐	☐	☐	☐
8.	I avoid speaking with groups of people	☐	☐	☐	☐	☐	☐
9.	My physical activities are affected	☐	☐	☐	☐	☐	☐
10.	My outdoor activities are affected	☐	☐	☐	☐	☐	☐
11.	My everyday housework is affected	☐	☐	☐	☐	☐	☐
12.	I find it hard to handle paper	☐	☐	☐	☐	☐	☐
13.	My career decisions are affected	☐	☐	☐	☐	☐	☐
14.	My work is affected	☐	☐	☐	☐	☐	☐
15.	I have difficulties using touch-technologies (e.g. computer-keyboard, smart phones)	☐	☐	☐	☐	☐	☐

Please turn to the next page

Hyperhidrosis Quality of Life Index

The following statements relate to how **your** life is being affected by your excessive sweating at **the moment.**

		Not at all	A little	Some what	Quite a bit	Very much	Not relevant
16.	My relationships with others are affected	☐	☐	☐	☐	☐	☐
17.	I feel embarrassed	☐	☐	☐	☐	☐	☐
18.	I do not socialize as much as I would like to	☐	☐	☐	☐	☐	☐
19.	I avoid meeting people	☐	☐	☐	☐	☐	☐
20.	I avoid going out	☐	☐	☐	☐	☐	☐
21.	I feel nervous	☐	☐	☐	☐	☐	☐
22.	I feel hopeless	☐	☐	☐	☐	☐	☐
23.	I feel sad	☐	☐	☐	☐	☐	☐
24.	I feel depressed	☐	☐	☐	☐	☐	☐
25.	I feel frustrated	☐	☐	☐	☐	☐	☐
26.	My confidence is affected	☐	☐	☐	☐	☐	☐
27.	My self-esteem is affected	☐	☐	☐	☐	☐	☐
28.	My whole life is affected	☐	☐	☐	☐	☐	☐
29.	Sweating is constantly on my mind	☐	☐	☐	☐	☐	☐
30.	I avoid taking on new challenges	☐	☐	☐	☐	☐	☐
31.	My summer activities are affected	☐	☐	☐	☐	☐	☐
32.	I feel more self-conscious	☐	☐	☐	☐	☐	☐
33.	My appearance is affected	☐	☐	☐	☐	☐	☐

Please turn to the next page

Hyperhidrosis Quality of Life Index

The following statements relate to how **your** life is being affected by your excessive sweating at **the moment.**

		Not at all	A little	Some what	Quite a bit	Very much	Not relevant
34.	I feel uncomfortable physically expressing affection	☐	☐	☐	☐	☐	☐
35.	I worry about people's reactions	☐	☐	☐	☐	☐	☐
36.	My sex life is affected	☐	☐	☐	☐	☐	☐
37.	I worry about leaving sweat marks in public	☐	☐	☐	☐	☐	☐
38.	I worry being in places close to other people	☐	☐	☐	☐	☐	☐
39.	My eating habits are affected	☐	☐	☐	☐	☐	☐
40.	I slide in and out of my shoes	☐	☐	☐	☐	☐	☐
41.	I have problems with being barefoot	☐	☐	☐	☐	☐	☐
42.	My eyes get irritated	☐	☐	☐	☐	☐	☐
43.	I feel my skin is hot all the time	☐	☐	☐	☐	☐	☐
44.	I worry about the extra demand on my finances	☐	☐	☐	☐	☐	☐
45.	I find it difficult to cope with my condition	☐	☐	☐	☐	☐	☐
46.	I find it hard to do things without planning in advance	☐	☐	☐	☐	☐	☐
47.	I feel that I need more time for hygiene chores	☐	☐	☐	☐	☐	☐

Please check that you have answered all questions

Thank you for your help!

avoidance of activities that would put the patient on the spot light, making the sweat more noticeable and; situations where self-consciousness, embarrassment and anxiety might be experienced explained most social life and life style impacts. For multi-factorial skin disorders, the psychological stress of living with the condition tends to exacerbate the condition (Beltraminelli and Itin 2008), which seems to be the case in hyperhidrosis. (Park et al. 2010) found that anxiety, stress and social relationships to be more influential aggravating factors for hyperhidrosis ahead of heat or summer season based on a Korean sample attending a university hospital centre. On the other hand, heighted concerns over how people would react to the sweat at the core of most impacts of the condition seems strongly rooted in a 'perfect body image' perpetuated in modern society (Beltraminelli and Itin 2008).

For the first time the findings of this study revealed perceptions of the patients on the management of hyperhidrosis. The patients perceived general practitioners as lacking knowledge on the condition; not showing empathy and understanding in their interaction with the patients. Such experiences were reported across the entire continuum of care, from diagnosis to treatment decision-making. Similar observations have been made in other skin conditions (Nelson et al. 2012). Further concerns of the patients such as the need for quality information on the condition; the lack of public awareness, leading to a lack of public sympathy have been previously reported (Golics et al. 2009) Managing excessive sweating and its impacts can be quite demanding in terms of time, effort and money. Patients spend 15 to 60 minutes in managing symptoms of the condition, 50 – 70% of patients change their clothes more than twice a day (Hamm et al. 2006), while slightly more than a fifth of patients have been noted to use some aides to make it possible to carry out their daily life as normal (Strutton et al. 2004). In this study patients reported carrying around towels, handkerchiefs for drying up; having a fan or air conditioner on when at home or at work; or carrying around a handheld fan or even using hand-dryers in public restrooms when on the go. Still, a good part of dealing with the condition involves disguising or concealing the sweating.

The complexity of QoL impacts of hyperhidrosis, particularly their interrelations and strong social rooting, has implications for their measurement. First, instruments that focus on severity of physical symptoms may underestimate the impact. Moreover, subjective experience of patients may be more relevant and important in determining QoL than the degree of objective severity (Russo et al. 2004; Jobling and Naldi 2006) such that even mild severity may still result in major

QoL impacts. On the other hand, use of a generic HRQoL instrument, in spite of their broadness might miss out on some issues specific to hyperhidrosis patients, while including some irrelevant materials. This calls for a measure that would reflect the appropriate emphasis on issues most important for hyperhidrosis patients while concurrently excluding irrelevant materials.

The second part of this study involved developing a new instrument for assessing the QoL impacts of hyperhidrosis, based on the information reflecting the experiences and perception of the patients. The implications of the approach taken were multi-fold: first, this ensured that the instrument being developed had high applicability for the intended measurement purpose and acceptability in the target patient population (patients with hyperhidrosis). Second, the involvement of the target patient population in item elicitation was essential to the content validity of the new measure (Rothman et al. 2009). Ultimately, this reflects the essence and nature of measures of QoL-impact as a vehicle for patients to express their voice in relation to the impacts of their condition on all aspects of daily life (Basra and Shahrukh 2009).

Drug regulatory authorities such as the FDA require documentation on the process followed in the development of PRO instruments used in making labeling claims, as evidence for content validity (US-F.D.A. 2009). This points towards the need for an organized and well thought out development process. The development of the new instrument, therefore, involved: first, defining measurement aims, the target population, the construct and level of specificity with which it would be measured; second, setting criteria to guide the selection of issues and drafting of content; and finally the actual drafting of the questionnaire.

The new instrument was intended for assessing impacts of QoL on individual patients in routine clinical practice and in clinical research. The target patient population includes all forms of hyperhidrosis, based on body-area affected. The construct was being measured at a level generic enough for the items to have relevance to all forms of hyperhidrosis. The items reflected aspects of QoL affected by hyperhidrosis based on the personal feelings and perceptions of the patients. Response categorization was chosen to reflect different levels of impairment in the concepts addressed in each item. Instructions were written to be clear, highly visible and offer useful guide to the patients in the questionnaire completion process. Frame of reference was chosen to minimise recall bias and match the aspects of hyperhidrosis-QoL. Formatting decisions were made to realize

a simple, natural and organized design ensuring easy navigation, minimal respondent burden and an attractive questionnaire.

The data collection in this study benefited from triangulation of several qualitative data collection methods including focus groups, semi-structured interviews and online open surveys. During the FGD, interactions among participants helped with stimulating new aspects or topics of discussion, generating additional data otherwise not realisable (Patrick et al. 2011b). The interviews, on the other hand, provided in-depth and detailed information about an individual's experience besides the relative ease of arranging appointments with the patients (Patrick et al. 2011b). The surveys with open ended questions were the low-hanging fruit, as they could be implemented with relative ease, while providing a good balance between ability to reach large numbers of patients relatively easily while still allowing respondents to give detailed description of their opinions (Bowling 2009), however, lack of opportunity for probing as is the case in FGD or interviews my limit the depth of information provided.

SUMMARY

- This study has provided deep insights into the major issues influencing the HRQoL of patients with hyperhidrosis. Further, the unmet health care needs of relevance to the patients HRQoL were also identified, including treatment and information related issues.

- The HRQoL issues identified from the qualitative study, which are based on the patient's own words provided a rich source for developing the content of a novel hyperhidrosis-specific QoL questionnaire for assessing QoL impacts of hyperhidrosis, ensuring that the new measure was indeed appropriate and had the right emphasis for the target patient population.

- The structured process followed in the development of the new instrument, including the development of a conceptual framework; having a clear criteria for the content; and subsequently drafting the content of the instrument in line with the criteria, further enhanced the appropriateness and suitability of the new measure for hyperhidrosis patients. Enhancing the new measure's content validity, acceptability and practicality

CHAPTER 4

Development of a Hyperhidrosis-Specific Health-Related Quality of Life Instrument: Content Validation

INTRODUCTION

Evidence supporting a particular interpretation of scores reflects claims that a PRO instrument measures what it purports to measure i.e. that it has construct validity (Lohr 2002). Crucially, such evidence is based on the presumption that the observable indicators are related to the underlying construct being assessed which ought to be established through content validation. Formally, the content validity of an instrument reflects the extent to which it represents the most relevant and important aspects of a concept in the context of a given measurement application (Magasi et al. 2012). Ensuring content validity, therefore, requires that the content domain is adequately sampled suggesting a rigorous instrument development process (Nunnally and Bernstein 1994, p.102). Similar to other psychometric properties, content validity relates to particular measurement aims, usage, construct being assessed and target population, hence the need for their clear articulation (Terwee et al. 2007). In particular, experts judge the appropriateness and relevance of the content in relation to the construct being measured and the considerations listed above (Streiner and Norman 2008, p.252). This process, then, is the first 'proof of concept' that the instrument's content is connected to the construct being measured. Without such evidence, construct validity and the meaning attached to the instrument's scores (interpretability) may not be established (Haynes et al. 1995). On the other hand an instrument demonstrating content validity is more likely to reflect strong construct validity, interpretability besides superior acceptability and practicality as it would tap into the most relevant issues for both the construct and patients, also rendering the measure more interesting to patients. The qualitative study in the previous chapter (chapter 3) reported on how QoL issues were collected from patients; the conceptualisation of HRQoL in hyperhidrosis; the subsequent transformation of QoL issues into content of the instrument providing initial evidence of content validity. In this study, content validity was formally assessed.

OBJECTIVES

The objectives of this study were to:

- Assess whether the content of the new instrument was relevant to patients with hyperhidrosis and the concept of quality of life in hyperhidrosis.
- Assess the adequacy with which the new instrument represents the concept of hyperhidrosis quality of life.
- Assess the appropriateness of the layout, recall period and technical quality of the new instrument for assessing hyperhidrosis quality of life.

METHODS

The scope of an instrument's content considered during content validation studies extends beyond just items and their responses. All elements of the instrument that would influence the data collected are included (Haynes et al. 1995). Responses and collected data may be influenced for example by the structure of the instrument (the instructions, response formats, frame of references), technical quality of the measure, apart from the relevance of the content to a given patient population (Patrick et al. 2011b). This study, therefore, considered the following aspects:

- The Layout: the graphical structure and design must lend themselves to a natural flow through the questionnaire (Mullin et al. 2000) including pagination, font size and font styles.
- Instructions provide important orientation to the patient regarding what is being measured, the frame of reference to apply when providing responses, and how to choose between response categories (Patrick et al. 2011b). Thus, the need that they are adequate, clear and appropriately located.
- Frame of reference: this defines the period of time patients need to refer to when providing their responses, the recall period (Norquist et al. 2011). It has to be suitable for the construct being measured (HRQoL impact), the characteristics of the disease, the treatment and duration of treatment-effect, the intended number of assessments and the target population (Norquist et al. 2011).

In addition, individual items were evaluated on the following criteria:

- Language-clarity: The sentence and wording of each item should be clear, understandable, straightforward and simple. Phrases and wording should be unambiguous and jargon free and should be understood by someone with a reading ability of a 12-year old (Streiner and Norman 2008).
- Completeness: The sentences should be complete, not broken and should end appropriately, comprehensively addressing the idea they are covering (Guyatt et al. 1993)
- Relevance: Each item should reflect an aspect of HRQoL of importance to the target patient population, thus also relevant to the construct being measured (Leidy et al. 1999)

- Scaling: This represents how the actual responses of the patients will be measured. The choice of the response format, number and labelling of response categories must fit the items and be appropriate to the construct being measured (DeVellis 2011).

Panels

Five dermatologists located at various centres in Germany, were invited to participate in the content validation of the hyperhidrosis. They were all multi-lingual, and were leading experts in hyperhidrosis. Two were heads of dermatology departments at University clinics. The other two were active in clinical research related to hyperhidrosis. An invitation was also made to seven patients who had lived with the condition for at least 10 years and who were English native speakers through the International Hyperhidrosis Society (IHHS). A minimum of 3 experts is accepted for such panels, although having five or more reduces chance agreement (Lynn 1986). Two separate panels were conducted for the dermatologists and the patients.

Three experts in outcomes research were invited to participate in the review process following content validation. This included a professor in health outcomes measurement, two clinical researchers with experience in hyperhidrosis, one with a medical degree, and the other with a pharmacy degree. Both panels were chaired by the principal investigator (P.K.).

Materials

Each panel member was provided a copy of the developmental HidroQoL and the content validation questionnaire. The developmental HidroQoL was conceptualised as an instrument for assessing the impacts of hyperhidrosis on the quality of life of patients. The instrument included 47 items addressing all aspects of quality of life considered important to patients solicited during previous qualitative work in hyperhidrosis patients (chapter 3) and six response options were provided: ‘No, not at all’; ‘A little’; ‘Somewhat’; ‘Quite a bit’; ‘Very much’; and ‘Not relevant’. The content validation questionnaire evaluated each of the 47 items of the HidroQoL on four aspects: language clarity, completeness, relevance, scaling, previously defined. Each of these aspects were rated on a 4 point-Likert scale as 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree, for all items. Additional space was provided for open-ended feedback or suggestions for each item as well as the entire questionnaire.

Procedure

An invitation to participate in the content validation study was sent to the experts by email. Subsequently, a copy of the HidroQoL and the content validation questionnaire were sent. The

purpose and intention of the study including guidance on how the experts were to assess the HidroQoL were included in the email communications and were also contained on the content validation form. Appointment dates for the panel session were also agreed after a number of email exchanges. Two sessions, for the dermatologists and the patients, were conducted in the form of a panel discussion, assessing the adequacy of all aspects of the HidroQoL. Decision-making during the panels was based on the consensus and agreement of all panel members. Completed content validation questionnaires were returned by the experts either before or at the beginning of the panel sessions. Decisions required the agreement of all panel members.

Data Processing and Analysis

Panel sessions were tape recorded and later transcribed. Transcripts were analysed for the major issues and decisions relating to each aspect of content assessed. Open ended feedback and suggestions on the content validation questionnaire were handled in the same way. Item ratings from the questionnaire were coded and then analysed using SAS Software, version 9.2 and MS Excel 2007.

Multiple approaches were used for quantifying and interpreting the ratings:

i) Mean item score was calculated for language clarity, completeness, relevance and scaling. A mean value of at least 3 is required for adequacy of an item, following previous work (Davidson 2003).

ii) Average deviation mean index calculated as standard deviation (SD) of individual item ratings was estimated to capture extent of disagreement on item ratings (Burke and Dunlap 2002). SD greater than 0.75 indicated disagreement (Davidson 2003).

iii) Content Validity Index (CVI) was calculated for each individual item and for the scale as a whole, for language clarity, completeness, relevance and scaling. An item CVI (I-CVI) is computed as the proportion of individuals giving an item a rating of 3 (agree) or 4 (strongly agree), while a scale CVI based on universal agreement (S-CVI/UA) is given by proportion of items receiving a rating of 3 or 4 from all raters (Polit and Beck 2006). Attaining endorsement requires a minimum I-CVI of 1 for a panel of 5 members or less and 0.78 where there are six or more (Lynn 1986; Polit and Beck 2006). Minimum threshold for S-CVI/UA of 0.8 was applied in judging validity at the scale level (Polit and Beck 2006)

iv) Finally, inter-rater agreement of the expert ratings was assessed at the scale level, using the Gwet's coefficient of agreement statistic, which is effective and free of the vulnerabilities of the multi-rater kappa coefficient (Gwet 2002). The Gwet coefficient

is interpreted in a similar way as the kappa coefficient: < .00 is considered poor; .00 – 0.2 is slight; .21 - .4 is fair; .41 - .6 is moderate; .61 - .8 is substantial, .81 – 1 is almost perfect (Landis and Koch 1977).

RESULTS

The results are presented under three sections. The first includes findings from the patient-expert panel, the second covers the dermatologist-expert panel and the third includes review panel which carried out the final revisions.

Part I: Patients-Expert Panel

Panel session

All aspects including the design and organisation of the instrument were thoroughly discussed during the panel session, in some instances, varied conclusions would be reached as recommendations. The general layout of the instrument, including the font-type, font-size and organisation of the instrument were considered appropriate and adequate. Instructions were thought to be understandable and helpful in completing the questionnaire. Suggestions were made on the questions related to patient demographics: i) to allow respondents to choose more than one dominant area affected ii) to add a question on previous treatments iii) to provide examples of activities patients needed to consider in addressing the question on additional money or time spent on hyperhidrosis, iv) to provide space where respondents would describe how the disease had impacted them, in addition to responding to the general impact question.

The panel had diverse views in relation to the frame of reference used as the recall period. They stated that ‘at the moment’ did not reflect their experiences. Members argued that their condition tends to fluctuate on longer time horizon than day to day. On the other hand, they said, the relief from treatments such as oral medications or iontopheresis tend to last over a couple of days to weeks. The panel suggested either not specifying time frame or to use ‘at peak when sweating is at its worst’ as recall period. In relation to response formatting the panel suggested reversing the initially proposed format (‘no, not at all’, ‘a little, somewhat’, ‘quite a bit’, ‘very much’, ‘not relevant’). One panel member argued that it seemed logical to have ‘no, not at all’ next to ‘not applicable’, than as initially arranged where ‘very much’ was placed next to ‘not applicable’. The suggested format was ‘very much’, ‘quite a bit’, ‘somewhat’, ‘a little’ ‘no, not at all’, ‘not relevant’. In order to assess whether the construct of hyperhidrosis-QoL was adequately covered by the developmental instrument, the panel was asked whether there were any gaps in the content or

whether they would make any additions on the items. They considered the HidroQoL to cover all important HRQoL issues for patients with hyperhidrosis, thus they suggested no additions.

Content Validation Questionnaire

All forty-seven items had mean scores of at least 3 for language clarity, completeness, relevance, scaling (Table 4.1). According to the Average Mean Deviation Index, item-level disagreement in the ratings was noted. The SD for language clarity ratings in 16 items exceeded 0.75 including for *I worry about being in places close to other people* (SD = 1.41), *I feel that I need more time for hygiene* (1.1), *I have problems speaking with groups of people* (1.1). Similarly, SD for completeness ratings for *I feel more self-conscious* (1.34), *I worry about people's reactions* (1.34) and *I worry being in places close to other people* (1.1) and an additional sixteen items were above threshold. Ratings for relevance and scaling had high SD in eight and two items, respectively. Six items had language-clarity I-CVI below 0.8, including '*my holiday is affected* (I-CVI=0.6)', '*I have problems speaking with groups of people* (I-CVI=0.6)', '*I worry being in places close to other people'*, and '*I slide in and out of my shoes* (I-CVI=0.6)*'*. I-CVI was below 0.8 for three items for completeness including *my holiday is affected*, *I worry being in places close to other people*, *I slide in and out of my shoes*, while only 1 item (*my eating habits are affected*) had relevance I-CVI below threshold. All items were endorsed for scaling. At the scale level, all aspects (language clarity, completeness, relevance and completeness) achieved content validity (S-CVI = 87% to 100%) (Table 4.2). Agreement on ratings at the scale level was also strong on all the four aspects assessed, the coefficient of agreement ranged from 0.7 to 1.

Suggestions

In addition to the individual item ratings, the experts also provided comments and suggestions pertaining to specific items as well as the whole questionnaire. Comments were given on 34 items (Table 4.3). For example respondents commented that they were not sure whether the item 'My holiday is affected was asking about the actual holidays or its planning. One expert thought the item 'my self esteem is affected' duplicates my self confidence is affected. In reference to the item "I feel my skin is hot all the time", one panel member commented that sweating would still occur even when they felt cold. Another comment made in relation to the same item was that it was not the skin was necessarily hot, but rather damp/wet.

Part II: Dermatologists panel

Panel session

The dermatologists' panel found the general layout of the instrument including the font-style, font-size and organisation of the instrument was appropriate and adequate. The panel found the font-size and font-type to be suitable and appropriate. The instruction *If a statement does not apply to you please mark 'not relevant***'** on the first/cover page of the draft HidroQoL (Appendix A4.1) were considered inappropriately placed on the first page. Concerns were raised that patients might also apply this instruction to the demographics question following immediately after on same page, resulting in confusion or mistakes in the completion of the demographic questions. The panel's recommendation was to remove this instruction from the first page, but to retain it on the rest of the pages. It was suggested that the instructions on the rest of the pages of the instrument be enclosed in a border. An additional change was suggested to the instruction *The statements in this questionnaire relate to how **your** life is being affected by your excessive sweating condition (hyperhidrosis) at the moment.* Instead of emphasising the words 'your' only, emphasis was to be placed on the entire clause '***your life is being affected by your excessive sweating*'.**

The recall period '*at the moment*' was considered to be too short and impractical. It was argued that when patients are asked about how they feel at the moment, they relate to events of the preceding days. They further stated that if the instrument were to be used for monitoring of response to treatments, a day may not be long enough to observe any meaningful changes. A recall period of 1 – 2 weeks was suggested instead. Several issues were raised regarding the response scaling: for the general impact question, '*in general, how would you rate the effect of excessive sweating on your life*' with response options: '*no effect at all'*, '*slight'*, *'moderate'*, *'quite a bit'* and *'extreme'*, the panel considered these to not appropriately reflect equal interval of increasing intensity. They suggested changing 'slight' to 'mild'; 'quite a bit' to 'strong' and extreme to 'very strong'. Furthermore, in relation to the response scaling used for the individual items, '*no, not at all', 'a little', 'somewhat', quite a bit' 'very much', 'not applicable*', the panel made a number of points. They considered '*a little'* and *'somewhat'* to lack a clear demarcation; '*quite a bit*' was seen as not reflecting midway between '*somewhat'* and *'very much'*. A strong case was made against including *'not relevant'*.

Table 4.1: Patient-panel ratings of language clarity, completeness, relevance, scaling of the HidroQoL

	Instrument Item	Language		Complete-ness		Relevance		Scaling		I-CVI			
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	Lan	Com	Rel	Scal
1	My choice of clothing is affected	4	0	4	0	4	0	4	0	1	1	1	1
2	My choice of footwear is affected	4	0	4	0	4	0	4	0	1	1	1	1
3	My holiday is affected	3.2	1.1	3.2	1.1	3.8	0.45	4	0	0.6	0.6	1	1
4	I have difficulties gripping objects	3.6	0.89	3.6	0.89	4	0	4	0	0.8	0.8	1	1
5	I have difficulties handling money	4	0	4	0	4	0	4	0	1	1	1	1
6	I have difficulties with physical contact with others	3.6	0.89	4	0	4	0	4	0	0.8	1	1	1
7	My hobbies are affected	3.8	0.45	4	0	4	0	4	0	1	1	1	1
8	I have problems speaking with groups of people	3.2	1.1	3.6	0.89	3.6	0.89	4	0	0.6	0.8	0.8	1
9	My physical activities are affected	3.8	0.45	4	0	4	0	4	0	1	1	1	1
10	My outdoor activities are affected	3.8	0.45	4	0	4	0	4	0	1	1	1	1
11	My everyday housework is affected	4	0	4	0	4	0	4	0	1	1	1	1
12	I find it hard to handle paper	4	0	4	0	4	0	4	0	1	1	1	1
13	My career decisions are affected	3.6	0.89	4	0	4	0	4	0	0.8	1	1	1
14	My work is affected	3.6	0.89	4	0	4	0	4	0	0.8	1	1	1
15	I have difficulties with using touch-technologies	4	0	4	0	4	0	4	0	1	1	1	1
16	My relationships with others are affected	3.6	0.89	4	0	4	0	4	0	0.8	1	1	1
17	I feel embarrassed	3.6	0.89	4	0	4	0	3.6	0.89	0.8	1	1	0.8
18	I do not socialize as much as I would like to	4	0	4	0	4	0	4	0	1	1	1	1
19	I avoid meeting people	4	0	4	0	4	0	4	0	1	1	1	1
20	I avoid going out	4	0	4	0	4	0	4	0	1	1	1	1
21	I feel nervous	4	0	3.6	0.89	3.6	0.89	4	0	1	0.8	0.8	1
22	I feel hopeless	4	0	3.6	0.89	3.6	0.89	4	0	1	0.8	0.8	1
23	I feel sad	4	0	3.6	0.89	3.6	0.89	4	0	1	0.8	0.8	1
24	I feel depressed	4	0	3.6	0.89	3.6	0.89	4	0	1	0.8	0.8	1
25	I feel frustrated	4	0	3.6	0.89	3.6	0.89	4	0	1	0.8	0.8	1
26	My confidence is affected	4	0	4	0	4	0	4	0	1	1	1	1

Note: Lan, Language clarity; Com, completeness; Rel, Relevance; Scal, Scaling.

Table 4.1 (continued)

Instrument Item	Language		Completeness		Relevance		Scaling		I-CVI			
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	**Lan**	**Com**	**Rel**	**Scal**
27 My self-esteem is affected	4	0	4	0	4	0	4	0	1	1	1	1
28 My whole life is affected	4	0	4	0	4	0	4	0	1	1	1	1
29 Sweating is constantly on my mind	4	0	4	0	4	0	4	0	1	1	1	1
30 I avoid taking on new challenges	3.8	0.45	4	0	4	0	4	0	1	1	1	1
31 My summer activities are affected	3.6	0.89	4	0	4	0	4	0	0.8	1	1	1
32 I feel more self conscious	3.8	0.45	3.4	1.34	4	0	4	0	1	0.8	1	1
33 My appearance is affected	3.8	0.45	3.6	0.89	4	0	4	0	1	0.8	1	1
34 I feel uncomfortable physically expressing affection	3.4	0.89	3.6	0.89	4	0	4	0	0.8	0.8	1	1
35 I worry about people's reactions	3.8	0.45	3.4	1.34	4	0	4	0	1	0.8	1	1
36 My sex life is affected	4	0	4	0	4	0	4	0	1	1	1	1
37 I worry about leaving sweat marks in public places	4	0	4	0	4	0	4	0	1	1	1	1
38 I worry being in places close to other people	3	1.41	3.2	1.1	4	0	4	0	0.6	0.6	1	1
39 My eating habits are affected	3.2	1.1	3.4	0.89	3.2	1.1	4	0	0.6	0.8	0.6	1
40 I slide in and out of my shoes	3	1	3.2	1.1	3.6	0.89	3.6	0.89	0.6	0.6	0.8	0.8
41 I have problems with being barefoot	3.8	0.45	3.8	0.45	4	0	4	0	1	1	1	1
42 My eyes get irritated	3.4	0.89	3.4	0.89	3.8	0.45	4	0	0.8	0.8	1	1
43 I feel my skin is hot all the time	3.4	0.89	3.6	0.89	3.6	0.55	4	0	0.8	0.8	1	1
44 I worry about the extra demands on my finances	3.8	0.45	3.6	0.89	4	0	4	0	1	0.8	1	1
45 I find it difficult to cope with my condition	4	0	4	0	4	0	4	0	1	1	1	1
46 I find it difficult to do things without planning in advance	3.8	0.45	4	0	4	0	4	0	1	1	1	1
47 I feel that I need more time for hygiene chores	3.2	1.1	3.4	0.89	4	0	4	0	0.6	0.8	1	1

Note: Lan, Language clarity; Com, completeness; Rel, Relevance; Scal, Scaling.

Table 4.2: Level of agreement and content validity index for the panel of patients

	CVI*	AC1, r**
Language clarity	87%	0.7
Completeness	94%	0.8
Relevance	98%	0.9
Scaling	100%	1

*Content Validity Index

** Gwet's AC1 Coefficient of agreement

Table 4.3: Comments from patients

Item	Suggestions
1 My choice of clothing is affected	Add 'by my excessive sweating'
2 My choice of footwear is affected	Add 'by my excessive sweating'
3 My holiday is affected	I wasn't sure whether this is asking if the holiday itself is affected, or if I plan my holidays to suit my hyperhidrosis. Would be better if this read 'my choice of holiday is affected'
4 I have difficulties gripping objects	..Would be better if read ' have difficulty holding onto objects' Its more than objects – hand rails on tube, steering wheel in care
6 I have difficulties with physical contact with others	.I understood this as: "I have difficulties in situations which involve physical contact with others". Not sure if that was correct?
7 My hobbies are affected	Is this asking "My choice of hobbies is affected"? Would be better if read 'my choice of hobbies are affected'. My choice of hobbies are affected
8 I have problems speaking with groups of people	-I wasn't sure what this was asking
9 My physical activities are affected	I took this to mean everyday physical activities, rather than sporting / recreational. . Maybe better if this read 'I avoid certain physical activities which would exacerbate/highlight my condition'
10 My outdoor activities are affected	Maybe better if this read 'I avoid certain physical activities which would exacerbate/highlight my condition'
12 I find it hard to handle paper	Could also add 'hard/embarrassing'
13 My career decisions are affected	I wasn't sure if this meant my choice of career. Could read 'my career choice has been affected by my condition' Mine have been affected in the past
14 My work is affected	Work = career?
16 My relationships with others are affected	Does this mean personal relationships / work relationships / intimate relationships / etc? Ok, although may want to distinguish between close/family relationships and friends/work colleagues
17 I feel embarrassed	Embarrassed all the time, or under certain conditions?
18 I do not socialize as much as I would like to	Could read 'my condition inhibits my social activities'
19 I avoid meeting people	Could read ' I avoid meeting people in circumstances where my condition may be obvious'
20 I avoid going out	as in Q19. Also, this may be worse at certain times of the year, so perhaps 'I avoid going out when my condition is at its worse'
21 I feel nervous	Could read 'I feel nervous that my condition will appear obvious to others' …

Table 4.3 (continued)

23 I feel sad	leave out
27 My self-esteem is affected	leave out, duplicates confidence question
30 I avoid taking on new challenges	Bit vague. 'I avoid taking on new challenges where my condition would be evident'
31 My summer activities are affected	leave out Ok although this underestimates the impact that warm weather has. 'summer is unbearable' springs to mind
32 I feel more self-conscious	leave out More than what? 'more' should be taken out
34 I feel uncomfortable physically expressing affection	leave out
10 My outdoor activities are affected	Maybe better if this read 'I avoid certain physical activities which would exacerbate/highlight my condition'
12 I find it hard to handle paper	Could also add 'hard/embarrassing'
13 My career decisions are affected	I wasn't sure if this meant my choice of career. Could read 'my career choice has been affected by my condition' Mine have been affected in the past
14 My work is affected	Work = career?
16 My relationships with others are affected	Does this mean personal relationships / work relationships / intimate relationships / etc? Ok, although may want to distinguish between close/family relationships and friends/work colleagues
17 I feel embarrassed	Embarrassed all the time, or under certain conditions?
18 I do not socialize as much as I would like to	Could read 'my condition inhibits my social activities'
19 I avoid meeting people	Could read ' I avoid meeting people in circumstances where my condition may be obvious'
20 I avoid going out	as in Q19. Also, this may be worse at certain times of the year, so perhaps 'I avoid going out when my condition is at its worse'
21 I feel nervous	Could read 'I feel nervous that my condition will appear obvious to others' …
23 I feel sad	leave out
27 My self-esteem is affected	leave out, duplicates confidence question
30 I avoid taking on new challenges	Bit vague. 'I avoid taking on new challenges where my condition would be evident'
31 My summer activities are affected	leave out Ok although this underestimates the impact that warm weather has. 'summer is unbearable' springs to mind
32 I feel more self-conscious	leave out More than what? 'more' should be taken out
34 I feel uncomfortable physically expressing affection	leave out
35 I worry about people's reactions	Reactions to what?

Table 4.3 (continued)

	Item	Suggestion
37	I worry about leaving sweat marks in public places	reword hyperhidrosis, eating habits 'that my sweating will be visible to members of the public '
38	I worry being in places close to other people	reword crowded places
39	My eating habits are affected	This isn't relevant to me, but could say: My eating habits are affected, i.e. avoid public places to eat. OR does this mean you eat certain foods that cause less sweating or something?
40	I slide in and out of my shoes	reword I'm generally uncomfortable in shoes
41	I have problems with being barefoot	add hyperhidrosis
42	My eyes get irritated	Could add 'by facial sweating'
43	I feel my skin is hot all the time	leave out Not necessarily 'hot' but 'damp/wet' Can be cold and still sweating though
44	I worry about the extra demands on my finances	I'm not sure about this. Personally, I haven't been worried about this
47	I feel that I need more time for hygiene chores	Reword Instead of 'hygiene chores', should read 'personal hygiene'

The panel raised concerns about ambiguity between the option *'no, no at all'* and '*not relevant'*. An example given was of the item '*my hobbies are affected*'. Although the expectation is that only those without hobbies would choose the '*not relevant*' option while those with hobbies but not affected choosing '*no, not at all'*, respondents may easily confuse the two.

Content validation questionnaire

The panel also assessed language clarity, completeness, relevance and scaling of each of the 47 individual items. Mean language clarity rating was below 3 in five items including *I feel uncomfortable physically expressing affection* (mean = 2.25), *I find it difficult to cope with my condition* (2.5), *I feel more self-conscious* (2.25). SD for sixteen items exceeded the minimum threshold (SD > 0.75), including '*I feel my skin is hot all the time'* (SD = 1.5), '*my career decisions are affected*' (1.15), '*my summer activities are affected*" (1.15) reflecting disagreement in the ratings. Mean relevance rating was below 3 for sixteen items. The same items also had a mean language clarity or completeness rating below 3. SD of relevance rating for 20 items was above the threshold, including my choice of footwear (SD 0 1.5), *my hobbies are affected* (SD = 1.15),

this also included all items showing mean score exceeding 3. All items had mean scaling rating of 3 or 4, there were no disagreements on any item. Further, ratings were analysed using content validity index. Sixteen items had I-CVI below 1 for language clarity, including '*I feel more self-conscious'* (CVI = 0.25), '*I find it difficult to cope with my condition'* (0.25) and *'I feel uncomfortable physically expressing affection'* (0.25). For completeness, 12 items were below the threshold (CVI => 1), eight of these had also been identified with language clarity problems. Twenty items did not achieve content validity for relevance, fifteen of which had shown problems for language clarity and completeness.

Table 4.4: Level of agreement and content validity index for the panel of dermatologists

	CVI*	AC1, r**
Language clarity	66%	0.5
Completeness	74%	0.6
Relevance	89%	0.2
Scaling	98%	1

*Content Validity Index
** Gwet's AC1 Coefficient of agreement

The items with optimal language clarity and completeness but lacking in relevance included *'my choice of footwear is affected'*, *'i have difficulties with physical contact with others*, ' *I worry about the addition demands on my finances'* were endorsed for language clarity and completeness. Only one item '*I feel my skin is hot all the time'* had I-CVI less than 1 for scaling.

Content validity indices were also estimated at the scale level (S-CVI/UA), for language clarity, completeness, relevance and scaling. S-CVI/UA for language clarity (66%) and completeness (74%) was below minimum threshold, while relevance and scaling aspects were above the content validity threshold. Inter-rater agreement was moderate for completeness (r = 0.5) and poor for relevance (r = 0.2) (Table 4.4). This hints on a number of challenges associated with ascertaining the relevance quality of life issues based merely based on observation as opposed to first hand experience from patients.

Suggestions

A rich set of comments were provided by the panel on 29 items. Suggestions were made to delete five items including *I feel my skin is hot all the time, I feel that I need more time for hygiene chores*, *I find it difficult to cope with my condition, My summer activities are affected.* In the case of the item *I find it difficult to cope with my condition,* it was argued that although the concept of coping

is closely related to QoL it relates to a different construct. The panel feared that the item '*My summer activities are affected'* would not reflect much sensitivity to change in clinical settings. A similar comment was made with regard to *My holidays are affected*. The item '*I have problems with speaking with groups of people*' was also thought to cause ambiguities in the sense that it was unclear what sort of group, whether it was the 'group factor' or the 'speaking'. More general comments were also made in relation to the level at which quality of life was being measured. Whether the instrument would focus on specific types of hyperhidrosis and the issues specific to each; or would assess HRQoL at a higher hierarchical level of the construct, common to all types of hyperhidrosis. The choice would have implications for the content, crafting of the items, structure of the measure and ultimately, its practicality. For example, if the instrument will aim to measure hyperhidrosis-QoL at a high level, then all items must be of relevance for all forms of hyperhidrosis. Paying no attention to this intricate decision risks development of a measure that would be biased against patients with one type of hyperhidrosis over another. The panel recommended assessing hyperhidrosis at a higher level where all items would apply to all forms of hyperhidrosis.

A consideration of the overall representativeness of the HidroQoL for quality of life in hyperhidrosis was also made by the panel. The panel identified various areas as being under-represented in the content: i) concerns related to bad odour ii) the burden related to extra effort involved in managing hyperhidrosis (e.g. carrying second bags, towel, air conditioning, washing clothes, treatment, personal-hygiene) iii) physical discomfort associated with hyperhidrosis (e.g. being wet, cracked skin, dampness, hot). Although these issues were not included in the 47-item version of the instrument, they were nonetheless, mentioned during qualitative study.

Part III: Review panel

The data collected during the content validation panels provided a wealth of information on the HidroQoL covering all aspects of the HidroQoL, for instance recommendations related to 'frame of reference', 'instructions', suggested items to be added and the ratings of the items of the HidroQoL. The review panel examined this data and made decisions based on the developmental goals of the HidroQoL. There was consensus to maintain the instrument's structure, the graphical design, font style and font size, presentation of the items, as originally intentioned, moreover no changes had been suggested by the expert panels. The review panel agreed to maintain the second instruction on the front page considering its relevance to the organisation of the entire instrument, with the argument that instructions on the first page relate to the entire instrument. Instructions were maintained on every page and were placed in borders. The recommendation maintained to

not specify recall period or to use 'at peak' was considered as not reflecting the intended use of the instrument, the assessment of impacts on quality of life in routine clinical practice or for research. This includes the assessment of change over time or making comparisons across patients. The review panel considered two other alternatives, 'in recent times', which was considered as lacking the necessary precision and containing some ambiguity; and 'over the last two weeks' which was thought to be too long. There were strong arguments for maintaining the initial proposal of 'at the moment', including the precision in assessing the patient's condition at the time of measurement. An additional consideration was the nature of the impacts of hyperhidrosis which may be felt on a longer time horizon. Than at the moment the review panel therefore agreed on 'in the last 7 days including today' as frame of reference.

Table 4.5: Dermatologists-panel ratings of language clarity, completeness, relevance, scaling of the HidroQoL

Instrument Item	Language		Complete-ness		Relevance		Scaling		CVI			
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Lan*	*Com*	*Rel*	*Scal*
1 My choice of clothing is affected	4	0	4	0	3.75	0.5	4	0	1	1	1	1
2 My choice of footwear is affected	3.75	0.5	4	0	3.25	1.5	4	0	1	1	0.75	1
3 My holiday is affected	2.75	1.5	3	1.41	2.75	1.5	4	0	0.5	0.75	0.5	1
4 I have difficulties gripping objects	4	0	4	0	4	0	4	0	1	1	1	1
5 I have difficulties handling money	4	0	4	0	3.5	1	4	0	1	1	0.75	1
6 I have difficulties with physical contact with others	3.5	0.58	3.75	0.5	4	0	4	0	1	1	1	1
7 My hobbies are affected	3.25	0.96	3.5	1	3	1.15	4	0	0.75	0.75	0.5	1
8 I have problems speaking with groups of people	3.25	0.96	3.5	1	3.5	1	4	0	0.75	0.75	0.75	1
9 My physical activities are affected	3.75	0.5	3.75	0.5	3.5	0.58	4	0	1	1	1	1
10 My outdoor activities are affected	3.5	0.58	3.75	0.5	3.25	0.96	4	0	1	1	0.75	1
11 My everyday housework is affected	4	0	4	0	3.5	1	4	0	1	1	0.75	1
12 I find it hard to handle paper	4	0	4	0	3.75	0.5	4	0	1	1	1	1
13 My career decisions are affected	3	1.15	3.5	1	3.75	0.5	4	0	0.5	0.75	1	1
14 My work is affected	4	0	4	0	4	0	4	0	1	1	1	1
15 I have difficulties with using touch-technologies	4	0	4	0	3.75	0.5	4	0	1	1	1	1
16 My relationships with others are affected	3.75	0.5	4	0	4	0	4	0	1	1	1	1

Table 4.5 (continued)

Instrument item	Language		Complete-ness		Relevance		Scaling		CVI			
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Lan*	*Com*	*Rel*	*Scal*
17I feel embarrassed	4	0	4	0	4	0	4	0	1	1	1	1
18I do not socialize as much as I would like to	4	0	4	0	4	0	4	0	1	1	1	1
19I avoid meeting people	3.75	0.5	3.75	0.5	3.75	0.5	4	0	1	1	1	1
20I avoid going out	4	0	4	0	3.75	0.5	4	0	1	1	1	1
21I feel nervous	4	0	4	0	3.75	0.5	4	0	1	1	1	1
22I feel hopeless	4	0	4	0	4	0	4	0	1	1	1	1
23I feel sad	3.5	1	3.5	1	2.75	0.96	4	0	0.75	0.75	0.5	1
24I feel depressed	4	0	4	0	3.75	0.5	4	0	1	1	1	1
25I feel frustrated	3.75	0.5	4	0	3.75	0.5	4	0	1	1	1	1
26My confidence is affected	4	0	4	0	3.75	0.5	4	0	1	1	1	1
27My self-esteem is affected	3	1.15	4	0	3.25	0.96	4	0	0.5	1	0.75	1
28My whole life is affected	4	0	4	0	3.75	0.5	4	0	1	1	1	1
29Sweating is constantly on my mind	4	0	4	0	3.75	0.5	4	0	1	1	1	1
30I avoid taking on new challenges	3.75	0.5	4	0	3.75	0.5	4	0	1	1	1	1
31My summer activities are affected	3	1.15	3.5	1	3.25	0.96	4	0	0.5	0.75	0.75	1
32I feel more self-conscious	2.25	1.26	2.25	1.26	2.5	1	4	0	0.25	0.25	0.25	1
33 My appearance is affected	3.75	0.5	3.75	0.5	3.75	0.5	4	0	1	1	1	1
34 I feel uncomfortable physically expressing affection	2.25	1.26	3.5	1	3	1.15	4	0	0.25	0.75	0.5	1
35 I worry about people's reactions	3.75	0.5	3.75	0.5	3.75	0.5	4	0	1	1	1	1
36 My sex life is affected	4	0	4	0	3.75	0.5	4	0	1	1	1	1
37 I worry about leaving sweat marks in public places	3.5	0.58	3.75	0.5	3.75	0.5	4	0	1	1	1	1
38 I worry being in places close to other people	3.5	1	3.5	1	3.25	0.96	4	0	0.75	0.75	0.75	1
39 My eating habits are affected	3.25	0.96	3.75	0.5	3.25	0.96	4	0	0.75	1	0.75	1
40 I slide in and out of my shoes	3.25	0.96	3.75	0.5	3.5	0.58	4	0	0.75	1	1	1
41 I have problems with being barefoot	3.5	0.58	3.75	0.5	3.25	0.96	4	0	1	1	0.75	1

Table 4.5 (continued)

	Language		Complete-ness		Relevance		Scaling		CVI			
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Lan*	*Com*	*Rel*	*Scal*
42 My eyes get irritated	3.75	0.5	3.75	0.5	3.25	0.96	4	0	1	1	0.75	1
43 I feel my skin is hot all the time	2.75	1.5	3.5	1	2.25	1.26	3.5	1	0.5	0.75	0.25	0.75
44 I worry about the extra demands on my finances	4	0	4	0	3.5	1	4	0	1	1	0.75	1
45 I find it difficult to cope with my condition	2.5	1.29	3.25	0.96	2.75	0.96	3.75	0.5	0.5	0.75	0.5	1
46 I find it difficult to do things without planning in advance	3.25	0.96	3.25	0.96	2.75	0.96	4	0	0.75	0.75	0.5	1
47 I feel that I need more time for hygiene chores	3.25	0.96	4	0	3.75	0.5	4	0	0.75	1	1	1

Additionally, as recommended by the dermatologist-panel the 'not relevant' response category was removed to minimise the risk of satisficing and measurement errors. In order to address the ambiguities surrounding the demarcation between 'Quite a bit' and 'very much' as pointed out by the experts, further consultations were made. Two experts on patient reported outcome instrument development were consulted on whether the response options were clear and represented equal intervals of increasing intensity. They considered the response categorisation to be appropriate and reflecting widely used response categorisation. On this basis, the review panel maintained the response categorisation.

The review panel agreed to delete one item '*I find it difficult to cope with my condition*', given the possibility that it might be tapping into a related yet different construct than quality of life (Table 4.7). Three new items were added i) *I worry about my body odour* ii) *I worry about my condition in the future* iii) *I worry about people's reaction* to improve coverage of the hyperhidrosis quality of life. A further seventeen items were revised, for instance: the item '*my holidays are affected'* was changed to '*my holidays are affected (e.g. planning, activities)*'; '*I have problems speaking with groups of people'* was amended to '*I avoid public speaking (e.g. doing presentations)*. With a developmental goal that the HidroQoL would be relevant for patients with hyperhidrosis of all forms (e.g. palmar, feet, axillary, facial) and with considerations of practicality and applicability, it was decided that the measure will assess hyperhidrosis-QoL at a higher hierarchical level, with

the implication that i) the items included would need to have relevance for all hyperhidrosis forms and ii) the actual crafting of the items would have to reflect the same.

Naming of the new disease-specific hyperhidrosis QoL instrument

Deciding on the name for the new instrument took a number of factors into consideration: to be capable of hinting on the underlying concept being measured by the instrument; the way that the construct was going to be measured; and ultimately to be easy to remember. It was agreed to include 'Quality of Life' in the name to reflect that the measure purports to measure this construct. It was agreed further to include 'hyperhidrosis' in the name to emphasise the focus of the instrument i.e. disease-specific quality of life of hyperhidrosis patients. Finally, the team debated on whether to use profile or index as a suffix. Index was chosen to reflect the intended measurement model, to hint on the availability of a single score that sums up the patients quality of life. Therefore the full name chosen for the new instrument was 'Hyperhidrosis Quality of Life Index'. The acronym HidroQoL was chosen as a combination of 'Hidro' reflecting water and 'QoL' reflecting quality of life. It was thought that this would also be easy to remember as a measure of HRQoL in hyperhidrosis. Following thorough consideration of the findings from the expert panels, the revisions decided by the review panel led to the developmental version of the new instrument, the HidroQoL (Table 4.1). This included 49 items scored on a common 5 point Likert scale. Field testing and further validation studies carried out later, used this version

Table 4.6: Suggestions made by the panel of dermatologists

Item		Suggestion
2	My choice of footwear is affected	– Either should be removed or item 1 should be changed to mention that "footwear" is included – Footwear is also clothing
3	My holiday is affected	– What exactly is affected? Choice of destination, what would happen if I holiday at the Antarctica – Would this question be expected to reflect change in a clinical trial setting; Should be removed - it is ambiguous
4	I have difficulties gripping objects	– For a subgroup of patients very important
6	I have difficulties with physical contact with others	– initially not understood "with others" can be dropped without loss of meaning; alternatively "touching others" can be used. – Most relevant for the majority o f patients
8	I have problems speaking with groups of people	– initially not understood; does this capture if you are asked to participate in a conference but when you are not the speaker - the current wording may narrow it down unnecessarily; must change to "I have problems interacting with groups of people" OR "I have problems with participating in gatherings "; "I avoid presenting in front of groups of people"
10	My outdoor activities are affected	– Indoor is more a problem
12	I find it hard to handle paper	– Only for one subgroup of patients
13	My career decisions are affected	– active or passive? Are you failing to rise up the career ladder because of others? Or because you don't want to be? Career decisions is more active - e.g. "I don’t want to be the boss". Solution: "My career is affected"

Table 4.6 (continued)

	Item	Suggestion
16	My relationships with others are affected	− suggestion: my relationships are affected
19	I avoid meeting people	− duplicates "i have problems speaking with groups of people"; Q.8 must be refined; for someone repairing cars Q.8 might be understood as involving something outside his working hours
23	I feel sad	− delete; sad, depressed, frustrated are similar ! - hyperhidrosis cannot make someone sad
24	I feel depressed	− Can someone be depressed without being sad? Is not frustration a mild form of depression
26	My confidence is affected	− add the words "self-" to confidence. For the majority of hyperhidrosis patients they might not see a difference with self-esteem.
27	My self-esteem is affected	− delete; patients will not distinguish this from "confidence"
28	My whole life is affected	− see how changing the place of this question affects the response - but this has been generally already covered; consistency check…can be correlated to the first question…
31	My summer activities are affected	− likely to not be sensitive to change; delete…same recommendation applies to holidays…not sensitive to therapy
32	I feel more self-conscious	− more than what ? How do you explain ? Remove "more". This is duplicate for self-confidence
34	I feel uncomfortable physically expressing affection	− how do you physically express affection without contact?
37	I worry about leaving sweat marks in public places	− "in public places" should be deleted; this makes it more inclusive…"I worry to disturb other people by leaving sweat marks"

Table 4.6 (continued)

Item		Suggestion
38	I worry being in places close to other people	− remove "in places", it should be clear "close" is referring to physical contact and not "emotional closeness"…may be examples are necessary…only use examples if best expressions cannot be found…
39	My eating habits are affected	− My eating and drinking habits are affected…"hyperhidrosis limits my eating and drinking habits"
40	I slide in and out of my shoes	− change as I see fit
41	I have problems with being barefoot	− add because of hyperhidrosis…
42	My eyes get irritated	− discuss how items specific to particular hyperhidrosis are included.
43	I feel my skin is hot all the time	− delete
44	I worry about the extra demands on my finances	− Perhaps interesting, too: "I worry about the side effects of the hyperhidrosis treatment" OR "I would pay 3000 € for a surgery if this would stop sweating.(JN)
45	I find it difficult to cope with my condition	− Delete 'I am worried about the extra effort dealing with my hyperhidrosis takes !' because coping as a construct is broad and means alot of things cannot be assessed just based on this item e.g. active coping, resignation, e.t.c. − I don't think patients know the meaning cope (JN)
46	I find it difficult to do things without planning in advance	− too general
47	I feel that I need more time for hygiene chores	− should be changed to I feel that I need more time for personal hygiene

Table 4.7: Revision to the items of the HidroQoL

	Before content validation	After content validation
1	My choice of clothing is affected	My choice of clothing is affected
2	My choice of footwear is affected	My choice of footwear is affected
3	My holiday is affected	My holidays are affected (e.g. planning, activities)
4	I have difficulties gripping objects	I have difficulties holding objects
5	I have difficulties handling money	I have difficulties handling money
6	I have difficulties with physical contact with others	I find it hard to touch other people
7	My hobbies are affected	My hobbies are affected
8	I have problems speaking with groups of people	I avoid public speaking (e.g. during presentations)
9	My physical activities are affected	My physical activities are affected
10	My outdoor activities are affected	My outdoor activities are affected
11	My everyday housework is affected	My everyday housework is affected
12	I find it hard to handle paper	I find it hard to handle paper
13	My career decisions are affected	My career decisions are affected (e.g. career choice)
14	My work is affected	My work is affected
15	I have difficulties with using touch-technologies	I have difficulties using touch-technologies (e.g. computer keyboard, smart-phones)
16	My relationships with others are affected	My personal relationships are affected
17	I feel embarrassed	I feel embarrassed
18	I do not socialize as much as I would like to	I do not socialise as much as I would like to
19	I avoid meeting people	I avoid meeting new people
20	I avoid going out	I avoid going out
21	I feel nervous	I feel nervous
22	I feel hopeless	I feel hopeless
23	I feel sad	I feel sad
24	I feel depressed	I feel depressed
25	I feel frustrated	I feel frustrated
26	My confidence is affected	My self-confidence is affected
27	My self-esteem is affected	My self-esteem is affected

Table 4.7 (continued)

	Before content validation	After content validation
28	My whole life is affected	My whole life is affected
29	Sweating is constantly on my mind	Sweating is constantly on my mind
30	I avoid taking on new challenges	I avoid taking on new challenges
31	My summer activities are affected	My summer activities are affected
32	I feel more self-conscous	I feel self-conscious
33	My appearance is affected	My appearance is affected
34	I feel uncomfortable physically expressing affection	I feel uncomfortable physically expressing affection (e.g. hugging and cuddling)
35	I worry about people's reactions	I worry about people's reactions
36	My sex life is affected	My sex life is affected
37	I worry about leaving sweat marks in public places	I worry about leaving sweating marks on things
38	I worry being in places close to other people	I find it hard to be near other people
39	My eating habits are affected	My choice of food and drinks is affected
40	I slide in and out of my shoes	I feel uncomfortable in my shoes
41	I have problems with being barefoot	I have problems with being barefooted
42	My eyes get irritated	My eyes feel irritated
43	I feel my skin is hot all the time	My skin feels uncomfortable
44	I worry about the extra demands on my finances	I worry about the additional money spent in dealing with my condition
45	I find it difficult to cope with my condition	[item deleted]
46	I find it difficult to do things without planning in advance	I find it hard to do things without planning in advance
47	I feel that I need more time for hygiene chores	I worry about the additional time spend in dealing with my condition
		I worry about my body odour
		I worry about my condition in the future
		I worry about the additional chores in dealing with my condition

DISCUSSION

Psychometric properties of PRO instruments such as validity and interpretability rely on the hypothesis that scores of an instrument reflect the ‘status’ of the underlying construct. This

assumes a link between the content of an instrument and the underlying construct (Wynd et al. 2003). Evidence of content validity, underpins this assumption by demonstrating the relevance of an instrument's content and how adequately it represents of the underlying construct (Rothman et al. 2009). Without such evidence, inferences drawn from the scores are not supported, as it is unclear what is actually being assessed, thus the definition of actual construct being measured becomes ambiguous (Haynes et al. 1995). This study, therefore, examined whether the HidroQoL adequately addressed important aspects of quality of life relevant to hyperhidrosis patients and whether the other aspects of content's structure and technical quality support this.

Content validation of the HidroQoL employed expert panels who evaluated all aspects of the HidroQoL. Panel discussions and a content validation questionnaire were used for collecting data. The latter enabled the systematic assessment of each item, enhancing the structure, objectivity and credibility of the process (Lynn 1986). The panel discussions provided further information on the structure and organisation of the instrument and other general issues, while permitting the collection of valuable insights. Focused discussions tend to generate rich and unique data as discussion members contribute and respond to each other's comments (Krueger 1994). In addition to the therapeutic expert's panel, a separate second panel was comprised of patients with hyperhidrosis. Apart from items and their corresponding response options, the structural aspects of the HidroQoL such as frame of reference, instructions and formatting were subjected to content validation. Not only are these aspects important in facilitating the definition -and understanding of the construct under measurement (Patrick et al. 2011b), but they also have a bearing on the acceptability and applicability of the instrument. Moreover, their role in the response generation process suggests that these have potential to cause substantial biases in the measurement process. For instance, a recall period that is too short may lead to understatement of impact being measured while the extreme opposite (too long recall period) may be associated with under- or overestimation of burden due to recall biases (Norquist et al. 2011). On recommendation of the experts, instructions were revised, for instance, changing the emphasis on particular phrases and; including instructions on each page throughout the instrument. A further recommendation was to change the frame of reference. The proposed 'at the moment', was thought to be too short and not a reflection of the nature of the impact of hyperhidrosis or its treatment. It was further argued that patients tend to think about the recent past even when asked about today. Recall period was, therefore, revised to 'in the last seven days including today'. Each of the 47 individual items was assessed on language clarity, completeness, relevance and scaling.

Figure 4.1: The developmental version of the HidroQoL with 49 items

Hyperhidrosis Quality of Life Index

HidroQoL

The statements in this questionnaire relate to how your life has been affected **by your excessive sweating condition (hyperhidrosis) in the last seven days including today.**

Please mark clearly one box for each statement.

Please answer the following questions:

Age: ☐ Years Gender: Male ☐ Female ☐

How long have you been experiencing excessive sweating: ☐ Years

Main body area(s) affected: Head/face ☐ Armpits ☐ Palms ☐ Feet ☐ Other ☐ Generalised ☐

If you selected more than one area, how important is each (in percentage) e.g. 70%, 30% ☐

Are you currently receiving treatment for the condition: Yes ☐ No ☐

If **Yes**, what treatment ☐

How much time per day do you spend dealing with your condition: (e.g. for personal hygiene, treatment, laundry) ☐ Minutes

How much extra money do you spend per month because of your condition: (e.g. for personal hygiene, treatment, clothing) ☐ (€ / $ / £)

What is your current employment status: Employed ☐ Unemployed ☐ Retired ☐ Full Time Student ☐

If unemployed or retired, is this related to your condition: Yes ☐ No ☐

In general, how would you rate the effect of excessive sweating on your life:

No, none at all ☐ Slight ☐ Moderate ☐ Quite a bit ☐ Extreme ☐

Please turn to the next page

Hyperhidrosis Quality of Life Index

The statements in this questionnaire relate to how your life has been affected **by your excessive sweating condition in the last seven days including today.**

Please mark clearly one box for each statement..

		Very Much ▼	Quite a bit ▼	Some what ▼	A little ▼	No, not at all ▼
1.	My choice of clothing is affected	☐	☐	☐	☐	☐
2.	My choice of footwear is affected	☐	☐	☐	☐	☐
3.	My holidays are affected (e.g. planning, activities)	☐	☐	☐	☐	☐
4.	I have difficulties holding objects	☐	☐	☐	☐	☐
5.	I have difficulties handling money	☐	☐	☐	☐	☐
6.	I find it hard to touch other people	☐	☐	☐	☐	☐
7.	My hobbies are affected	☐	☐	☐	☐	☐
8.	My physical activities are affected	☐	☐	☐	☐	☐
9.	My outdoor activities are affected	☐	☐	☐	☐	☐
10.	My summer activities are affected	☐	☐	☐	☐	☐
11.	My everyday housework is affected	☐	☐	☐	☐	☐
12.	I avoid public speaking (e.g. presentations)	☐	☐	☐	☐	☐
13.	I find it hard to handle paper	☐	☐	☐	☐	☐
14.	My work is affected	☐	☐	☐	☐	☐
15.	My career decisions are affected (e.g. career choice)	☐	☐	☐	☐	☐
16.	I have difficulties using touch-technologies (e.g. computer-keyboard, smart phones)	☐	☐	☐	☐	☐
17.	I do not socialise as much as I would like to	☐	☐	☐	☐	☐

Please turn to the next page

Hyperhidrosis Quality of Life Index

The statements in this questionnaire relate to how your life has been affected **by your excessive sweating condition in the last seven days including today.**

Please mark clearly one box for each statement..

		Very much	Quite a bit	Some what	A little	No, not at all
18.	I avoid meeting new people	☐	☐	☐	☐	☐
19.	I avoid going out	☐	☐	☐	☐	☐
20.	My personal relationships are affected	☐	☐	☐	☐	☐
21.	I feel embarrassed	☐	☐	☐	☐	☐
22.	I feel nervous	☐	☐	☐	☐	☐
23.	I feel hopeless	☐	☐	☐	☐	☐
24.	I feel sad	☐	☐	☐	☐	☐
25.	I feel depressed	☐	☐	☐	☐	☐
26.	I feel frustrated	☐	☐	☐	☐	☐
27.	My self-confidence is affected	☐	☐	☐	☐	☐
28.	My self-esteem is affected	☐	☐	☐	☐	☐
29.	Sweating is constantly on my mind	☐	☐	☐	☐	☐
30.	I avoid taking on new challenges	☐	☐	☐	☐	☐
31.	I feel self-conscious	☐	☐	☐	☐	☐
32.	My appearance is affected	☐	☐	☐	☐	☐
33.	I worry about my body odour	☐	☐	☐	☐	☐
34.	I worry about leaving sweat marks on things	☐	☐	☐	☐	☐

Please turn to the next page

Hyperhidrosis Quality of Life Index

The statements in this questionnaire relate to how your life has been affected **by your excessive sweating condition in the last seven days including today.**

Please mark clearly one box for each statement.

		Very much	Quite a bit	Some what	A little	No, not at all
35	I worry about people's reactions					
36.	I find it hard to be near other people	☐	☐	☐	☐	☐
37.	I feel uncomfortable physically expressing affection (e.g. hugging others)	☐	☐	☐	☐	☐
38.	My sex life is affected	☐	☐	☐	☐	☐
39.	My choice of food and drinks is affected	☐	☐	☐	☐	☐
40.	I feel uncomfortable in my shoes	☐	☐	☐	☐	☐
41.	I have problems with being barefooted	☐	☐	☐	☐	☐
42.	My skin feels uncomfortable	☐	☐	☐	☐	☐
43.	My eyes feel irritated	☐	☐	☐	☐	☐
44.	I worry about the additional **money** spent in dealing with my condition	☐	☐	☐	☐	☐
45.	I worry about the additional **chores** in dealing with my condition	☐	☐	☐	☐	☐
46.	I worry about the additional **time** spent in dealing with my condition	☐	☐	☐	☐	☐
47.	I find it hard to do things without planning in advance	☐	☐	☐	☐	☐
48.	I worry about my condition in future	☐	☐	☐	☐	☐
49.	My whole life is affected	☐	☐	☐	☐	☐

Please check that you have answered all questions

Thank you for your help!

The two methods of analysing the items, based on mean scores versus estimating an item-content validity index, yielded slightly differing results. The latter was more stringent and led to the endorsement of less items relative to the former across all aspects. On the other hand, the CVI offers a systematic approach for estimating the scale level validity (Polit and Beck 2006).

The HidroQoL met content validity criteria for language clarity, completeness, relevance and scaling based on the judgement of the patient-panel. Relevance and scaling were also endorsed in the dermatologist-panel, while language clarity and completeness were not supported in this panel. Language clarity and completeness were closely related: items identified as having completeness issues in the dermatologists panel (12-items) and the patient's panel (1-item) also had language clarity problems. Free comments provided by the experts may elucidate on this issue. For instance the comments i) *I wasn't sure whether this is asking if the holiday itself is affected or if I plan my holidays to suit my hyperhidrosis* (for: my holiday is affected'), ii) *reword crowded place; remove 'in places'...maybe examples are necessary* (for: I worry being in places close to other people) points towards ambiguities surrounding certain word, rendering the sentences unclear and incomplete.

The two-panels rated relevance differently, the patient's panel endorsed all items except '*my holiday is affected*', while the dermatologists endorsed 27 out of the 47 items. Differences in how medical practitioners and patients with skin disease evaluate their own quality of life have been observed before (Jemec 1996). An additional and important consideration stems from the fact that the physician makes judgements based on observations while the patients relate to first hand experiences. Moreover, as observations cannot be divorced from the observer they tend to be liable to observer biases (Streiner and Norman 2008). This to a great extent reflects the poor inter-rater agreement related to the rating of relevance in the dermatologists' panel. In view of this, there is a strong case for involving patients in the content validation process, especially in considerations of the relevance of content. Out of the twenty items not endorsed for relevance by the panel of dermatologists, fourteen had language clarity problems. The single item considered lacking relevance by the patient-panel (my holidays are affected) also had language problems. This highlights the importance of making serious consideration of technical quality issues during instrument development (Hambleton and Rogers 1991).

SUMMARY

- The format (e.g. font size and font style) and design of the HidroQoL was appropriate and facilitated natural flow through the questionnaire.
- Instructions were revised to enhance their clarity and visibility in the instrument.
- The recall period of the HidroQoL was revised to '*in the last 7 days including today*' to reflect the characteristics of the conditions and its treatment as well as the experiences of patients.
- Items of the HidroQoL were judged to be relevant to patients with hyperhidrosis and the concept of quality of life as it relates to them.
- Sixteen items had language clarity issues, twelve of which were also considered lacking completeness. At the scale level content validity of language clarity and completeness were not supported.
- Seventeen items were revised, which involved replacement of words or addition of examples.
- One item was deleted and four were added in order to adequately cover the construct of quality of life as it relates to hyperhidrosis.
- Methodological decisions such as involvement of patients, whether the study includes a qualitative component and the choice of quantitative method for evaluating the ratings has a bearing on findings of content validation.

CHAPTER 5

Development of a Hyperhidrosis-Specific Quality of Life Instrument (HidroQoL): Factor Analysis

Item reduction and construct validation using classical test theory

INTRODUCTION

Demonstrating that an instrument measures what it intends to measure involves providing "evidence for, as well as potential consequences of, score interpretation (Messick 1995). This presumes the existence of a robust vehicle for assigning scores to the construct being measured i.e. a measurement model. Ascertaining construct validity in a newly developed instrument requires initially developing and testing a measurement model. This interprets the conceptual framework of an instrument in measurement terms providing a basis for subsequent hypothesis testing based on the instrument's scores e.g. correlations with other existing instruments. Reflecting the central role of the measurement model, scientific advisory committee on medical outcome trust regarded it as one of the eight key attributes of a measurement instrument(Lohr 2002).

Various multivariate analyses, including inter-item and item-total score correlations, step-wise regression analysis and factor analysis are typically utilised for developing and validating measurement models (Frost et al. 2007b). This technique seeks to identify the least number of variables accounting for covariation among items, which then allows a large number of items to be simplified into a few variables (Kline 1994). Utilising inter-item correlations, factors group together items sharing more correlation. On the other hand less correlation is expected between items in different factors. In turn, this method allows the development of scales comprised of items that tap into the intended construct and share homogeneity. Such scales are likely to exhibit strong internal consistency. The usefulness of FA during instrument development also relates to the process of item reduction. The grouping together of variables that share a greater correlation also allows the identification of items that share a weak relationship with the construct being measured which can be candidates for review. This can also identify aspects of the construct that are underrepresented. Factor analysis is useful for evaluating other properties of instruments, including testing hypotheses regarding factorial structure of existing instruments, measurement invariance, and testing for response shift (Visser et al. 2005; Meredith and Teresi 2006). In view of the perennial nature of the instrument validation process, FA seems versatile to support hypothesis tests of different forms carried out in an instrument's life cycle.

OBJECTIVES

The objectives of this study were to:

- Explore the general functioning of the HidroQoL in the intended target population.
- Explore and test the factorial structure of the HidroQoL
- Identify poor items and perform an item reduction on the HidroQoL
- Evaluate the internal consistency of the revised version of the HidroQoL
- Evaluate the replicability of the factorial structure of the HidroQoL

METHODS

Study design

Field-testing studies require the involvement of a large heterogeneous group of patients that represents the full range of the target population in terms of demographic and disease characteristics (Fayers and Machin 2007). For the intended methods of analysis, rooted in correlation analysis, the sample used needs to provide responses covering the entire response range(Gorsuch 1997). This study, therefore, followed a cross-sectional design where respondents completed the developmental HidroQoL questionnaire on a single time period. A large sample of patients (n > 250) representing all types and severity levels of hyperhidrosis was targeted.

Rules of thumb on sample size requirements for correlation analysis and factor analysis vary in their guidance, ranging from 5 to 20 observations per variable with more suggestion above and below this ratio (Costello and Osborne 2005). However, the minimum sample size required for accurate recovery of population factor pattern matrix is influenced by many factors including the distribution and reliability of the variables, degree of association among variables, communalities, degree to which factors are over identified (Reise et al. 2000; Schmitt 2011). Thus power and precision ought to be core consideration in parametric estimation based factor methods (Schmitt 2011), while in non-parametric approaches when communalities are high, sample size of 100 may be adequate (Reise et al. 2000).

Study population

This analysis included patients with self-reported hyperhidrosis, fulfilling the following criteria:

- Aged 18 years or above.
- With a score of 2 or higher on the HDSS
- With onset of hyperhidrosis in teenage years or early adult years.

Patients meeting the following criteria were excluded from the study:

- Below the age of 18
- With onset of hyperhidrosis after age of 30 and reporting a co-morbidity (hypertension, diabetes, pm hormonal disorders, psychological disorders)
- With HDSS score of 1.

Patients were recruited through hyperhidrosis online social networking communities, mainly the *International Hyperhidrosis Society* (IHHS) and the *UK Hyperhidrosis support group*, from May to September 2012. A detailed description of the study population and procedures is available in chapter 2.

Outcome measures

The developmental Hyperhidrosis Quality of Life instrument (HidroQoL) was used for collecting data on the QoL impact of hyperhidrosis. The instrument has 49 items, each scored on a 5 point scale, including '*no, not at all*', '*a little'*, '*somewhat'*, '*quite a bit'* and *'very much'* (
Figure **4.1**). Disease severity was assessed using the Hyperhidrosis Disease Severity Scale, scored as follows: 1. *My sweating is never noticeable and never interferes with my daily activities* 2. *My sweating is tolerable but sometimes interferes with my daily activities* 3. *My sweating is barely tolerable and frequently interferes with my daily activities* 4. *My sweating is intolerable and always interferes with my daily activities*. Additional questions related to patient demographics, characteristics of disease and treatment related characteristics were also included.

Data Processing And Analysis

Data analysis was carried out using MPLUS 6 and STATA. First, the data was explored using frequencies, for categorical variables; and mean, standard deviation and median for continuous variables. Further, the distribution of the responses to the HidroQoL were established calculating the proportion of endorsement for each item. Correlation analysis, based on polychorric correlations, was carried out to identify items that were multi-collinear. Correlation > .08 indicates multi-collinearity (Fayers and Machin, 2007). Exploratory factor analysis (EFA) was carried out to explore the factorial structure of the HidroQoL as follows:

- Optimal number of factors was determined by parallel analysis and confirmed using scree plots and goodness of fit statistics.

The following criteria were used for identifying poor items during EFA:

- Items with highest loading below 0.4;
- Items with more than one loading above 0.4, with none above 0.5;
- Items with residual variance (uniqueness) of 0.7 or more; and
- Items whose content does not match their factor (Lackey et al. 2003; Costello and Osborne 2005; Nijsten et al. 2006a)

Hypothesis testing of factorial structure used confirmatory factor analysis (CFA). Goodness of fit statistics were applied in determining how well the data fit to the model:

- Chi-square test of model fit: p-value < 0.05 indicates lack of fit;
- Tucker Lewis Index, Comparative Fit Index: Values greater than >.90 show acceptable fit, while >.95 indicate good fit;
- Root Mean Square Error of Approximation (RMSEA): Values > .1 show poor fit, 0.08 - .1 shows moderate fit; < 0.05 shows good fit;
- Standardised Root Mean Square Residual (SRMR): < 0.06 show good fit;
- Weighted Root Mean Squared Residual (WRMR): < 0.95 shows good fit (Hays et al. 2005; Byrne 2011)

RESULTS

In the purpose of clarity the results will be presented in three parts. Part I, reports on the correlation analysis carried out to identify redundant items; Part II, addresses exploratory analysis of the factorial structure of the HidroQoL; and Part III, covers the CFA analysis carried out to test the replicability of the proposed structure

Sociodemographic characteristics of study participants

A total of 674 patients with hyperhidrosis were recruited for the study, through the online patient communities of the IHHS and the UK hyperhidrosis support group. Out of the total, 559 patients were from the U.S. and 115 were from the UK. The subsample including patients from the USA was employed in the correlation analysis and subsequent exploratory factor analyses. The group including patients from the UK, was utilised in the hypothesis testing of the proposed factorial structure.

US subsample - 'sample 1'

Sample 1 comprised of a total of 559 patients with hyperhidrosis from the U.S. This included 106 males (19%) and 453 females (81%), with a mean age of 41 (±14years) (Table 5.1). Those aged 50 to 59 (n = 156, 28%) made up the largest age group (Figure 5.1). For the majority of the patients, the condition affects multiple areas. Patients whom the hands, feet and axillary were all affected made up twenty seven per cent of the sample. Thirty five patients had lived with condition for less than 10 years. Forty-six per cent of the patients reported 'their sweating as intolerable and always interfering with daily activities', the highest level of severity (Figure 5.3) Forty-nine percent reported the impact of the sweating on their life to be extreme (Figure 5.4). The majority of patients (89%) had seen a doctor for their condition, and sixty-six per cent reported receiving treatment within the last 6 months while those receiving treatment currently were thirty-per cent (Table 5.2)

UK subsample -'Sample 2'

Sample 2 comprised of a total of 115 patients with hyperhidrosis from the UK Thirty-seven (32%) were male and 78 (81%) were female (Table 5.1). The mean age was 40.2 years, with those aged between 30 and 39 making up the largest age group (n = 36). Similar to the US group, participants reported multiple areas in 36 participants whose sweating was generalized. Forty-five percent of the patients reported no co-morbidity. The proportion of patients reporting currently receiving treatment was 36%, while those who had previously received surgical treatment was 17%, in both instances, greater than the proportion in the U.S. sample.

Distribution of item responses

As an initial step, the distribution of the responses was assessed. For all items there was no response category accounting for more than 80% of responses, showing reasonable variability (Table 5.3) Nevertheless, the items showed a negative skew reflecting some ceiling effects. In forty-four items the highest response category 'very much' was chosen by 20% of participants. The ceiling effects were worse in seventeen items where 50% of participants chose 'very much'. Item Q1, Q21, Q31 and Q35 showed excessively large kurtosis. Thus the data shows some minor departure from normality. Nine items showed very low use (below 5%) of the lowest response category, 'no, not at all', this includes Q1, Q10, Q21, Q25, Q27, Q9, Q31, Q35 and Q49. This

raises questions related to the utility of this category for these items. Missing data occurred at random, and for the items affected (Q18 – Q49), was not more than 2% of responses.

The incidence of missing data increased with successive items, starting from item Q18, reflecting drop-outs, people who started responding to the questionnaire but stopped along the way. Withstanding the ceiling effects and the underuse of response category 'no, not at all' the distribution of responses was encouraging.

Table 5.1: Sociodemographic characteristics of study participants

	Sample 1 (n = 559)	Sample 2 (n =115)
Gender, n (%)		
Male	106 (19%)	37 (33%)
Female	453 (81%)	78 (67%)
Age (years)		
Mean (SD)	40.7(14.19)	40.2(13.3)
Median	39	39
Range	18 -74	18-74
Age (years), n		
18 to 29	143	27
30 to 39	142	36
40 to 49	112	24
50 to 59	156	16
≥ 60	5	12
Duration of condition, years		
Mean (SD)	27.7 (14.1)	24.1 (13)
Median	25	21
Range	2 to 69	2 to 60
Duration of condition (years), n		
< 10	35	12
10 to 19	134	36
20 to 29	149	26
30 to 39	109	25
40 to 49	75	10
≥ 50	57	6
Body area affected, n (%)		
Head*	123 (22%)	32
Axilla*	51 (9%)	14
General	121 (22%)	36
Axilla, Palms, Feet	150 (27%)	15
Palms and Feet	114 (20%)	19

Table 5.1 (continued)

	Sample 1 (n = 559)	Sample 2 (n =115)
Severity of disease (HDSS score), n		
1	0	0
2	73	16
3	227	46
4	259	53
Global impact of hyperhidrosis (GQ score), n		
No, none at all	0	0
Slight	2	2
Moderate	46	11
Quite a bit	214	39
Extreme	276	58
Co-morbidity, n (%)		
None	307 (55%)	52 (45%)
Menopausal complaints	58 (10%)	8 (7%)
Diabetes	29 (5%)	4 (3%)
Hypertension	48 (9%)	10 (9%)
Neurological disorders	57(10%)	9 (8%)
Thyroid disorders	60 (11%)	23 (20%)
Employment status, n (%)		
Employed	360 (64%)	74(64%)
Unemployed	98 (18%)	24(21%)
Retired	67 (12%)	11(10%)
Full-time student	34 (6%)	6(5%)

Variability in the data was maintained, reflecting the usefulness of the HidroQoL's response categories as well as the relevance of the items in discriminating among patients. Thirty item pairs had a correlation of 0.8 or greater reflecting multicollinearity problems (Table 5.4, Table 5.5, Table 5.6, Table 5.7). This included, I *feel embarrassed* (Q21) against *I feel nervous* (Q22) and *I feel self-conscious* (Q31)*; My self-confidence is affected* (Q27) against *My self-esteem is affected* (Q28) and *I feel self-conscious* (Q31); and *I have difficulties holding objects* (Q4) against *I have difficulties handling money* (Q5), *I find it hard to touch other people* (Q6), *I find it hard to handle paper* (Q13) and *I have difficulties using touch-technologies* (e.g. computer-keyboard, smart phones) (Q16).

Figure 5.1: Age distribution of the study participants

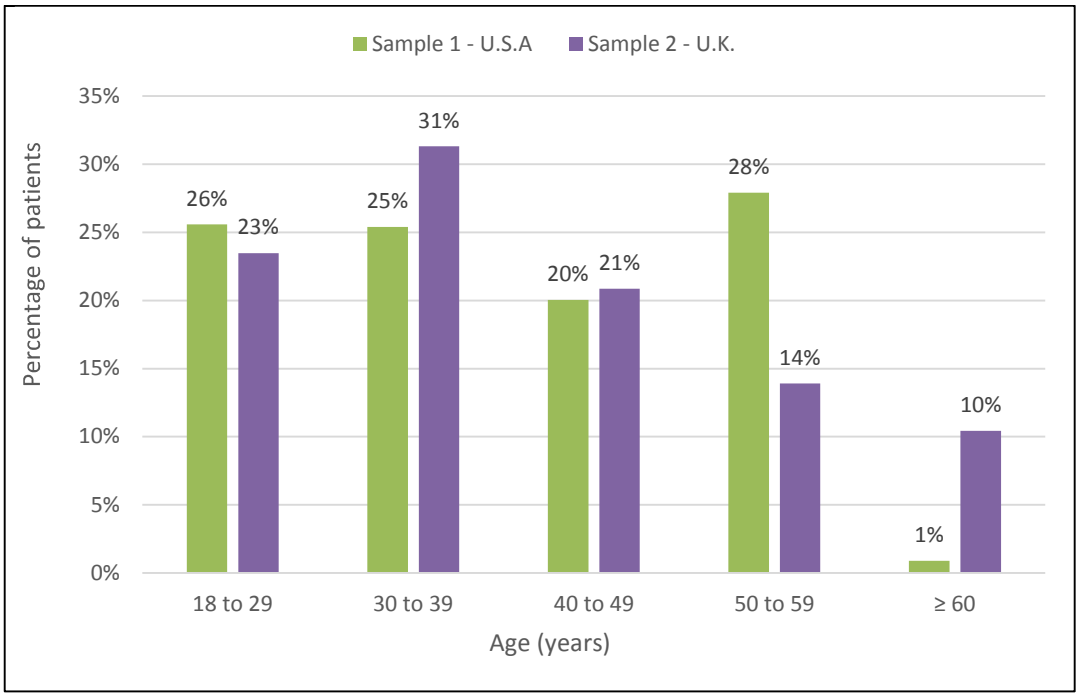


Figure 5.2: Duration of disease

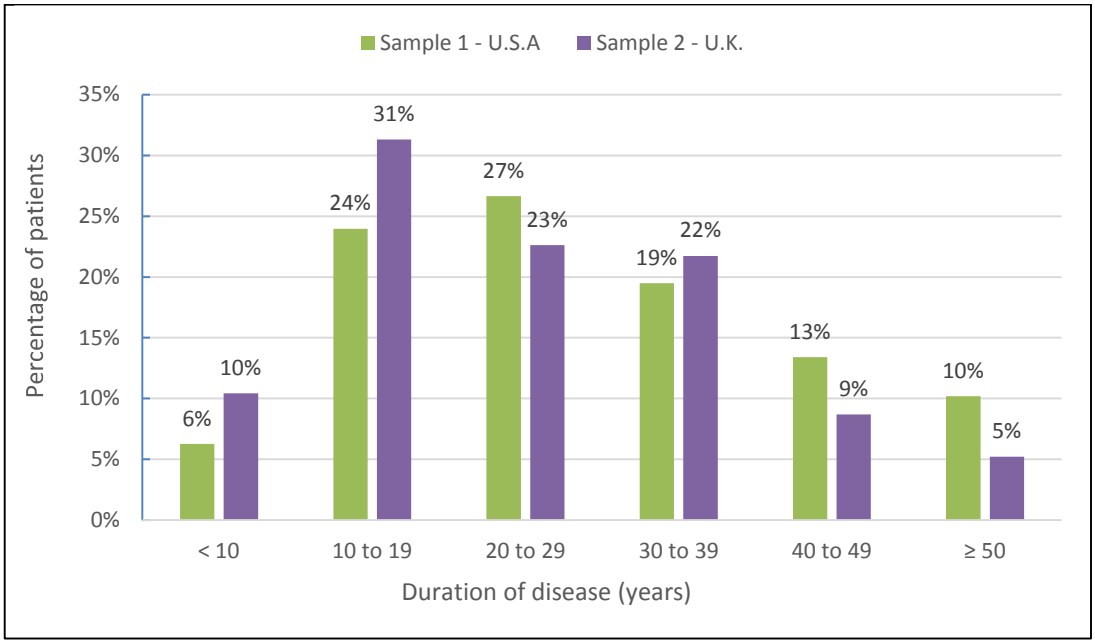


Figure 5.3: Patient's self-reported disease severity

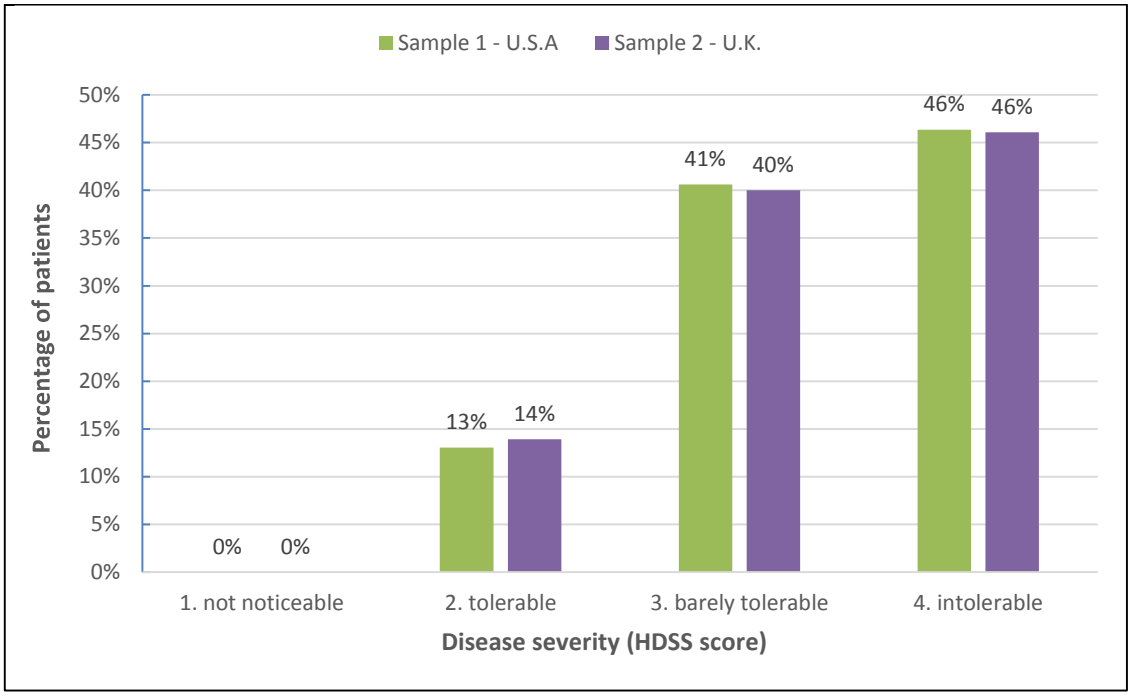


Figure 5.4: General impact of disease on patient's life

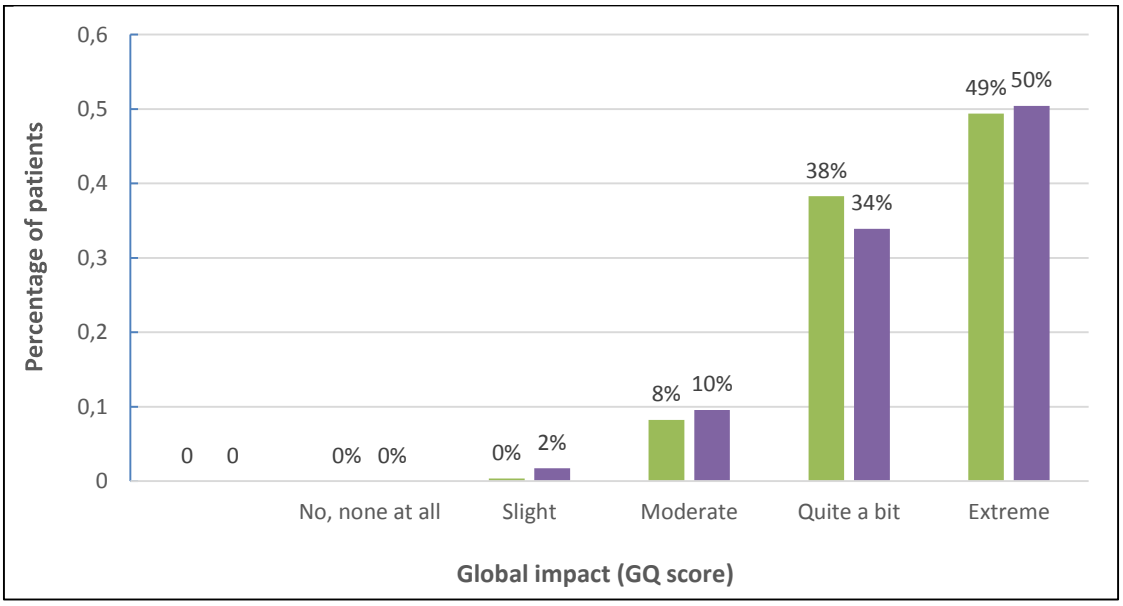


Table 5.2: Access to and use of treatment

	Sample 1- USA		*Sample 2 – UK*	
	n	**%**	**n**	**%**
Seen a doctor in relation to hyperhidrosis	485	87%	103	90%
Treated within last 6 months	368	66%	73	63%
Has received Botox within last 6 months	47	8%	7	6%
Surgical treatment	61	11%	19	17%
Currently being treated	169	30%	41	36%

Table 5.3: Frequency of endorsement to the HidroQoL

	Number of participants with each score						**Proportion of participants (%) with each score**						**Mean**	**Med**	**Skw.**	**Kurt**
	1	***2***	***3***	***4***	***5***	***M***	***1***	***2***	***3***	***4***	***5***	***M***				
Q1	16	28	51	103	361	0	3%	5%	9%	18%	65%	0%	4.37	5	-1.69	5.07
Q2	98	36	62	66	297	0	18%	6%	11%	12%	53%	0%	3.77	5	-0.83	2.07
Q3	69	59	137	112	182	0	12%	11%	25%	20%	33%	0%	3.5	4	-0.47	2.05
Q4	185	68	89	104	113	0	33%	12%	16%	19%	20%	0%	2.81	3	0.12	1.5
Q5	244	69	94	72	80	0	44%	12%	17%	13%	14%	0%	2.42	2	0.52	1.8
Q5	91	39	61	87	281	0	16%	7%	11%	16%	50%	0%	3.77	5	-0.84	2.15
Q7	50	43	93	141	232	0	9%	8%	17%	25%	42%	0%	3.83	4	-0.89	2.68
Q8	29	48	73	121	288	0	5%	9%	13%	22%	52%	0%	4.06	5	-1.12	3.15
Q9	30	49	75	102	303	0	5%	9%	13%	18%	54%	0%	4.07	5	-1.12	3.08
Q10	20	37	62	86	354	0	4%	7%	11%	15%	63%	0%	4.28	5	-1.47	4.1
Q11	100	83	119	115	142	0	18%	15%	21%	21%	25%	0%	3.21	3	-0.21	1.74
Q12	80	72	69	86	252	0	14%	13%	12%	15%	45%	0%	3.64	4	-0.62	1.87
Q13	178	46	66	82	187	0	32%	8%	12%	15%	33%	0%	3.1	3	-0.13	1.34
Q14	61	57	122	140	179	0	11%	10%	22%	25%	32%	0%	3.57	4	-0.58	2.22
Q15	102	48	63	104	242	0	18%	9%	11%	19%	43%	0%	3.6	4	-0.65	1.88
Q15	184	43	85	101	146	0	33%	8%	15%	18%	26%	0%	2.97	3	-0.04	1.41
Q17	57	66	90	105	241	0	10%	12%	16%	19%	43%	0%	3.73	4	-0.7	2.16
Q18	112	70	105	127	142	3	20%	13%	19%	23%	25%	1%	3.21	3	-0.26	1.7
Q19	130	81	111	124	110	3	23%	14%	20%	22%	20%	1%	3.01	3	-0.08	1.66
Q20	83	85	130	110	148	3	15%	15%	23%	20%	26%	1%	3.28	3	-0.25	1.82
Q21	9	14	40	85	408	3	2%	3%	7%	15%	73%	1%	4.56	5	-2.22	7.71
Q22	29	33	65	119	310	3	5%	6%	12%	21%	55%	1%	4.17	5	-1.35	3.83
Q23	105	68	87	74	222	3	19%	12%	16%	13%	40%	1%	3.43	4	-0.4	1.63
Q24	110	104	94	72	176	3	20%	19%	17%	13%	31%	1%	3.18	3	-0.11	1.52

Notes: M, missing; Med., median; Skw., Skewness; Kurt., Kurtosis.

Table 5.3 (continued)

	Number of participants with each score						Proportion of participants (%) with each score						Mean	Med	Skw.	Kurt
	1	*2*	*3*	*4*	*5*	*M*	*1*	*2*	*3*	*4*	*5*	*M*				
Q25	147	104	95	62	148	3	26%	19%	17%	11%	26%	1%	2.93	3	0.12	1.52
Q26	20	40	54	95	346	4	4%	7%	10%	17%	62%	1%	4.27	5	-1.47	4.1
Q27	27	50	75	112	291	4	5%	9%	13%	20%	52%	1%	4.06	5	-1.1	3.07
Q28	46	49	72	103	285	4	8%	9%	13%	18%	51%	1%	3.96	5	-1.02	2.76
Q29	17	43	66	137	292	4	3%	8%	12%	25%	52%	1%	4.16	5	-1.23	3.6
Q30	89	72	123	110	161	4	16%	13%	22%	20%	29%	1%	3.33	3	-0.32	1.82
Q31	9	24	47	106	369	4	2%	4%	8%	19%	66%	1%	4.45	5	-1.79	5.64
Q32	30	52	85	109	279	4	5%	9%	15%	19%	50%	1%	4	5	-1	2.84
Q33	95	92	100	98	170	4	17%	16%	18%	18%	30%	1%	3.28	3	-0.24	1.65
Q34	29	32	47	77	370	4	5%	6%	8%	14%	66%	1%	4.31	5	-1.63	4.52
Q35	7	24	56	118	348	6	1%	4%	10%	21%	62%	1%	4.4	5	-1.59	4.95
Q35	74	66	119	130	164	6	13%	12%	21%	23%	29%	1%	3.44	4	-0.45	1.99
Q37	48	53	93	129	229	7	9%	9%	17%	23%	41%	1%	3.79	4	-0.81	2.48
Q38	164	89	114	61	121	10	29%	16%	20%	11%	22%	2%	2.79	3	0.22	1.62
Q39	246	86	96	55	67	9	44%	15%	17%	10%	12%	2%	2.29	2	0.69	2.1
Q40	130	66	60	84	211	8	23%	12%	11%	15%	38%	1%	3.33	4	-0.33	1.48
Q41	178	47	35	55	237	7	32%	8%	6%	10%	42%	1%	3.23	4	-0.23	1.26
Q42	80	62	91	107	210	9	14%	11%	16%	19%	38%	2%	3.55	4	-0.56	1.92
Q43	285	69	82	59	54	10	51%	12%	15%	11%	10%	2%	2.14	1	0.84	2.29
Q44	162	121	100	68	98	10	29%	22%	18%	12%	18%	2%	2.67	2	0.36	1.76
Q45	147	118	97	103	88	6	26%	21%	17%	18%	16%	1%	2.76	3	0.21	1.7
Q45	112	103	108	103	124	9	20%	18%	19%	18%	22%	2%	3.04	3	-0.03	1.66
Q47	56	57	73	134	232	7	10%	10%	13%	24%	42%	1%	3.78	4	-0.83	2.41
Q48	87	57	103	115	190	7	16%	10%	18%	21%	34%	1%	3.48	4	-0.5	1.9
Q49	21	48	68	110	303	9	4%	9%	12%	20%	54%	2%	4.14	5	-1.19	3.32

Out of the thirty pairs showing multi-collinearity (correlation > 0.8), thirteen items were removed. Items *my outdoor activities are affected* (Q9) and *my summer activities are affected* (Q10) were dealing with similar issues. An aspect of both of these was addressed in *my physical activities are affected* (Q8). The social elements of the items can be argued to have been taken care of by items on '*going out*' and '*socializing*'.

Table 5.4: Polychoric correlations between items of the HidroQoL (part 1)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Q2	0.17													
Q3	0.45	0.03												
Q4	0.16	0.78	0.20											
Q5	0.13	0.76	0.24	0.92										
Q6	0.12	0.67	0.18	0.84	0.83									
Q7	0.36	0.40	0.50	0.58	0.59	0.54								
Q8	0.45	0.07	0.68	0.27	0.32	0.22	0.73							
Q9	0.43	0.02	0.68	0.20	0.24	0.16	0.66	0.92						
Q10	0.41	0.07	0.69	0.19	0.24	0.17	0.65	0.87	0.95					
Q11	0.36	0.02	0.59	0.22	0.28	0.13	0.48	0.70	0.74	0.72				
Q12	0.38	0.07	0.53	0.30	0.34	0.29	0.44	0.44	0.41	0.38	0.30			
Q13	0.03	0.79	0.09	0.91	0.91	0.85	0.52	0.16	0.08	0.08	0.15	0.26		
Q14	0.33	0.26	0.48	0.44	0.49	0.44	0.54	0.54	0.45	0.43	0.39	0.48	0.44	
Q15	0.35	0.27	0.48	0.41	0.50	0.45	0.53	0.49	0.38	0.36	0.25	0.60	0.42	0.75
Q16	0.10	0.80	0.15	0.88	0.87	0.81	0.54	0.19	0.14	0.13	0.16	0.27	0.90	0.42
Q17	0.37	0.02	0.63	0.26	0.31	0.36	0.47	0.50	0.48	0.47	0.36	0.66	0.20	0.53
Q18	0.27	0.10	0.55	0.31	0.36	0.42	0.41	0.41	0.38	0.36	0.29	0.68	0.29	0.49
Q19	0.38	-0.09	0.64	0.14	0.15	0.18	0.42	0.54	0.54	0.53	0.46	0.59	0.06	0.50
Q20	0.33	0.09	0.55	0.21	0.22	0.30	0.44	0.48	0.44	0.45	0.33	0.50	0.17	0.51
Q21	0.38	0.05	0.57	0.21	0.22	0.36	0.48	0.54	0.53	0.53	0.41	0.57	0.17	0.47
Q22	0.30	0.16	0.50	0.32	0.37	0.44	0.46	0.47	0.40	0.38	0.26	0.63	0.30	0.49
Q23	0.32	0.07	0.58	0.28	0.34	0.36	0.43	0.50	0.49	0.52	0.36	0.52	0.22	0.46
Q24	0.34	0.09	0.56	0.24	0.31	0.31	0.41	0.50	0.48	0.50	0.35	0.52	0.19	0.46
Q25	0.33	0.06	0.54	0.21	0.26	0.26	0.37	0.48	0.48	0.46	0.39	0.46	0.15	0.44
Q26	0.35	0.06	0.51	0.24	0.32	0.35	0.45	0.50	0.50	0.50	0.36	0.48	0.19	0.41
Q27	0.33	0.02	0.53	0.23	0.30	0.33	0.45	0.50	0.44	0.43	0.31	0.60	0.20	0.48
Q28	0.34	0.05	0.54	0.25	0.30	0.34	0.43	0.50	0.45	0.44	0.33	0.58	0.21	0.48
Q29	0.49	0.06	0.59	0.23	0.25	0.31	0.47	0.52	0.51	0.51	0.41	0.54	0.17	0.45
Q30	0.41	0.06	0.61	0.29	0.33	0.31	0.50	0.56	0.53	0.48	0.40	0.65	0.22	0.55
Q31	0.44	0.04	0.53	0.22	0.26	0.31	0.48	0.51	0.50	0.49	0.37	0.57	0.16	0.45
Q32	0.53	-0.16	0.63	0.02	0.04	0.05	0.43	0.59	0.60	0.58	0.50	0.49	-0.09	0.41
Q33	0.43	0.14	0.38	0.18	0.20	0.18	0.28	0.32	0.32	0.33	0.25	0.33	0.11	0.26
Q34	0.39	0.42	0.31	0.57	0.58	0.60	0.53	0.34	0.31	0.32	0.24	0.33	0.55	0.43
Q35	0.32	0.14	0.43	0.29	0.31	0.46	0.43	0.40	0.33	0.33	0.21	0.50	0.26	0.52
Q36	0.34	0.10	0.57	0.32	0.36	0.41	0.43	0.52	0.47	0.46	0.34	0.59	0.28	0.53
Q37	0.35	0.13	0.47	0.40	0.40	0.57	0.46	0.47	0.44	0.43	0.36	0.48	0.34	0.48
Q38	0.34	0.14	0.48	0.21	0.26	0.31	0.40	0.45	0.40	0.42	0.32	0.43	0.21	0.42
Q39	0.30	0.08	0.44	0.13	0.18	0.14	0.29	0.42	0.45	0.45	0.42	0.30	0.07	0.34
Q40	0.08	0.90	0.06	0.77	0.74	0.68	0.41	0.12	0.05	0.10	0.04	0.13	0.77	0.32
Q41	-0.03	0.89	-0.01	0.77	0.74	0.72	0.40	0.03	-0.04	0.01	-0.02	0.14	0.81	0.35
Q42	0.22	0.47	0.36	0.47	0.50	0.50	0.49	0.43	0.36	0.43	0.33	0.29	0.44	0.39
Q43	0.23	-0.07	0.42	0.07	0.08	0.04	0.24	0.41	0.44	0.43	0.51	0.25	-0.01	0.15
Q44	0.36	0.16	0.50	0.28	0.32	0.24	0.36	0.35	0.33	0.34	0.31	0.42	0.24	0.43
Q45	0.34	0.16	0.56	0.27	0.30	0.25	0.43	0.45	0.46	0.46	0.50	0.41	0.21	0.46
Q46	0.33	0.13	0.53	0.23	0.27	0.24	0.44	0.46	0.43	0.42	0.42	0.40	0.17	0.44
Q47	0.28	0.06	0.45	0.17	0.17	0.26	0.36	0.39	0.39	0.37	0.33	0.40	0.16	0.39
Q48	0.46	0.04	0.65	0.17	0.18	0.22	0.45	0.55	0.56	0.55	0.46	0.50	0.10	0.43
Q49	0.46	0.08	0.63	0.25	0.31	0.32	0.52	0.59	0.56	0.54	0.47	0.50	0.21	0.59

Table 5.5: Polychoric correlations between the items of the HidroQoL (part 2)

	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28
Q16	0.43													
Q17	0.63	0.26												
Q18	0.63	0.33	0.85											
Q19	0.53	0.09	0.85	0.81										
Q20	0.53	0.22	0.72	0.72	0.75									
Q21	0.51	0.13	0.64	0.6	0.63	0.61								
Q22	0.55	0.26	0.64	0.65	0.6	0.59	0.8							
Q23	0.51	0.24	0.63	0.6	0.64	0.62	0.76	0.76						
Q24	0.53	0.24	0.66	0.63	0.66	0.65	0.76	0.76	0.9					
Q25	0.51	0.20	0.63	0.61	0.66	0.66	0.71	0.71	0.84	0.93				
Q26	0.45	0.26	0.56	0.53	0.55	0.54	0.72	0.66	0.78	0.79	0.75			
Q27	0.58	0.20	0.71	0.7	0.68	0.66	0.78	0.79	0.76	0.78	0.75	0.76		
Q28	0.57	0.22	0.71	0.72	0.67	0.67	0.77	0.78	0.76	0.78	0.77	0.74	0.98	
Q29	0.47	0.18	0.56	0.51	0.57	0.53	0.73	0.62	0.67	0.66	0.6	0.71	0.69	0.68
Q30	0.66	0.24	0.76	0.73	0.73	0.63	0.71	0.68	0.67	0.71	0.71	0.64	0.73	0.73
Q31	0.51	0.19	0.63	0.59	0.63	0.59	0.83	0.73	0.74	0.74	0.69	0.78	0.83	0.82
Q32	0.36	-0.02	0.57	0.46	0.64	0.54	0.62	0.5	0.6	0.6	0.59	0.59	0.63	0.64
Q33	0.25	0.15	0.36	0.3	0.35	0.37	0.43	0.42	0.45	0.43	0.41	0.41	0.44	0.45
Q34	0.43	0.58	0.4	0.4	0.32	0.42	0.51	0.51	0.45	0.44	0.41	0.47	0.47	0.47
Q35	0.53	0.27	0.62	0.61	0.5	0.6	0.77	0.73	0.7	0.68	0.61	0.67	0.75	0.74
Q36	0.56	0.28	0.73	0.74	0.72	0.72	0.68	0.72	0.74	0.72	0.71	0.63	0.73	0.74
Q37	0.45	0.34	0.62	0.63	0.63	0.63	0.65	0.61	0.68	0.64	0.62	0.63	0.65	0.65
Q38	0.47	0.24	0.53	0.49	0.52	0.7	0.52	0.47	0.53	0.55	0.57	0.52	0.57	0.59
Q39	0.31	0.15	0.45	0.38	0.45	0.46	0.34	0.29	0.36	0.42	0.43	0.36	0.37	0.4
Q40	0.38	0.77	0.09	0.18	-0	0.14	0.13	0.24	0.21	0.2	0.17	0.22	0.2	0.21
Q41	0.35	0.80	0.09	0.17	-0.1	0.14	0.07	0.24	0.16	0.16	0.12	0.11	0.11	0.11
Q42	0.35	0.46	0.33	0.35	0.29	0.39	0.41	0.41	0.45	0.44	0.43	0.43	0.4	0.4
Q43	0.11	0.02	0.3	0.25	0.39	0.23	0.3	0.19	0.26	0.31	0.38	0.25	0.27	0.28
Q44	0.45	0.23	0.43	0.4	0.44	0.48	0.45	0.46	0.44	0.48	0.45	0.4	0.45	0.47
Q45	0.45	0.23	0.51	0.46	0.53	0.56	0.5	0.5	0.55	0.55	0.56	0.46	0.54	0.53
Q46	0.45	0.19	0.52	0.42	0.53	0.55	0.46	0.51	0.53	0.55	0.57	0.46	0.53	0.52
Q47	0.45	0.14	0.48	0.44	0.48	0.53	0.54	0.55	0.63	0.64	0.62	0.6	0.57	0.57
Q48	0.51	0.13	0.64	0.53	0.64	0.57	0.58	0.58	0.61	0.64	0.63	0.55	0.59	0.6
Q49	0.59	0.21	0.64	0.6	0.66	0.65	0.64	0.61	0.7	0.71	0.68	0.66	0.7	0.71

Table 5.6: Polychoric correlations between items of HidroQoL (part 3)

	Q29	Q30	Q31	Q32	Q33	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47	Q48
Q30	0.66																			
Q31	0.77	0.75																		
Q32	0.62	0.61	0.71																	
Q33	0.45	0.46	0.47	0.51																
Q34	0.57	0.43	0.56	0.32	0.48															
Q35	0.66	0.62	0.77	0.51	0.37	0.54														
Q36	0.62	0.72	0.72	0.57	0.44	0.5	0.75													
Q37	0.63	0.61	0.66	0.55	0.44	0.54	0.7	0.78												
Q38	0.45	0.54	0.5	0.49	0.42	0.43	0.44	0.54	0.53											
Q39	0.34	0.43	0.32	0.42	0.36	0.22	0.25	0.42	0.34	0.56										
Q40	0.22	0.15	0.16	-0	0.2	0.5	0.22	0.2	0.25	0.26	0.17									
Q41	0.12	0.15	0.11	-0.2	0.08	0.52	0.23	0.18	0.21	0.23	0.09	0.88								
Q42	0.43	0.41	0.41	0.34	0.35	0.48	0.4	0.46	0.44	0.38	0.25	0.53	0.55							
Q43	0.25	0.35	0.31	0.44	0.23	0.05	0.14	0.29	0.24	0.24	0.39	-0.04	-0.10	0.31						
Q44	0.47	0.45	0.44	0.37	0.39	0.35	0.43	0.5	0.44	0.49	0.39	0.20	0.22	0.37	0.30					
Q45	0.5	0.51	0.49	0.46	0.46	0.44	0.45	0.57	0.46	0.52	0.44	0.22	0.22	0.43	0.34	0.73				
Q46	0.5	0.53	0.47	0.45	0.45	0.4	0.47	0.54	0.47	0.53	0.39	0.18	0.19	0.41	0.30	0.71	0.87			
Q47	0.57	0.5	0.53	0.43	0.41	0.38	0.57	0.58	0.53	0.48	0.35	0.16	0.13	0.38	0.28	0.59	0.66	0.73		
Q48	0.62	0.67	0.62	0.59	0.45	0.35	0.53	0.63	0.55	0.49	0.45	0.10	0.03	0.38	0.37	0.55	0.63	0.64	0.72	
Q49	0.73	0.69	0.68	0.62	0.39	0.46	0.64	0.67	0.63	0.56	0.40	0.20	0.12	0.44	0.31	0.58	0.62	0.63	0.66	0.76

Therefore Q9 and Q10 were removed. In the case of the collinearity between item *I have difficulties using touch technologies (e.g. computer-keyboard, smart-phones)* (Q16) and items *I have difficulties holding objects* (Q4) and *I find it hard to handle paper* (Q13), item Q16 was removed. This was based on its lesser prevalence during qualitative research (chapter 3) and the narrower conceptual breadth. The item *My whole life is affected* (Q49) was unique, as a general impact question it reflected a general view of respondents condition summing up all aspects already addressed by the rest of the items. This suggests that it was overlapping with the rest of the instrument's items. Therefore, item Q49 was also removed, despite showing no correlation above 0.8 with any of the remaining items. This stage led to a 36-item version of the developmental HidroQoL (HidroQoL-36). The final version of the HidoroQoL following the first item reduction contained 36 items.

Table 5.7: Multicollinear items (correlations of at least 0.8)

Item	Related item
Q2	Q40 Q 41
Q4	Q5 Q6 Q13 Q16
Q5	Q6 Q13 Q16
Q6	Q13 Q16
Q8	Q9 Q10
Q9	Q10
Q13	Q16 Q41
Q16	Q41
Q17	Q18 Q19
Q18	Q19
Q21	Q22 Q31
Q23	Q24 Q25
Q24	Q25
Q27	Q28 Q31
Q28	Q31
Q40	Q41
Q45	Q46

Part II: Exploratory Factor Analysis of The Hidroqol-36

Following correlation analysis, exploratory factor analysis was carried out on the HidroQoL to explore its dimensional structure as well as to perform item reduction. First, the optimal number of factors to be extracted was determined, then the factors were estimated. The factor solution was rotated to yield interpretable results (DeVellis 2011). The poorly performing items were dropped in subsequent iterations, until a 'simple structure was achieved'. According to Thurstone's criteria a simple structure is characterised by a few high loadings on each factor with the rest of the loadings being zero or close to zero with variables having significant multiple loadings being at a minimum (Kline 1994). Three factors were extracted from the EFA of the HidroQoL-36, based on Horn's parallel analysis criterion (Table 5.9). This is supported by the scree-plot criterion (Figure 5.5). Three factors lied to the left side of the elbow on the plot, the rest of the factors from the 4th going to the right were rubble, the fourth factor also marked a change in the slope of the curve. Goodness of fit indices criterion showed mixed results. Although the Chi-square test of model fit was significant (chi-square = 2316.34, d.f. = 525, p = 0) indicating poor fit of the 3 factor solution, practical fit indices suggested otherwise (RMSEA = .078, SRMR = 0.51, CFI = 0.934, TLI = .921) (Analysis 1, Table 5.10). The factor pattern matrix was analysed to determine the performance of the individual items. Twenty-eight items had a clear strong loading on at least one factor; two items showed poor loadings, below 0.4, on all factors; five items showed crossed loadings (the highest loading on these items was at least 0.4 but below 0.5, and the difference with the smaller loading was less than 0.2) (Table 5.11). A single item had strong loadings on two factors. Three of the five items loading onto the first factor, *I find it hard to handle paper (Q13)*, *I have difficulties holding objects (Q5),* and *I find it hard to touch other people* (Q6) and *my hobbies are affected (Q7)* reflected limitations associated with palmar sweating.

The item *My choice of footwear (Q2)* seemed misplaced under this factor. The second factor had strong loadings from eight items, including *my everyday housework is affected* (Q11), *my holidays are affected (Q3), my choice of clothing is affected (Q1*), *I worry about the additional chores in dealing with my condition* (Q45). This factor captured an array of effects experienced in everyday life activities. Nonetheless, one item, *My eyes feel irritated (Q43)* did not fit the conceptual focus of the factor. Sixteen items loaded into the third factor. This was comprised of items focusing on the emotional impacts of hyperhidrosis including *I feel embarrassed (Q21)*, *I feel nervous (Q22) I feel sad (Q24)*, *Sweating is constantly on my mind (Q29)*. This factor also included items addressing more of the social aspect of life, I *avoid public speaking (Q12), I avoid meeting new*

people (Q18), I find it difficult to be near other people (Q36). Nevertheless, items such as *I worry about people's reactions (Q35)* demonstrate the strong connection between the two aspects. Taken together the two aspects of this factor were assessing the psychosocial impacts of hyperhidrosis.

Table 5.8: Item review based on correlation matrix (addressing multicollinearity)

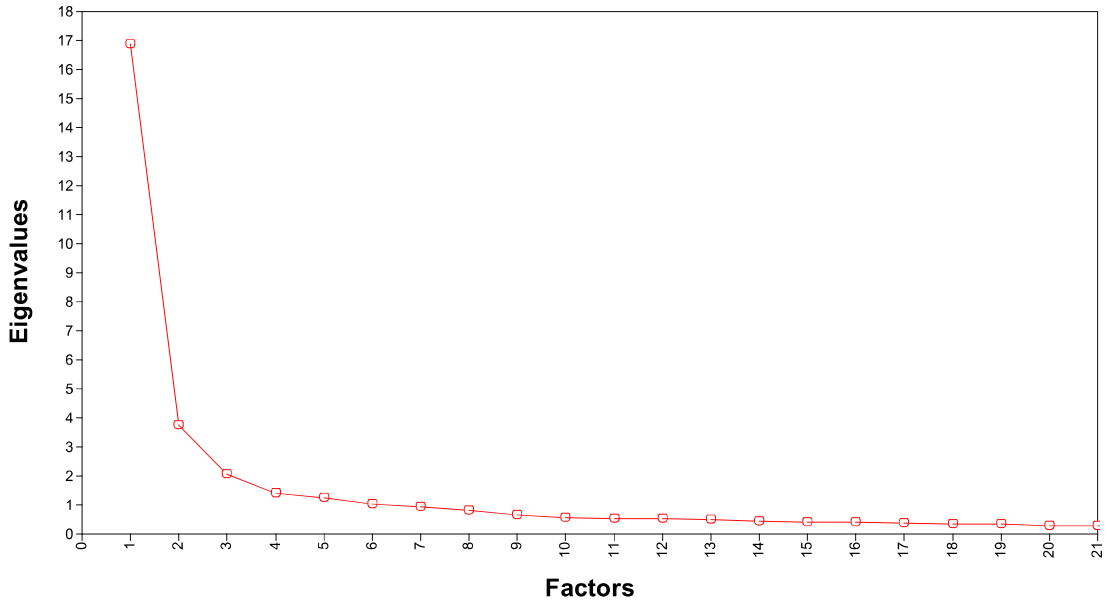
Item removed	Justification
I feel uncomfortable in my shoes (Q40) I have problems with being barefoot (Q41)	Adequately captured in: my choice of footwear is affected (Q2). Q2 had higher prevalence than Q40; Q41 in qualitative study (QS).
I worry about the additional time spent in dealing with my condition (Q46)	Amount of time may be a reflection of activities involved in: I worry about additional chores in dealing with my condition (Q45). Q45 has higher prevalence in Q46.
I feel self-conscious (Q 31)	May raise ambiguity issues Less prevalently reported than Q21, Q27, Q28, which are more straight forward
My self-esteem is affected (Q28)	Respondents not able to distinguish from: my self-confidence is affected (Q27); Q28 less prevalent than Q27 in QS.
My whole life is affected (Q49)	Seen as duplicating the entire instrument. Was included for consistency purposes
I feel hopeless (Q23)	Seen as ambiguous in comparison to Q24 and Q25, both of which adequately address issues in Q23.
I feel depressed (Q25)	Is a complex construct compared to "feeling sad (Q24)"
I avoid going out (Q19)	Content overlap with Q17
I have difficulties handling money (Q5)	Content overlap with: I have difficulties holding objects (Q4)
I have difficulties using touch technologies (e.g. computer-keyboard, smart-phones (Q16)	Content overlap with Q4. Predictable with Q13 Less prevalent than Q4 or Q13.
My outdoor activities are affected (Q9) My summer activities are affected (Q10)	Both Q9 & Q10 ask for an overlapping set of activities; My physical activities are affected (Q8) seems to be capturing some of that. Social aspects of summer/outdoor activities is addressed items on 'social' life.

Table 5.9: Eigen values for all 36 items*

Factor	Based on WLSMV	Based on Principal Factoring		
	Observed	*Adj.fact*	*Unadj*	*Bias*
1	16.88	13.33	13.88	0.54
2	3.76	2.55	3.04	0.49
3	2.05	1.05	1.52	0.47
4	1.40	0.60	1.01	0.41
5	1.26	0.51	0.91	0.40
6	1.04	0.16	0.53	0.37
7	0.95	0.13	0.49	0.35
8	0.80	0.10	0.39	0.29
9	0.64	0.03	0.30	0.27
10	0.57	0.00	0.24	0.24

* observed and Horn's parallel analysis

Figure 5.5 Scree-plot based on the Eigenvalues from the WLSMV for all 36 items



The two items with poor loadings were *I worry about my body odour (Q33)* and *my work is affected (Q14)*, while the items showing weak-moderate loading on multiple factors included *My skin feels uncomfortable (Q42)*, *My appearance is affected (Q32)*, *I worry about leaving sweat marks on objects (Q34), I worry about my condition in future (Q47).* These results seem to suggest that issues related to the physical aspects of the sweating play a minimal role in the overall quality of life in hyperhidrosis.

Table 5.10: Steps during EFA analysis: removal of poorly performing items

Analysis	Poor-load	Cross-load	Res_var >0.7	Factor# -K's rule	Factor# -Paral	Factor# extract	RMSEA	SRMR	CFI	TLI	Chi-1	Chi-2
1, All 36 items	Q33, Q14	Q42, Q32, Q47, Q34, Q48, Q7	_	6	3	3	.078 (.075, .081) p=0.000	.051	.934	.921	2316.34 df=525 p=0	27716.08 df=630 p=0
2, remove Q33	Q14	Q42, Q32, Q47, Q34, Q48, Q7	Q1	6	3	3	.079 (.076, .083) p=0	.051	0.933	0.922	2233.99 df=493 p=0	27396.52 595 p=0
3, remove Q14	_	Q42, Q32, Q47, Q34, Q48, Q7	Q1	5	3	3	.078 (.075, .082) p=0	.05	.94	.93	2039.42 df=462 p=0	26958.234 df=561 p=0
4, remove Q32 and Q42	_	Q47, Q34, Q48, Q7	Q43, Q1	5	3	3	0.82 (.078, .085) p=0	.05	.942	.928	1908.4 df=403 p=0	26319.9 df=496 p=0
5, remove Q34	_	Q47, Q48, Q7	Q43, Q1	5	3	3	.084 (.08, .088) p=0	0.05	.942	.929	1859.1 df=375 p=0	26230.1 df=465 p=0
6, remove Q47, Q48	_	Q38, Q7	Q43*, Q1	5	3	3	.086 (.082, .090) P=0	.049	.945	.931	1646.39 df = 322 p = 0	24607.0 df = 406 p = 0
7, remove Q38, Q1		Q7	-	5	2	3	.089 (.085, .094) P=0	0.48	.948	.933	1491.1 P = 273 P = 0	23804.98 df =351 p = 0
8, remove Q2, Q4, Q13		_	Q43, Q6	4	2	2	.099 (.094, .104) P=0	.059	.937	.924	1490.17 df = 229 p=0	20221.98 df = 276 p = 0
9, remove Q43, Q6	_	_	Q39 (=0.7)	4	1	2	.102 (.096, .107) p=0	.055	.944	.932	1273.01 df=188 p=0	19726.8 df=231 p=0
10, remove Q39	_	_	_	3	1	2	.106 (.101, .112) p=0	.056	.944	.93	1239.92 df=169 p=0	19305.35 df=210 p=0

Table 5.11: Factor pattern and factor structure matrices for the 36 items of the HidroQoL

	Factor pattern								Factor structure		
	F1	*SE*	*F2*	*SE*	*F3*	*SE*	*Res.Var*	*Issue*	*F1*	*F2*	*F3*
Q13	0.98	0.02	-0.12	0.05	-0.02	0.02	0.07		0.96	0.04	0.27
Q4	0.96	0.02	0.02	0.04	-0.04	0.04	0.10		0.95	0.15	0.32
Q2	0.88	0.03	0.00	0.01	-0.22	0.06	0.31		0.80	0.01	0.10
Q6	0.82	0.03	-0.21	0.05	0.27	0.05	0.17		0.89	0.10	0.44
Q7	0.53	0.04	0.51	0.04	0.02	0.03	0.35		0.62	0.61	0.52
Q11	0.11	0.05	0.84	0.05	-0.17	0.06	0.42		0.19	0.75	0.38
Q8	0.20	0.05	0.79	0.03	0.00	0.02	0.28		0.34	0.83	0.56
Q3	-0.02	0.03	0.58	0.04	0.31	0.05	0.37		0.19	0.76	0.65
Q45	0.12	0.04	0.58	0.05	0.23	0.07	0.40		0.30	0.74	0.62
Q43	-0.07	0.06	0.58	0.06	-0.03	0.07	0.69		0.02	0.55	0.30
Q44	0.14	0.05	0.50	0.05	0.20	0.06	0.52		0.30	0.65	0.56
Q39	-0.01	0.04	0.45	0.05	0.19	0.06	0.66		0.14	0.57	0.46
Q1	0.04	0.06	0.44	0.06	0.15	0.07	0.69		0.17	0.54	0.44
Q18	0.01	0.02	-0.18	0.05	0.96	0.03	0.25		0.33	0.41	0.85
Q27	-0.08	0.04	-0.04	0.05	0.94	0.03	0.22		0.26	0.52	0.88
Q35	0.06	0.05	-0.18	0.05	0.91	0.04	0.29		0.36	0.39	0.83
Q17	-0.04	0.03	-0.03	0.04	0.90	0.03	0.25		0.28	0.52	0.87
Q22	0.06	0.04	-0.11	0.05	0.89	0.04	0.28		0.36	0.45	0.84
Q21	-0.05	0.05	0.01	0.05	0.88	0.04	0.25		0.28	0.54	0.86
Q36	0.06	0.03	0.00	0.03	0.85	0.03	0.25		0.36	0.53	0.87
Q24	-0.06	0.04	0.07	0.05	0.83	0.04	0.28		0.25	0.57	0.85
Q26	-0.04	0.04	0.07	0.05	0.77	0.04	0.36		0.25	0.53	0.80
Q30	0.01	0.03	0.14	0.05	0.75	0.04	0.29		0.30	0.60	0.84
Q20	-0.01	0.03	0.11	0.05	0.73	0.04	0.37		0.27	0.55	0.79
Q37	0.17	0.04	0.00	0.03	0.72	0.03	0.36		0.43	0.47	0.79

Further, the interconnection among the various aspects of quality of life being assessed was quite apparent. The first factor 'limitations related to palms' shared weak correlation with the second factor (rho = .171), a moderate correlation (rho = 0.364) with the second factor. On the other hand, the second factor had a rather strong correlation with the third factor, psychosocial impact.

Table 5.11 (continued)

	Factor pattern								Factor structure		
	F1	*SE*	*F2*	*SE*	*F3*	*SE*	*Res.Var*	*Issue*	*F1*	*F2*	*F3*
Q12	0.05	0.04	0.04	0.06	0.66	0.05	0.50		0.30	0.46	0.71
Q29	0.01	0.04	0.21	0.05	0.63	0.04	0.39		0.28	0.60	0.76
Q15	0.33	0.04	0.08	0.05	0.53	0.05	0.42		0.53	0.46	0.70
Q38	0.06	0.04	0.27	0.05	0.47	0.05	0.53		0.28	0.56	0.65
Q42	0.43	0.04	0.26	0.05	0.16	0.06	0.59	CL	0.53	0.43	0.48
Q32	-0.22	0.05	0.47	0.05	0.47	0.05	0.36	CL	0.03	0.71	0.67
Q47	-0.01	0.04	0.33	0.05	0.47	0.06	0.48	CL	0.22	0.62	0.67
Q34	0.47	0.04	0.05	0.06	0.36	0.06	0.49	CL	0.61	0.35	0.56
Q48	-0.06	0.04	0.45	0.05	0.47	0.05	0.34	CL	0.19	0.73	0.72
Q33	0.04	0.05	0.29	0.06	0.31	0.06	0.70	LL	0.20	0.49	0.50
Q14	0.37	0.04	0.21	0.05	0.35	0.05	0.48	LL	0.54	0.49	0.61

Note: CL, crossloading; LL, low-loading

This raises questions whether the two can be combined as one factor.

Item reduction process

A further step following EFA of the 36-item set involved removing items that showed poor performance, to enhance the structure of the instrument. By examining the relationship between the individual items and their related constructs, the best items could be identified and selected (Gorsuch 1997). Poorly performing items were iteratively removed, in total this included fourteen items (Table 5.10). Initially (Analysis 2 – 3) two items *I worry about my body odour* (Q33) and *my work is affected* (Q14) showing weak loading on all factors were removed. This did not seem to affect the factor loadings of the remaining items, although overall fit of the 3 factor solution was enhanced according to the practical fit statistics.

Next, an additional five items including *My appearance is affected* (Q32), *my skin feels uncomfortable* (Q42), *I worry about leaving sweat marks on things* (Q34), *I worry about my condition in future* (Q47) and *I find it hard to do things in advance* (Q48) which showed cross-loading and lacked strong loading on any factor were sequentially removed. The items *my sex life is affected* (Q38) and *my choice of clothing is affected* (Q1) were subsequently removed due to cross loading and high singularity, respectively. At each step of the EFA the optimal number of factors to be extracted was assessed, during the initial steps three factors were extracted. After the removal of the 9th item, Q1, two factors were extracted.

In view of the hierarchical level at which the construct of hyperhidrosis quality of life was being measured and the intended target population of the instrument, all hyperhidrosis patients, the first factor was considered to be a group-specific domain, with relevance only to patients with palmar sweating. Measuring quality of life impacts at the hyperhidrosis-type hierarchical level would have required items specific to the different areas affected. Ultimately such an approach would results in a much longer instrument, likely to have unfavourable applicability in a routine busy clinic. Thus items *I have difficulties holding objects* (Q4) and *I find it hard to handle paper* (Q13) were removed. Additionally *My choice of footwear is affected* (Q2) also loading onto the first factor was removed. Subsequently, items *my eyes feel irritated* (Q43), *I find hard to touch other people* (Q6) and *My choice of food and drinks is affected* (Q39) due to excessive uniqueness i.e. these items shared too little co-variation with the rest of the items.

Revised HidroQoL: 21 item set

Ultimately, the iterative item reduction process yielded a set of twenty one items which fitted to a two-factor solution (Table 5.10, Analysis 10; Table 5.12). Six items loading onto the first factor, were related to 'daily life activities', for example *My physical activities are affected* (Q8), *My everyday housework is affected* (Q11), *I worry about the additional chores in dealing with my condition* (Q45). Fifteen items loaded onto the second factor, these were related to psycho-social impact and included *I worry about people's reactions* (Q35), *I feel embarrassed* (Q21), *I feel nervous* (Q22), *I feel sad* (Q24), *I avoid public speaking* (Q12), and *I do not socialise as much as I would like to* (Q17). The two factors correlated strongly (rho = 0.645), suggesting that a single factor solution might fit the data. Moreover, the Horn's parallel analysis and scree-plot were in support of a single factor solution (Table 5.13, Figure 5.6) However, the two-factor solution showed much better fit based on goodness of fit statistics.

On the other hand, this would be indicative of the amount of measurement error with which the instrument was measuring the intended constructs (Lackey et al. 2003, p.162). The item-test correlation for the 'daily life activities impact' factor ranged from 0.68 (Q44) to 0.78 (Q3), for the psychosocial impacts this ranged from 0.7 (Q12) to 0.83 (Q36) (Table 5.14). The coefficient alpha for the two factors was 0.83 and 0.94, respectively.

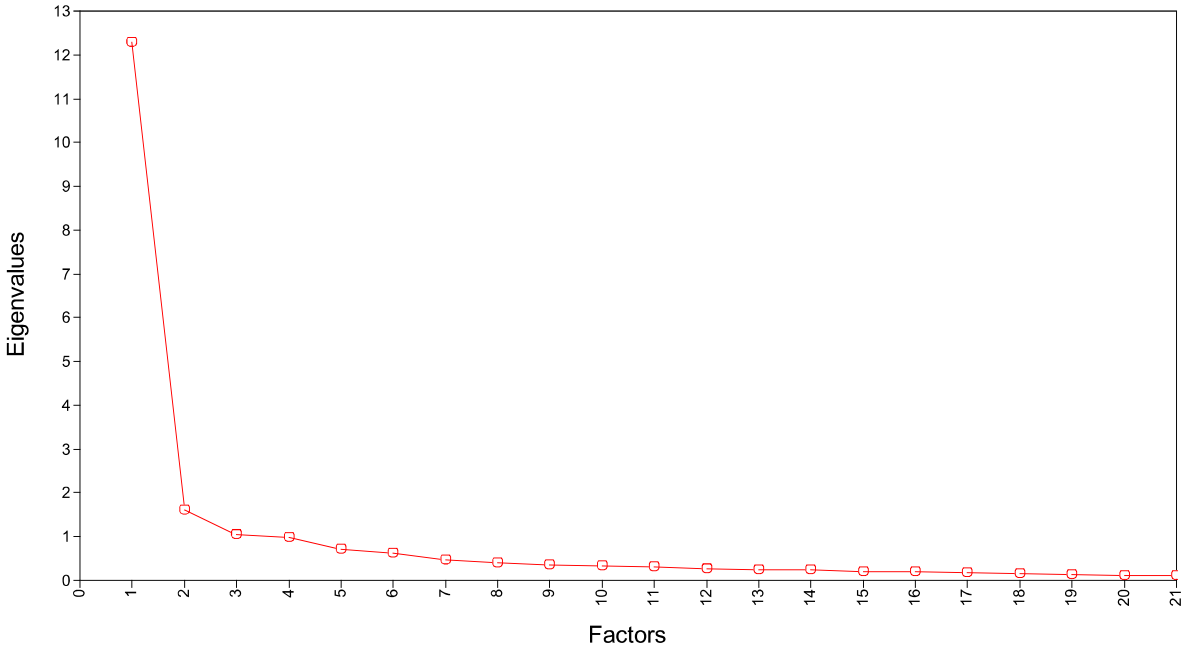
Table 5.12: Factor pattern matrix and residual variances for the 21 items of the HidroQoL

Item	Factor pattern					Factor structure	
	F1	*SE*	*F2*	*SE*	*Res.Var*	*F1*	*F2*
Q8	**0.89**	0.019	0.004	0.003	0.204	**0.892**	0.578
Q11	**0.832**	0.043	-0.144	0.053	0.441	**0.74**	0.393
Q7	**0.632**	0.04	0.133	0.046	0.475	**0.718**	0.541
Q45	**0.593**	0.052	0.235	0.061	0.414	**0.744**	0.618
Q3	**0.563**	0.037	0.292	0.042	0.385	**0.752**	0.655
Q44	**0.529**	0.051	0.215	0.059	0.527	**0.668**	0.556
Q29	0.176	0.048	**0.645**	0.041	0.406	0.592	**0.759**
Q15	0.168	0.048	**0.578**	0.044	0.513	0.541	**0.686**
Q30	0.129	0.04	**0.756**	0.034	0.286	0.616	**0.839**
Q12	0.082	0.052	**0.66**	0.045	0.489	0.507	**0.712**
Q20	0.079	0.042	**0.729**	0.035	0.388	0.55	**0.78**
Q26	0.027	0.051	**0.782**	0.04	0.362	0.531	**0.799**
Q37	0.015	0.045	**0.773**	0.037	0.387	0.513	**0.783**
Q24	0.011	0.04	**0.838**	0.032	0.285	0.552	**0.845**
Q21	0.004	0.046	**0.864**	0.034	0.249	0.561	**0.866**
Q36	0.001	0.03	**0.87**	0.024	0.242	0.562	**0.871**
Q17	0	0.033	**0.867**	0.026	0.248	0.56	**0.867**
Q27	-0.069	0.038	**0.926**	0.027	0.22	0.528	**0.882**
Q22	-0.083	0.041	**0.898**	0.029	0.283	0.497	**0.845**
Q18	-0.115	0.041	**0.933**	0.03	0.254	0.487	**0.859**
Q35	-0.171	0.049	**0.935**	0.034	0.303	0.432	**0.825**

Table 5.13: Eigenvalues for all 21 items: observed and Horn's parallel analysis

Factor	Observed	Adjusted	Unadjusted	Estimated bias
1	12.292	9.700	10.137	0.437
2	1.604	0.814	1.130	0.316
3	1.045	0.503	0.799	0.295
4	0.986	0.333	0.571	0.238
5	0.71	0.144	0.342	0.197
6	0.628	0.086	0.252	0.166
7	0.466	-0.020	0.128	0.148
8	0.404	-0.020	0.103	0.123
9	0.361	-0.003	0.073	0.076
10	0.343	-0.054	-0.001	0.053

Figure 5.6: Scree-plot based on the Eigen-values from the WLSMV for all 21 items



Part III: Confirmatory Factor Analysis of the 21 item HidroQoL

Confirmatory factor analysis technique was used to test the factorial structure of the HidroQoL obtained from the analysis in the previous section (the exploratory factor analysis results). The following hypotheses were tested:

- The HidroQoL has two factors, '*impact on daily life activities*' and '*psychosocial impacts*'. Six items, Q3, Q7, Q8, Q11, Q44, Q45 load on the first, and 15 items on the second, including Q12, Q15, Q17, Q18, Q20, Q21, Q24, Q26, Q27, Q29, Q35, Q37. There is no cross-loading of the items, although the two factors are hypothesised to be correlated.
- A single factor , into which all items have a loading, underlies the HidroQoL

First, the distribution of the item responses was explored, given that departures from multivariate normality distort goodness of fit indices (Ozer et al. 2009). There was no response category with more than 70% of the responses, among all items (Table 5.15). In eighteen out of the 21 items the upper-extreme response tended to have the highest frequency of endorsement. Nonetheless normality assumptions were not violated, skewness ranged from -1.75 to 0.82, while kurtosis ranged from 1.5 to 5.5 (six items had kurtosis exceeding 3). Item 'I *worry about the additional*

money in dealing with my condition (Q44) had missing data on 2% of the responses, thirteen items had missing data on 1% of responses, the other 7 items had no missing responses

Two factor model

The two-factor CFA model was statistically rejected based on a significant Chi-test (Chi-statistic = 316.18, df=188, $p < 0.001$) (Table 5.16). Practical fit indices were also considered, these reflected contrary findings. The RMSEA was 0.077 (.064, .092) indicating acceptable fit. The CFI and TLI at 0.98 and 0.977, respectively, both reflected excellent fit, further supported by the WRMR (.897). The item loadings for both factors were all significant, large, and had the expected signs.

Table 5.14: Correlation measures of internal consistency for the 21-item HidroQoL

Item	Item-test correlation	Item-rest correlation	ave. inter-item covariance	alpha
Daily life activities impact				
q8	0.76	0.65	0.84	0.79
q11	0.74	0.59	0.81	0.80
q7	0.69	0.55	0.88	0.81
q45	0.77	0.63	0.79	0.79
q3	0.78	0.66	0.79	0.79
q44	0.68	0.51	0.86	0.82
scale			**0.83**	**0.83**
Psychosocial impacts				
q29	0.71	0.67	0.93	0.94
q15	0.67	0.60	0.90	0.94
q30	0.82	0.79	0.88	0.94
q12	0.70	0.64	0.90	0.94
q20	0.75	0.70	0.90	0.94
q26	0.70	0.66	0.93	0.94
q37	0.75	0.70	0.91	0.94
q24	0.79	0.75	0.87	0.94
q21	0.71	0.68	0.95	0.94
q36	0.83	0.80	0.88	0.94
q17	0.81	0.78	0.88	0.94
q27	0.83	0.80	0.90	0.94
q22	0.77	0.73	0.91	0.94
q18	0.80	0.76	0.88	0.94
q35	0.74	0.71	0.94	0.94
scale			**0.9**	**0.94**

Table 5.15: Distribution of responses to the HidroQoL in Sample 2

	Number of patients						Proportion of patients									
Items	***1***	***2***	***3***	***4***	***5***	***M***	***1***	***2***	***3***	***4***	***5***	***M***	**Mean**	**Med**	**Skew**	**Kurt**
Q3	12	12	17	28	46	0	10%	10%	15%	24%	40%	0%	3.73	4	-0.77	2.33
Q7	7	14	20	34	40	0	6%	12%	17%	30%	35%	0%	3.75	4	-0.71	2.48
Q8	6	8	15	29	57	0	5%	7%	13%	25%	50%	0%	4.07	4	-1.18	3.44
Q11	23	19	19	27	27	0	20%	17%	17%	23%	23%	0%	3.14	3	-0.17	1.65
Q12	16	19	12	21	47	0	14%	17%	10%	18%	41%	0%	3.56	4	-0.52	1.75
Q15	24	11	14	11	55	0	21%	10%	12%	10%	48%	0%	3.54	4	-0.53	1.63
Q17	23	13	11	15	53	0	20%	11%	10%	13%	46%	0%	3.54	4	-0.54	1.64
Q18	24	18	11	23	38	1	21%	16%	10%	20%	33%	1%	3.29	4	-0.29	1.52
Q20	13	20	20	25	36	1	11%	17%	17%	22%	31%	1%	3.45	4	-0.38	1.84
Q21	0	3	8	24	79	1	0%	3%	7%	21%	69%	1%	4.57	5	-1.75	5.49
Q22	4	15	14	21	60	1	3%	13%	12%	18%	52%	1%	4.04	5	-0.96	2.63
Q24	18	17	20	13	46	1	16%	15%	17%	11%	40%	1%	3.46	4	-0.38	1.65
Q26	2	8	14	17	73	1	2%	7%	12%	15%	63%	1%	4.32	5	-1.41	3.91
Q27	4	7	15	22	66	1	3%	6%	13%	19%	57%	1%	4.22	5	-1.33	3.85
Q29	5	9	5	30	65	1	4%	8%	4%	26%	57%	1%	4.24	5	-1.54	4.41
Q30	17	13	20	27	37	1	15%	11%	17%	23%	32%	1%	3.47	4	-0.51	1.94
Q35	2	6	12	17	77	1	2%	5%	10%	15%	67%	1%	4.41	5	-1.66	4.87
Q36	18	16	20	25	35	1	16%	14%	17%	22%	30%	1%	3.38	4	-0.38	1.79
Q37	12	12	17	28	45	1	10%	10%	15%	24%	39%	1%	3.72	4	-0.75	2.31
Q44	54	20	13	10	16	2	47%	17%	11%	9%	14%	2%	2.24	2	0.82	2.19
Q45	39	23	17	13	22	1	34%	20%	15%	11%	19%	1%	2.61	2	0.41	1.7

The standardised loadings ranged from .63 for item Q3, to .87 for Q8 (Figure 5.7, Arrows from the boxes 'indicator variables' to the two eclipses '*factors*'). The item residuals ranged from .075 (Q3) to .014 (Q8) for Q3 and Q8, reflecting optimal fit. Based on the r-squared values, the '*daily life activities impact'* factor explains from 40% (Q3) to 78% (Q8) of variance in its related items (Table 5.17). The *psycho-social impact* factor on the other hand explains from 53% (Q15) to 87% (Q36) of variance in its related items.

DISCUSSION

It is pertinent that new HRQoL instruments are tested in the target population and under the conditions in which they will ultimately be applied prior to being brought into use. This permits the evaluation of how well the items function and an assessment of the adequacy with which the

conceptual framework has been translated into a measurement model. Further validation work and application of an HRQoL measure makes the implicit assumption that the instrument is internally valid. This study, therefore, set out to explore the general functioning of the HidroQoL in patients with hyperhidrosis. In addition, the measurement model of the HidroQoL was explored and tested. Also, item reduction was carried out. The initial item reduction was carried out using correlation analysis to remove multicollinear items. This often reflects redundancy in content suggesting the little contribution in measurement that such items actually make. The decision to remove items took into account the importance of the issue reflected in an item based on previous qualitative research (reported in Chapter 3). Thirteen items were removed at this stage.

Further item reduction was carried out based on EFA, leading to the removal of fifteen items. The process proceeded systematically, where removal of underperforming items was followed by further iterations of EFA, until a set of optimally performing items was achieved. Physical discomfort related issues, seemed nominal to the hyperhidrosis-quality of life, for instance 'my skin feels uncomfortable' (Q42), and 'my eyes feel irritated' (Q3).

Table 5.16: Goodness of fit of the CFA models estimated

Goodness of fit test	Threshold	2-factor	1-factor	Difference
Chi-square Test of Model fit				
Value		316.167	489.52	44.7
DF		188	189	1
P-value	> 0.05	0.0000	0.0000	0.000
Root Mean Square Error of Approximation				
Estimate	< 0.05	0.077	.118	
90 Percent C.I.		(0.064 0.092)	(0.105, 0.130)	
Probability RMSEA <= 0.05		0.002	0.000	
CFI	**> 0.95**	0.980	0.952	
TLI	**> 0.95**	0.977	0.947	
Chi-square Test of Model Fit for the Baseline Model				
Value		6487.1	6487.996	
DF		210	210	
P-value	> 0.05	0.0000	0.0000	
Weighted Root Mean Square Residual	**< 0.90**	0.897	1.216	

Table 5.17: Correlation between variables and construct (r-squared) and uniqueness (residual variance)

Item	2-factor model		1-factor model	
	R-squared	*Residual Variance*	*R-squared*	*Residual Variance*
Q8	0.78	0.22	0.55	0.45
Q11	0.47	0.53	0.33	0.67
Q7	0.57	0.43	0.42	0.59
Q45	0.74	0.26	0.54	0.46
Q3	0.40	0.61	0.27	0.73
Q44	0.75	0.25	0.57	0.43
Q35	0.61	0.39	0.59	0.41
Q18	0.90	0.10	0.90	0.11
Q27	0.82	0.18	0.81	0.19
Q22	0.62	0.38	0.60	0.40
Q36	0.76	0.24	0.75	0.26
Q17	0.83	0.17	0.82	0.18
Q21	0.74	0.26	0.73	0.27
Q24	0.82	0.18	0.81	0.19
Q26	0.71	0.29	0.70	0.30
Q37	0.66	0.34	0.65	0.36
Q30	0.87	0.13	0.86	0.14
Q20	0.54	0.46	0.52	0.48
Q12	0.59	0.42	0.57	0.43
Q29	0.56	0.44	0.55	0.45
Q15	0.53	0.47	0.52	0.48

The item on ‘clothing choices’ (Q1), among the most prevalent issues in a previous qualitative work, surprisingly performed poorly. This might be an artefact of the reliance of EFA on covariance matrices such that, in lowly endorsed and highly endorsed items, the item will exhibit minimal variation in its score. Although the item reduction process presents an opportunity for reducing respondent burden and enhancing a scale’s measurement attributes, caution is needed on how and when items can be removed. Coste et al. (1997) notes a lack of conceptualisation of the process and an overreliance on statistical approaches as common pitfalls during item reduction. Reise et al. (2000) emphasises the need for sound planning of the process, clarity on the construct being measured and a clear rationale reasons for removing any items during item reduction.

.

Figure 5.7: Path-diagram of the CFA Model with two correlated factors.

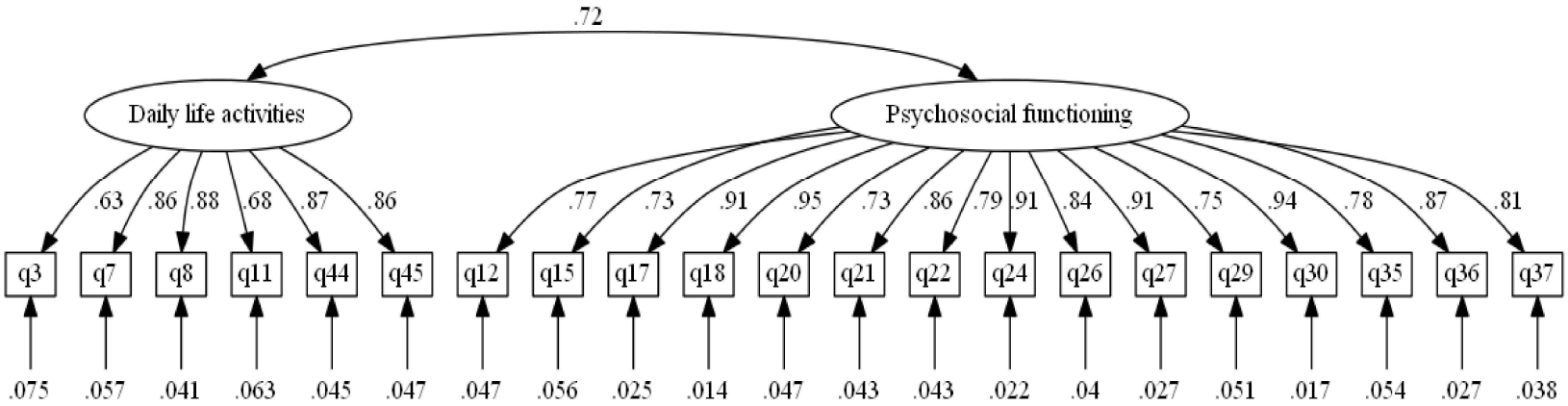
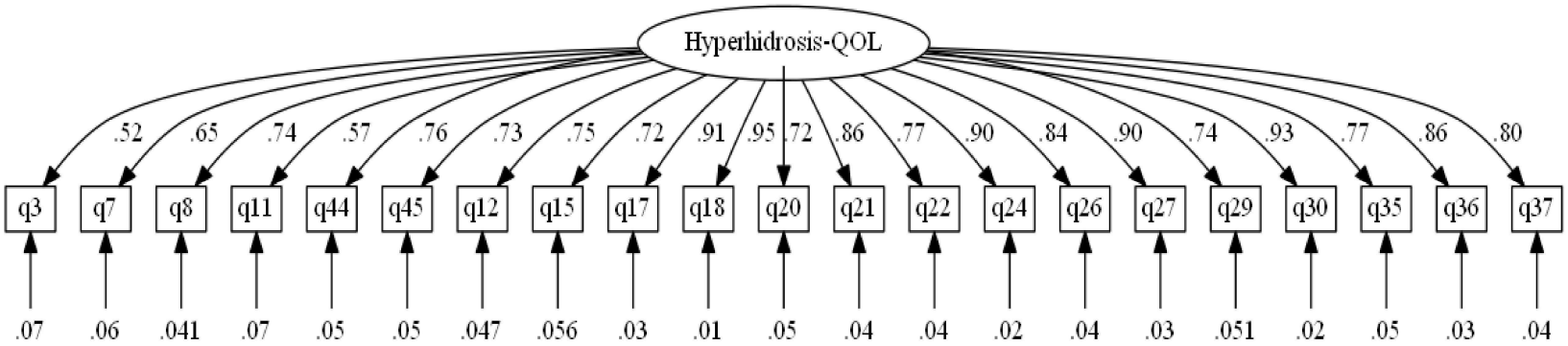


Figure 5.8: Path-diagram for CFA model with 1 –factor



Juniper et al. (1997) compared application of FA against 'patient importance ratings' in performing item reduction. They found the instruments to share much content although there were difference in some items. The FA method left out some items considered of highest importance to patients. Similarly, some items with very strong measurement attributes were not important to patients. This indicates that ultimately optimal measurement of HRQoL requires a clear method for reconciling such frictions, marrying statistical objectives with clinical sensibility. An EFA of the final 21 item set showed a two-factor solution, where six items loaded onto the first factor, 'impact on daily life activities' and fifteen items loaded onto the second factor 'psychosocial impact'. Both the factor pattern and factor structure matrices, reflecting the partial and non-adjusted correlations of the items with the factors showed that the final set of items were tapped strongly into their respective factors. There is some debate surrounding the factor loadings to be interpreted. Kline (Kline 1994) recommends the factor pattern matrix, citing that they best reflect the relationship between an item and the construct, given that the effects of other variables are accounted for, while Gorsuch (Gorsuch 1997) based on the same reasons argues that this undervalues the correlations which should be considered when evaluating such relationships. Given that both matrixes provide important information, Pat (Pat 2004) recommends considering both.

The two factors showed strong correlation, raising the question whether the two could have been combined into one factor. How large correlations ought to be before combining them into a single factor does not lend itself to a simple or clear cut answer (DeVellis 2011, p.146). Some suggestions indicate thresholds of 0.6 (Lackey et al. 2003). In this case, the researcher is left to his own devices. Still, this issues reflects a potent issue in PRO measurement, that of unidimensionality. First, a parallel analysis of the 21 items supported a single factor solution, although this was not followed on consideration of the goodness of fit indices, which were poor with a single factor. Second, the ratio of the eigenvalues of the first and second factors exceeded the minimum threshold of 4 (Hattie 1985; Basra et al. 2006). Third, the single factor model of the CFA showed fit on practical indices. Finally, in spite of the clear better fit that the 2 factor model had over the 1-factor model, the residuals when the single factor model was implemented were not inflated. These findings suggest that unidimensionality property is tenable for hyperhidrosis-QoL as a construct and the HidroQoL as an instrument for its measurement.

Dimensional structure has been tested and reported on one hyperhidrosis quality of life measure, the 'Hyperhidrosis Scale'. Kuo et al. (Kuo et al. 2004) based on a Korean population of patients awaiting surgery obtained 5 factors from an EFA of their newly developed 29-item questionnaire for hyperhidrosis-quality of life, which included: 'functional domain', 'psychological domain', 'social domain', 'affective domain' and a 'physical domain'. Reliance on eigenvalues for factor extraction in their study implies that over-extraction was quite likely. Moreover, the Kuo et al measure contains severity and symptom related questions which might be causal rather than indicator variables. Moreover in the current study these were found to make little contribution to the measurement of hyperhidrosis-related QoL and to lack fit in the dimensional structure of the HidroQoL, respectively.

SUMMARY

- A total of 559 patients from the USA completed the 49-item HidroQoL.
- In all items no response category accounted for more than 80% of responses, although 17 items had more than 50% of participants on the extreme category ("very much").
- Correlation analysis showed 30 pairs of items to be multi-collinear. Thirteen items were removed to resolve the content redundancy, leading to a 36-item version of the HidroQoL.
- Exploratory factor analysis on the 36-item measure showed that 3 factors best captured variability in scores; 28 items showed optimal fit, based on their factor loadings. Two items had poor fit (no loading >= 0.4), five items were cross-loading.
- Poorly performing items were sequentially removed, through iterations of EFA. An optimal "simple" factor structure was realised after eliminating 15 items. Two alternative interpretations were permitted by the results, a 1 factor solution and a two factor solution. The 2 factors in the latter were interpretable as 'impact on daily life activities' and 'psychosocial impacts'.
- The two alternative factor solutions for the HidroQoL were tested on a new group of patients (UK sample, N = 115) using confirmatory factor analysis.
- Although both solutions performed poorly on absolute goodness of fit tests, practical fit indices indicated good fit for both of them. Nonetheless the two factor solution provided a significantly better fit to the data.

CHAPTER 6

Development of a Hyperhidrosis-specific quality of life instrument (HidroQoL): Rasch analysis

Item reduction and construct validation using
Item Response Theory (IRT)

INTRODUCTION

In skin disease, the assessment of HRQoL plays a particularly central role, due to, among other factors, skin's high visibility and the strong connection between impacts on social life and self-perception, and severity of disease (Grob et al. 2005). This suggests a key role for HRQoL measures in skin disease, with the implication for the need to ensure their rigor in order to guarantee accurate, efficient and reliable assessment of HRQoL impact. For this, a clear conceptual model reflected in a robust measurement model are essential ingredients. The latter has been traditionally achieved through classical test theory (CTT) based methods such as item-total correlation analysis, stepwise regression and factor analysis (Coste et al. 1997). An increasingly applied method is based on Rasch model, a variation of the IRT. This is regarded by some as setting new rules in measurement and thus creating a new 'gold standard' in HRQoL instrument development (Reise and Henson 2003; Nijsten 2012).The Rasch model generates a linear metric scaled in logit-units, representing the construct being measured, on which both the items and persons are located hierarchically reflecting their levels on the construct (Prieto et al. 2003). Further, the probability of a particular response on an item by an individual is then given by a logistic function of the difference between the item location and person location and nothing else (Twiss et al. 2011).Then, following a prescriptive approach, items and persons are assessed for conformity to the model applying fit statistic (Nijsten et al. 2006a). Ultimately, demonstrating conformity to the Rasch model gives an instrument a number of advantages. First, ordinal scores into interval level scores, a requisite property for the calculation of effect sizes and other statistics in clinical research usually taken for granted (Reise and Haviland 2005). Second, by conceptualizing measurement error as an item level property, high reliability can be attained even with a shorter questionnaire, making it possible to minimize patient burden without compromising precision (Reeve et al. 2007). Additionally, the property of invariance of the item and person parameters facilitates a variety of highly useful analyses and validation hypotheses including, equating of measures, the testing for differential item functioning, among others (DeMars 2010). Whether an instrument based on the Rasch Model (RM) is significantly better than another based on CTT is still a hot topic. Nonetheless, for the practitioner the range of tools accessible for understanding the psychometric properties of an instrument and its items seems enough justification for the application to scale development, testing for properties otherwise taken for granted in CTT (Reise and Henson 2003).

OBJECTIVES

The objectives of this study were to:

- Evaluate the extent to which the new instrument (HidroQoL-36 version) conforms to the RM, assessing whether:
  - the scale was unidimensional
  - targeting of the items to the population of hyperhidrosis patients was optimal
  - response categories were functioning optimally
  - the items were invariant across groups according to gender, age, site of hyperhidrosis, severity of disease i.e. whether DIF was present.
  - item estimates remained invariant across patient populations, comparing UK and North-America.
- Identify and remove from the HidroQoL poorly performing items based on the RM.
- Evaluate the construct validity of the final version of the HidroQoL.

METHODS

Study design

This study followed a prospective cross-section design. The major design consideration in Rasch analysis study is ensuring that respondents reflect the entire continuum of the construct, from the highest possible quality of life impairment and to the minimum possible impairment (Bond 2004). To ensure this, a large and heterogeneous patient population reflecting varying levels of disease severity and different types of hyperhidrosis was targeted. The RM analyses can be carried on a sample as small as 100, nevertheless a sample size of at least 243 is large enough to achieve precision of ± 0.5 logits within at 99% level of confidence even in heavily skewed data (Linacre 1999). For stable estimation of category thresholds, at least 10 observations are needed in each response category of an instrument. In view of this, the recruitment targeted 400 patients, representing the full range of disease severity according to the Hyperhidrosis Disease Severity Scale (HDSS).

Study population

Inclusion criteria:

- Self-reported hyperhidrosis.
- Aged 18 years or above.

- With a score of 2 or higher on the HDSS
- With onset of hyperhidrosis in teenage years or early adult years.

***Exclusion criteria*:**

- Below the age of 18
- With onset of hyperhidrosis after age of 30 and reporting a co-morbidity (hypertension, diabetes, pm hormonal disorders, psychological disorders)
- With HDSS score of 1.

Recruitment and data collection procedure

Patients were recruited through hyperhidrosis online social networking communities, mainly the International Hyperhidrosis Society (IHHS) and the UK Hyperhidrosis support group, from May to September 2012. A detailed description of the study population and procedures is provided in Chapter 2.

Data processing and analysis

Data analysis based on the Rasch model was carried out using RUMM 2030. Factor analysis was carried out using MPLUS-6 and STATA 11. A detailed description of the Rasch model and its application in scale development is available in Chapter 2.

RESULTS

For purposes of clarity the results will be presented in three parts. Part I: calibration of the HidroQoL, Part II: item reduction and refinement, and Part III: cross-validation of the refined HidroQoL. Patients recruited for this study were grouped as follows: sample 1 comprised of study participants from the U.S and Canada, and sample 2 which included respondents from the UK Analysis reported in Part I and Part II utilized sample 1, while part III employed sample 2.

Part I: Calibration of the HidroQoL using the Rasch Model

Sociodemographic characteristics of study participants (sample 1)

A total of 595 patients with hyperhidrosis were recruited for this study. Of the total, 113 (19%) were male and 482 (81%) were female (Table 6.1). The mean age was 25 years, with those aged between 18 and 39 making up 52% of the sample (Figure 6.1). Forty patients reported living with hyperhidrosis for a period of less than 10 years, on the other hand those with duration of disease

from 20 to 29 years comprised the largest number (Figure 6.2). For most patients, hyperhidrosis affected multiple areas: for example in 27% of the participants hyperhidrosis affected the hands, feet and axillar. Forty-six percent of the patients (N = 274) had sweating that was intolerable and always interfered with daily activities (HDSS score = 4) (Figure 6.3).

Forty-nine percent of the patients (N = 294) perceived their sweating to have an extreme negative impact on their overall life (Figure 6.4). Eighty-six percent of the patients (n=513) reported previously visiting their doctor in relation to their sweating condition, while 14% had not (Table 6.2)**.** A lesser number (n = 324) reported getting treatment within the last 6 months. An even smaller number (n = 176) reported receiving treatment currently. Forty-eight patients had received a Botox injection within the last 6 months.

HidroQoL affirmation responses

Further exploratory analysis of the data involved analyzing the distribution of responses for each item. The items showed a positive skew towards the higher response categories (Table 6.3). All items except Q4, Q5, Q39, Q43 and Q45 had ceiling effects. In contrast, 13 items (Q4, Q5, Q13, Q16, Q19, Q25, Q38 – Q41, and Q43 – Q45) showed floor effects. Floor or ceiling effects are seen if either of the items extremities has at least 20% of responses (Both et al. 2007). Nevertheless, this did not compromise meaningful variability in the data, with 80% considered as the upper limit of endorsement for categories (Streiner and Norman 2008, p.84). Noteworthy is the connection between ceiling and floor effects in the response data, all items without ceiling effects show floor-effects. The influence of the severity of disease and its impact, as reported in this sample, on the response pattern should not be ignored. The majority of patients had severe HH. On the other hand, questions of appropriateness of the response categorisation and the level of difficulty of the items can still be raised.

For example in relation the former, items Q21, Q31 and Q35 show less than 10 observations in the response category “no, not at all”. This can be a problem in the context of Rasch modeling, particularly in estimating stable threshold values (Bond and Fox 2007, p.222). A general recommendation is to have a minimum of 10 observations per category (Linacre 1999). There was no pattern to the missing data, thus data is missing at random. Nevertheless the rates of missing data increase towards the end of the questionnaire, for instance before Q18, no item has missing information. Further, missing data only reaches 2% at Q38. This item deals with a very personal

aspect of life, which some respondents may not be willing to talk about. On the other hand, the higher rates of missing data notable among items with floor effect might also be a reflection of their relevance. Moreover, there was no choice provided for responses that were not applicable.

Table 6.1: Sociodemographic characteristics of study participants

Gender, n (%)	
Male	113 (19%)
Female	482(81%)
Age (years)	
Mean (SD)	40.5 (14.2)
Median	39
Range	18 - 74
Age (years), n	
18 to 29	154
30 to 39	155
40 to 49	117
50 to 59	95
≥ 60	74
Duration of condition (years)	
Mean (SD)	27.5 (14.1)
Median	25
Range	2 - 69
Duration of condition (years), n	
< 10	40
10 to 19	144
20 to 29	156
30 to 39	116
40 to 49	79
50 to 59	49
60 to 69	11
Body are affected	
Head*	129 (22%)
Axilla*	54 (9%)
General	130 (22%)
Axilla, Palms, Feet	158 (27%)
Palms and Feet	124 (21%)

Table 6.1 (continued)

Severity of disease (HDSS score), n	
1	0
2	79
3	242
4	274
Global impact of hyperhidrosis (GQ score), n	
No, none at all	0
Slight	3
Moderate	50
Quite a bit	230
Extreme	290
Co-morbidity	
None	327 (55%)
Menopausal complaints	61(10%)
Diabetes	30 (5%)
Hypertension	47 (8%)
Neurological disorders	64 (11%)
Thyroid disorders	66 (11%)
Employment status	
Employed	380(64%)
Unemployed	107 (18%)
Retired	70(12%)
Full-time student	30(6%)

* This category included patients reporting other areas, the indicated area thus reflects the predominant body area affected.

Calibrating the HidroQoL on the Rasch model

Rasch analysis was carried out on the HidroQOL-36, a version of the HidroQoL containing 36 items, developed in the previous chapter following resolution of item redundancy in the initial developmental HidroQoL (HidroQoL-49). The Likelihood ratio test for choice of appropriate RM supported the use of the partial credit model, the Chi-statistic was 529.47 (degrees of freedom =

1.04) and was significant ($p < 0.001$). Partial credit model allows the differences between category thresholds of items to vary across items (Masters and Wright 1997). In this case, restricting such differences to be equal as is assumed in the Rating Scale model would lead to a loss of information (Tennant and Conaghan 2007).

Figure 6.1: Age distribution of study participants

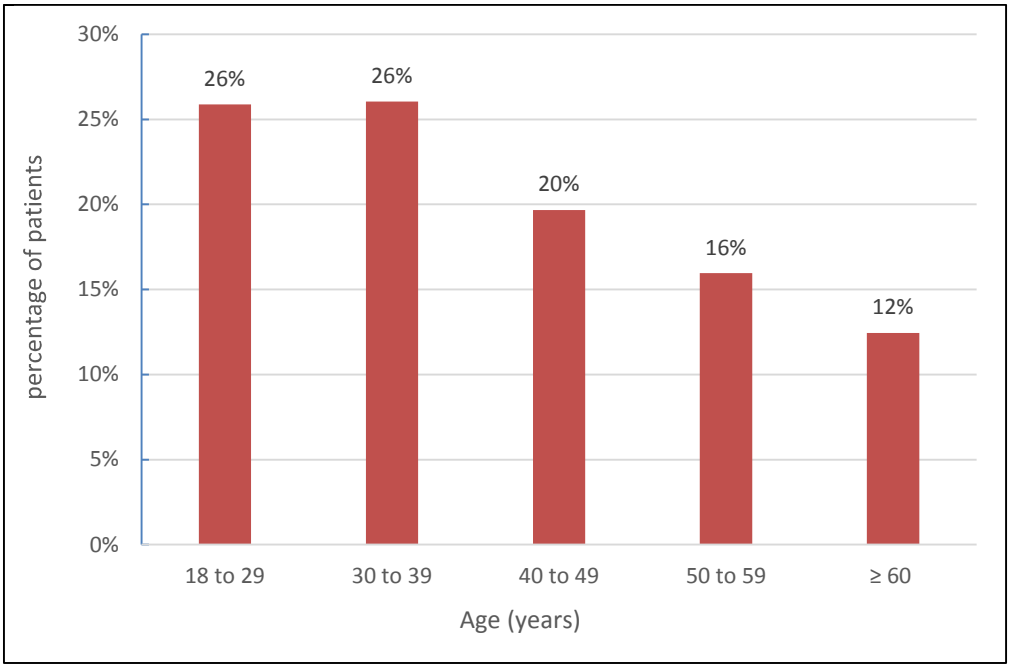
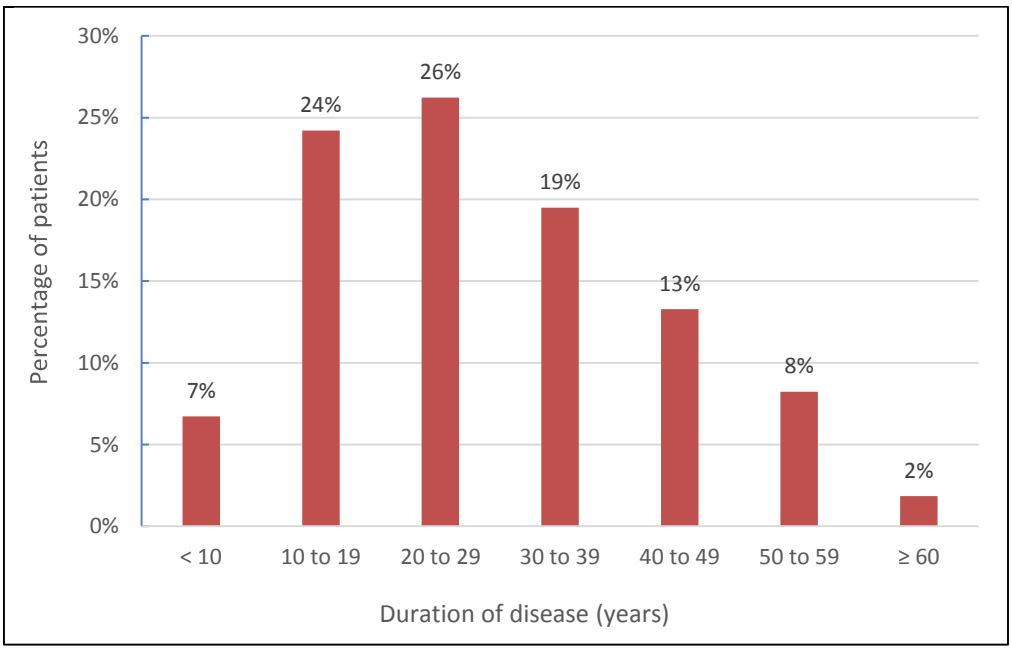


Figure 6.2: Duration of disease of the study participants



HidroQoL-36

Overall fit of the HidroQOL-36 to the RM was assessed using the Item Trait Interaction Chi-Squared Statistic (ITICS) and the mean fit residuals. The former was 1642.32 (df = 324) and was significant ($p < 0.001$), indicating a lack of fit of HidroQoL-36 to the RM (Table 6.4, Analysis 1).

Figure 6.3: Patient reported disease severity based on the HDSS

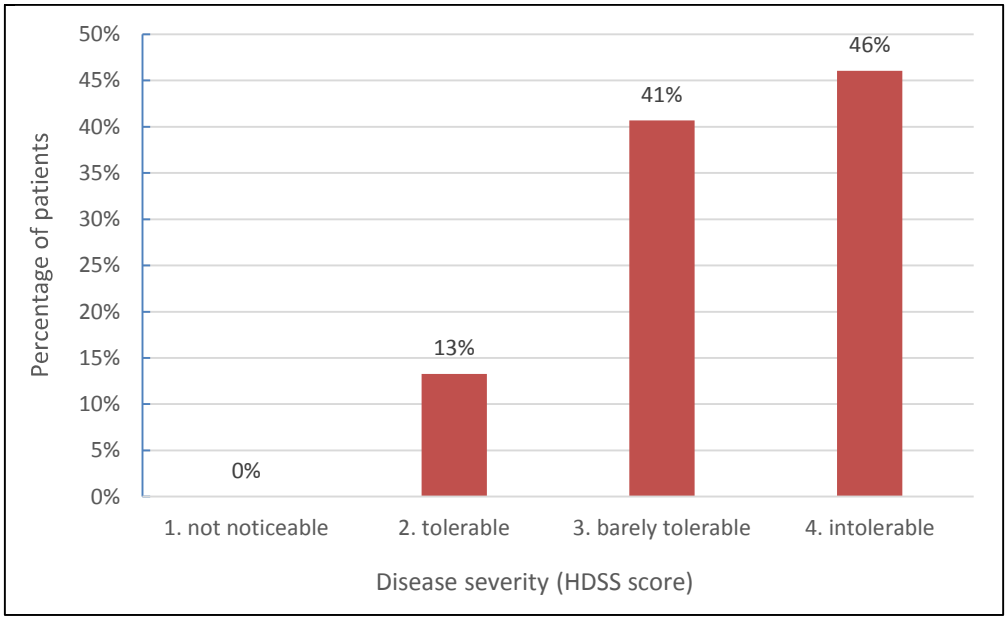
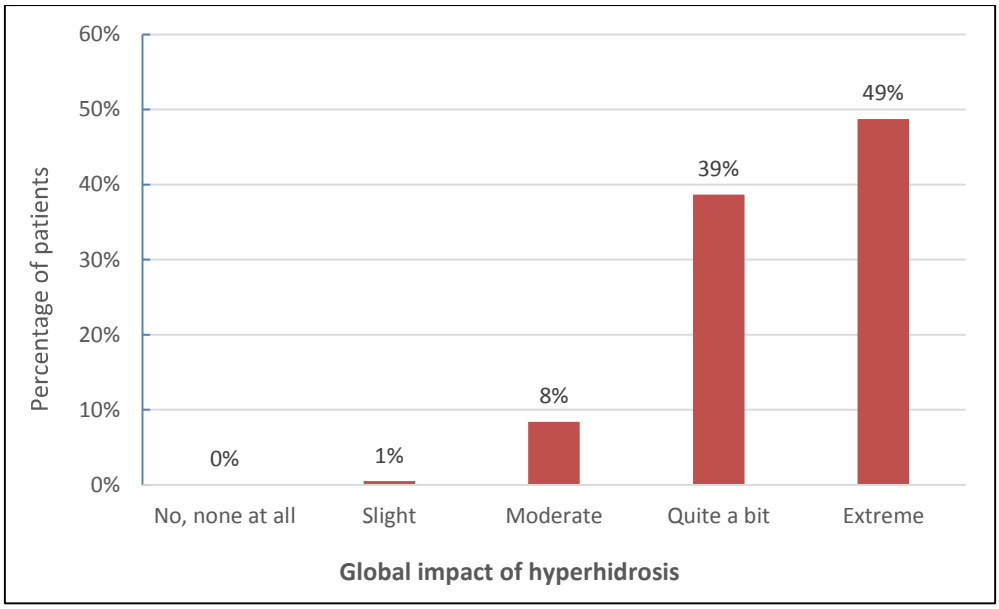


Figure 6.4: Patient reported general impact of disease



The item and person residual mean values reflected same conclusion (items, residual mean = 0.22, SD = 3.96; and persons, residual mean = -0.00, SD = 1.48). Thus, although the items showed poor

fit, the sample largely responded in conformity to the RM. Furthermore, model fit was also explored at the individual item and person level, using fit residuals and ITICS. The fit residuals for 15 items fell between -2.5 and 2.5, indicating optimal fit (Table 6.5). Eleven items had fit residuals exceeding 2.5 and an additional ten had fit residuals below -2.5, indicating underfit and overfit. The under-fitting group (fit residuals > 2.5) largely included items relevant to effects related to particular body areas.The over-fitting group (fit residuals < -2.5) primarily included items relating to negative emotions and the social impacts of hyperhidrosis, giving hints on the sources of the poor fit. Suboptimal response categorization is one cause of poor item fit to the RM (Linacre 1999). For an optimally functioning response categorisation, the choice of categories is expected to conform to the rasch probabilistic pattern. Thus, functioning of the response categorisation for the HidroQoL was tested. Three items (Q8, Q29 and Q35) showed appropriately ordered category thresholds, where consecutive category thresholds increased with increasing levels of the latent variable (quality of life impact) (**Figure 6.5**). The rest of the items had disordered thresholds, where the monotonicity of the thresholds was violated. This implies that for these items response categories was used inconsistently for example respondents struggling to distinguishing between response categories. Measures applied in clinical practice need to be optimally targeted for the intended population (Pallant and Tennant 2007). The mean location parameters for persons and items were, therefore, compared. Furthermore, the spread of the items along the latent variable was also analysed. The mean person location was 0.5 (± 0.82) in comparison to that of 0 (± 0.6) for items.

Table 6.2: Access to and utilization of treatment

	Number of patients (%)
Seen a doctor in relation to hyperhidrosis	513 (86%)
Treated within last 6 months	201 (34%)
Has received Botox within last 6 months	48 (8%)
Surgical treatment	64 (11%)
Currently being treated	176 (30%)

Table 6.3: Patients responses to the HidroQoL

Item	Descriptor	No, not at all		A little		Somewhat		Quite a bit		Very much		Missing	
		n	%	n	%	n	%	n	%	n	%	n	%
Q1	My choice of clothing is affected	18	3%	30	5%	51	9%	110	18%	386	65%	0	0%
Q2	My choice of footwear is affected	106	18%	38	6%	66	11%	72	12%	313	53%	0	0%
Q3	My holidays are affected (e.g. planning, activities)	70	12%	67	11%	143	24%	121	20%	194	33%	0	0%
Q4	I have difficulties holding objects	201	34%	71	12%	96	16%	112	19%	115	19%	0	0%
Q5	I have difficulties handling money	262	44%	71	12%	102	17%	79	13%	81	14%	0	0%
Q6	I find it hard to touch other people	98	16%	43	7%	67	11%	89	15%	298	50%	0	0%
Q7	My hobbies are affected	52	9%	51	9%	95	16%	153	26%	244	41%	0	0%
Q8	My physical activities are affected	29	5%	52	9%	77	13%	129	22%	308	52%	0	0%
Q9	My outdoor activities are affected	30	5%	52	9%	81	14%	110	18%	322	54%	0	0%
Q10	My summer activities are affected	21	4%	37	6%	67	11%	96	16%	374	63%	0	0%
Q11	My everyday housework is affected	107	18%	89	15%	127	21%	124	21%	148	25%	0	0%
Q12	I avoid public speaking (e.g. presentations)	84	14%	75	13%	72	12%	96	16%	268	45%	0	0%
Q13	I find it hard to handle paper	193	32%	49	8%	69	12%	91	15%	193	32%	0	0%
Q14	My work is affected	65	11%	60	10%	132	22%	149	25%	189	32%	0	0%
Q15	My career decisions are affected (e.g. career choice)	107	18%	49	8%	69	12%	111	19%	259	44%	0	0%
Q16	I have difficulties using touch-technologies (e.g. computer-keyboard, smart phones)	198	33%	47	8%	90	15%	105	18%	155	26%	0	0%
Q17	I do not socialise as much as I would like to	60	10%	69	12%	98	16%	113	19%	255	43%	0	0%
Q18	I avoid meeting new people	117	20%	75	13%	114	19%	138	23%	148	25%	3	1%
Q19	I avoid going out	135	23%	88	15%	118	20%	133	22%	118	20%	3	1%
Q20	My personal relationships are affected	89	15%	87	15%	139	23%	120	20%	157	26%	3	1%
Q21	I feel embarrassed	9	2%	15	3%	44	7%	89	15%	435	73%	3	1%
Q22	I feel nervous	30	5%	35	6%	70	12%	126	21%	331	56%	3	1%
Q23	I feel hopeless	110	18%	73	12%	95	16%	79	13%	235	39%	3	1%
Q24	I feel sad	117	20%	109	18%	99	17%	79	13%	188	32%	3	1%
Q25	I feel depressed	152	26%	115	19%	102	17%	66	11%	157	26%	3	1%
Q26	I feel frustrated	22	4%	42	7%	61	10%	99	17%	367	62%	4	1%
Q27	My self-confidence is affected	28	5%	51	9%	81	14%	117	20%	314	53%	4	1%
Q28	My self-esteem is affected	47	8%	51	9%	78	13%	110	18%	305	51%	4	1%
Q29	Sweating is constantly on my mind	17	3%	47	8%	70	12%	146	25%	311	52%	4	1%
Q30	I avoid taking on new challenges	93	16%	77	13%	131	22%	118	20%	172	29%	4	1%

Table 6.3 (continued)

Item		No, not at all		A little		Somewhat		Quite a bit		Very much		Missing	
		n	%	n	%	n	%	n	%	n	%	n	%
Q31	I feel self-conscious	9	2%	27	5%	50	8%	108	18%	397	67%	4	1%
Q32	My appearance is affected	33	6%	52	9%	95	16%	113	19%	298	50%	4	1%
Q33	I worry about my body odour	102	17%	98	16%	105	18%	105	18%	181	30%	4	1%
Q34	I worry about leaving sweat marks on things	30	5%	35	6%	49	8%	82	14%	395	66%	4	1%
Q35	I worry about people's reactions	7	1%	27	5%	58	10%	126	21%	370	62%	7	1%
Q36	I find it hard to be near other people	79	13%	69	12%	129	22%	136	23%	175	29%	7	1%
Q37	I feel uncomfortable physically expressing affection (e.g. hugging others)	49	8%	57	10%	103	17%	135	23%	243	41%	8	1%
Q38	My sex life is affected	172	29%	92	15%	128	22%	67	11%	125	21%	11	2%
Q39	My choice of food and drinks is affected	260	44%	88	15%	108	18%	58	10%	71	12%	10	2%
Q40	I feel uncomfortable in my shoes	138	23%	70	12%	66	11%	89	15%	223	37%	9	2%
Q41	I have problems with being barefooted	188	32%	49	8%	38	6%	61	10%	251	42%	8	1%
Q42	My skin feels uncomfortable	83	14%	69	12%	99	17%	113	19%	221	37%	10	2%
Q43	My eyes feel irritated	305	51%	72	12%	87	15%	61	10%	59	10%	11	2%
Q44	I worry about the additional money spent in dealing with my condition	173	29%	130	22%	105	18%	75	13%	101	17%	11	2%
Q45	I worry about the additional chores in dealing with my condition	155	26%	126	21%	103	17%	110	18%	94	16%	7	1%
Q46	I worry about the additional time spent in dealing with my condition	119	20%	108	18%	116	19%	111	19%	131	22%	10	2%
Q47	I worry about my condition in future	61	10%	61	10%	77	13%	144	24%	244	41%	8	1%
Q48	I find it hard to do things without planning in advance	90	15%	64	11%	112	19%	117	20%	204	34%	8	1%
Q49	My whole life is affected	23	4%	52	9%	71	12%	118	20%	321	54%	10	2%

This indicates that the HidroQoL was at a slightly lower level of HRQoL impairment in comparison to the sample. The item-person distribution map, shows an even distribution of the items across the latent variable (Figure 5.6). Reliability was assessed using the PSI. The HidroQOL-36 showed a PSI of 0.94, reflecting capability to distinguish up to 4 levels of QOL impairment patient groups. This is way ahead of the minimum levels needed for individual level use.

Part II: Item Reduction and Refinement

Following the initial calibration of the HidroQoL-36 which showed poor fit both for the overall model as well as for the individual items and persons, revisions were made to the instrument, based on the RM (Pallant and Tennant 2007). First, the category threshold disordering was addressed and then the misfit in the items. Subsequently, unidimensionality and local independence assumptions and the invariance property were tested.

Revision of response categorisation

One way of addressing disordered response category thresholds is to combine adjacent categories (Bond and Fox 2007). Thus, for the HidroQoL-36, first, categories 'somewhat' and 'a little' were combined; leading to a rescoring of the categories as 0-1-1-2-3. This resolved disordering in six items. Overall fit to the model showed a slight improvement, although the total-ITICS statistic was still significant (Table 6.4, Analysis 2). Further, the response category 'quite a bit' was combined with 'a little' and "somewhat', leading to a 3 point scaling, score as 0-1-1-1-2. This resolved the disordering of thresholds in the rest of the items as well as improving overall fit to the RM (Figure 6.7, Figure 6.8). The total ITICS statistic declined from 1642.64 (df = 324) before any rescoring to 1087.4 (df = 324) post collapsing to 3 categories. This had minimal consequences on reliability (PSI = 0.94). Seven items (Q44, Q6, Q1, Q21, Q37, Q48, Q18) showed improved fit. Thus in total twenty-one items showed optimal RM fit, while the remaining fifteen lacked fit.

Removal of misfitting items

The last resort for items not fitting the RM is to remove them from an instrument. Such items may lead to biased estimates of latent variables as well as item parameters in case of overfitting or worse still, ability to accurately measure may be degraded in the case of overfits (Baghaei 2008b; Smith et al. 2008). Still, removal of items ought to consider impact on the entire scale. Therefore, misfitting items were sequentially removed from the HidroQoL-36, iteratively assessing their impact on overall model fit as well as the remainder of the items. Four underfitting items, Q2- *My choice of footwear is affected*; Q4- *I have difficulties holding objects*; Q13- *I find it hard to handle paper*; Q43- *My eyes feel irritated* relevant only to hyperhidrosis affecting particular body areas were the first candidates for removal. The ITICS statistic declined to 768.26 (df = 288), reflecting improvement in fit (Table 6.6, Analysis 4). The item fit residual mean (SD), at -0.12(3.35) was still

suboptimal. The person fit residual (SD), -0.31(1.59) had slightly improved, though still out of optimal range. The PSI of 0.94 reflected that reliability remained strong.

Other items were impacted: specifically, fit of Q1 and Q6 deteriorated, while fit of Q26 improved. Such unexpected/unwanted consequences reflect a shift in the instrument's frame of reference. In the subsequent nine steps, 14 items were removed sequentially as shown in Table 5.6 (Analysis 5 – 12). At each stage, the impact of removing misfitting items on overall fit, including total ITICS statistic, the mean item and person residual fit statistics, reliability (according to the PSI) as well as the individual item fit of the remaining items was evaluated. Overall fit showed a steady improvement: the total ITICS statistic decreased to 194 (df=162) although it remained significant ($p = 0.04$). The mean fit residual and its SD for items and persons both fell to acceptable ranges, at -0.12 (1.22) and -0.31(1.37), respectively. Reliability was still adequate for individual level analyses despite declining to 0.89. Fit of item Q22 showed changes in both ways during 2 iteration stages: improving when Q39, Q6 and Q11 were removed (at stage 5), worsening at stage 6, when Q36, Q27 and Q30 were removed, then improving again when Q33 and Q42 were removed. In the end, 18 eighteen items showing good fit to the RM were achieved (Table 6.7).

Table 6.4: Overall model fit statistics for the 36 items HidroQoL and subsequent versions after rescoring

Action	Overall Model Fit			Item Fit Residuals		Person Fit Residuals		Dimension-ality	PSI
	Chi	*df*	*p*	*Mean*	*SD*	*Mean*	*SD*	*Sign. t-test(%)*	
1. All 36 items included	1642.64	324	0.00	0.22	3.96	-0.01	1.5	26.39%[1]	0.94
2. Revise scoring to 01123	1404.5	324	0.00	0.00	3.9	-0.089	1.55	33.28%[2]	0.94
3. Revise scoring to 01113	1087.4	324	0.001	-0.05	3.48	-0.234	1.64	28.57%	0.94

Note:

1. Items 13, 4, 2, 6, 7, 42, 34 had positive loadings of 0.3 and above. Items 29, 18, 26, 20, 48, 36, 32, 30, 17, 21, 24 and 27 had negative loadings.
2. Items 13, 4, 2, 6, 7, 42, 34 had positive loadings of 0.3 and above. Items 26, 35, 18, 29, 22, 48, 32, 20, 17, 36, 21, 30, 24, 36 and 27 had negative loadings below – 0.3.

Table 6.5: Rasch model item parameters for the 36-items of the HidroQoL

	Fitting items	Location	SE	Fit Resid	ChiSq.	Prob	Thresholds			
							1	2	3	4
I36	I find it hard to be near others	0.28	0.07	-4.95	48.65	0.00	-0.33	-0.62	0.25	0.70
I30	I avoid taking on new challenges	0.41	0.07	-4.12	33.99	0.00	-0.27	-0.58	0.37	0.49
I27	My self-confidence is affected	-0.76	0.07	-4.10	30.05	0.00	-0.72	-0.01	0.37	0.35
I37	I feel uncomfortable physically expressing affection	-0.28	0.07	-3.27	30.11	0.00	-0.34	-0.40	0.23	0.51
I20	My personal relationships are affected	0.52	0.07	-3.17	31.20	0.00	-0.45	-0.50	0.33	0.61
I17	I do not socialize as much as I would like to	-0.17	0.07	-3.10	19.20	0.02	-0.38	-0.17	0.33	0.22
I24	I feel sad	0.58	0.06	-3.04	25.97	0.00	-0.37	0.08	0.43	-0.13
I22	I feel nervous	-0.94	0.08	-3.02	33.21	0.00	-0.16	-0.27	0.10	0.33
I18	I avoid meeting new people	0.64	0.07	-2.76	22.36	0.01	-0.09	-0.54	-0.07	0.70
I21	I feel embarrassed	-2.08	0.09	-2.62	28.34	0.00	-0.38	-0.28	0.45	0.20
I29	Sweating is constantly on my mind	-0.90	0.08	-2.37	24.52	0.00	-0.95	0.12	0.14	0.70
I48	I find it hard to do things without planning in advance	0.17	0.07	-2.37	19.89	0.02	-0.01	-0.51	0.24	0.28
I26	I feel frustrated	-1.05	0.08	-1.92	16.08	0.07	-0.47	0.16	0.28	0.03
I03	My holidays are affected (e.g. planning, activities)	0.12	0.07	-1.74	12.81	0.17	-0.29	-0.66	0.51	0.44
I14	My work is affected	0.08	0.07	-1.41	12.32	0.20	-0.17	-0.72	0.19	0.70
I35	I worry about people's reactions	-1.59	0.09	-1.40	22.01	0.01	-1.47	0.10	0.54	0.84
I07	My hobbies are affected	-0.29	0.07	-1.29	13.99	0.12	-0.14	-0.40	0.01	0.54
I08	My physical activities are affected	-0.67	0.07	-1.07	8.68	0.47	-0.60	0.07	0.17	0.36

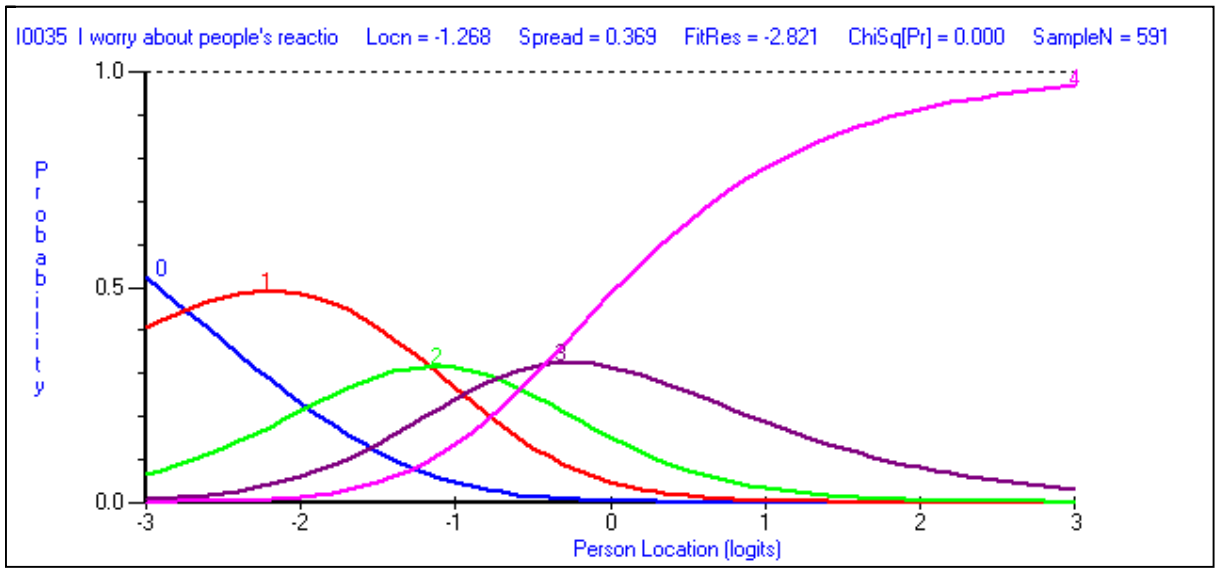
Note: SE., Standard error; FitResid., Fit Residual from the Rasch model; ChiSq., Chi-squared; Prob., p-value of the Chi-squared statistic.

Table 6.5: (continued)

No	Fitting Items	Location	SE	Fit Resid	ChiSq	Prob	Threshold			
I15	My career decisions are affected (e.g. career choice)	-0.02	0.06	-0.76	7.75	0.56	0.56	-0.27	-0.22	-0.07
I45	I worry about the additional chores in dealing with my condition	1.28	0.07	-0.72	8.27	0.51	-0.49	-0.06	-0.16	0.71
I38	My sex life is affected	0.99	0.07	-0.33	6.23	0.72	0.06	-0.53	0.59	-0.12
I32	My appearance is affected	-0.60	0.07	-0.22	3.32	0.95	-0.45	-0.23	0.49	0.20
I34	I worry about leaving sweat marks on things	-1.05	0.08	-0.18	16.68	0.05	0.05	0.18	0.19	-0.41
I12	I avoid public speaking (presentations)	-0.03	0.06	0.32	8.93	0.44	-0.07	0.23	0.05	-0.21
I42	My skin feels uncomfortable	0.09	0.07	0.53	11.23	0.26	0.01	-0.27	0.15	0.11
I47	I worry about my condition in the future	-0.16	0.07	0.67	5.31	0.81	-0.18	-0.03	-0.21	0.42
I04	I have difficulties holding objects	1.01	0.07	2.66	28.31	0.00	0.52	-0.58	-0.29	0.35
I01	My choice of clothing is affected	-1.14	0.08	3.09	25.03	0.00	-0.13	0.13	0.05	-0.05
I44	I worry about the additional money in dealing with my condition	1.23	0.07	3.38	6.93	0.64	-0.33	-0.02	0.25	0.10
I06	I find it hard to handle paper	-0.21	0.06	3.44	40.61	0.00	0.74	-0.23	0.00	-0.51
I33	I worry about my body odour	0.46	0.06	3.63	25.46	0.00	-0.23	-0.02	0.18	0.07
I11	My everyday housework is affected	0.57	0.07	4.13	45.41	0.00	-0.18	-0.40	0.20	0.38
I43	My eyes feel irritated	1.63	0.07	4.22	83.73	0.00	0.71	-0.69	-0.07	0.06
I39	My choice of food and drinks is affected	1.52	0.07	4.94	35.83	0.00	0.35	-0.66	0.35	-0.05
I13	I find it hard to handle paper	0.51	0.06	5.97	105.73	0.00	1.09	-0.45	-0.25	-0.39
I02	My choice of footwear is affected	-0.18	0.06	9.69	256.28	0.00	1.11	-0.30	0.14	-0.95

Figure 6.5: Category Probability Curves of the original 5-category HidroQoL

a. *Item 35 (I worry about people's reactions) had appropriately ordered category thresholds: the curve for each response category shows a unique peak.*



b. *Item 7 (My hobbies are affected) shows disordered category thresholds: response category curve for score '1' reflecting the option 'a little' does not have a unique peak.*

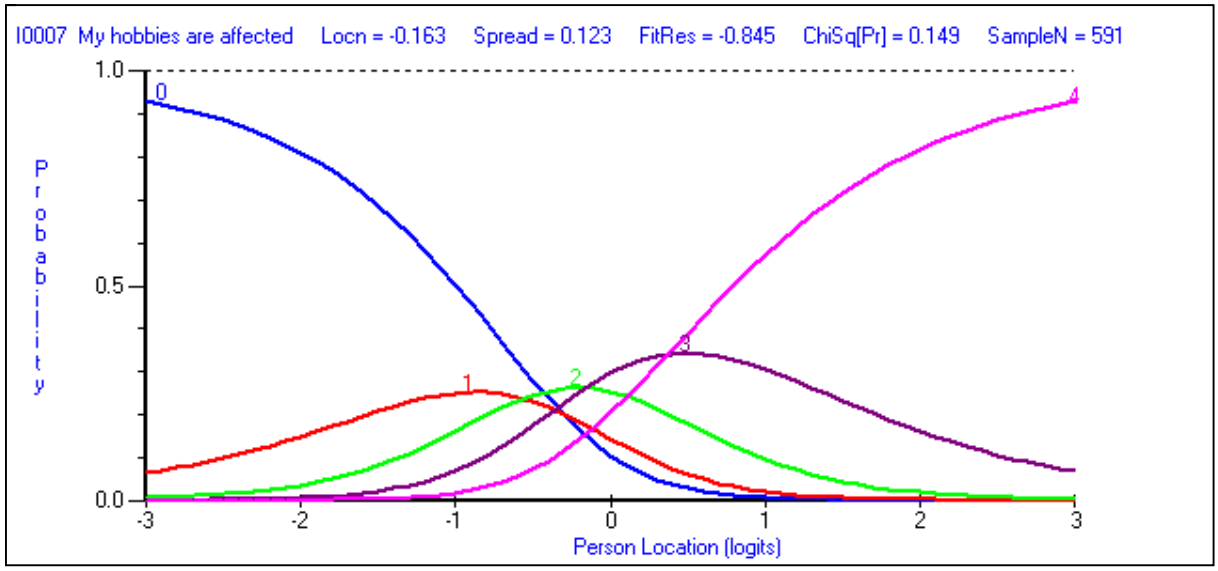
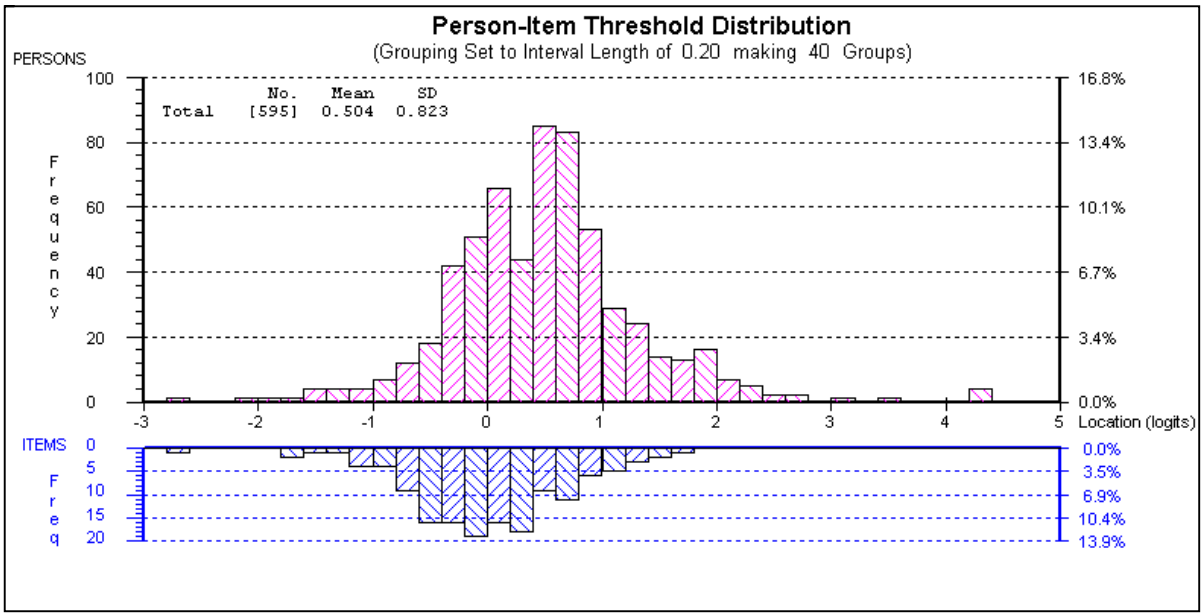


Figure 6.6:Person-item distribution map of the 36-items of the HidroQoL showing an even spread of the items across the latent variable.



The impact of removing items on the targeting of the HidroQoL against the sample was also assessed. The difference between the person mean location and item mean location had increased (Figure 5.9). From a person mean location of 0.87(±1.35) for the HidroQOL-36, person mean location had shifted to 1.25 (±1.6) following item reduction (estimates related to Analysis 12, Table 6.6, Figure 6.9). This is also reflected in number of person's with extreme scores which had increased to 19 persons from 4 prior to item reduction. This suggests a shift in the scale.

Testing the assumption of unidimensionality

A key underlying assumption of the Rasch model is that the latent variable is unidimensional. This was therefore formally tested on the HidroQoL-18. A principal component analysis (PCA) was carried out on the residuals after extraction of the Rasch component. The first PC had an eigenvalue of 2 and accounted for 11.1 % of variance in the residuals. Both demonstrate a clear single dimension of the data.

Further, a more stringent test of unidimensionality (Smith 2002) was used. Person estimates generated from two pairs of item subsets, with the highest positive and negative loading above |±0.3| on the first principal component, respectively, were compared using a t-test. Items with high

positive loadings included Q8, Q7, Q14, Q3, and Q15, while those with the highest negative loadings included Q29, Q22, Q37, Q47, Q35 and Q26. The proportion of significant t-tests was 5.3% [4.8%, 5.7%] confirming unidimensionality of the instrument.

Figure 6.7: Item threshold map of the HidroQoL-36 after rescoring to 3 point response categories

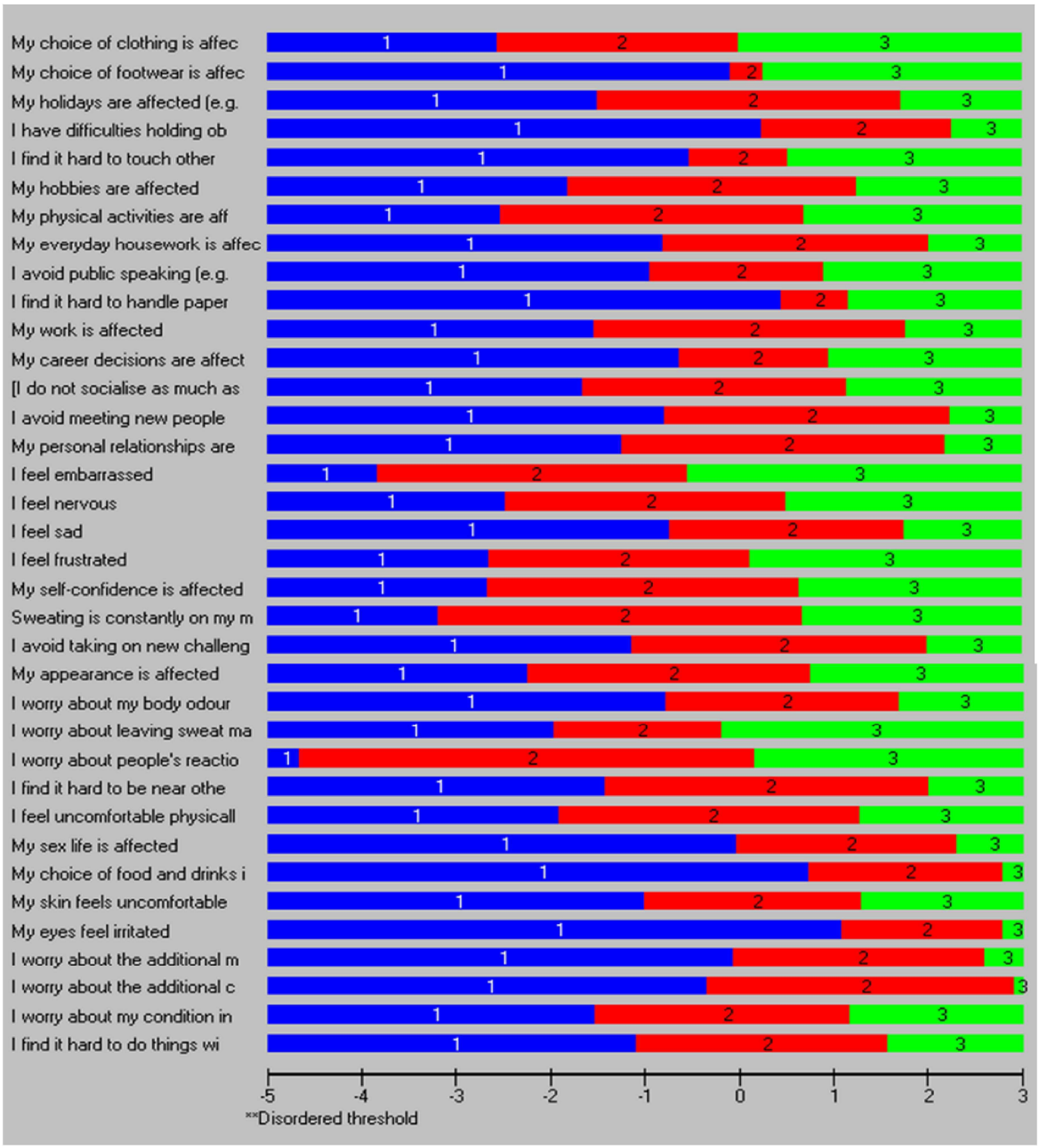
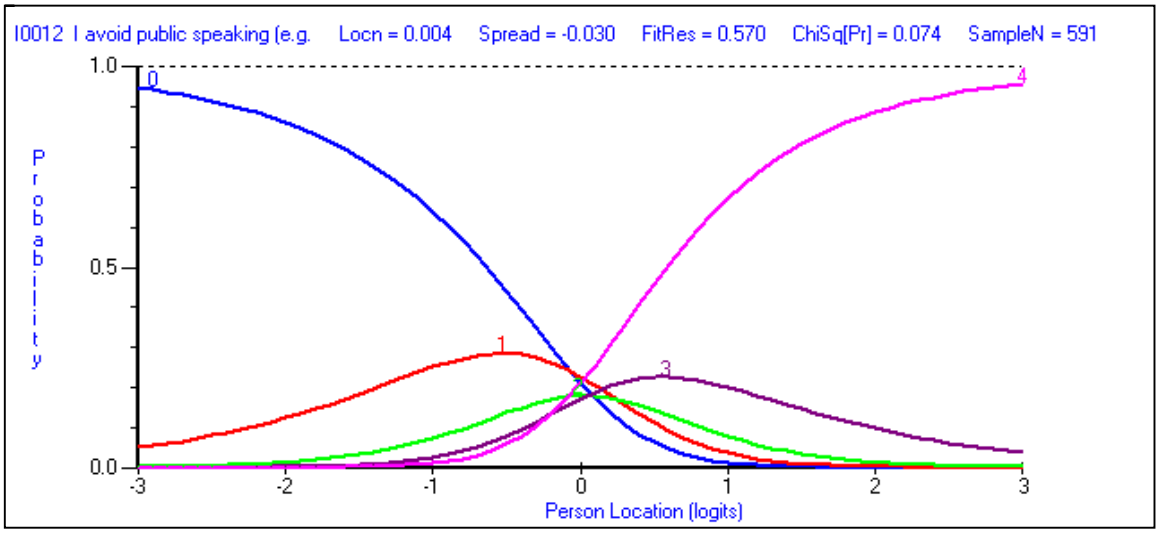
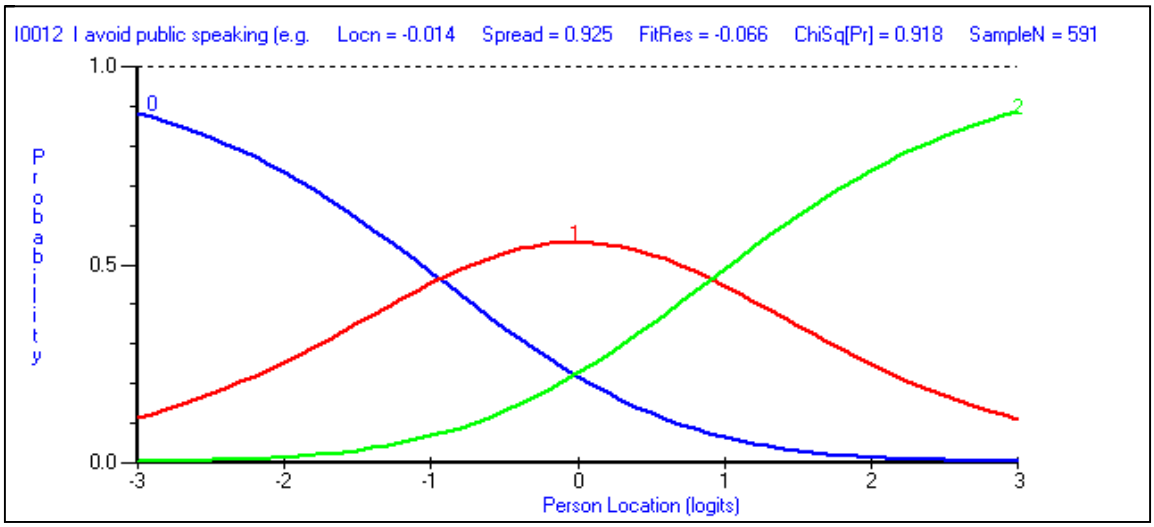


Figure 6.8: An illustration of the impact of rescoring on category probability curves

a. *Item 12 (I avoid public speaking) showing disordered category probability curves, prior to rescoring*



b. *Item 12 (I avoid public speaking) showing appropriately ordered category probability curves after rescoring from 5 to 3 categories.*



Misfitting persons

Considering that the RM treats persons in a similar way to items, lack of overall fit may be a consequence of a few misfitting persons (Pallant and Tennant 2007). Given that all items in the HidroQOL-18 were fitting the model, misfitting persons were suspected to be behind the lack of overall model fit. Sequentially removing 6 respondents showing the largest fit residuals led to a

good overall model fit (ITICS = 178.49, df = 162, p = 0.18). Item and person fit residuals both fell to acceptable ranges, at -0.25 (1.14) and -0.31 (1.3), respectively (Table 6.6, Analysis 13).

Analysis of Differential Item Functioning (DIF)

Once an instrument demonstrates fit the RM item calibrations should remain invariant across populations and testing situations (Reeve and Mâsse 2004). This means that when people with different demographic characteristic, for example, females and males, complete an instrument scores obtained should not differ, holding underlying variable constant. This assumption was explored by testing for differential item functioning for country, gender, age, body are affected, disease severity and co-morbidity. Items Q3, Q7, Q8, Q15, Q32 and Q34 showed uniform DIF for body area affected. Item Q3 showed uniform DIF for disease severity; and item Q8 showed DIF for co-morbidity. In the second step, the purification phase, all items with DIF were removed, in order to remain with a set of 'pure' items showing no DIF, as means of addressing the issue of compensatory DIF (Teresi 2006). This proceeded sequentially by removing one item at a time and iteratively assessing for DIF. Two items (Q26 and Q35) showing no DIF during the initial assessment had DIF for body area affected. In contrast, two items (Q47 and Q7) previously showing DIF no longer showed DIF. This on one hand, implies that, not all DIF observed is real, and on the other, that DIF could also be induced in items due to its presence in others (Tennant and Pallant 2007).

In the end a set of 8 pure items without DIF for any of the patient characteristics considered were realised. In the final step, DIF was reassessed while anchoring the scale on the 'pure' set of items (Table 6.8). There were no differences in the resultant items identified as showing DIF following the -purification process.

Table 6.6: Impact of item reduction steps on overall fit of the RM and individual items.

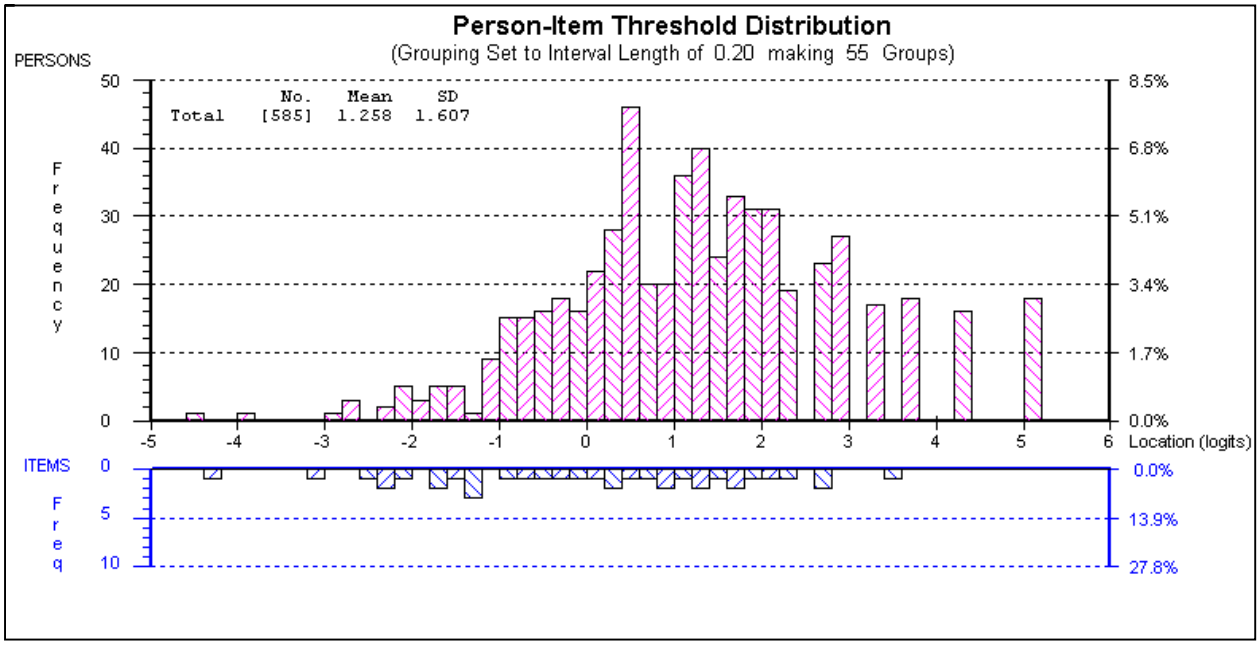
Action	Overall Model Fit			Item Fit Residuals		Person Fit Residuals		PSI	Fit Resid > 2.5	Fit Resid < -2.5
	Chi	*df*	*p*	*Mean*	*SD*	*Mean*	*SD*			
3. Revise scoring to 01113	**1087.40**	324.00	0.001	-0.05	3.48	-0.23	1.64	0.94	39, 2, 13, 43, 33, 11, 4	26, 20, 22, 17, 30, 24, 27, 36
4. Q2, Q4, Q13 and Q43 removed	**768.26**	288.00	0.001	-0.12	3.35	-0.31	1.59	0.94	39, 6, 11, 33, 42, 1, 44	22, 20, 24, 17, 30, 27, 36
5. Q39, Q6 and Q11 removed	**612.89**	261.00	0.001	-0.29	2.69	-0.35	1.55	0.93	33, 42, 1, 44	20, 24, 17, 30, 27, 36
6. Q36, Q27 and Q30	**453.34**	234.00	0.001	-0.22	2.26	-0.34	1.48	0.92	33, 42, 1, 44	22, 20, 24, 17
7. Q33 and Q42 removed	**346.51**	216.00	0.001	-0.23	1.96	-0.34	1.43	0.92	44, 1	24, 20, 17
8.Q17 and Q20 removed	**290.29**	198.00	0.001	-0.21	1.78	-0.33	1.37	0.91	44, 1	21, 24
9. Item 1 deleted	**277.60**	189.00	0.001	-0.23	1.68	-0.34	1.37	0.91	44	21, 24
10. Q24 deleted	**233.68**	180.00	0.001	-0.19	1.54	-0.34	1.35	0.90	44	-
11. Q44 deleted	**239.49**	171.00	0.001	-0.27	1.39	-0.34	1.33	0.89	-	21
12. Q21 deleted	**194.00**	162.00	0.04	-0.25	1.22	-0.33	1.37	0.89		
13. 6 respondents removed	**178.49**	162.00	0.18	-0.25	1.16					
14. 6 respondents removed	**169.57**	162.00	0.33	-0.25	1.14	-0.31	1.30	0.89		
15. 2 respondents removed	**171.47**	162.00	-0.25	-0.25	1.13	-0.31	1.29	0.89		

Table 6.7: Parameter estimates for the final 18 items fitting the Rasch model following item reduction

Fitting Item		Location	SE	FitRes.	Chi-Sq	p-value	THreshold 1	THreshold 2
Q3	My holidays are affected	0.47	0.09	-1.10	7.64	0.57	-1.70	1.70
Q7	My hobbies are affected	-0.01	0.09	0.83	5.31	0.81	-1.56	1.56
Q8	My physical activities are affected	-0.67	0.09	-1.30	4.41	0.88	-1.64	1.64
Q12	I avoid public speaking (e.g. presentations)	0.33	0.08	1.28	14.09	0.12	-0.97	0.97
Q14	My work is affected	0.51	0.09	-0.20	16.18	0.06	-1.74	1.74
Q15	My career decisions are affected (e.g. career choice)	0.49	0.08	-0.26	3.23	0.95	-0.81	0.81
Q18	I avoid meeting new people	1.16	0.08	0.07	3.03	0.96	-1.59	1.59
Q22	I feel nervous	-0.76	0.09	-1.80	16.52	0.06	-1.53	1.53
Q26	I feel frustrated	-1.08	0.09	-1.93	10.16	0.34	-1.42	1.42
Q29	Sweating is constantly on my mind	-1.04	0.09	-0.92	5.89	0.75	-2.00	2.00
Q32	My appearance is affected	-0.45	0.09	0.91	4.90	0.84	-1.55	1.55
Q34	I worry about leaving sweat marks on things	-1.07	0.09	1.40	9.02	0.44	-0.87	0.87
Q35	I worry about people's reactions	-2.05	0.10	-0.29	16.66	0.05	-2.40	2.40
Q37	I feel uncomfortable physically expressing affection (e.g. hugging others)	-0.03	0.09	-0.50	13.66	0.13	-1.65	1.65
Q38	My sex life is affected	1.60	0.08	1.19	14.60	0.10	-1.22	1.22
Q45	I worry about the additional chores in dealing with my condition	1.80	0.09	0.08	5.10	0.83	-1.74	1.74
Q47	I worry about my condition in future	0.17	0.08	0.38	8.12	0.52	-1.42	1.42
Q48	I find it hard to do things without planning in advance	0.63	0.08	-2.24	12.96	0.16	-1.43	1.43

DIF was still seen in all items which previously showed problems (Table 6.9, Figure 6.10). Nonetheless, the purification process gave useful hints on items whose DIF was likely to be compensatory rather than real.

Figure 6.9: Person-item distribution showing targeting of the HidroQoL following item reduction



Magnitude of DIF

Showing statistically significant DIF, only rules out the possibility that such observation is a result of chance (Linacre 2009). Of practical relevance for measurement is the actual size of the observed DIF. This was, thus, evaluated by measuring the differences between the general and group specific-item difficulty estimates across each patient characteristic. The magnitude of differences across all DIF items exceeded 0.5 logits indicating non-trivial DIF. For example, item Q32 (my appearance is affected) was 3.3. logits easier for patients with generalised hyperhidrosis in comparison with the estimate based on all patients. This item was 2.06 logits more difficult for patients with palmar and plantar hyperhidrosis (Table 6.10). DIF by age was most severe in item Q34: respondents the age group 40 to 49 found this item 2.9 logits easier than the average patient.

Impact of DIF

Having established that the DIF was not due to chance and was not trivial, the ultimate question was whether it indeed mattered, in the actual evaluation of persons (Tennant and Pallant 2007). This was assessed by looking at the impact of the item level DIF on the final scale scores. First, the direction of the DIF for the groups was assessed to see whether the DIF cancelled out, where some items favoured one group and the other items favoured the other groups.

Table 6.8: Pure set of items showing no DIF following the purification

	Location	Threshold_1	Threshold_2
Q7	-0.498	-1.554	1.554
Q14	-0.010	-1.700	1.700
Q18	0.607	-1.536	1.536
Q29	-1.513	-1.983	1.983
Q37	-0.511	-1.634	1.634
Q38	1.061	-1.226	1.226
Q45	1.222	-1.750	1.750
Q47	-0.359	-1.417	1.417

Table 6.9: DIF in items according to patient characteristics

Item	Age		Body area affected		Disease severity		Co-morbidity	
	F-Stat.	*p*	*F-Stat*	*p*	*F-Stat.*	*p*	*F-tat.*	*p*
Q3	16.16	0.00002	12.7	0	7.48	0.000629		ns
Q7	1.55	ns	5.9	0.000571		ns		ns
Q8	14.5	0.00005	12.7	0.000001		ns	4.8	0.000265
Q12	1.67	ns		ns		ns		ns
Q14	0.41	ns		ns		ns		ns
Q15	1.15	ns	5.9	0.000584		ns		ns
Q18		ns		ns		ns		ns
Q22		ns		ns		ns		ns
Q26		ns		ns		ns		ns
Q29		ns		ns		ns		ns
Q32	10.29	0.000001	28.21	0		ns		ns
Q34	11.74	0.00001	14.79	0.00001		ns		ns
Q35		ns		ns		ns		ns
Q37		ns		ns		ns		ns
Q38		ns		ns		ns		ns
Q45		ns		ns		ns		ns
Q47	6.25	0.000365		ns		ns		ns
Q48		ns		ns		ns		ns

Notes: F, F-statistic; P, p-value; ns, not significant.

For DIF according to age, two sets of items seemed to cancel out: Q47 against Q3 and; Q34 against Q32. However, DIF according to body area does not show clear balancing out. Person estimates obtained from the eight 'pure' items during the purification stage were compared with person

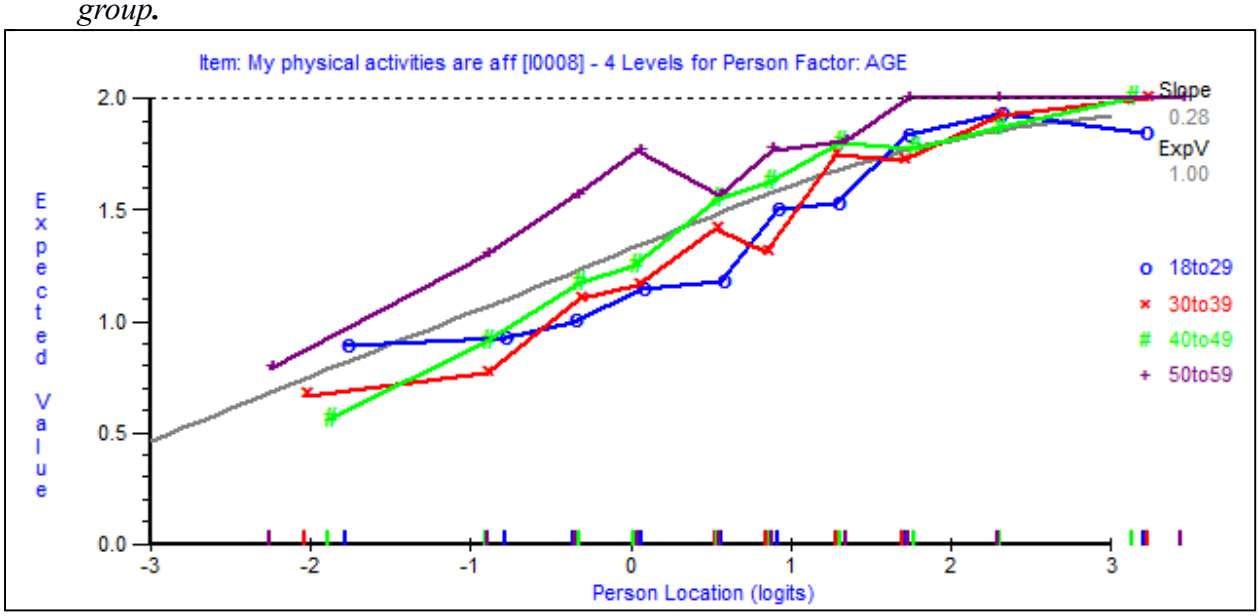
estimates from the full item set with the DIF-free items anchored at their values obtained during the purification phase. The two sets of person estimates correlated highly (Pearson correlation = 0.94), 83.2% of sets of person estimates showed a difference of below 0.5 logits [comparison with standard deviation]; 97.3% of person estimate sets had a difference below 1 logit. Based on a paired t-test the person estimates from the two measures were not statistically different (p=0.29). This indicates that the DIF in the different items simultaneously had no impact on the scale scores. Further, the DIF impact at the scale level was also assessed graphically using Test Characteristic Curves. Group specific test characteristic curves (TCC) were estimated for each patient characteristic to assess whether the relationship between raw score and latent variable remained invariant across groups. TCCs were estimated for groups across age, body area affected and disease severity. The results showed largely invariant TCCs across all patient characteristics (age, severity of disease, co-morbidity), nonetheless, the TCC for site of hyperhidrosis showed marginal variance (≤ 0.5 logits) (Figure 6.11). The findings indicate that the DIF observed at the item level did not affect the optimal functioning of the overall scale.

Local Independence

Rasch model's assumption of local independence requires that any set of items should not share any meaningful correlation, once the Rasch component is accounted for (Baghaei 2008a). This assumption was evaluated for the HidroQOL-18 by examining the residual correlations between items. Four item pairs showed a correlation greater than 0.2: Q15 and Q32; Q3 and Q34; Q14 and Q15; and Q7 and Q8. Similar to DIF analysis the size and practical implications of local dependence are not clear based on correlations alone. Thus the size of the local dependence was measured. The locally dependent items (Q15, Q8, and Q3) were split according to responses on the corresponding independent items (Q32, Q7 and Q34). Overall fit to the RM was poor following the split, the ITICS statistic was 267.93 and was significant (Table 6.11, Analysis 1). Two items showed poor fit (Q48; and Q7- 0, for those who had a score of zero on item 8). The PSI was not affected by splitting. The magnitude of response dependence was calculated using the category thresholds of the new set of split items.

Figure 6.10: An illustration of DIF as reflected in empirical group-specific Item characteristics curves

a. Item 8, ICC showing expected score for each level of the latent variable traced for each age group.



b. Item 34, ICC showing expected score for each level of the latent variable traced for the body area involved in the hyperhidrosis

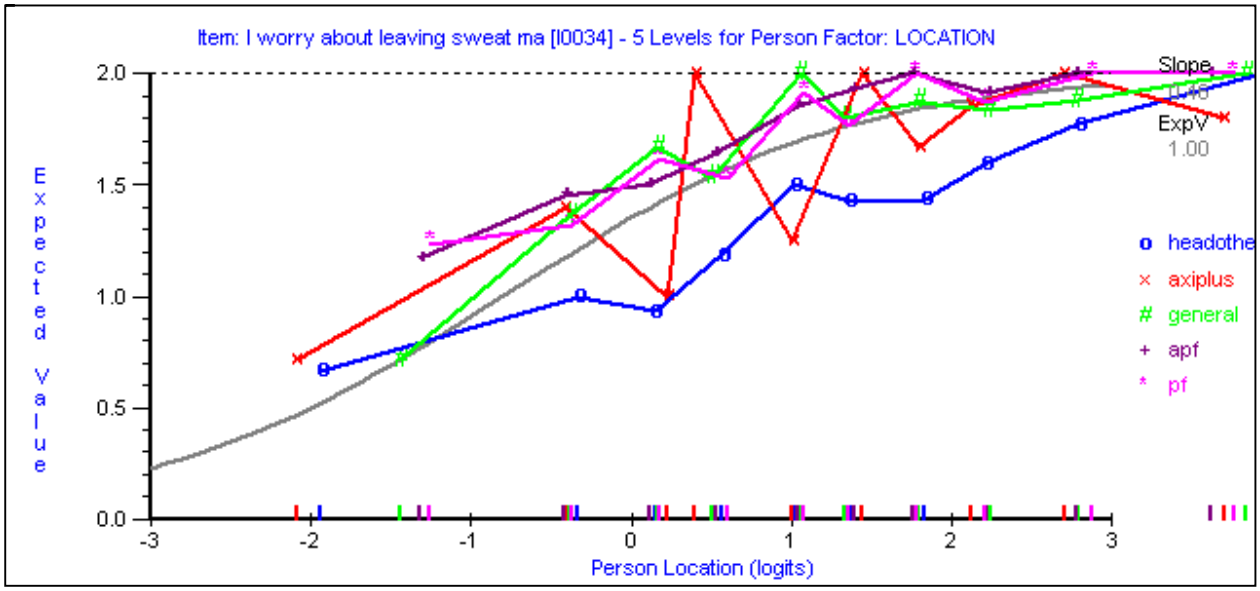


Table 6.10: Magnitude of DIF

Items split according to body area affected			Items split according to age-group			Items split according to comorbidity		
Item	*Difficulty Estimate*	*DIF Size*	*Item*	*Difficulty Estimate*	*DIF Size*	*Item*	*Difficulty Estimate*	*DIF Size*
Q3_original	-0.06							
Q3_head	0.00	-0.06	Q3_18to29	1.60	-1.66			
Q3_axilar	1.31	-1.36	Q3_30to39	0.95	-1.01			
Q3_generic	0.22	-0.28	Q3_40to49	0.73	-0.78			
Q3_p&f	1.23	-1.29	Q3_50+	-0.06	0.01			
Q8_Original	-1.13							
Q8_head	-0.79	-0.34	Q8_18to29	-0.13	-1.00	Q8_none	0.07	-1.20
Q8_axilar	0.22	-1.35	Q8_30to39	0.26	-1.39	Q8_men	-4.51	3.38
Q8_generic	-0.70	-0.43	Q8_40to49	-0.24	-0.89	Q8_diab	-1.21	0.08
Q8_p&f	-0.38	-0.75	Q8_50to59	-1.36	0.23	Q8_hyper	-1.28	0.15
Q12_original	-0.21							
Q12_head	0.64	-0.84						
Q12_axilar	0.24	-0.44						
Q12_generic	0.36	-0.57						
Q12_p&f	1.21	-1.42						
Q32 Original	*-0.97*							
Q32_head	-1.76	0.79	Q32_18to29	0.60	-1.57			
Q32_axilar	-0.28	-0.69	Q32_30to39	-0.50	-0.47			
Q32_generic	-4.25	3.28	Q32_40to49	-0.11	-0.86			
Q32_p&f	1.09	-2.06	Q32_50+	-0.67	-0.30			
Q34_Original	*-1.53*							
Q34_head	0.14	-1.67	Q34_18to29	-1.86	0.33			
Q34_axilar	-1.14	-0.38	Q34_18to29	-0.79	-0.74			
Q34_generic	-0.88	-0.65	Q34_40to49	-4.42	2.89			
Q34_p&f	-4.49	2.96	Q34_50to59	-0.12	-1.40			
Q48_original	*0.09*							
Q48_head	0.68	-0.59						
Q48_axilar	1.20	-1.10						
Q48_generic	0.31	-0.22						
Q48_p&f	1.42	-1.32						

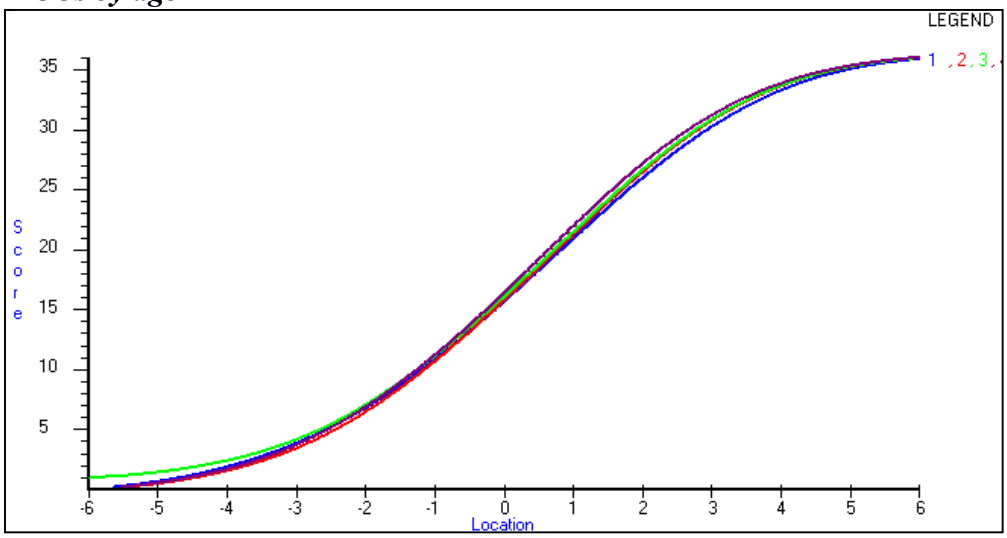
Note: Size of DIF for each item is [item difficulty estimate, whole sample (*Qxx_original*) – groupspecific item estimate].

The largest dependence was seen between item Q7 and Q8 at 1.18 logits, while that between Q14 and Q15 was 0.826. On the other hand the dependence between Q4 and Q34 was trivial, at 0.182. Using the magnitude of response dependence as a basis for decision-making regarding how to address the dependence, items Q15, Q8, and Q48 were sequentially removed (Table 6.11, Analysis

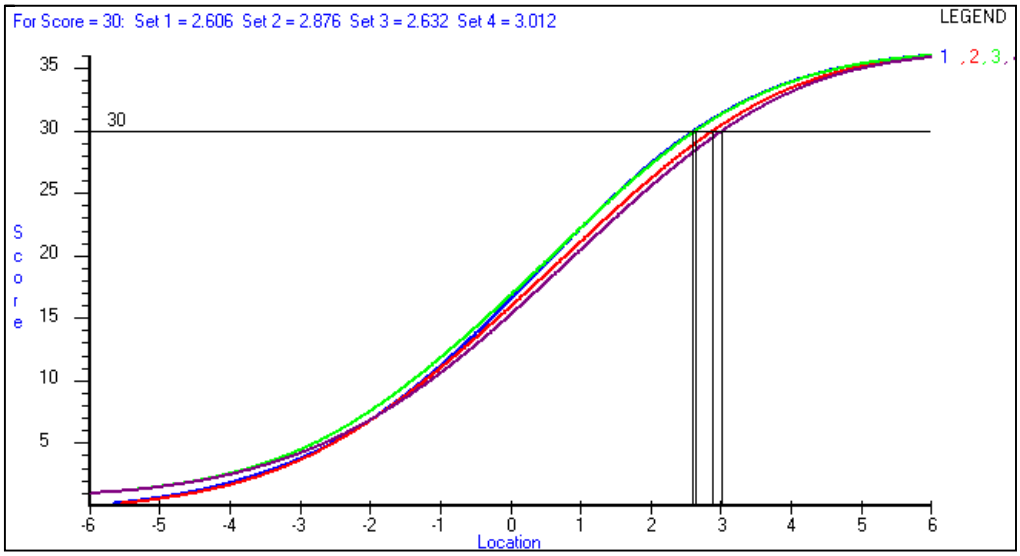
2- 4). The HidroQoL showed lack of overall fit to the RM after the removal of the three items, although the remaining items fitted the RM. The HidroQoL also failed the formal test of unidimensionality, the proportion of significant t-tests of person estimates for subtests of the instrument, exceeded 5%. Following the removal of misfitting respondents (n = 22), fit of the HidroQoL to the RM was improved. Unidimensionality was also achieved (Analysis 5).

Figure 6.11: Test Characteristic Curves for the HidroQoL-18

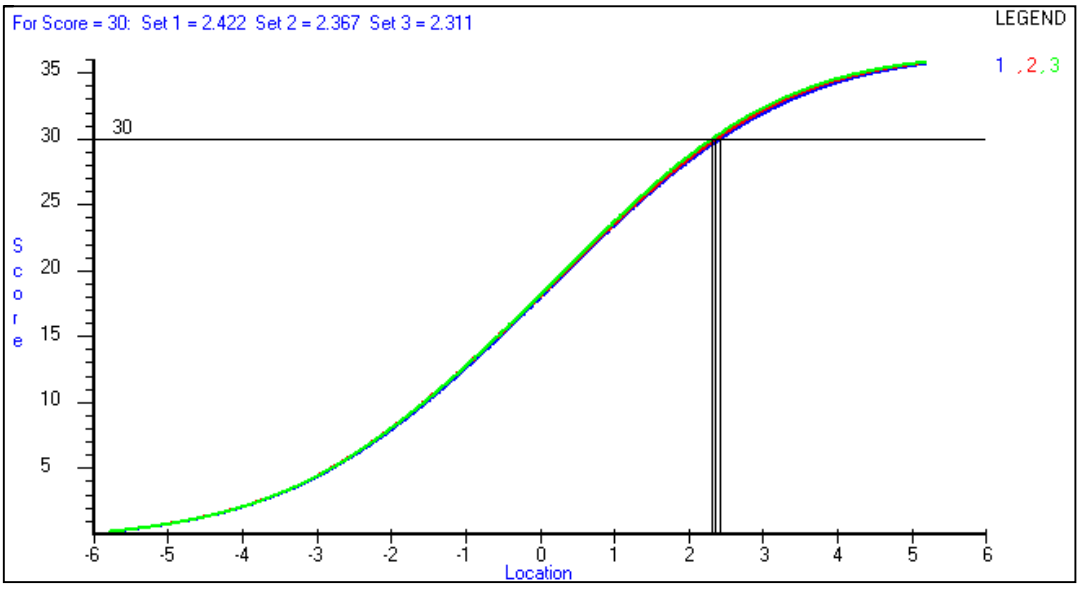
TCCs by age



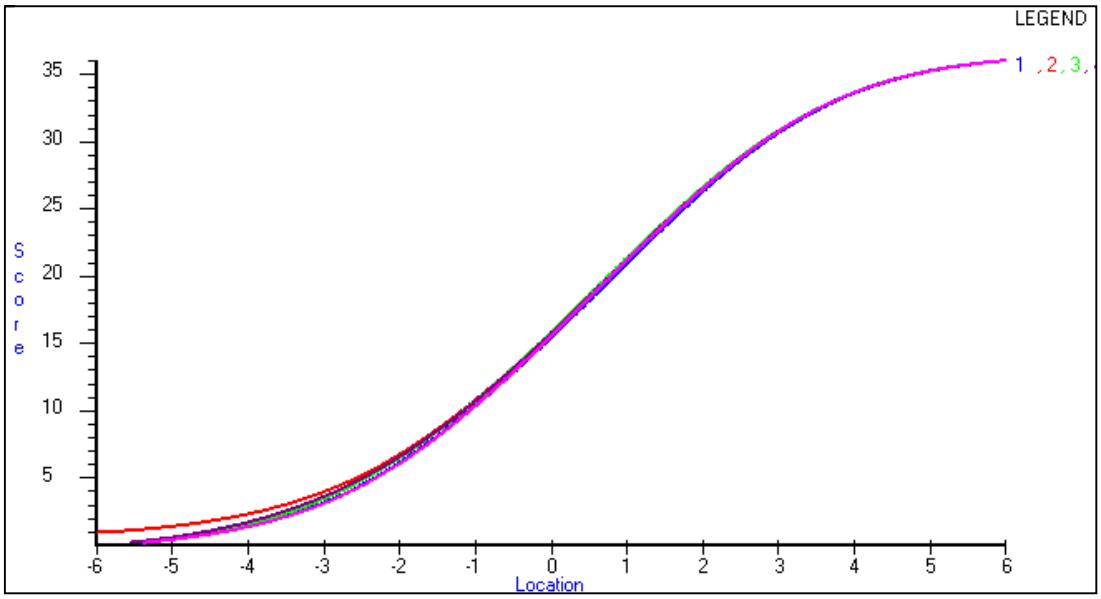
TCCs by body area affected



TCCs by severity of disease



TCCs by co-morbidity



Applying classical exploratory factor analysis to the Rasch calibrated HidroQoL

Following the calibration of the HidroQoL as a unidimensional measure using Rasch analysis, classical exploratory factor analysis (EFA) was employed to cross-validate the calibration; and to further confirm unidimensionality. The number of factors to be extracted was determined based on multiple criteria. According to the Kaiser's rule, a single factor achieved an eigenvalue greater

than 1, for both the 15 and 18 item HidroQoL versions (Table 6.12). This was supported by scree plots: only a single factor lay to the left of the elbow in the plots for both the 15 and 18 item versions (Figure 6.12).

This means that variance in the HidroQoL scores can be effectively captured using a single latent variable. All items of the HidroQoL also showed strong loading onto the single factor (range 0.592 to 0.782) (Figure 6.13). The HidroQoL's items also showed a high level of shared variance, the highest item residual was 0.65 and 0.62, for the 18-item and 15-item versions. These results strongly suggest that the HidroQoL (both 18-item and 15-item versions) is unidimensional, supporting the results of the Rasch analysis.

Part III: Testing the Invariance Of The Hidroqol

Rasch calibrated measures are expected to show invariance in item parameters across different sub-populations (Bond and Fox 2007). Thus, the HidroQoL was recalibrated on a new sample to determine if it would replicate original calibration.

Table 6.11: Impact of adjusting for response dependence on overall model fit.

Action	Overall Model Fit			Item Fit Residuals		Person Fit Residuals		Person Location		Unid	PSI
	Chi	*df*	*p*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*sig*	
1. Splitting Q7, Q15 and Q3	267.93	188	<.001	-0.29	1.29	-0.3	1.18	1.1	1.66		0.872
2 Removing item 15, from HidroQoL-18	168.94	153	0.179	-0.27		-0.3	1.24	1.3	1.67		0.882
3. Removing Q8 from HidroQoL	187.96	144	< 0.001	-.23	1.33	-.3	1.2	1.31	1.68		0.88
4. Removing Q48	165.72	135	0.039	-0.2	1.21	-0.3	1.18	1.34	1.66	7.09	
5. Removing 22 persons	159.64	135	0.07	-0.09	1.19	-0.24	1.05	1.365	1.65	3.42	0.869

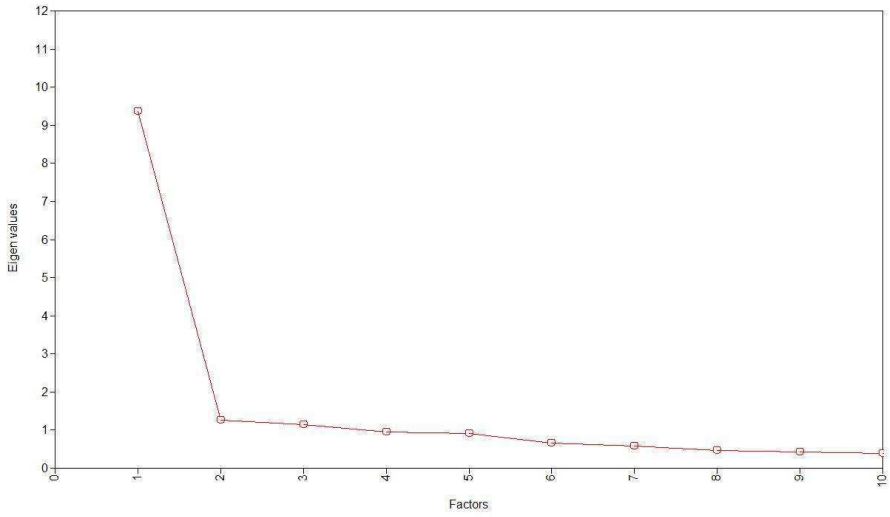
Note: Chi: Chi-squared; df: degrees of freedom;

Table 6.12: Eigen values of the HidroQoL

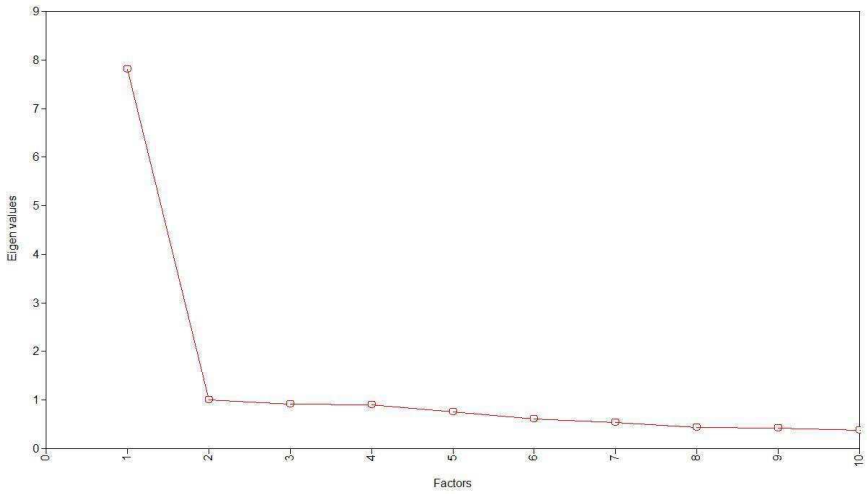
Instrument	Factors	Adjusted Eigen value	observed Eigen value	Estimated Bias
18-items	1	6.536	6.834	0.298
15-items	1	5.244	5.544	0.300

Figure 6.12: Scree plots of the HidroQoL

a. scree plot of 18 item version of the HidroQoL



b. scree plot of the 15-item version of the HidroQoL



Specifically the following hypothesis were tested:

- the HidroQoL (18-item and 15-item versions) fits the RM;
- the hierarchical ordering of the items would not change; and
- the item parameters remained invariant within linear transformation across the two populations (US/Canada versus UK).

Moreover, these aspects entail assessing the cross cultural validity of the instrument, an increasingly demanded property especially in instruments applied in international clinical trials (Nijsten et al. 2007).

Table 6.13: Factor loadings and residual variances of the HidroQoL

	18- item version			15-item version		
Item	*Loading*	*S.E.*	*Res. Var*	*Loading*	*S.E*	*Res. Var*
Q3	0.724	0.023	0.476	0.694	0.026	0.518
Q7	0.696	0.027	0.516	0.619	0.033	0.617
Q8	0.77	0.022	0.407			
Q12	0.658	0.027	0.568	0.663	0.028	0.561
Q14	0.684	0.026	0.532	0.612	0.033	0.626
Q15	0.736	0.024	0.459			
Q18	0.715	0.025	0.488	0.724	0.025	0.475
Q22	0.793	0.023	0.372	0.808	0.023	0.347
Q26	0.768	0.025	0.41	0.799	0.024	0.361
Q29	0.739	0.025	0.454	0.758	0.025	0.426
Q32	0.682	0.028	0.535	0.682	0.029	0.535
Q34	0.592	0.036	0.649	0.615	0.035	0.622
Q35	0.759	0.027	0.424	0.79	0.026	0.376
Q37	0.729	0.024	0.469	0.756	0.024	0.428
Q38	0.636	0.03	0.596	0.648	0.03	0.579
Q45	0.699	0.027	0.511	0.693	0.029	0.52
Q47	0.72	0.025	0.481	0.691	0.028	0.522
Q48	0.782	0.021	0.389			

Sociodemographic characteristics of the study participants (sample 2)

A total of 115 patients with hyperhidrosis were included for analysis of part III of the results. Thirty-seven participants (32%) were male and 78 (68%) were male (Table 6.14) The mean age was 40.2 years, with 36 patients being aged between 30 and 39, making up the largest age group. Patients reporting generalised hyperhidrosis made up the largest group, with 36 participants. Forty-five percent of the patients reported no co-morbidity (n = 54).

HIdroQoL-18

All eighteen items showed good fit to the RM with fit residuals lying between -2.5 and 2.5. The average mean person location estimate of 1.26 (1.57) indicated that the HidroQoL was at a lower QoL impact, relative to this patient population. This is similar to estimates obtained during initial calibration. The PSI (0.889) replicated the strong reliability obtained earlier (Table 6.15).

Table 6.14: Sociodemographic characteristics of study participants

Gender , N (%)	
Male	37(33%)
Female	78 (67%)
Age (years)	
Mean (SD)	40.2 (13.3)
Median	39
Range	18-74
Duration of condition (years)	
Mean (SD)	24.1
Median	21
Range	2 to 60
Body area affected, n	
Head*	28
Axilla*	14
General	36
Axilla, Palms, Feet	15
Palms and Feet	19
Co-morbidity	
None	52 (45%)
Menopausal complaints	8 (7%)
Diabetes	4 (3%)
Hypertension	10 (9%)
Neurological disorders	9 (8%)
Thyroid disorders	23 (20%)
Employment status	
Employed	74 (64%)
Unemployed	24 (21%)
Retired	11 (10%)
Full-time student	6 (5%)

***This only indicates this site as a predominant site affected.

Table 6.15: Overall model fit statistics for the HidroQoL on participants from UK

Action	Overall Model Fit			Item Fit Residuals		Person Fit Residuals		Unid. test	PSI
	Chi	*df*	*p*	*Mean*	*SD*	*Mean*	*SD*	*Sign. t-test (%)*	
1. 18 items*	49.47	36	0.07	-0.02	1.1	-0.2	1.18	7.14%	0.889
2. Q8 and Q15 removed	40.32	32	0.15	-0.014	1.17	-0.22	1.12	5.36%	0.876
3. Q48 removed	45.05	30	0.04	0.03	1.12	-0.21	1.08	2.68%	0.864

The three level categorization seemed to function well, all items had appropriately ordered thresholds, increasing monotonically along the latent variable. Unidimensionality was tested using multiple approaches, based on the residuals. A principal component analysis (PCA) of the residuals showed that the first component explained 12% of variation in the residuals and had an eigen-value of 2, supporting unidimensionality. Strict unidimensionality was assessed by comparing person estimates from a subset of items with positive loading to the first residual PC (Q48, Q8, Q26, Q47) against a subset of items with negative loading (Q38, Q37, Q15, Q14). 7.14% (CI: 2.9%, 12.7%) of person estimates showed statistically significant differences, indicating unidimensionality. To test the assumption of local independence a correlation analysis of the residuals was carried out. This showed a number of item pairs exceeding the average: Q8 and Q7; Q48 against Q14; Q48 and Q38. The next step was to examine the magnitude of the observed response dependence. The response dependence between Q8 against Q7 was of magnitude of 2.23 logits; that between Q48 and Q14 was -.33 logits while that between Q48 and Q38 was -0.69. The response dependence between Q8 and Q7 was similarly observed during original calibration; moreover item Q48 also showed worse fit following the resolution of the response dependence during the original calibration. This suggests that the composition of the HidroQoL as a unidimensional metric for measuring hyperhidrosis-specific QoL is invariant.

Comparison of model fit: HidroQOL-15

During the initial/original calibration of the HidroQoL three items (Q8, Q15 and Q48) were removed from the HidroQOL-18 (18 item set of the instrument) to resolve local dependence problems, resulting into HidroQOL-15. The 15-item HidroQOL was also assessed for its fit to the RM based on the new patient population (UK, N = 115). The ITICS statistic decreased to 45.05

(df=32) and indicated poor fit to the RM (p=0.04). The mean fit residuals for persons and items showed no change from the HidroQoL-18 (persons, -0.21±1.08; items, 0.03±1.12). Reliability had declined (PSI = 0.86). The targeting of the instrument also showed no change from the HidroQoL-18. According to the Smith's t-test the HidroQoL's 15-item version was also unidimensional: 2.68% (CI: 1.2%, 9.3%) of person estimates derived from two subsets of the HidroQoL showed significant t-tests. This suggests that although the individual items remain optimal and comprise a unidimensional scale, its functioning across (the ordering of items) across levels of impairment is not invariant.

Testing for invariance

Further, the hierarchical ordering of the items of the HidroQoL was compared, between the patient populations. The item maps for the 18-item version of the HidroQoL was generated and compared with the one generated from the US and Canadian sample (Figure 6.13). The hierarchy of the items had changed on account of items Q37, Q35, Q34, Q29 and Q26 which had shifted their order. Items Q37, Q35, Q34, Q29 seemed more difficult for the UK sample, while Q26 was regarded easier than as originally calibrated.

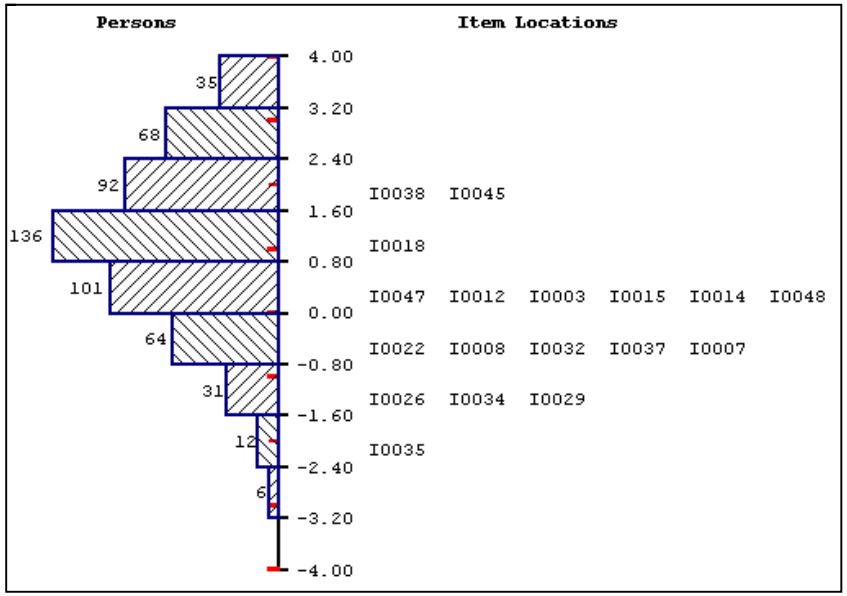
Next, the invariance of the actual item difficulty estimates of the HidroQoL was assessed using a scatter plot of item estimates from the original calibration against those from the U.K sample (Figure 6.14). The imprecision of each estimate was taken account of by overlaying 95% control lines based on standard errors for each set of item-difficulty as suggested by (Bond and Fox 2007). The graph shows that one out of the eighteen items fell outside the control lines. With 94.5% of the estimates falling within the control lines, the argument for invariance is, therefore, supported.

Revision of the HidroQoL

Following the item reduction process reported in this and the previous chapter the developmental HidroQoL was revised to 18 items, loosing 31 items (Figure 6.15). The response scoring was revised to a 3 point scale. Fifteen of the 18 items are based on the results of the Rasch item reduction process (in this chapter). The other three are retained for their importance to patients (according to results of the qualitative study in chapter 3). This was the final version taken for reliability, validity and responsiveness assessment, reported in the subsequent three chapters.

Figure 6.13: Comparison of item hierarchical order between original calibration of the HidroQoL and from the UK sample

a) Original item hierarchical order based on US and Canada participants



b) Item hierarchical order based on calibration on participants from the UK.

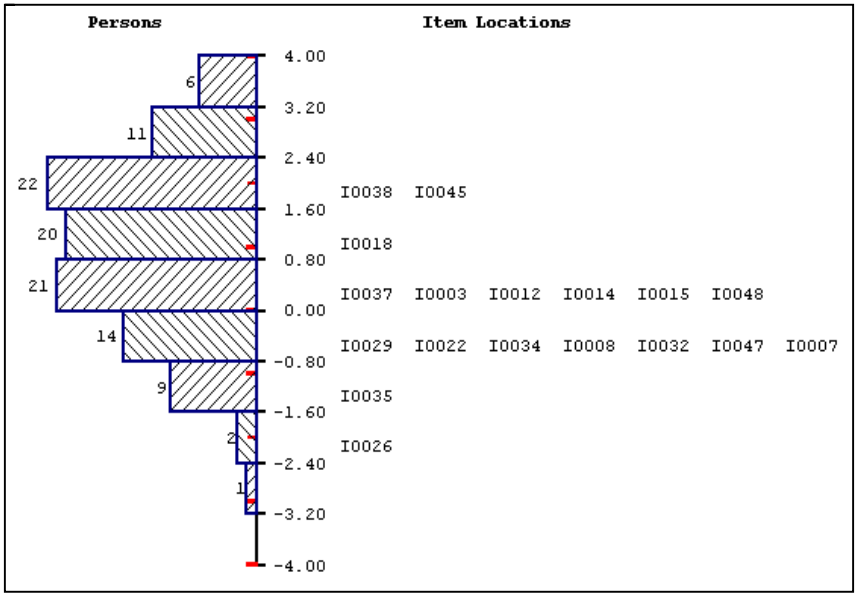
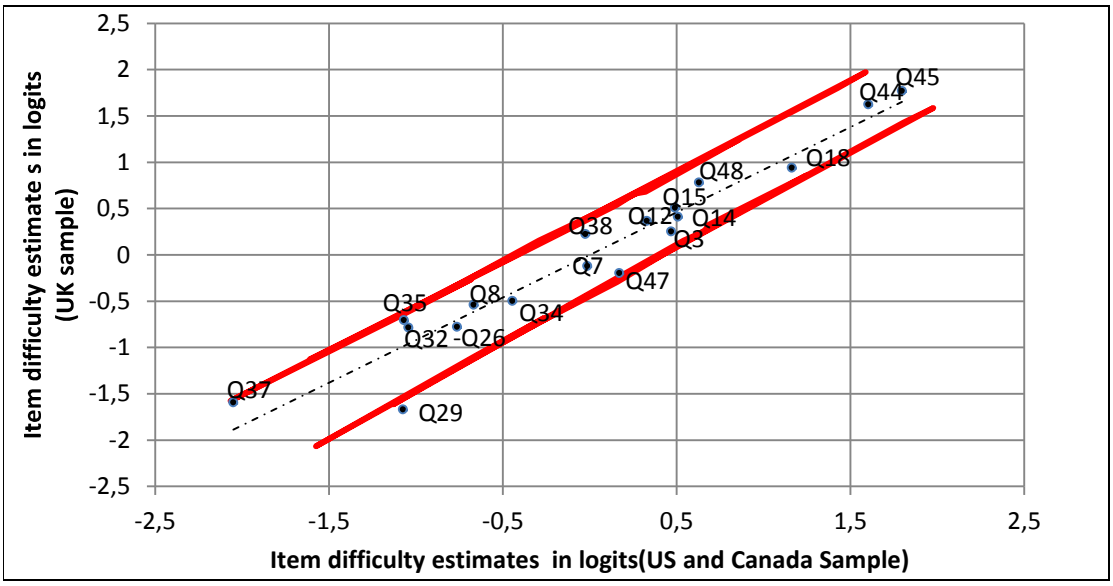


Figure 6.14: Scatter plot of item difficulty estimates of the HidroQoL, plotting estimates from the original calibration against those from the UK samples

a) 18-item set



b) 15-item set

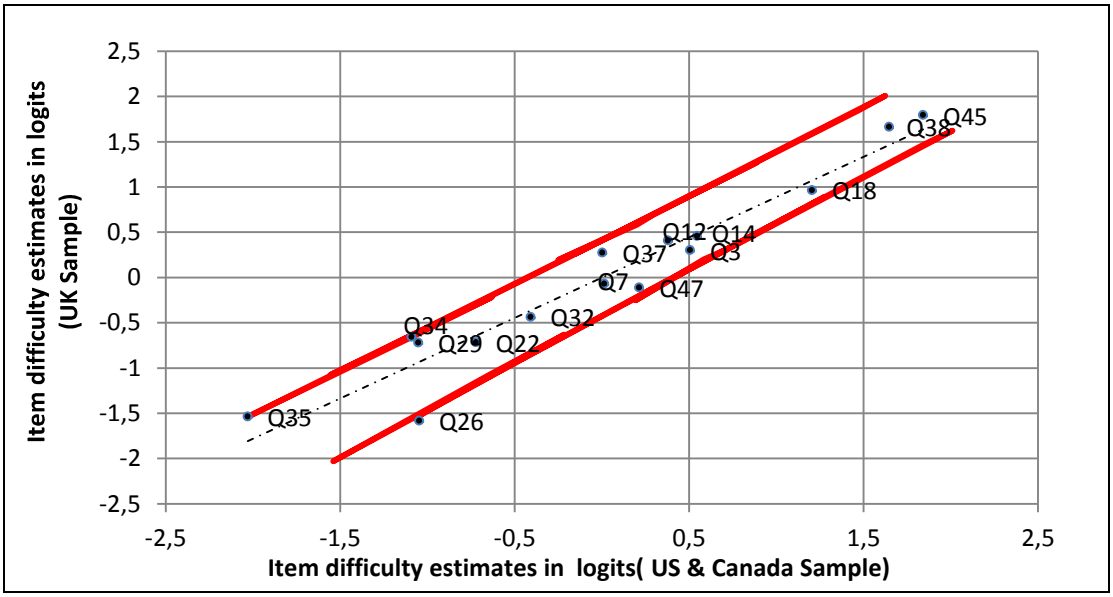


Figure 6.15: The final version of the HidroQoL with 18-items

Hyperhidrosis Quality of Life Index

The statements in this questionnaire relate to how your life has been affected **by your excessive sweating condition (hyperhidrosis) in the last seven days including today.**

Please choose one box for each statement.

If a statement does not apply to you please choose 'No, not at all'.

	Very much ▼	A little ▼	No, not at all ▼
1. My choice of clothing is affected	☐	☐	☐
2. My physical activities are affected	☐	☐	☐
3. My hobbies are affected	☐	☐	☐
4. My work is affected	☐	☐	☐
5. I worry about the additional activities in dealing with my condition	☐	☐	☐
6. My holidays are affected (e.g. planning, activities)	☐	☐	☐
7. I feel nervous	☐	☐	☐
8. I feel embarrassed	☐	☐	☐
9. I feel frustrated	☐	☐	☐
10. I feel uncomfortable physically expressing affection (e.g. hugging)	☐	☐	☐
11. I think about sweating	☐	☐	☐
12. I worry about my future health	☐	☐	☐
13. I worry about people's reactions	☐	☐	☐
14. I worry about leaving sweat marks on things	☐	☐	☐
15. I avoid meeting new people	☐	☐	☐
16. I avoid public speaking (e.g. presentations)	☐	☐	☐
17. My appearance is affected	☐	☐	☐
18. My sex life is affected	☐	☐	☐

Figure 6.15 (continued)

Hyperhidrosis Quality of Life Index

Please complete the following additional questions

Age: ☐ Years Gender: Male ☐ Female ☐

How long have you been experiencing excessive sweating: ☐ Years

Main body area(s) affected: Head/face ☐ Armpits ☐ Palms ☐ Feet ☐ Other ☐
Generalised ☐

If you selected more than one area, how important is each (in percentage) e.g. 70%, 30% ☐

Are you currently receiving treatment for the condition: Yes ☐ No ☐

If **Yes**, what treatment ☐

How much time per day do you spend dealing with your condition: (e.g. for personal hygiene, treatment, laundry) ☐ Minutes

How much extra money do you spend per month because of your condition ? (e.g. for personal hygiene, treatment, clothing) ☐ (£)

What is your current employment status: Employed ☐ Unemployed ☐ Retired ☐
Full time student ☐

If unemployed or retired, is this related to your condition: Yes ☐ No ☐

Please check that you have answered all questions

Thank you for your help!

DISCUSSION

Before a QOL instrument is considered ready for use, it is imperative to undertake field testing, similar to the large "phase III" trials in clinical trial language, in order to determine whether its measurement model is functioning as envisaged and if necessary to undertake any revisions. This constitutes an important aspect of Messick's substantive aspect of validity, *that theoretical rationales relating to both item content and processing models explain observed consistencies among items* (Wolfe and Smith 2007a). On the other hand this demonstrates the internal validity of the instrument, the organization of items into scales and subscales and how they relate to each other. Among other approaches for developing and testing the scaling of an instrument, the unidimensional Rasch model offers a number of advantages. It allows the hierarchical ordering of the items based on the level of the underlying construct (QOL impairment) they assess (Prieto et al. 2003). In addition, the model is consistent with requirements of conjoint measurement permitting the transformation of raw scores into interval scaled measures once fit to the model is confirmed (Bond 2004). Therefore, the aim of this study was to assess the extent to which the HidroQoL conforms to the RM. In addition, to use the RM and its properties as a basis for the item reduction process and construct validation of the HidroQoL. This is based on the consideration of the RM as a template for an instrument satisfying fundamental axioms of measurement.

The version of the HidroQoL used in this study, the HidroQoL-36, contained 36-items and was a result of earlier item reduction steps reported in the previous chapter. The HidroQOL-36 lacked overall fit to the RM, which means that the hierarchical ordering of the items was not invariant across different levels of QOL impairment along the latent variable continuum (Pallant and Tennant 2007). This may be a consequence of any number of problems including dysfunctional response categories, items or person whose response patterns are inconsistent with the RM's probabilistic structure, violations of model assumptions such as local independence, presence of additional dimensions or differential item functioning (Tesio 2003). Further analysis explored each of these elements.

An analysis of the individual items showed that 21 items had poor fit to the RM. Underfitting items (fit residuals < 2.5) by and large included items specific to types of hyperhidrosis (body area affected). For example, the following items, *I have difficulties holding objects (Q4), I find it hard*

to handle paper (Q13), I find it hard to touch other people (Q6) are all linked to palmar hyperhidrosis. Items *my choice of footwear is affected* (Q2) and *My eyes feel irritated (Q43)*, are linked to plantar hyperhidrosis and craniofacial hyperhidrosis, respectively. The lack of fit suggests these items were considered to be measuring a different construct, other than the one captured by the Rasch model, indicating multidimensionality problems. On the other hand, nearly all items that overfit the model (with fit residuals < -2.5) were related to psychosocial QOL impacts: *I feel uncomfortable physically expressing affection (e.g. hugging others)* (*Q37), I find it hard to be near others (Q36)*, *I worry about people's reactions (Q35)*, *I feel frustrated (Q29)*, *My self-confidence is affected (Q27)*, *I feel sad (Q24)*, *I feel nervous (Q22)* and *I feel embarrassed (Q21)*. The overlap and interrelation in the content among this set of items meant that a response on one item predicted a choice of a particular response on another. Overfit presents a challenge to the assumption of local independence (Whalley et al. 2004). In addition, the functioning of the response categories was assessed. Optimally functioning response categorization is not only ideal in the context of attaining good fit to the RM at the item level, but also lends support to the substantive aspect of validity, that the response categorization is understood as intended (Wolfe and Smith 2007b). Three items, *My physical activities are affected (Q8), Sweating is constantly on my mind* (Q29) and *I worry about people's reactions* (Q35) had optimally functioning response categories, with category thresholds increasing monotonically. The rest of the items (N = 33) had disordered category thresholds, reflecting inconsistencies in how the sample used the response categories. This was most seen among the categories 'a little' 'somewhat' and 'quite a bit', which may stem from the wide range of overlapping degrees of mildness to which they all relate to. Consequently, subjects might have been unable to distinguish between categories (Pallant and Tennant 2007).

The assumption of unidimensionality was not supported on the HidroQOL-36, based on both PCA analysis of residuals and based on Smith's t-test (Smith 2002). This was already hinted in the misfitting of hyperhidrosis-specific items. In the Rasch framework, multidimensionality may also result from persons whose pattern of response departed from the RM probabilistic prediction. Revision of the HidroQOL-36 based on Rasch analysis, first addressed the issue of dysfunctional response categories. Combining categories '*a little*' and '*somewhat*' resolved response categorization problem in 6 items, 27 items still had disordered thresholds. Collapsing the three categories, '*a little'*, '*somewhat*' and '*quite a bit*' achieved optimally ordered response category thresholds in all items. This sheds light on where most of the inconsistent use of response

categorization occurred, i.e. between '*somewhat*' and '*quite a bit*'. This indicates the importance of descriptors for options on the utility of scales.

Further, misfitting items were removed from the HidroQoL with the intention of achieving greater conformity to the Rasch model (Nijsten et al. 2006a). Relying on statistical considerations for decision-making regarding which items to include and which ones not to, has its own drawbacks, for example the elimination of items important to patients. Nonetheless maintaining items which add little to the measurement aims, either because they are assessing a different construct other than the intended construct or because they provide redundant information (2008a) would not ensure adequate and efficient measurement of the underlying construct. Therefore, eighteen misfitting items were sequentially removed from the HidroQoL. Fit to the RM was enhanced, while reliability and targeting slightly worsened. Nonetheless, reliability remained adequate for individual level comparisons (above 0.85), the difference between item mean location and person mean location increased by more than 0.5 logits. The number of persons with extreme scores increased. The removal of the items had an impact on the scope and range of the continuum of the latent variable in turn affecting reliability and leading to increase in extreme scores (Hagquist et al. 2009). In addition, the removal of misfitting items affected how the remaining items fit the model for example, some overfitting saw an improved fit (Bond and Fox 2007, p.240). This highlights the importance of a sequential approach in the item-reduction process, to allow careful assessment of the impacts of removing items on the scale as a whole and on other individual items. Thus, the sequential process followed in the item reduction of the HidroQoL safeguarded against adversely affecting the goals of development and measurement for the HidroQoL.

During the Rasch analysis six misfitting persons were removed from the sample. Misfitting persons shows response patterns that are inconsistent with a person's level of ability, representing departures from the Rasch probabilistic pattern (Tesio 2003). Underfit of persons can be a case of 'lucky guessing' where people of low ability unexpectedly get a difficult question right, or 'carelessness', where a person of high ability gets an item on a low ability level wrong. While such behaviors tell nothing of the actual person's ability, the risk is that they might be mistakenly included in the calibration of person abilities leading to false conclusions (Bond and Fox 2007, p.64). On the other hand, there seems to be a real danger of removing persons misfitting person from the sample to the extent of compromising the generalisability of findings all in the name of RM (Pallant and Tennant 2007) ultimately rendering the instrument not usable in practice.

An implication of the RM property of invariance of item parameters, is that conditioning on the latent trait, the items are expected to function in the same way in people belonging to different demographic groups (Reeve and Mâsse 2004, p.272). An assessment of DIF on the items of the HidroQoL found no DIF for patient's gender or country. Uniform DIF was detected with respect to disease severity, age, body area affected and co-morbidity. Several items faced DIF across more than one demographic factor, for instance, *My holidays are affected (Q3)*, *My physical activities are affected (Q8)*, *My appearance is affected (Q32)* and *I worry about leaving sweat marks on things (Q34)*, had DIF for age and body area affected. Items Q3 and Q8 also showed DIF for disease severity and co-morbidity, respectively.

Although removing items affected by DIF is seen as a solution, blindly doing so does not always yield intended results. In some situations this may make the DIF worse in the remaining items (Tennant and Pallant 2007). Moreover, removing items may also impact on the definition of the underlying construct (Bond and Fox 2007). This calls for a 'cost-benefit' analysis of the DIF in terms of its measurement implications. All DIF identified was of a moderate to large size, with items Q32 and Q34 showing the largest DIF for body area affected and age. DIF equal to or above 0.64 is considered large, while that equal to or above 0.43 is considered to be slight or moderate (Tristan 2006). Further, an investigation into the effects of the item level DIF on the functioning of the scale as a whole was carried out using the Test Characteristic Curve. This showed that the DIF had minimal impact on the scale as whole. Therefore all items showing DIF were retained. These findings suggest that the HidroQoL can be used across patients with varying levels of disease severity, across different ages, hyperhidrosis affecting different body areas, patients with different co-morbidities, without worrying that the results will be biased against one group for each of the factors. However, cross sectional comparisons of individual patients varying in body area affected using items Q3, Q7, Q32 and Q34 individually would be discouraged due to the non-trivial DIF associated with these items (Edelen et al. 2006). This recommendation extends to items showing DIF for age, disease severity, co-morbidity.

With these findings the HidroQoL is the first disease-specific QOL instrument in hyperhidrosis in which invariance across types of hyperhidrosis has been explicitly assessed. Other measures such as the Hyperhidrosis Scale use a modular approach, where sections of the instrument apply to a specific type of hyperhidrosis. Moreover, in addressing the issues of bias and unidimensionality concurrently, the substantive challenge associated with the optimal scoping of the construct

especially where the target population for the instrument has diverse characteristics is more effectively addressed.

The DIF for body area affected observed on the HidroQoL reflects a fundamental challenge in the design and development of HRQoL measures for hyperhidrosis, that of developing a measure that would be relevant across the different types of hyperhidrosis. None of the previous measures has been assessed for DIF by body area affected, although the design of one measure, the Hyperhidrosis Scale, shows consideration of this aspect by containing subsections relevant to hyperhidrosis of different areas (Keller et al. 2001).

Where an instrument achieves sufficient fit to the RM, item calibrations are expected to remain invariant across patient populations (Bond and Fox 2007). This property was explicitly assessed for the HidroQOL, by comparing initial calibrations (based on a sample from the US & Canada) and a recalibration based on a fresh sample (UK). Hierarchical ordering of items changed for 5 items, including I *feel uncomfortable physically expressing affection (e.g. hugging others) (Q37), I worry about people's reactions (Q35), I worry about leaving sweat marks on things (Q34)*. These items touch upon issues reflecting social norms core to self-image and public life that may differ across any two cultures. Nonetheless, the importance of these difference should not be overemphasized, as the hierarchical ordering is based on point estimates and does not into account measurement errors. Indeed, seventeen of the HidroQoL-18's items and fourteen of the HidroQoL-15's items were invariant between the initial calibration and re-calibration samples, once measurement error was taken into account. This confirms the construct validity of the HidroQoL, by supporting the initial definition of the underlying construct in patient populations from the USA and Canada. On the other hand, this provides compelling evidence for the crucial property of cultural equivalence of the HidroQoL, between the US & Canada and the UK. Ultimately, whether the impact of Rasch model is regarded as 'evolutionary' or 'revolutionary' cannot obviate the unique properties and advantages brought to the scale development and construct validation processes, particularly in achieving well defined constructs. In this respect, the RM allowed thorough understanding and evaluation of the items of HidroQoL finally leading to an optimally defined construct, otherwise not feasible in the CTT.

SUMMARY

- Field testing was carried out on the developmental version of the HidroQOL in patient with hyperhidrosis (N = 595) from the U.S.A, Canada and the UK

- The measurement model of the HidroQOL (HidroQOL-36), following the resolution of item redundancy, was tested using the Rasch model. Further, item reduction and revision was performed on the HidroQOL to achieve conformity to the model. The desirability of this lies in that that the Rasch model fulfils the requirements of conjoint measurement, allowing transformation of the raw scores into interval scaled measures.
- The HidroQOL-36 had overall poor fit to the Rasch model, 10 items underfitted (fit-residuals > 2.5), another 10 items showed overfit (fit-residuals < -2.5). Reliability was strong (PSI = 0.94). Three items had optimally functioning response categories, the remaining 33 items showed disordered category thresholds. The HidroQOL-36 showed lack of unidimensionality.
- Optimal functioning of response categories was achieved by collapsing the categories *quite a bit*, *somewhat* and *a little* as one.
- Eighteen misfitting items were sequentially removed, resulting in a set of 18-items fitting the model (HidroQOL-18). The assumption of unidimensionality was also supported.
- DIF was assessed on the HidroQOL-18: 7 items showed uniform DIF for age, 6 items showed uniform DIF for body-area affected; 1 item showed DIF for disease-severity and another items showed DIF for co-morbidity. DIF for age balanced out for two pairs of items (Q47 and Q3) and (Q34 and Q32), but not across any other factor. Although all DIF identified was non-trivial, its impact on the functioning of the scale as a whole was negligible.
- Non-trivial response dependence was detected in two item pairs Q7 and Q8 and Q14 and Q15 while one other pair (Q4 and Q34) showed trivial response dependence. Three items were removed to resolve this, Q15, Q8 and Q48, resulting in the HidroQOL-15.
- Results of factor analysis for both the 18-item and 15-item versions of the HidroQOL showed a single-factor solution supporting findings from the Rasch model.
- The item calibrations of HidroQOL-18 and HidroQOL-15 were tested for invariance by recalibration in a fresh patient population. For both versions one item, Q29, and Q26, respectively, showed lack of invariance, supporting that the instrument as a whole was largely invariant across patient populations.

CHAPTER 7

Evaluation of the Reliability of the Hyperhidrosis Quality of Life Index (HidroQoL)

INTRODUCTION

The centrality of HRQoL as the ultimate measure of disease impact and efficacy of drug therapies is clear. The current challenge, however, is in how to transform the process of measuring, collecting and applying HRQoL within the clinic, from guesswork into science (Finlay 2011). This is particularly relevant in skin disease where the impairment in HRQoL is profound (Finlay 1998) and represents a key indicator of disease activity. Part of the task entails ensuring that measurement instruments produce valid and reliable results. The latter means that an instrument produces measurements that are free of measurement error (Lohr 2002). In multi-item scales measuring unidimensional concepts, where items are assumed to be indicators of a single underlying construct reliability is demonstrated in internal consistency. The degree to which the different items forming the scales are homogenous or whether they tap into different components of differing constructs (Fayers and Machin 2007). Overall, internal consistency shows that the instrument is capable of identifying variability in patients condition (Streiner and Norman 2008) and that each of the included items contributes to measuring the underlying concept. Where an instrument is used across time, reliability can be demonstrated by reproducibility of scores, test-retest reliability. This reflects the degree to which an instrument yields stable scores over time, with repeated administration, among respondents who are assumed not to have changed (Lohr 2002). This entails that test-retest reliability can only be determined in a longitudinal context and that it relies on the assumption that the patient's condition has indeed not changed.

Reliability is central to the measurement process such that it has an impact on other attributes of an instrument. For example poor reliability may obscure correlation of a measure with other measures, in the assessment of convergence validity. On the other hand, an instrument's ability to detect change over time, responsiveness, is equally affected by poor reliability. Fundamentally, reliability not a property of an instrument, but only an indication of the degree of reliability related to the use of an instrument in specific target populations and in a specific setting (Streiner and Norman 2008). This means that reliability may vary with target population and application of an instrument, indicating the need for establishing reliability each time a measure is put to a new use.

OBJECTIVES

The objectives of this study were to:

- Assess the internal consistency of the scores for the impact on daily life activities and psychosocial impact domains of the HidroQoL; and the overall scale score.
- Assess the test-retest reliability of the individual items of the HidroQoL; the scores for the impact on daily life activities and psychosocial impact domains; and the overall scale score.

METHODS

Study design

This study followed a prospective longitudinal study design with patient's assessed on two occasions, at baseline (assessment 1) and followed-up (assessment 2) at least 7 days after initial assessment. This interval has been recommended (Salek and Luscombe 1992) as offering a good balance between avoiding 'learning effects' in the second assessment and 'ensuring that change in the construct being measured does not take place. In addition, even though patients may experience much variability in their sweating on a day to day basis the overall impacts on their life are relatively stable over a number of days. Moreover, effects of hyperhidrosis treatments such as oral systemic drugs or Iontophoresis last 5 – 14 days, during which time little change may be expected.

Patient population

The study population was recruited through the UK Hyperhidrosis support group and the International Hyperhidrosis Society (IHHS). These two organisations rely largely on social networking for communications with their members. Further details about the patient population are available in chapter 2.

Inclusion criteria:

- Patients with self-reported excessive sweating problems;
- Experiencing some interference in their daily life (HDSS > 1);
- Onset of hyperhidrosis in teenage or early adulthood years;
- Aged 17 or above.

Exclusion criteria:

- Patients not experiencing excessive sweating problems;
- Experiencing no interference in their daily life (HDSS score = 1);

- Onset of hyperhidrosis after age of 30; and reporting a co-morbidity (hypertension, diabetes, PM hormonal disorders, psychological disorders);
- Aged below 17.

Outcome Measures

Apart from the HidroQoL questionnaire, patients were also asked to complete the HDSS, a validated single item scale for assessing the severity of sweating and its interference on patient's daily life (Kowalski et al. 2004). This instrument has been reviewed in chapter 1.

Procedures

Following completion of the first assessment, patients received communication regarding their follow-up assessment due in 7 days and were informed that they would subsequently receive appropriate communication containing the access details for the follow-up questionnaire. On the 5th day following their initial assessment patients were sent an email with the link to the follow-up (second assessment) questionnaire (web-HidroQoL). Although the plan was for all patients to complete the questionnaire on the 7th day, once the email with the details to the second assessment was sent out patients could complete at any time, before or after the 7th day.

Data Processing And Analysis

Data entry was automated; information was directly gathered into a database as patients completed the web HidroQoL. Cleaning and coding of data was performed prior to analysis using statistical software which included SPSS for Windows version 20 (SPSS Inc., Illinois, U.S.A) and STATA Version 11. The analyses involved the estimation of reliability coefficients. Internal consistency of the HidroQoL was assessed using the Cronbach's alpha coefficient, which gives the average inter-item correlation (DeVellis 2011). Internal consistency was also assessed using inter-item and item-total partial correlation, based on Spearman's rank sum correlation. In an internally consistent scale a moderate correlation ($r = 0.3$) is expected between items; and correlation ranging of 0.2 – 0.8 is optimal for item-total partial correlation (Streiner and Norman 2008). A very high inter-item correlation may reflect content redundancy between items.

To assess reproducibility of the HidroQoL scores, the level of agreement between scores from the first (baseline) and second (follow-up) assessments was assessed using Intra Class Correlation

(ICC). This shows the absolute agreement between the two scores, after accounting for both systematic bias and measurement error and is based on a decomposition of the variance in scores using ANOVA (Terwee et al. 2007). Reliability coefficients of 0.7 are optimal for group analyses, while using in individual comparisons requires rates of 0.9 (Nunnally and Bernstein 1994; Lohr 2002).

RESULTS

Socio-demographic Characteristics of Study Participants

A total of 260 participants completed the study questionnaires. One hundred and forty two patients (54.6%) were from the USA and 73 (28.1%) from the UK. The remaining participants (n =45, 17.3%) came from 20 other countries. The mean age of the patients from the US was 38 (±15) with a range of 17 – 73, those from the UK had similar mean age and age range (Table 7.1). The largest age group was those 17 to 30 for both the USA (n = 50, 35.2%) and the UK patient populations (n = 101, 38.8%). Eighty five percent (N = 120) of US sample and sixty three percent (n = 72) of the UK sample were female. The majority of patients had an HDSS score of 3 their sweating was barely tolerable and frequently interfered with their daily activities (USA, n = 63, 44%; UK, n = 32, 44) (Table 7.1).

The majority of patients had seen a doctor before in relation to their sweating (94%, USA; 88%, UK). A total of 65 patients (43%) from the USA and 33 patients (47%) from the UK had received some treatment for their sweating within the last six months, a smaller number (n = 53, 37.5%, USA; n = 26, 35.6%, UK) were being treated currently (Table 7.2). The majority of the patients did not have co-morbidities, the most prevalent among those listed were psychiatric or neurologic disorders (N = 22, 16%, USA; n = 7, 9.6%, UK). The majority considered the effects of the condition on their life as ‘large’ (GQ score = 3) (USA sample, n = 53, 37.3%; UK Sample, n = 33, 45.2%, Pooled, N = 106, 40.8%).

Part I: Internal Consistency

Internal consistency of the HidroQoL was assessed for the UK and the USA samples separately; and for the pooled patient population combining patients from all countries; using the baseline and

follow-up scores. In the pooled sample, the Cronbach's alpha estimates of the HidroQoL overall scale were 0.89 and 0.93, for test 1 and test 2, respectively (Table 7.3).

Table 7.1: Sociodemographic characteristics of the study participants (during assessment 1)

Characteristic	USA Sample (N =142)	UK Sample (N =73)	Pooled Sample* (N = 260)
Gender, n (%)			
Male	22 (15.5%)	27 (37%)	65 (25%)
Female	120 (84.5%)	46 (63%)	195 (75%)
Age, years			
Mean, SD	38, 15	38, 14	37, 14
Median	33	37	33
Mode	31	25	25
Range	57	44	57
Age groups			
below 30	50 (35.2%)	26 (35.6%)	101 (38.8%)
31 - 40	33 (23.2%)	23 (31.5%)	69 (26.5%)
41 - 50	29 (20.4%)	10 (13.7%)	43 (16.5%)
51 - 60	17 (12 %)	4 (5.5%)	23 (8.8%)
61 plus	13 (9.2%)	10 (13.7%)	24 (9.2%)
Body site involved, n (%)			
General	31	15	52
Palms, feet & armpits	50	11	73
Palms and feet	29	20	65
Armpits only	7	4	15
Armpits plus other	8	3	13
Head	11	11	24
Palms	3	3	8
Feet	3	2	6
Trunk/other	1	2	4
Employment, n (%)			
Employed	91 (64.1%)	42 (57.5%)	160 (61.5%)
Unemployed	22 (15.5%)	9 (12.3%)	37 (14.2%)
Retired	10 (7%)	11 (15.1%)	21 (8.1%)
Full time student	19 (13.4%)	11 (15.1%)	42 (16.2%)
***Country, n (%)**			
USA			142 (54.6%)
UK			73 (28.1%)
Australia			11 (4.2%)
Canada			11 (4.2%)
other			23 (9%)

Table 7.1 (continued)

	USA sample	UK sample	Pooled* sample
HDSS Score, n (%)			
1	0	0	0
2	28 (19.7%)	17 (23.3%)	51 (19.6%)
3	63 (44.4%)	32 (43.8%)	120 (46.2%)
4	51 (35.9%)	24 (32.9%)	89 (34.2%)
GQ Score, n (%)			
0		2 (2.74%)	2 (0.77%)
1	9 (6.34%)	5 (6.85%)	18 (6.92%)
2	37 (26.06%)	17 (23.29%)	65 (25%)
3	53 (37.32%)	33 (45.21%)	106 (40.77%)
4	43 (30.28%)	16 (21.92%)	69 (26.54%)

Table 7.2: Disease-related characteristics of study participants

Characteristic	USA Sample (n = 142)		UK Sample (n = 73)		Pooled Sample* (n = 260)	
Treated by a medical practitioner regarding hyperhidrosis	133	93.7%	64	87%	233	89.6%
Have received Surgical treatment	19	13.4%	8	11%	33	12.7%
Received Botox within last 6 months	19	13.4%	5	6.8%	27	10.4%
Received treatment within last 6 months	65	45.8%	33	45.2%	114	43.8%
Currently receiving treatment	53	37.5%	26	35.6%	89	34.2%
Oral-systemic drugs (pill-form)	31	21.8%	13	17.8%	47	18.1%
Iontophoresis	10	7%	13	17.8%	25	9.6%
Aluminium Chloride Topical treatment	15	10.6%	4	5.5%	22	8.5%
Non-prescription/cosmetic preparations	17	12%	7	9.6%	27	10.4%
Co-morbidities, n (%)						
Thyroid disorders	10	7%	3	4.1%	13	5%
Psychiatric or neurologic disorders	22	15.5%	7	9.6%	30	11.5%
Menopausal related complaints	13	9.2%	2	2.7%	16	6.2%
Diabetes	10	7%	1	1.4%	11	4.2%
Hypertension	20	14.1%	5	6.8%	29	11.2%
Other	25	17.6%	15	20.5%	47	18%

Note: * The pooled sample includes patients from 11 other countries in addition to the USA and the UK

Coefficient estimates for the *impact on daily life activities* domain (H-DA) were 0.76 and 0.86; and for the *psychosocial impact* domain (H-PS) they were 0.86 and 0.90, for test 1 and test 2, respectively. Estimates obtained from the US sample were larger, while those from the UK sample were the smallest, although all within a percentage point margin of difference. Optimal homogeneity is reflected in moderate inter-item correlation and moderate-to-strong corrected item-total correlations (Streiner and Norman 2008). This was, therefore, also examined for each of the HidroQoL's items. In the pooled sample, corrected item-total correlation ranged from 0.376 to 0.618 (Table 7.4). The lowest correlation was seen on item 1 (My choice of clothing is affected, rs = 0.376), while that for item 15, $r_s = 0.618$, was the highest. In the US sample corrected item-total correlations ranged from 0.410 to 0.664. Item 'I feel frustrated' (item 9) had the highest correlation while the lowest was seen on 'my sex life is affected' (item 18). In the UK group the values of item-total correlation ranged from 0.24 to 0.739 (Table 7.5). The item 'I avoid meeting new people' (item 15) had the highest item-correlation value, while the lowest value was seen on item 'my choice of clothing is affected' (item # 1). This indicates that the HidroQoL is well balanced, as no item carried too much weight; each of the included items tapped an aspect of the underlying construct (hyperhidrosis QoL), including the item 'my choice of clothing is affected'. On the other hand, the items 'I feel frustrated' and 'I avoid meeting new people' seem to highlight the experiences of having hyperhidrosis, in summing up the emotional and social responses of the disease by the patient.

Inter-item correlations

In classical test theory a core assumption is that items of the instrument reflect a sampling from the universe of indicators of a given underlying construct (Nunnally and Bernstein 1994). As all items are expected to be tapping into the underlying construct, items are expected to at least share a moderate correlation with each other. Therefore, the correlation among the items of the HidroQoL was assessed.

Table 7.3: Internal Consistency* of the HidroQoL

HidroQoL score	US Sample		UK Sample		Pooled Sample	
	Test 1	*Test 2*	*Test 1*	*Test 2*	*Test 1*	*Test 2*
Total scale, 18 items	0.89	0.94	0.89	0.92	0.89	0.93
Impact on daily life activities	0.78	0.89	0.72	0.78	0.76	0.86
Psychosocial impact	0.87	0.91	0.85	0.89	0.86	0.90

Note: *Cronbach alpha coefficient

Table 7.4: Item-total scale correlations, for the HidroQoL, pooled/international, test 1 (n = 260)

	SVID	CITC	CAID
My choice of clothing is affected	42.71	.376	.885
My physical activities are affected	40.93	.493	.882
My hobbies are affected	41.18	.483	.882
My work is affected	41.0	.535	.881
I worry about the additional activities in dealing with my condition	41.0	.574	.879
My holidays are affected (e.g. planning, activities)	40.5	.496	.882
I feel nervous	40.7	.588	.879
I feel embarrassed	42.1	.578	.881
I feel frustrated	42.2	.571	.881
I feel uncomfortable physically expressing affection (e.g. hugging)	40.9	.550	.880
I think about sweating	42.1	.566	.881
I worry about my future health	40.0	.535	.881
I worry about people's reactions	42.4	.526	.882
I worry about leaving sweat marks on things	42.5	.450	.883
I avoid meeting new people	38.9	.618	.877
I avoid public speaking (e.g. presentations)	39.2	.592	.879
My appearance is affected	39.8	.574	.879
My sex life is affected	40.7	.417	.886

Note: CAID, Cronbach's alpha if item is deleted; **CITC**, Corrected Item-Total Correlation; **SVID**, Scale Variance if Item Deleted

In the pooled patient population, item 1 (*my choice of clothing is affected*) had correlations greater than 0.3 with two out of the 17 other items. The remaining items belonging to the *impact on daily life activities* domain (item 2 to item 6) showed an inter-item correlation ranging from 0.27 to 0.48 (Table 7.7). The lowest correlation was between my *My holidays are affected e.g. planning, activities* and *My work is affected* (r = 0.27). Item 18 showed correlations greater than 0.3 with three other items in the full HidroQoL scale. The remaining items in the psychosocial domain (items 7 to 17), showed correlations ranging from 0.20 to 0.63, with six item pairs below 0.3. Item 17 (*my appearance is affected*) showed moderate to strong correlation with all items in the *impact on daily life activities* domain and seven items in the *psychosocial impact* domain. This suggests that 'effects on appearance' might be an underlying issue in understanding impact on daily life activities resulting from HH.

Table 7.5: Item-total scale correlations, for the HidroQoL, U.S. Sample, test 1 (n = 142)

	SVID	CITC	CAID
My choice of clothing is affected	42.428	.426	.888
My physical activities are affected	41.177	.475	.886
My hobbies are affected	41.607	.484	.886
My work is affected	41.208	.536	.884
I worry about the additional activities in dealing with my condition	41.003	.568	.883
My holidays are affected (e.g. planning, activities)	40.696	.480	.887
I feel nervous	40.204	.636	.881
I feel embarrassed	42.055	.583	.884
I feel frustrated	41.608	.664	.882
I feel uncomfortable physically expressing affection (e.g. hugging)	40.407	.606	.882
I think about sweating	41.869	.621	.883
I worry about my future health	40.051	.542	.884
I worry about people's reactions	41.941	.560	.884
I worry about leaving sweat marks on things	42.340	.446	.887
I avoid meeting new people	39.284	.580	.883
I avoid public speaking (e.g. presentations)	39.669	.541	.885
My appearance is affected	40.180	.546	.884
My sex life is affected	40.850	.410	.890

Note: **CAID**, Cronbach's alpha if item is deleted; **CITC**, Corrected Item-Total Correlation; **SVID**, Scale Variance if Item Deleted

On the other hand, item 5 (*I worry about the additional activities in dealing with my condition*) showed moderate to strong correlation with 10 of the 12 items in the psycho-social domain. Thus, chores in dealing with hyperhidrosis, may not only represent a challenge to daily life activities but may also be taxing on patient's psycho-social life.

Inter-item correlations were also explored for the US patient population. Correlations of item 1 with item 3 and item 4 were 0.28 and 0.27, otherwise the rest of the items in domain 1 had correlation ranging from 0.31 to 0.41. Items 17 and 18 had a pattern noted previously in the pooled patient population. Item 17 had moderate correlations with all items in domain 1 (items 1 to 6) but only four of the domain 2 items. On the other hand, item 18 had correlations of at least 0.3 with 2

items only in the whole scale. The correlation of item 7 (*I feel nervous*) with eight items in the psychosocial impact was greater than 0.4, hinting that across the different psychosocial impacts experienced by the patients 'feeling nervous' was a cross-cutting impact. The highest correlation across the items was seen between item 8 and item 9 (r = 0.675). In the UK group, item 1 had a correlation of a moderate magnitude with only two other items in the instrument (item 9 and item 17). On the other hand, 4 items had a correlation of at least moderate magnitude with item #18. The rest of the items in domain 1 had a correlation ranging from 0.12 to 0.57, among each other. On the other hand, inter-item correlation of items in domain 2 ranged from 0.168 to 0.648. Item #5 showed a moderate strong correlation with 9 items in the psychosocial domain, as previously noted in the US patient population.

Table 7.6: Item-total scale correlations, for the HidroQoL, UK Sample, test 1 (n = 73)

	SVID	CITC	CAID
My choice of clothing is affected	46.618	.240	.889
My physical activities are affected	42.651	.587	.879
My hobbies are affected	42.981	.566	.879
My work is affected	42.818	.568	.879
I worry about the additional activities in dealing with my condition	43.709	.575	.879
My holidays are affected (e.g. planning, activities)	43.541	.478	.883
I feel nervous	43.898	.497	.882
I feel embarrassed	45.527	.553	.882
I feel frustrated	46.166	.434	.884
I feel uncomfortable physically expressing affection (e.g. hugging)	43.454	.553	.880
I think about sweating	45.097	.522	.882
I worry about my future health	43.139	.480	.883
I worry about people's reactions	45.301	.520	.882
I worry about leaving sweat marks on things	44.784	.521	.881
I avoid meeting new people	40.581	.739	.872
I avoid public speaking (e.g. presentations)	41.965	.595	.878
My appearance is affected	42.136	.604	.878
My sex life is affected	43.427	.423	.886

Note: **CAID**, Cronbach's alpha if item is deleted; **CITC**, Corrected Item-Total Correlation; **SVID**, Scale Variance if Item Deleted

Pattern noted for item 7 and 17 was not maintained. The results from the different samples consistently show strong reliability for the HidroQoL. The small differences were likely a consequence of differences in sample sizes.

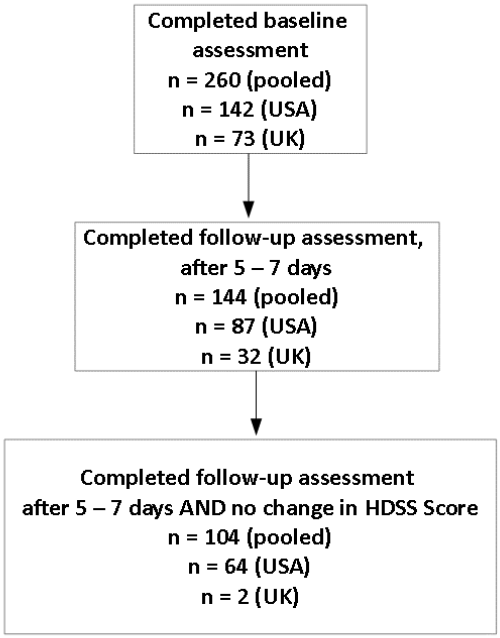
Part II: Inter-temporal stability of the HidroQoL Scores

The reproducibility of the HidroQoL scores in repeated administration was tested. Patients completed the HidroQoL on two occasions, at baseline (test 1) and follow up assessment (test 2). A central issue in the reliability relates to ensuring that the patients condition has not indeed changed. One approach is to use a reasonably short time frame, to ensure that the underlying condition of the patient does not change but not too short to risk the patients recalling the prior responses. In this study patients took the follow-up assessment 5 to 7 days after initial assessment. On the other hand, patients also completed the HDSS scale, a self-assessment disease severity scale. Test-retest reliability of the HidroQoL was assessed only in patients whose underlying disease severity had not changed.

A total of 144 patients (pooled population) completed the second assessment out of the 260 patients completing the initial assessment, 104 patients showed no change on their HDSS score between the first and second assessments, therefore only these were considered in the analysis (Figure 7.1). The level of agreement between the baseline (test 1) and follow-up scores (test 2) was assessed using ICC. In the pooled sample, the level of agreement in the HidroQoL scores was strong (ICC: Overall scale score, 0.92; H-DA, 0.8; H-PS, 0.91) (Table 7.10).

The individual items scores also showed a strong reproducibility (ICC range, 0.74 – 0-88). The ICC for item 5 (*I worry about additional activities in dealing with my condition*) was the lowest (ICC = 0.59). In the USA sample, similar results were observed (ICC: Overall scale scores, 0.92; H-DA, 0.89; H-PS, 0.90) (Table 7.11). The ICC of the individual item scores ranged from 0.654 to 0.88. Item 5 (*I worry about the additional activities in dealing with my condition*) showed the lowest ICC (0.456). The UK patient population was small (n = 22) thus the obtained estimates may be considered only as preliminary. The HidroQoL showed strong test-retest reliability in the UK patient population. The HidroQoL total score had an ICC of 0.93 (Table 7.12).

Figure 7.1: Patients included in the test-retest reliability study



On the other hand, the ICC values for the domain scores were 0.87 for *impact on daily life activities* domain and 0.92 for the *psychosocial impact* domain. At the individual item level, the item *I worry about the additional activities in dealing with my condition* also showed the lowest ICC (0.59). The rest of the items had an ICC ranging from 0.65 to 0.93.

DISCUSSION

The broader impacts of disease on the individual patient's life, particularly the impairment in daily life and limitations associated with psychosocial functioning, need to be addressed in patient management as well as clinical studies as a corner-stone to a truly patient centred care. Moreover, in skin conditions such as hyperhidrosis, symptoms alone might be insufficient to provide an accurate picture of disease activity. In such situations QoL impairment functions as an important predictor of disease activity (Chren 2005). This means that the measurement of HRQoL including the related processes of collecting, analysing and using such information ought to proceed scientifically to ensure that measures are free from any measurement-error (Guyatt et al. 1993).

Table 7.7: Correlations[§] among HidroQoL's items, test 1, pooled sample (n = 260)

	Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	My choice of clothing is affected	1																	
2	My physical activities are affected	.282	1																
3	My hobbies are affected	.245	.479	1															
4	My work is affected	.207	.328	.450	1														
5	I worry about the additional activities in dealing with my condition	.293	.355	.354	.447	1													
6	My holidays are affected (e.g. planning, activities)	.266	.419	.359	.266	.368	1												
7	I feel nervous	.100	.214	.314	.398	.379	.219	1											
8	I feel embarrassed	.218	.188	.213	.234	.305	.325	.507	1										
9	I feel frustrated	.377	.182	.194	.243	.348	.265	.449	.627	1									
10	I feel uncomfortable physically expressing affection (e.g. hugging)	.261	.331	.276	.342	.349	.220	.495	.424	.382	1								
11	I think about sweating	.143	.262	.189	.358	.363	.263	.481	.549	.482	.340	1							
12	I worry about my future health	.279	.266	.167	.246	.361	.361	.355	.312	.370	.249	.350	1						
13	I worry about people's reactions	.078	.216	.152	.305	.313	.146	.495	.522	.399	.402	.437	.307	1					
14	I worry about leaving sweat marks on things	.202	.114	.182	.227	.260	.112	.377	.419	.426	.331	.424	.232	.605	1				
15	I avoid meeting new people	.165	.281	.268	.362	.359	.287	.480	.360	.309	.383	.385	.457	.408	.377	1			
16	I avoid public speaking (e.g. presentations)	.159	.284	.348	.413	.384	.331	.438	.353	.375	.322	.331	.372	.304	.231	.619	1		
17	My appearance is affected	.373	.399	.329	.311	.325	.513	.233	.375	.353	.315	.332	.376	.267	.205	.343	.328	1	
18	My sex life is affected	.195	.274	.271	.249	.230	.209	.185	.178	.227	.271	.231	.321	.212	.172	.313	.284	.350	1

[§] Spearman's rank correlations

Table 7.8: Correlations[§] among HidroQoL's items, test 1, USA Sample (n = 142)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. My choice of clothing is affected	1																	
2. My physical activities are affected	.389	1																
3. My hobbies are affected	.281	.512	1															
4. My work is affected	.269	.354	.410	1														
5. I worry about the additional activities in dealing with my condition	.382	.345	.334	.398	1													
6. My holidays are affected (e.g. planning, activities)	.313	.433	.339	.394	.432	1												
7. I feel nervous	.167	.241	.244	.303	.379	.256	1											
8. I feel embarrassed	.243	.169	.260	.250	.309	.299	.522	1										
9. I feel frustrated	.396	.208	.306	.321	.386	.315	.566	.675	1									
10. I feel uncomfortable physically expressing affection (e.g. hugging)	.307	.293	.187	.358	.424	.246	.589	.453	.501	1								
11. I think about sweating	.201	.272	.274	.392	.387	.267	.513	.522	.500	.501	1							
12. I worry about my future health	.310	.227	.204	.252	.410	.346	.419	.364	.387	.266	.368	1						
13. I worry about people's reactions	.098	.221	.147	.233	.278	.133	.583	.543	.490	.480	.470	.323	1					
14. I worry about leaving sweat marks on things	.208	.113	.192	.159	.246	.053	.398	.433	.472	.347	.394	.206	.606	1				
15. I avoid meeting new people	.158	.233	.241	.272	.306	.216	.481	.329	.334	.416	.380	.490	.464	.375	1			
16. I avoid public speaking (e.g. presentations)	.106	.231	.299	.352	.297	.276	.476	.280	.437	.372	.359	.389	.321	.215	.638	1		
17. My appearance is affected	.382	.393	.397	.376	.301	.446	.277	.346	.415	.265	.356	.333	.298	.225	.248	.230	1	
18. My sex life is affected	.212	.219	.294	.349	.218	.154	.209	.209	.244	.328	.299	.239	.222	.218	.287	.209	.350	1

[§] Spearman's rank correlations

Table 7.9: Correlations[§] among HidroQoL's items, test 1, UK (n = 72)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	My choice of clothing is affected	1																	
2	My physical activities are affected	.156	1																
3	My hobbies are affected	.112	.380	1															
4	My work is affected	.052	.263	.569	1														
5	I worry about the additional activities in dealing with my condition	.046	.401	.306	.526	1													
6	My holidays are affected (e.g. planning, activities)	.245	.511	.450	.122	.226	1												
7	I feel nervous	-.031	.174	.533	.575	.393	.111	1											
8	I feel embarrassed	.187	.293	.220	.293	.253	.304	.437	1										
9	I feel frustrated	.310	.246	.172	.199	.309	.120	.182	.501	1									
10	I feel uncomfortable physically expressing affection (e.g. hugging)	.181	.506	.426	.276	.257	.251	.420	.455	.302	1								
11	I think about sweating	.073	.332	.198	.411	.406	.213	.338	.503	.364	.205	1							
12	I worry about my future health	.127	.347	.184	.177	.313	.385	.168	.194	.323	.257	.300	1						
13	I worry about people's reactions	.076	.296	.235	.462	.434	.178	.339	.420	.273	.274	.417	.283	1					
14	I worry about leaving sweat marks on things	.226	.193	.291	.374	.300	.260	.309	.443	.378	.353	.518	.274	.648	1				
15	I avoid meeting new people	.199	.476	.464	.524	.518	.254	.518	.411	.284	.422	.484	.395	.400	.444	1			
16	I avoid public speaking (e.g. presentations)	.081	.382	.474	.550	.413	.240	.372	.417	.204	.282	.299	.197	.280	.231	.624	1		
17	My appearance is affected	.432	.524	.290	.250	.362	.549	.118	.359	.270	.452	.291	.436	.227	.257	.378	.389	1	
18	My sex life is affected	.063	.274	.198	.136	.283	.293	.135	.125	.206	.261	.131	.451	.201	.105	.451	.392	.374	1

[§] Spearman's rank correlations

Table 7.10: Test-retest reliability for individual items of the HidroQoL, international Sample (n = 104)

		ICC	95% CI		
			Lower	*Upper*	Sig
1	My choice of clothing is affected	.741	.620	.824	.0001
2	My physical activities are affected	.799	.704	.863	.0001
3	My hobbies are affected	.831	.747	.886	.0001
4	My work is affected	.740	.619	.823	.0001
5	I worry about the additional activities in dealing with my condition	.592	.402	.722	.0001
6	My holidays are affected (e.g. planning, activities)	.768	.660	.842	.0001
7	I feel nervous	.860	.793	.904	.0001
8	I feel embarrassed	.874	.816	.914	.0001
9	I feel frustrated	.760	.648	.836	.0001
10	I feel uncomfortable physically expressing affection (e.g. hugging)	.770	.663	.843	.0001
11	I think about sweating	.718	.587	.808	.0001
12	I worry about my future health	.822	.739	.879	.0001
13	I worry about people's reactions	.741	.610	.826	.0001
14	I worry about leaving sweat marks on things	.779	.673	.850	.0001
15	I avoid meeting new people	.879	.823	.918	.0001
16	I avoid public speaking (e.g. presentations)	.798	.702	.863	.0001
17	My appearance is affected	.848	.777	.897	.0001
18	My sex life is affected	.876	.814	.916	.0001
	HidroQoL - Daily life activities	.883	.828	.921	.0001
	HidroQoL - Psychosocial domain	.914	.868	.943	.0001
	HidroQoL-total	.926	.885	.952	.0001

Otherwise, distinguishing patients with different levels of impairment or assessing change in the patients' condition may be obscured (Terwee et al. 2003). This study, therefore, was aimed at assessing the reliability of the HidroQoL in patients with hyperhidrosis. Reliability was assessed for the total scale as well as for the two domains, impact on daily life activities and psychosocial domains of the HidroQoL. The results showed adequate internal consistency according to

Cronbach's alpha for the HidroQoL total score, as well as for the two domains, daily life activities domain and for the psychosocial domain. This suggests that the domains (sub-scales) and the overall scale are homogeneous. The observed moderate-strong corrected item-total correlations provide further evidence that the items optimally tap into the same construct. This means that each of the scales have been optimally defined and that each of the items reflect a different aspect of the core construct.

Table 7.11: Test-retest reliability for individual items of the HidroQoL, USA Sample (n = 64)

	Item	ICC	95% CI		P-value
			Lower	*Upper*	
1	My choice of clothing is affected	.759	0.604	0.853	.0001
2	My physical activities are affected	.806	0.682	0.882	.0001
3	My hobbies are affected	.844	0.74	0.906	.0001
4	My work is affected	.769	0.622	0.859	.0001
5	I worry about the additional activities in dealing with my condition	.456	0.105	0.669	0.008
6	My holidays are affected (e.g. planning, activities)	.806	0.681	0.882	0.000
7	I feel nervous	.845	0.746	0.905	0.000
8	I feel embarrassed	.889	0.819	0.932	0.000
9	I feel frustrated	.798	0.668	0.877	0.000
10	I feel uncomfortable physically expressing affection (e.g. hugging)	.750	0.59	0.848	0.000
11	I think about sweating	.709	0.523	0.823	.0001
12	I worry about my future health	.857[c]	0.766	0.913	.0001
13	I worry about people's reactions	.697	0.505	0.814	.0001
14	I worry about leaving sweat marks on things	.654	0.437	0.789	.0001
15	I avoid meeting new people	.864	0.776	0.917	.0001
16	I avoid public speaking (e.g. presentations)	.761	0.607	0.854	.0001
17	My appearance is affected	.771	0.624	0.861	.0001
18	My sex life is affected	.868	0.783	0.919	.0001
	HidroQoL - Daily life activities	**.892**	**0.824**	**0.934**	**.0001**
	HidroQoL - Psychosocial domain	**.904**	**0.843**	**0.941**	**.0001**
	HidroQoL-total	**.919**	**0.868**	**0.951**	**.0001**

Table 7.12: Test-retest reliability for individual items of the HidroQoL, UK Sample (N = 22)

	Item	ICC	95% CI		p-value
			Lower	*Upper*	
1	My choice of clothing is affected	.784	.485	.909	.000
2	My physical activities are affected	.779	.471	.907	.001
3	My hobbies are affected	.814	.559	.922	.000
4	My work is affected	.826	.597	.926	.000
5	I worry about the additional activities in dealing with my condition	.593	.084	.824	.016
6	My holidays are affected (e.g. planning, activities)	.645	.177	.848	.009
7	I feel nervous	.885	.732	.951	.000
8	I feel embarrassed	.878	.713	.948	.000
9	I feel frustrated	.738	.398	.888	.001
10	I feel uncomfortable physically expressing affection (e.g. hugging)	.809	.504	.922	.000
11	I think about sweating	.721	.342	.881	.001
12	I worry about my future health	.744	.395	.892	.001
13	I worry about people's reactions	.820	.542	.926	.000
14	I worry about leaving sweat marks on things	.917	.801	.965	.000
15	I avoid meeting new people	.927	.829	.969	.000
16	I avoid public speaking (e.g. presentations)	.749	.350	.898	.000
17	My appearance is affected	.901	.768	.958	.000
18	My sex life is affected	.870	.596	.951	.000
	HidroQoL - Daily life activities	.866	.689	.943	.000
	HidroQoL - Psychosocial domain	.919	.649	.973	.000
	HidroQoL-total	.932	.740	.976	.000

Although the general guide in scale development is to delete those items that contribute little to variance of the scale i.e. items which when deleted do not result in a major change in scale variance (Fayers and Machin 2007), the fact that all items of the HidroQoL affected total variance of the scale within comparable magnitudes reflects balance in the instrument, that all items made a largely similar contribution to the scale.

The correlation of the item '*my choice of clothing is affected*' with the rest of the items was low, particularly in the UK population and for the pooled sample. This mirrors results from the Rasch analysis reported in the previous chapter, where this item also performed sub-optimally. This represents a friction given that this was one of the most prevalent items during the qualitative research done during the content development of the measure. Elimination of such an item might compromise the content validity and clinical relevance of the instrument. Similarly the item '*My sex life is affected*' while showing a low correlation with the rest of the instrument's items, was placed on the upper end of the impairment continuum during Rasch analysis. Omitting this item might compromise the instrument's ability to measure patients experiencing extreme effects. Together, these results show that the HidroQoL is indeed sensitive to variability in the patient's condition, with minimal measurement error.

The reproducibility of the HidroQoL has been established. This involved repeated administration of the instrument, 5 to 7 days following baseline (first) assessment. This period was considered long enough to ensure patients do not recall their initial answers but short enough for the condition to have remained stable (Salek and Luscombe 1992). To ensure this, a disease severity scale (the HDSS) was administered. Only patients with stable disease severity were included during the analysis.

The results showed a strong level of agreement between the baseline and follow-up scores in patients whose condition had not changed, for the HidroQoL total score as well as for the two domain scores, impact on daily life activities and psychosocial impacts. This indicates that the HidroQoL appropriately distinguishes clinically relevant change from measurement error. Moreover, the magnitude of observed ICC would support use of the measure for QoL measurement in individual patients.

SUMMARY

- The internal consistency and reproducibility of the HidroQoL were established in differing patient populations, including a group from the UK and U.S.A and a pooled international patient population.

- The internal consistency Cronbach's alpha was estimated for the HidroQoL scale and for the two domains (impact on daily life activities and psychosocial domain) in the pooled patient population as well as for the UK and USA patient populations, separately.
- Reproducibility of the HidroQoL was assessed for HidroQoL total score, the two domain scores (impact on daily life activities impact and psychosocial impact domain) and the individual items by estimating the level of agreement between scores from the baseline and the follow-up assessments using ICC.
- Cronbach's alpha values ranged from 0.89 to 0.94 for the HidroQoL total score, 0.72 – 0.88 for the impact on daily life activities domain and 0.87 – 0.9 for the psychosocial domain.
- Therefore, scales of the HidroQoL (overall scale, impact on daily life activities domain and psychosocial impact domain) showed optimal homogeneity, reflecting clear definition of the scales as well as the inclusion of relevant items.
- The ICC between the baseline and follow scores were, 0.88 – 0.87 for the daily life activities, 0.90 – 0.92 for the psychosocial domain and 0.92 – 0.93 for the HidroQoL total scores. This provides strong support for the application of the measure in evaluating QoL in individual patients
- The results obtained support the longitudinal as well as cross-sectional application of the HidroQoL scores in USA and UK patient populations as well as in international patient population.

CHAPTER 8

Evaluation of the Validity of the Hyperhidrosis Quality of Life Index (HidroQoL)

INTRODUCTION

Validity encompasses the evaluative judgement of the degree to which empirical evidence and theoretical rationales support the trustworthiness of interpretations and actions based on scale scores (Messick 1988). This indicates a focus on the participants and their responses; and the inferences that can be drawn about them, based on their scale scores (responses). Although a delineation is made among different types of validity (content validity, criterion validity and construct validity) a unified perspective of validity considers all forms of validity to be encompassed by construct validity (Streiner and Norman 2008). Construct validity relates to the extent to which theoretically derived hypothesis relating to the construct being measured by an instrument are supported by empirical evidence (Terwee et al. 2007). Although there is no prescription regarding type, form and nature of such empirical evidence, the need to demonstrate construct validity, arises each time a measure is used in a new situation or where different inference will be drawn, reflecting on the continuous nature of the validation process (Streiner and Norman 2008). For this reason, there is an even greater imperative to generate such evidence for new instruments. Evidence demonstrating the adequacy with which the content of the new instrument, the HidroQoL, covers and represents the full content domain of HRQoL issues in HH was presented in chapter four. Additional construct validation data based on the internal structure of the new measure, applying both the EFA and CFA as well as modern test theory's Rasch model were presented in chapter's five and six. In the current study further construct validation of the HidroQoL was undertaken including: testing for group differences in the scores of the HidroQoL across gender, age groups and disease severity; and testing the relationship between scores of the HidroQoL and those of other established instruments i.e. *convergent and divergent validities.*

OBJECTIVES

The objectives of this study were to:

- Explore for differences in HidroQoL scores in patients with different characteristics with respect to; demographic factors; level of disease severity; overall impact of disease; daily time spent in managing symptoms of the condition and their impacts; values of Willingness to Pay (WTP) for cure of condition.

- Assess the relationship between the scores of the HidroQoL and other measures of disease impact, including: EQ-5D score, Skindex-17 score, DLQI score, HDSS score, GQ score, WTP values, daily time spent managing condition.

METHODS

Study design

This study followed a cross-sectional design where a heterogeneous group of patients from the USA and the UK were assessed on a single occasion. The absence of a recommended design for validation studies means that the choice of study design is dependent on the hypotheses assessed. The study population was recruited through the UK Hyperhidrosis support group and the International Hyperhidrosis Society (IHHS). Further details regarding the patient population and recruitment process are available in Chapter 2.

Inclusion criteria:

- Patients with self-reported excessive sweating problems;
- Experiencing some interference in daily life (HDSS > 1);
- With onset of hyperhidrosis in teenage or early adulthood years;
- Aged 17 or above.

Exclusion criteria:

- Patients not experiencing any excessive sweating problems;
- Experiencing no interference in daily life (HDSS = 1);
- With onset of hyperhidrosis after age of 30 and reporting a co-morbidity (hypertension, diabetes, PM hormonal disorders, psychological disorders);
- Aged below 17.

Outcome Measures

Apart from the HidroQoL, data were also collected on: patient's disease severity using the HDSS; dermatology-specific QoL using the DLQI and the Skindex-17; and generic HRQoL using the EQ-5D. In addition, the following questions were administered:

- Global question (GQ) on overall impact of hyperhidrosis, scored on a 5 point Likert scale;
- Patient's willingness to pay for a complete cure in hyperhidrosis;
- Time spent daily in dealing with HH; and
- Additional monthly expenditures arising from hyperhidrosis.

The HRQoL instruments have been reviewed and presented in more detail in Chapters 1 and 2 (Table 8.1).

Data Processing And Analysis

The use of a web-system for the data collection made it possible to have the data directly entered into a database automatically during completion, eliminating the need for manual data entry thus avoiding potential errors (Dillman 2006). All data analyses were carried out using STATA 11.2 and SPSS. Further, descriptive analysis was carried out to explore the distribution of variables. Hypothesis testing used a conventional level of significance of 0.05 (Munro 2005). Specifically the following tests were carried out:

- the Mann-Whitney U test and K-Wallis were utilised to test for group differences in HidroQoL scores (for instance across patient's socio-demographic characteristics and disease characteristics; HDSS scores; GQ scores).
- Spearman's Rank correlation analysis was to assess the relationship between the scores of the HidroQoL and other measures. A correlation coefficient greater than 0.3 – 0-4 supports convergence validity (Fayers and Machin 2007).
- Univariate OLS regression was used to assess the relationship between the HidroQoL scores and the scores of the HDSS; Skindex-17; DLQI; and EQ-5D. The coefficient of determination, R^2, provided a measure of how much variance in the independent variable was being explained by the predictor variable (Norman and Streiner 2007).
- Multivariate OLS regression was used to determine the predictors of HRQoL in hyperhidrosis patients, HidroQoL score was the independent variable; and patient's socio-demographic and disease characteristics were dependent variables. Post-hoc diagnostic tests were carried out to assess model fit and assumptions, including tests for normality of residuals (Shapiro Wilk Test; Kernel density Plots); heteroskedasticity (Breuch Pagan and White's tests; scatter plot of fitted vs. residuals) and multi-colinearity (Gujarati 2003)

RESULTS

Sociodemographic Characteristics Of Study Participants

A total of 163 participants completed the HidroQoL questionnaire, out of 204 initially enrolled for the study, representing 80% completion rate. One hundred and twenty seven patients (78%) were from the USA and thirty six (22%) the UK (Table 8.2).

Table 8.1: Attributes of outcome measures used in data collection

	DLQI	**Skindex - 17**	**HDSS**	**EQ_5D**
Concept measured	Impact of skin disease on patient's QoL (based on intensity of effects)	effects of skin-disease on HRQoL: (based on frequency of effects)	hyperhidrosis severity and degree of interference in daily life	Generic QoL (health status)
Target population	Adults with skin disease	Adults with skin disease	Adults with hyperhidrosis	Adults
Number of items	10	17	1	5 domains plus a VAS scale
Number of domains	*8:* Symptoms, Daily activities, Leisure Work/school, Personal relationships, Treatment	*2:* Symptom, Psychosocial	Na	*5 (for descriptive part):* mobility, self-care, usual activities, pain/discomfort, anxiety
Scoring format	A total score is calculated by summing the item scores	*Total score* and *domain scores* calculated by summing individual scores.	Na	*Descriptive part*: analysis at individual item level only; item scores forming a 5-digit number can be read off a reference values to obtain health status preference values
Total score range	0 (no impact) - 30 (maximum impact on QoL)	0 (no impact) - 34 (maximum impact on QoL)	0 (lowest severity) - 4 (highest severity)	*Descriptive part:*for individual items, 1 (no problems) - 5 (maximum problems). *VAS scale:* 0 (worst health imaginable) - 100 (best health imaginable)

The mean age of the patients from the USA was 38.8 (14.1) with a range of 17 – 73, while those from the UK had a mean of 42.8 (16.13), ranging from 20 to 74. The age group 17 to 29 was the largest (n = 37, 29%) in the USA sample, while the age group 30 to 39 was the largest in the UK sample (n = 12, 33.3%) (Figure 8.1). One hundred and twenty patients (84%) from the US and seventy two patients (72%) from the UK sample were female. The majority of USA patients (N = 57) had an HDSS score of 3, their sweating was barely tolerable and frequently interfered with their daily activities; while the majority in the UK sample (n=13) had the most extreme HDSS score (representing a score of 4 out of 4), their sweating was intolerable and always interfered with daily activities (Figure 8.2). The majority (n = 46, USA sample; n = 15, UK sample) considered the effects of the condition on their life as 'large' (representing a score 4 out of 5) (Figure 8.3). The majority of patients in both samples had seen a doctor before in relation to their sweating (94%, USA; 97%, UK) for their sweating (Table 8.3). 43% of those from the USA and 47% of the UK sample had received treatment for their sweating within the last six months while; 36% (USA) and 39% (UK) were currently receiving treatment. Patients on average spent 121.9 minutes (USA) and 42.6 minutes (UK) per day in managing their condition and they incurred an additional monthly expenditure of £ 55 GBP and £ 22 GBP, respectively (Table 8.4). The most prevalent range of sum of money patients are willing to pay for a complete cure was 50 - 99 GBP for USA patients (n = 37) and 1- 49 GBP (N = 12) for the UK patients (Figure 8.4).

Items of the HidroQoL receiving affirmation

An item was affirmed if an answer other than *No, not at all* was chosen and; was considered missing if patients did not provide a response. Item mean scores are presented in Figure 8.5. The item 'thinking about sweating' received the highest affirmation from US patients (99%), while the 'choice of clothing' item was the most affirmed in the UK sample (97%) (Table 8.5). In both groups the item 'my sex life is affected' was the least affirmed, suggesting that this issue might not be important for most patients. Otherwise, the sensitivity associated with issues pertaining to sex cannot be ruled out as an influence. No data were missing.

Table 8.2: Sociodemographic characteristics of the patients

	USA Sample (n = 127)	UK Sample (n = 36)
Gender, n (%)		
Male	20 (16%)	10 (28%)
Female	107 (84%)	26 (72%)
Age, years		
Mean, SD	38.8 (14.1)	42.8
Median	37	39.5
Mode, n	25	24
Range	17 – 73	20 - 74
Age (years), n		
≤29	37	6
30 to 39	34	12
40 to 49	28	7
50 to 59	14	2
≥ 60	14	9
Body site involved, n (%)		
Generalised	28 (22%)	10 (28%)
Palms, feet & axilla	48 (38%)	6 (17%)
Palms and feet	26 (20%)	5 (14%)
Head, Face	14 (11%)	7 (19%)
Axilla	8 (6%)	4 (11%)
Palms	2 (2%)	2 (6%)
Feet	1 (1%)	1 (3%)
Other		1 (3%)
Disease severity (HDSS score), n		
1	0	0
2	25	11
3	57	12
4	45	13
Global impact of hyperhidrosis (GQ score), n		
No effect	0	2
Small effect	9	5
Moderate effect	34	7
Large effect	46	15
Extremely large effect	38	7

Table 8.2 (continued)

	USA Sample (n = 127)	UK Sample (n = 36)
Co-morbidities, n (%)		
Thyroid disorders	10 (8%)	2 (6%)
Psychiatric of neurologic disorders	21(17%)	4(11%)
Menopausal related complaints	13 (10%)	1(3%)
Diabetes	10 (8%)	1(3%)
Hypertension	20 (16%)	5 (14%)
Other	22 (18%)	10 (28%)
Employment, n (%)		
Employed	84 (66%)	22(61%)
Unemployed	18 (14%)	3 (8%)
Retired	10 (8%)	9 (25%)
Full time student	15 (12%)	2 (6%)

Table 8.3: Patients' treatment history

Characteristic	USA	UK sample
Seen doctor in relation to sweating	119 (94%)	35 (97%)
Received treatment in last 6 months	55 (43%)	17 (47%)
Have received surgical treatment before	18 (14%)	6 (17%)
Received Botox Injection in last 6 months	17 (13%)	4 (11%)
Currently receiving Treatment	46 (36%)	14 (39%)
Current treatments, n (% of sample)		
Aluminium Chloride (topical cream)	14 (11%)	2 (6%)
Systemic oral medication	26 (20%)	8 (22%)
Iontopheresis	9 (7%)	10 (19%)
Cosmetic preparations	16 (13%)	5 (14%)

Scores of the HidroQoL

The mean HidroQoL total score was 25.64 (±6.95) for the USA and 26.96 (±7.52) for the UK sample (Figure 8.5). The mean scale item scores are presented in Figure 8.6. The range for the HidroQoL total score was 2 to 36, in the USA group and 1 to 33, in the UK group. Five patients (4%) in the USA group and none of the patients in the UK sample achieved the maximum HidroQoL overall scale (36).

Part I: Comparison of HidroQoL scores along socio-demographic characteristics of patients

The first step in evaluating the validity of the HidroQoL involved making comparisons of the scores of the HidroQoL across important patient characteristics.

Gender

Mann Whitney (MW) test was used to compare HidroQoL scores across males and female patients. There were no significant gender differences in HidroQoL score ($p = 0.53$) in the US sample (Table 8.7) Analysis of the domain scores also indicated non-significant gender differences for the *impact on daily life activities* domain (H-DA) ($p = 0.08$) as well as the *psychosocial impact* domain (H-PS) ($p = 0.808$).

Table 8.4: Patient's level of disease burden: time[a] and money[b] spent in managing the condition and willingness to pay[c]

	USA Sample	UK Sample
Time spent in managing the sweating, Minutes		
Mean	121.9	42.64
Median	45	37.5
Mode	60	60
Range	0 - 1440	0 - 120
IQR	15 - 90	20-60
Money spent in managing the sweating per month, GBP (£)		
Mean	55.4	22
Median	25	15.5
Mode	0	0
Range	0 – 1000	0-100
IQR	0 - 50	0-30
Willingness To Pay (GBP - £), n		
£0	4	6
£1 - £49	36	12
£50 - £99	37	9
£100 - £199	19	2
£200 - £299	6	1
≥ £300	25	6

a. For example extra time spent on personal hygiene and treatment due to disease.
b. For example extra money spent on personal hygiene, treatment or new clothes.
c. Willingness to pay for a treatment that would provide complete cure of the hyperhidrosis in British Pounds (£).

Figure 8.1: Patients' age distribution

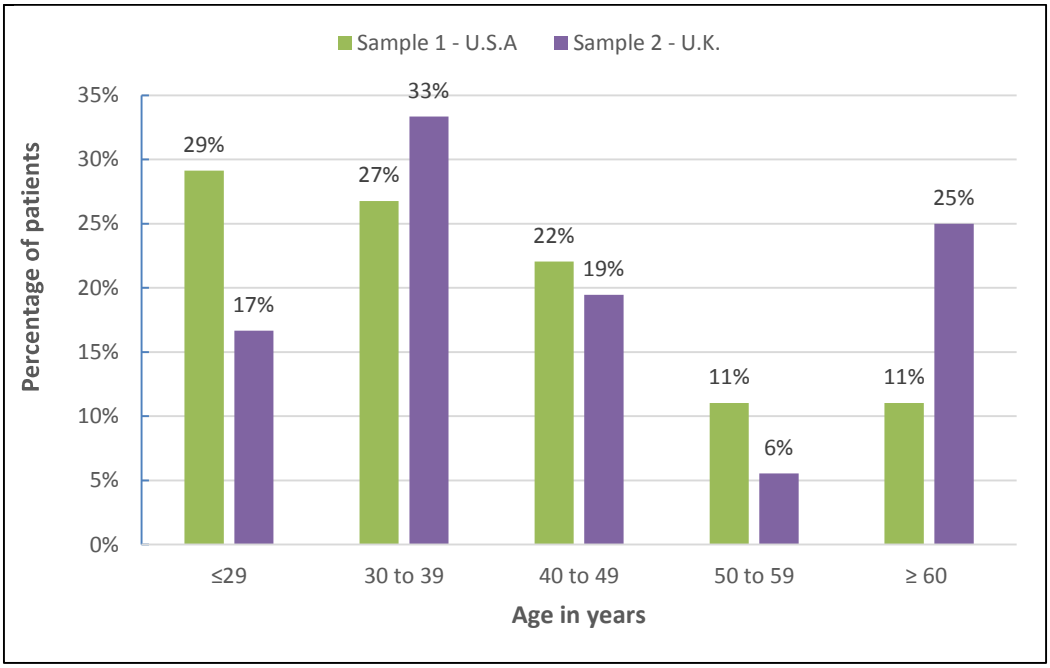


Figure 8.2: Patients' disease severity based on HDSS score

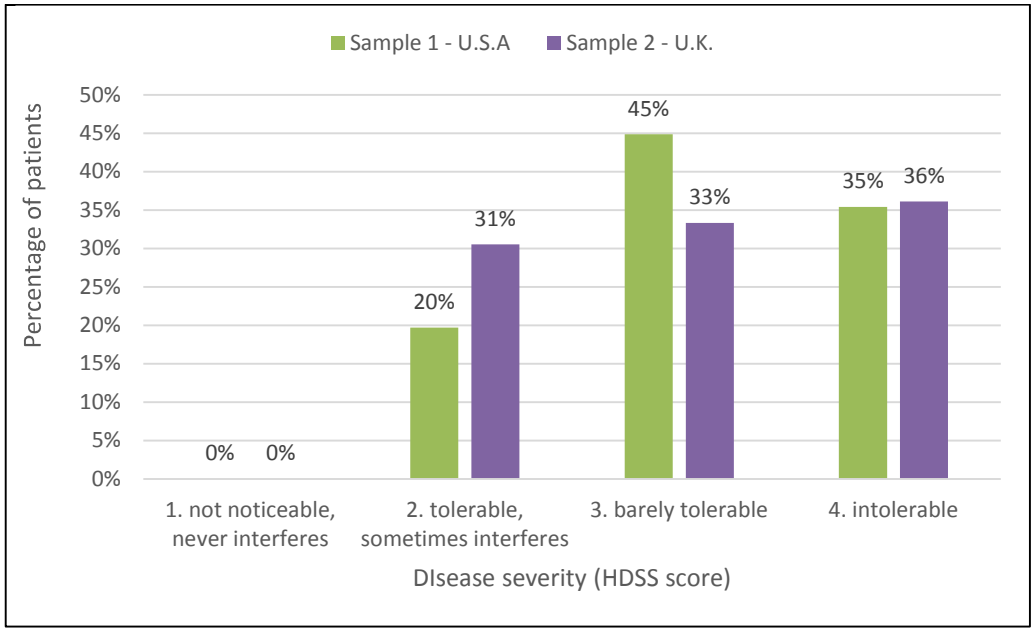


Figure 8.3: Overall health related quality of life impairment

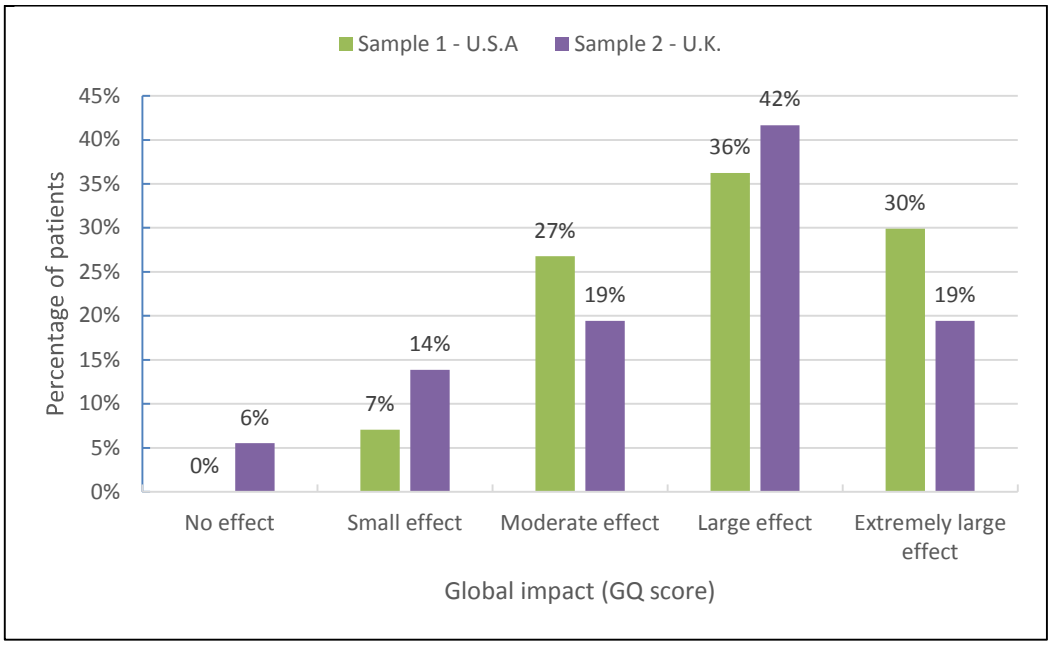


Figure 8.4: Amount of money patients are willing to pay (WTP) for a permanent cure for their condition.

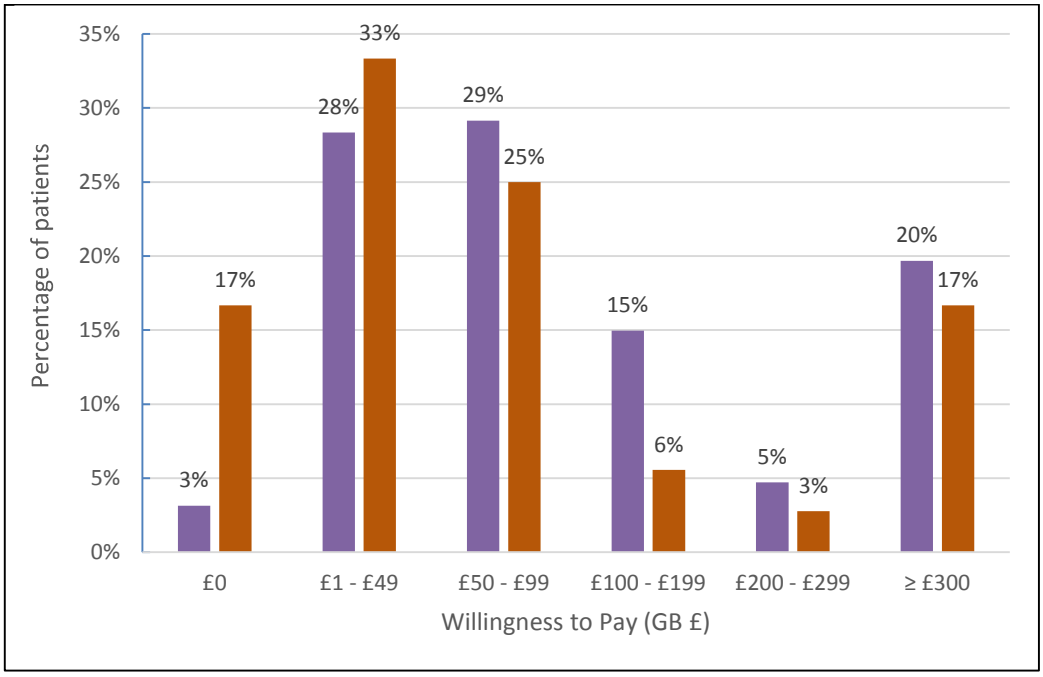


Table 8.5: Items of the HidroQoL receiving affirmation and missing responses

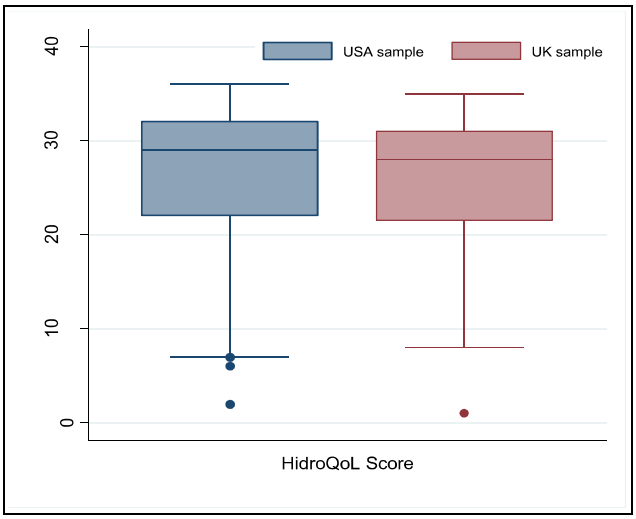
Item	USA Sample		UK Sample	
	Affirmative responses	*Missing*	*Affirmative responses*	*Missing*
Impact on daily life activities				
My choice of clothing is affected	119(94%)	0 (%)	35(97%)	0(0%)
My physical activities are affected	113(89%)	0 (%)	32(89%)	0(0%)
My hobbies are affected	119(94%)	0 (%)	30(83%)	0(0%)
My work is affected	119(94%)	0 (%)	28(78%)	0(0%)
I worry about the additional activities in dealing with my condition	118(93%)	0 (%)	32(89%)	0(0%)
My holidays are affected (e.g. planning, activities)	106(83%)	0 (%)	32(89%)	0(0%)
Impact on psychosocial impact				
I feel nervous	117(92%)	0 (%)	31(86%)	0(0%)
I feel embarrassed	123(97%)	0 (%)	35(97%)	0(0%)
I feel frustrated	124(98%)	0 (%)	35(97%)	0(0%)
I feel uncomfortable physically expressing affection (e.g. hugging)	116(91%)	0 (%)	32(89%)	0(0%)
I think about sweating	126(99%)	0 (%)	34(94%)	0(0%)
I worry about my future health	91(72%)	0 (%)	23(64%)	0(0%)
I worry about people's reactions	123(97%)	0 (%)	34(94%)	0(0%)
I worry about leaving sweat marks on things	120(94%)	0 (%)	33(92%)	0(0%)
I avoid meeting new people	90(71%)	0 (%)	22(61%)	0(0%)
I avoid public speaking (e.g. presentations)	101(80%)	0 (%)	26(72%)	0(0%)
My appearance is affected	109(86%)	0 (%)	30(83%)	0(0%)
My sex life is affected	79(62%)	0 (%)	18(50%)	0(0%)

At the individual level, one item (my hobbies are affected) showed significant gender differences ($p < 0.01$, Females, Median score = 2; Male, Median score = 1). In the UK group, the total HidroQoL score also showed non-significant differences between males and females ($p = 0.31$) (Table 8.8). Comparisons involving the two domains also showed non-significant differences (H-DA, $p = 0.91$, H-PS, $p = 0.16$).

Table 8.6: Frequency of the HidroQoL Scores

HidroQoL total score	USA Sample			UK Sample		
	Freq.	*%*	*Cum.*	*Freq.*	*%*	*Cum.*
1				1	2.78	2.78
2	1	0.79	0.79			
6	1	0.79	1.57			
7	1	0.79	2.36			
8				1	2.78	5.56
12	1	0.79	3.15			
14	2	1.57	4.72			
15	4	3.15	7.87			
16	1	0.79	8.66	1	2.78	8.33
17	3	2.36	11.02	1	2.78	11.11
18	1	0.79	11.81			
19	3	2.36	14.17	1	2.78	13.89
20	4	3.15	17.32	3	8.33	22.22
21	6	4.72	22.05	1	2.78	25
22	5	3.94	25.98	3	8.33	33.33
23	5	3.94	29.92	1	2.78	36.11
24	3	2.36	32.28	2	5.56	41.67
25	1	0.79	33.07	2	5.56	47.22
26	7	5.51	38.58			
27	7	5.51	44.09	1	2.78	50
28	6	4.72	48.82			
29	7	5.51	54.33	3	8.33	58.33
30	8	6.3	60.63	2	5.56	63.89
31	11	8.66	69.29	5	13.89	77.78
32	9	7.09	76.38	3	8.33	86.11
33	9	7.09	83.46	3	8.33	94.44
34	9	7.09	90.55			
35	7	5.51	96.06	2	5.56	100
36	5	3.94	100			
Total	**127**	**100**		**36**	**100**	

Figure 8.5: Distribution of the HidroQoL total Scores using box and whisker plot (USA, n = 127; UK, n = 36).



Only two items showed significant gender differences (*I avoid public speaking e.g. presentations,* median score: M = 1, F = 2; *my appearance is affected,* median score: M = 1, F = 2). This indicates that the effects of hyperhidrosis patients are largely similar between males and females, although females seemed to suffer greater impairment in those aspects related to the public; reflecting women's greater concern for their 'looks' especially in social life.

Age

HidroQoL scores of patients belonging to different age groups were compared, using the Kruskall Wallis (KW) test. Patients were divided into five age groups: 17 to 29; 30 to 39; 40 to 49; 50 to 59 and; above 60. No statistically significant differences in the overall HidroQoL score across the age-groups were seen in the USA group (p = 0.7). The median HidroQoL score of the 17 to 29 age-group was the highest (30) while that of the 40-49 group, was the lowest (26) (Table 8.9). A comparison of the scores for the daily life activities and psychosocial domain scores also turned no significant differences (H-DA, p = 0.95, H-PS, p = 0.55). None of the individual items showed a statistically significant difference across the age-groups. Similar analysis in the UK sample also showed non-significant age-group differences in the overall HidroQoL score (p = 0.98) and in the domain scores (H-DA, p = 0.75, H-PS, p = 0.96) (Table 8.10). At the individual item level, only one item (*I feel nervous*) showed statistically significant differences across the age groups. The age groups 17 – 29 and 30 to 39 had the highest item median score (2). The relationship between age and the HidroQoL score was further explored using regression analysis of HidroQoL total

score on age (measured in years) and by drawing a scatter plot of the two variables. This analysis was carried out on the sample only, due to the small size of the UK sample The results support the finding obtained earlier, there was no significant relationship between age and HidroQoL scores (model F-statistic = 2.9, p = 0.09, B-coefficient = 0.008) (Table 8.11). The scatter plot showed no clear pattern of association (Figure 8.7)

Figure 8.6: Mean scores for the HidroQOL's individual items

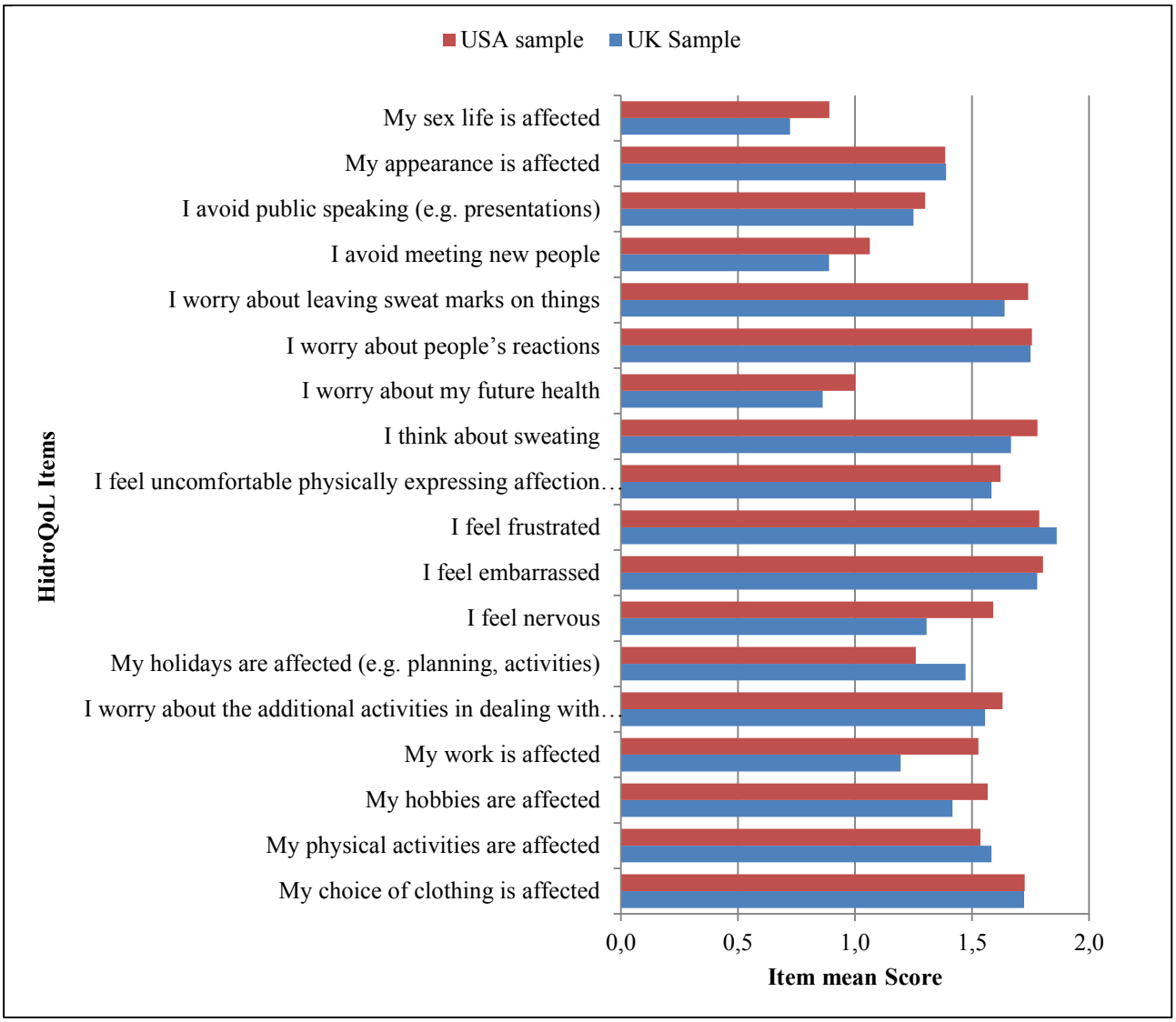


Table 8.7: Comparison of HidroQoL scores by patient's gender (USA sample)

Item	Median		MW	
	Male (n=20)	*Female (n= 107)*	*z-score*	*p-value*
Impact on daily life activities				
My choice of clothing is affected	2	2	-0.93	0.35
My physical activities are affected	2	2	-1.61	0.108
My hobbies are affected	1	2	-2.64	0.008
My work is affected	2	2	-0.32	0.75
I worry about the additional activities in dealing with my condition	2	2	-0.76	0.45
My holidays are affected (e.g. planning, activities)	1	1	-1.40	0.16
Psychosocial impact				
I feel nervous	2	2	-0.38	0.7
I feel embarrassed	2	2	-1.74	0.082
I feel frustrated	2	2	-0.08	0.938
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	0.84	0.401
I think about sweating	2	2	0.76	0.45
I worry about my future health	1	1	-0.97	0.334
I worry about people's reactions	2	2	-0.97	0.426
I worry about leaving sweat marks on things	2	2	0.07	0.947
I avoid meeting new people	1	1	1.42	0.156
I avoid public speaking (e.g. presentations)	2	1	1.18	0.239
My appearance is affected	2	2	-0.26	0.797
My sex life is affected	1	1	0.74	0.459
Impact on daily life activities	**9**	**10**	**-1.75**	**0.08**
Psychosocial impact	**19**	**19**	**0.25**	**0.803**
Overall scale	**27**	**29**	**-0.63**	**0.529**

Table 8.8: Comparison of HidroQoL scores by patient's gender (UK sample)

Item	Median		MW	
	Male (n=10)	*Female (n= 26)*	*z-score*	*p-value*
Impact on daily life activities				
My choice of clothing is affected	2	2	-1.17	0.24
My physical activities are affected	2	2	-0.39	0.69
My hobbies are affected	2	2	0.3	0.76
My work is affected	1	1	-0.13	0.89
I worry about the additional activities in dealing with my condition	2	2	1.19	0.23
My holidays are affected (e.g. planning, activities)	1.5	2	-0.8	0.42
Psychosocial impact				
I feel nervous	1	1	-0.5	0.62
I feel embarrassed	2	2	-1.8	0.07
I feel frustrated	2	2	-0.97	0.33
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	-0.87	0.38
I think about sweating	2	2	-1.04	0.3
I worry about my future health	1	1	0.57	0.57
I worry about people's reactions	2	2	0.05	0.96
I worry about leaving sweat marks on things	1.5	2	-1.69	0.09
I avoid meeting new people	1	1	-1.2	0.23
I avoid public speaking (e.g. presentations)	1	2	-1.73	0.04
My appearance is affected	1	2	-2.52	0.01
My sex life is affected	1	0	0.81	0.42
Impact on daily life activities	**9.5**	**9**	**-0.11**	**0.91**
Psychosocial impact	**16**	**18.5**	**-1.42**	**0.16**
Overall scale	**24**	**29**	**-1.01**	**0.31**

Notes: z, z-score; p, p-value

Table 8.9: Comparison HidroQoL scores by patient's age (USA sample)

	Median					KW Test
	Group 1 (n = 37)	*Group 2 (n = 34)*	*Group 3 (n = 28)*	*Group 4 (n = 14)*	*Group 5 (n = 14)*	*p-value*
Impact on daily life activities						
My choice of clothing is affected	2	2	2	2	2	0.17
My physical activities are affected	2	2	2	2	2	0.53
My hobbies are affected	2	2	2	2	1.5	0.36
My work is affected	2	1	2	2	2	0.62
I worry about the additional activities in dealing with my condition	2	2	2	2	1.5	0.1
My holidays are affected (e.g. planning, activities)	1	1	1	1	2	0.5
Psychosocial Impact						
I feel nervous	2	2	2	2	2	0.39
I feel embarrassed	2	2	2	2	2	0.6
I feel frustrated	2	2	2	2	2	0.45
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	2	2	2	0.23
I think about sweating	2	2	2	2	2	0.84
I worry about my future health	1	1	1	1	1	0.11
I worry about people's reactions	2	2	2	2	2	0.63
I worry about leaving sweat marks on things	2	2	2	2	2	0.11
I avoid meeting new people	1	1	1	1	1	0.99
I avoid public speaking (e.g. presentations)	2	1	2	1	1	0.68
My appearance is affected	1	2	1	1.5	2	0.28
My sex life is affected	1	1	1	1	0.5	0.91
Impact on daily life activities	10	9	9	10	10	**0.95**
Psychosocial impact	19	19	17.5	18.5	18.5	**0.55**
Overall scale	30	29	26	27.5	28	**0.7**

Table 8.10: Comparison of individual item and total HidroQoL scores by patient's age (UK Sample)

Item	Median					KW
	Group 1 (n = 6)	*Group 2 (n= 12)*	*Group 3 (n= 7)*	*Group 4 (n= 2)*	*Group5 (n= 9)*	*p-value*
Impact on daily life activities						
My choice of clothing is affected	2	2	2	2	2	0.69
My physical activities are affected	2	2	2	2	2	0.4
My hobbies are affected	2	2	2	1	1	0.28
My work is affected	2	2	1	2	0	0.21
I worry about the additional activities in dealing with my condition	2	2	2	2	2	0.67
My holidays are affected (e.g. planning, activities)	2	2	2	2	2	1
Psychosocial impact						
I feel nervous	2	2	1	1	1	0.04
I feel embarrassed	2	2	2	2	2	0.45
I feel frustrated	2	2	2	2	2	0.5
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	1	2	2	0.17
I think about sweating	2	2	2	2	2	0.38
I worry about my future health	1	1	1	2	1	0.65
I worry about people's reactions	2	2	2	2	2	0.17
I worry about leaving sweat marks on things	2	2	2	2	1	0.08
I avoid meeting new people	1	1	1	1	0	0.78
I avoid public speaking (e.g. presentations)	1	2	2	1	2	0.95
My appearance is affected	1	1	1	2	2	0.17
My sex life is affected	0	1	1	0	0	0.31
Impact on daily life activities	**11**	**9.5**	**9**	**9**	**9**	**0.75**
Psychosocial impact	**17.5**	**19**	**18**	**18**	**16**	**0.96**
Overall scale	**29**	**30**	**27**	**27**	**24**	**0.98**

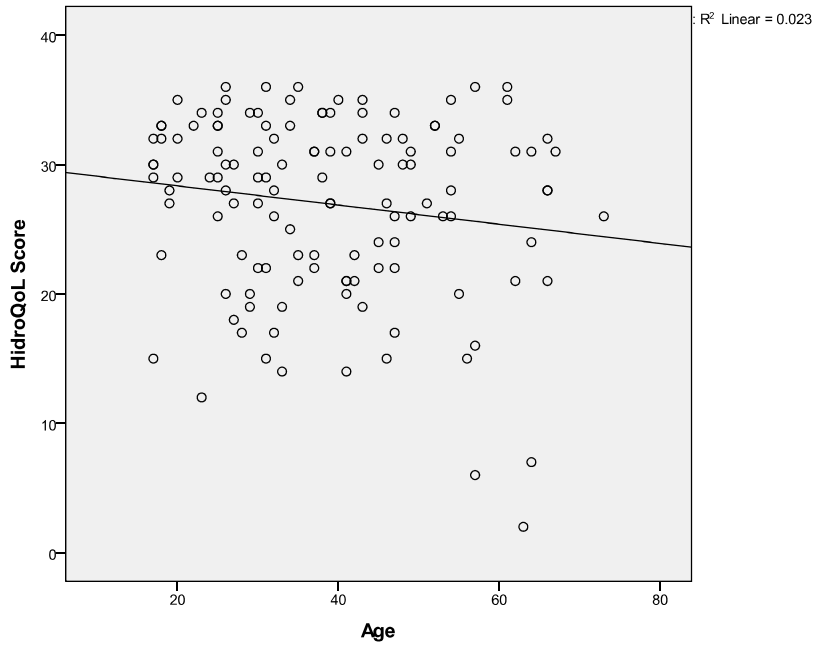
Table 8.11: Univariate regression analyses of HidroQoL score against patient's age

Ind. Var	N	F(df)	p	R2	Adj R2	Dep Var	ß	SE	t	p	Beta
HidroQoL Score	127	2.9 (1, 125)	0.091	0.023	0.015	Age	-0.074	0.044	-1.7	0.1	-0.151
						C	29.836	1.796	16.61	0	

Note:

M: model; **Ind. Var.**: Independent Variable; **p:** p-value**; Adj. R2**: Adjusted r-squared; **Dep. Var**.: dependent variable; **ß:** Coefficient for dependent variable; **DMT**: Daily time spent dealing with hyperhidrosis; **AME**: Additional monthly expenditures due to hyperhidrosis; ***C*:** constant.

Figure 8.7: Scatter plot showing relationship between HidroQoL Score and age.



HDSS score

A comparison of HidroQoL scores across patients with different levels of disease severity, according to the HDSS score, was carried out using the KW test. Patients were grouped according to their HDSS score: group 1, HDSS = 2; group 2, HDSS = 3; group 3, HDSS = 4. Patient with HDSS = 1 were excluded from the study. In the USA sample statistically significant differences ($p < 0.001$) were observed in the HidroQoL overall score across the disease-severity groups (Table 8.12). The median overall score increased with HDSS score from 21 (group 1), 27 (group 2) to 33 (group 3). Domain scores also showed statistically significant differences ($p < 0.001$ for both *impact on daily life activities* and *psychosocial impact* domains). The item scores also showed statistically significant differences across the severity groups ($p < 0.01$).

Table 8.12: Comparison of individual items and total HidroQoL scores by HDSS score (level of disease severity): USA sample

HidroQoL Score	Median			KW
	HDSS 2 (n = 25)	*HDSS 3 (n = 57)*	*HDSS 4 (n = 45)*	*p*
Impact on daily life activities				
My choice of clothing is affected	1	2	2	0
My physical activities are affected	1	2	2	0
My hobbies are affected	1	2	2	0
My work is affected	1	2	2	0
I worry about the additional activities in dealing with my condition	1	2	2	0
My holidays are affected (e.g. planning, activities)	1	1	2	0
Psychosocial impact				
I feel nervous	1	2	2	0
I feel embarrassed	2	2	2	0
I feel frustrated	2	2	2	0
I feel uncomfortable physically expressing affection (e.g. hugging)	1	2	2	0
I think about sweating	1	2	2	0
I worry about my future health	1	1	1	0
I worry about people's reactions	2	2	2	0
I worry about leaving sweat marks on things	2	2	2	0.01
I avoid meeting new people	0	1	2	0
I avoid public speaking (e.g. presentations)	1	1	2	0
My appearance is affected	1	1	2	0
My sex life is affected	1	1	1	0.01
Impact on daily life activities	**8**	**9**	**12**	**0**
Psychosocial impact	**14**	**18**	**22**	**0**
Overall scale	**21**	**27**	**33**	**0**

Similar analyses were carried out on the UK sample. Overall HidroQoL score and the scores for the two domains showed statistically significant differences across the disease severity groups ($p < 0.01$ for all scores) (Table 8.13). Further, significant differences were seen in six out the eighteen items of the HidroQoL including *my holidays are affected (e.g. planning, activities)* ($p = 0.02$), *my work is affected* ($p = 0.01$), *I worry about the additional activities in dealing with my condition* ($p = 0.03$), *I feel frustrated* ($p = 0.01$), *I think about sweating* ($p < 0.001$) and *I avoid meeting new*

people ($p < 0.001$). These results show that the HidroQoL is capable of distinguishing between patients experiencing different levels of self-reported disease severity, based on the HDSS scores. Patients with higher disease severity showed greater impairment in HRQoL.

Table 8.13: Comparison of individual items and total HidroQoL scores by HDSS score (level of disease severity) (UK sample)

	Median			KW p-value
	HDSS 2 (n=11)	*HDSS 3 (n=12)*	*HDSS4 (n=13)*	
Impact on daily life activities				
My choice of clothing is affected	2	2	2	0.45
My physical activities are affected	1	2	2	0.02
My hobbies are affected	1	2	2	0.13
My work is affected	1	1	2	0.01
I worry about the additional activities in dealing with my condition	1	2	2	0.03
My holidays are affected (e.g. planning, activities)	1	2	2	0.1
Psychosocial impact				
I feel nervous	1	1	2	0.1
I feel embarrassed	2	2	2	0.19
I feel frustrated	2	2	2	0.01
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	2	0.68
I think about sweating	1	2	2	0
I worry about my future health	0	1	1	0.09
I worry about people's reactions	2	2	2	0.08
I worry about leaving sweat marks on things	2	2	2	0.08
I avoid meeting new people	0	1	1	0
I avoid public speaking (e.g. presentations)	1	2	2	0.12
My appearance is affected	1	2	2	0.17
My sex life is affected	0	1	1	0.38
Impact on daily life activities	**7**	**9**	**12**	**0.01**
Psychosocial impact	**13**	**17**	**20**	**0**
Overall scale	**21**	**26**	**31**	**0**

Site of Hyperhidrosis

The site of hyperhidrosis varied across patients, for instance, the armpits, the feet, the palms and the head. HidroQoL scores of patients with different sites of hyperhidrosis were compared using the KW test. As there were numerous combinations of affected sites among the patients. Five groups were created as follows: generalised sweating (group 1); palms, feet and axillary (group 2); palms or/and feet (group 3); face or/and head (group 4) and; axillary (or plus other) (group 5) (Table 8.14). In the USA sample, significant differences were observed in the total HidroQoL score across the different groups of patients based on body site ($p < 0.01$). Patients with generalised sweating (group 1) suffered the greatest impairment (median overall score = 32); while those with hyperhidrosis affecting the *palms and feet* (group 3) had the lowest impairment (median overall score = 26). Significant differences were also observed in the domain scores ($p < 0.01$, for both *impact on daily life activities* domain and the *psychosocial impact* domain). An analysis at the individual item level showed significant differences in twelve items. The remaining six showing non-significant differences included *my hobbies are affected*, *my work is affected*, *I feel nervous*, *I feel embarrassed*, *I feel frustrated*, *I worry about leaving sweat marks on things*. This suggests that aspects of psychosocial impairment resulting from hyperhidrosis might be common across patients with different types of hyperhidrosis.

Similar analyses were carried out on the UK sample. The KW test showed non-significant differences in the overall HidroQoL score and the scores for the two domain scores (overall HidroQoL score, $p = 0.13$; H-DA, $p = 0.17$; H-PS, $p = 0.2$) (Table 8.15). At the individual item level, two items, *my physical activities are affected* and *my appearance is affected* showed significant differences.

Global impact Score

HidroQoL scores of patients experiencing different levels of overall impact were compared using the KW test. Patients were divided into four groups according to their level of overall impact (based on their GQ score): Group 1, GQ=1, small effect; Group 2, GQ=3, moderate effect; Group 3, GQ=4, large effect; Group 4, GQ=5, extremely large effect. No patient in either the US or the UK samples reported GQ = 1 (no effect at all). The overall scale score showed statistically significantly differences across the four patient groups ($p < 0.001$) (Table 8.16).

Table 8.14: Comparison of individual items and total HidroQoL scores by site of hyperhidrosis (USA sample)

	Median					KW
	Group1 (n=28)	*Group2 (n= 48)*	*Group3 (n=29)*	*Group 4 (n=14)*	*Group 5 (n=8)*	*p*
Impact on daily life activities						
My choice of clothing is affected	2	2	1	2	2	0
My physical activities are affected	2	2	1	2	2	0.01
My hobbies are affected	2	2	2	2	1	0.15
My work is affected	2	2	2	2	2	0.27
I worry about the additional activities in dealing with my condition	2	2	2	2	2	0.03
My holidays are affected (e.g. planning, activities)	2	1	1	2	2	0
Psychosocial impact						
I feel nervous	2	2	2	2	2	0.08
I feel embarrassed	2	2	2	2	2	0.07
I feel frustrated	2	2	2	2	2	0.2
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	2	2	2	0.1
I think about sweating	2	2	2	2	2	0.03
I worry about my future health	1	1	1	1	2	0.04
I worry about people's reactions	2	2	2	2	2	0.49
I worry about leaving sweat marks on things	2	2	2	2	2	0.02
I avoid meeting new people	1	1	1	0	1	0.01
I avoid public speaking (e.g. presentations)	2	2	1	1	2	0.03
My appearance is affected	2	1	1	2	2	0
My sex life is affected	1	1	1	0	1	0
Impact on daily life activities	**12**	**10**	**9**	**11**	**10**	**0**
Psychosocial impact	**22**	**18**	**18**	**17**	**19**	**0**
Overall scale	**32**	**27**	**26**	**27**	**30**	**0**

Note: Group 1: generalised sweating; **group 2**: Axillar, palmar and feet; **group 3: palms and/or feet**; **Group 4:** Head or face**; Group 5:** Axillary plus other areas

Table 8.15: Comparison of individual items and total HidroQoL scores by site of hyperhidrosis (UK Sample)

	Median					KW p-value
	Group1 (n=28)	*Group2 (n= 48)*	*Group3 (n=29)*	*Group 4 (n=14)*	*Group 5 (n=8)*	
Impact on daily life activities						
My choice of clothing is affected	2	2	2	2	2	0.67
My physical activities are affected	2	2	1	2	2	0
My hobbies are affected	2	2	2	1	2	0.49
My work is affected	1	2	1	1	2	0.43
I worry about the additional activities in dealing with my condition	2	2	1	2	2	0.19
My holidays are affected (e.g. planning, activities)	2	2	1	2	2	0.58
Psychosocial impact						
I feel nervous	1	2	1	1	2	0.16
I feel embarrassed	2	2	2	2	2	0.23
I feel frustrated	2	2	2	2	2	0.12
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	2	2	2	0.42
I think about sweating	2	2	1	2	2	0.04
I worry about my future health	1	1	0	1	1	0.04
I worry about people's reactions	2	2	2	2	2	0.53
I worry about leaving sweat marks on things	2	2	2	2	2	0.85
I avoid meeting new people	1	2	0	0	1	0.28
I avoid public speaking (e.g. presentations)	1	2	1	2	2	0.49
My appearance is affected	2	2	1	2	2	0.03
My sex life is affected	0	2	0	0	1	0.41
Impact on daily life activities	**10**	**9**	**6.5**	**9**	**12**	**0.17**
Psychosocial impact	**18**	**20.5**	**14**	**16**	**19**	**0.2**
Overall scale	**30**	**30**	**21**	**25**	**31**	**0.13**

Table 8.16: Comparison of individual items and total HidroQoL scores by GQ Score (global life impact): USA sample

Items	Median				KW
	Group 2 (n = 9)	*Group 3 (n = 34)*	*Group 4 (n = 46)*	*Group 5 (n = 38)*	p
Impact on daily life activities					
My choice of clothing is affected	1	2	2	2	0
My physical activities are affected	1	1	2	2	0
My hobbies are affected	1	1	2	2	0
My work is affected	1	1	2	2	0
I worry about the additional activities in dealing with my condition	1	2	2	2	0
My holidays are affected (e.g. planning, activities)	1	1	1	2	0
Psychosocial impact					
I feel nervous	1	2	2	2	0
I feel embarrassed	2	2	2	2	0
I feel frustrated	1	2	2	2	0
I feel uncomfortable physically expressing affection (e.g. hugging)	1	2	2	2	0.01
I think about sweating	1	2	2	2	0
I worry about my future health	0	1	1	1	0
I worry about people's reactions	1	2	2	2	0.01
I worry about leaving sweat marks on things	1	2	2	2	0
I avoid meeting new people	0	1	1	2	0
I avoid public speaking (e.g. presentations)	0	1	1	2	0
My appearance is affected	1	1	2	2	0
My sex life is affected	0	1	1	2	0
Impact on daily life activities	**6**	**8**	**9**	**12**	**0**
Psychosocial impact	**11**	**16**	**19**	**22**	**0**
Overall scale	**17**	**23**	**28**	**33**	**0**

The median total score was highest for the group with *extremely large effect* (33) and declined with lower levels of impact; and was lowest for patients experiencing a *small effect* (17). An analysis of the two domain scores also showed significant differences in both ($p < 0.001$, for both H-DA and H-PS). All items also showed statistically significant differences in scores across levels of overall impact. Similar analyses were carried out using the UK sample (Table 8.17).

Table 8.17: Comparison of individual items and total HidroQoL scores by GQ Score (overall HRQoL impact) (UK sample)

	Median		MW test	
	Group1 (n=14)	*Group2 (n=22)*	*z*	*p*
Impact on daily life activities				
My choice of clothing is affected	2	2	-1.25	0.21
My physical activities are affected	2	2	-1.42	0.16
My hobbies are affected	1	2	-1.38	0.17
My work is affected	1	2	-0.85	0.39
I worry about the additional activities in dealing with my condition	2	2	-1.8	0.07
My holidays are affected (e.g. planning, activities)	1	2	-1.68	0.09
Psychosocial impact				
I feel nervous	1	2	-1.79	0.07
I feel embarrassed	2	2	-1.98	0.05
I feel frustrated	2	2	-2.62	0.01
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	0.08	0.94
I think about sweating	1	2	-3.86	0
I worry about my future health	0	1	-2.28	0.02
I worry about people's reactions	2	2	-1.86	0.06
I worry about leaving sweat marks on things	2	2	-1.82	0.07
I avoid meeting new people	0	1	-3.14	0
I avoid public speaking (e.g. presentations)	1	2	-0.98	0.33
My appearance is affected	2	2	-0.91	0.37
My sex life is affected	0	1	-1.66	0.1
Impact on daily life activities	**8**	**9.5**	**-1.64**	**0.1**
Psychosocial impact	**14**	**19**	**-2.8**	**0.01**
Overall scale	**22**	**30**	**-2.47**	**0.01**

Due to the smaller number of patients in this group, participants were divided into two groups: Group 1, GQ score = 1-2, small to moderate effect; and Group 2, GQ = 4 - 5, large to extremely large effect. The overall HidroQoL score showed statistically significant differences between the two groups ($p = 0.01$). As expected, the median score for group 2 (30) was larger than that for group 1 (22). One domain showed significant differences between the two patient groups (H-DA, $p = 0.1$, H-PS, $p = 0.01$). Four items showed significant differences across overall impact groups, including: I feel frustrated, I think about sweating, I worry about my future health and I avoid meeting people. These results demonstrate the ability of the HidroQoL to tap into the overall impact associated with hyperhidrosis.

Willingness to Pay

Patients were asked for their willingness to pay (WTP) for a treatment that would cure their sweating, with the following choices: £ 0; £ 1 - 49; £ 50 - 99; £ 100 - 199; £ 200 - 299; £ 300 or more. HidroQoL scores of patients choosing different WTP values were compared using the KW test. Three groups were formed based on patient's WTP: group 1, £ 0 – 49; group 2, £ 50 - £199; and group 3, £ 200 or more. In the USA sample, median HidroQoL total score of 32 was observed in group 3; 29 in group two and 27 in group 1, reflecting decreases consistent with declining WTP values (Table 8.18). The KW test showed that the scale total score differences were statistically significant ($p < 0.01$). At the domain level, scores for both the *impact on daily life activities* and *psychosocial impact* domains showed statistically significant differences across the WTP-patient groups ($p < 0.05$ for both). Score differences were also explored at the individual item level. Four items showed statistically significant differences across patients based on their WTP group, including: *my choice of clothing is affected* ($p < 0.05$); *my holidays are affected* ($p < 0.05$); *I worry about my future health* ($p < 0.01$); and *I avoid meeting new people* ($p < 0.01$).

Similar analyses were carried out on the UK sample. Due to the small size of the sample two groups were formed: Group 1,WTP = £0-99; and Group 2, WTP = £100 or more. No statistically significant differences in the HidroQoL scores were observed at the scale or domain levels between the two groups (Table 8.19). An analysis of the item scores also showed non-significant differences between WTP patient groups.

Table 8.18: Comparison of individual items and total HidroQoL scores by patient's WTP for complete cure for the sweating (USA sample)

	Median			KW
	Group 1 (n =40)	*Group 2 (n=56)*	*Group 3 (n=31)*	p
Impact on daily life activities				
My choice of clothing is affected	2	2	2	0.02
My physical activities are affected	2	2	2	0.33
My hobbies are affected	2	2	2	0.167
My work is affected	2	2	2	0.088
I worry about the additional activities in dealing with my condition	2	2	2	0.031
My holidays are affected (e.g. planning, activities)	1	1	2	0.036
Psychosocial impact				
I feel nervous	2	2	2	0.668
I feel embarrassed	2	2	2	0.675
I feel frustrated	2	2	2	0.18
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	2	0.907
I think about sweating	2	2	2	0.625
I worry about my future health	1	1	2	0.001
I worry about people's reactions	2	2	2	0.175
I worry about leaving sweat marks on things	2	2	2	0.208
I avoid meeting new people	1	1	2	0.004
I avoid public speaking (e.g. presentations)	1	2	2	0.098
My appearance is affected	1	2	2	0.524
My sex life is affected	1	1	1	0.175
Impact on daily life activities	**9**	**9**	**11**	**0.011**
Psychosocial impact	**18**	**19**	**20**	**0.017**
Overall scale	**27**	**29**	**32**	**0.005**

Table 8.19: Comparison of individual items and total HidroQoL scores by patient's WTP for complete cure for the sweating (UK sample)

Item	Median		MW test	
	Group1 (n=27)	*Group2 (n=9)*	*z*	*p*
Impact on daily life activities				
My choice of clothing is affected	2	2	-0.27	0.79
My physical activities are affected	2	2	0.9	0.37
My hobbies are affected	2	2	-0.93	0.35
My work is affected	1	2	-1.1	0.27
I worry about the additional activities in dealing with my condition	2	2	-0.26	0.79
My holidays are affected (e.g. planning, activities)	2	2	0.48	0.63
Psychosocial impact				
I feel nervous	1	2	-1.12	0.26
I feel embarrassed	2	2	-0.74	0.46
I feel frustrated	2	2	-0.03	0.97
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	0.43	0.67
I think about sweating	2	2	-1.07	0.28
I worry about my future health	1	1	0.86	0.39
I worry about people's reactions	2	2	0.13	0.89
I worry about leaving sweat marks on things	2	2	-0.33	0.74
I avoid meeting new people	1	1	-0.97	0.33
I avoid public speaking (e.g. presentations)	2	2	-0.64	0.52
My appearance is affected	2	2	0.78	0.44
My sex life is affected	0	1	-1.19	0.23
Impact on daily life activities	**9.00**	**12.00**	**-0.34**	**0.74**
Psychosocial impact	**16.00**	**18.00**	**-0.82**	**0.41**
Overall scale	**25**	**29**	**-0.44**	**0.66**

The apparent difference in the relationship between WTP and QoL between the two patient groups, the USA and UK patient groups has numerous explanations. The first relates to the

practical limitations with the UK sample, one group had 9 observations only reducing power for observing any differences. Secondly, the UK and USA healthcare systems differ in the scope and magnitude of out of pocket health expenditures, leading to differences in nominal values patients are willing to spend as well as a different valuation of those nominal monetary figures. For example, an expenditure amounting to a 100 pounds in the UK might be valued differently in the USA, more or less than its equivalent dollar value. Lastly, willingness to pay tends to be influenced by a subject's ability to pay (Drummond et al. 2005)

Daily time spent in managing the condition (DMT)

The patients were asked to report the amount of time they spent in dealing with the condition each day (reported in minutes). Three patient groups were created based on the reported time: 0 to 59 minutes (group 1); 60 – 90 minutes (group 2) and those spending greater than 90 minutes (group 3). Differences in HidroQoL scores across these groups were explored using KW test. Statistically significant differences were observed in the overall score across the three groups ($p < 0.01$), in the US sample (Table 8.20). Group 3 (32) had the largest median overall scale score; while that for group 1 was the smallest (24). Comparisons were also made in scores at the domain level. Significant differences were seen in the scores of both the *impact on daily life activities* (H-DA) and the *psychosocial impact* (H-PS) domains ($p < 0.01$ for both).

Similar comparisons were carried out on the UK sample. Patients were divided into two groups: group 1, DMT below 30 minutes; and group 2, DMT of 30 minutes or greater. The overall score and a single domain showed statistically significant differences (overall score, $p = 0.01$, H-DA, $p = 0.06$, H-PS, $p = 0.00$) (Table 8.21). At the individual item level, five showed differences of statistical significance: I worry about the additional activities in dealing with my condition ($p = 0.03$); I feel nervous ($p = 0.01$); I think about sweating ($p = 0.02$); I avoid meeting new people ($p = 0.01$); My appearance is affected ($p = 0.03$). A univariate regression analysis was also performed for further insights into the nature of relationship between DMT and the HidroQoL score. The analysis was performed in the US sample only, as the UK sample was considered small for this analysis.

Table 8.20: Comparison of individual items and total HidroQoL scores by daily time spent in managing sweating (USA sample).

	Median			KW
	Group 1(n=61)	*Group 2(n=30)*	*Group 3(n=36)*	*p-value*
Impact on daily life activities				
My choice of clothing is affected	2	2	2	0
My physical activities are affected	1	2	2	0
My hobbies are affected	1	2	2	0
My work is affected	1	2	2	0.01
I worry about the additional activities in dealing with my condition	2	2	2	0
My holidays are affected (e.g. planning, activities)	1	1	2	0
Psychosocial impact				
I feel nervous	2	2	2	0.04
I feel embarrassed	2	2	2	0.07
I feel frustrated	2	2	2	0
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	2	0.03
I think about sweating	2	2	2	0.03
I worry about my future health	1	1	1	0
I worry about people's reactions	2	2	2	0.43
I worry about leaving sweat marks on things	2	2	2	0.06
I avoid meeting new people	1	1	1	0.02
I avoid public speaking (e.g. presentations)	1	2	2	0.01
My appearance is affected	1	2	2	0
My sex life is affected	1	0.5	1	0.06
Impact on daily life activities	**9**	**11**	**12**	**0**
Psychosocial impact	**17**	**19.5**	**20.5**	**0**
Overall scale	**24**	**30.5**	**32**	**0**

Table 8.21: Comparison of individual items and total HidroQoL scores by daily time spent in managing sweating (UK Sample).

	Median		MW	
	Group 1 (N = 21)	*Group 2 (N = 15)*	*z*	*p*
Impact on daily life activities				
My choice of clothing is affected	2	2	-0.64	0.52
My physical activities are affected	2	2	-2	0.05
My hobbies are affected	2	1	-1.63	0.1
My work is affected	1	1	-0.602	0.55
I worry about the additional activities in dealing with my condition	2	2	-2.24	0.03
My holidays are affected (e.g. planning, activities)	2	2	-1.77	0.08
Psychosocial impact				
I feel nervous	1	2	-2.73	0.01
I feel embarrassed	2	2	-1.63	0.1
I feel frustrated	2	2	-1.77	0.08
I feel uncomfortable physically expressing affection (e.g. hugging)	2	2	-1.37	0.17
I think about sweating	2	2	-2.38	0.02
I worry about my future health	1	1	-1.34	0.18
I worry about people's reactions	2	2	-1.49	0.14
I worry about leaving sweat marks on things	2	2	-0.37	0.71
I avoid meeting new people	1	1	-2.77	0.01
I avoid public speaking (e.g. presentations)	2	1	-1.61	0.11
My appearance is affected	2	2	-2.15	0.03
My sex life is affected	0	2	-0.59	0.55
Impact on daily life activities	**9**	**9**	**-1.9**	**0.06**
Psychosocial impact	**17**	**20**	**-2.88**	**0**
Overall scale	**27**	**31**	**-2.69**	**0.01**

The regression model was significant (F-statistic = 11.23; p = 0.001). DMT explained 8.2% of variability in the HidroQoL scores. The positive B-coefficient (B = 0.008) indicates a positive

relationship between HidroQoL scores and DMT (Table 8.22). A scatter plot of the two variables shows a wide variation in HidroQOL scores below 200 minutes and a clear pattern indicating a positive relationship above that cut-off (Figure 8.8). The greater HRQoL impairment in patients spending more time in managing their condition reflects the lost utility from the time taken from enjoyable activities that the patients are not able to undertake. It is also possible that time spent managing the condition is associated with a level of discomfort e.g. patients being reminded of their condition. These results suggest that the HidroQoL can distinguish patients according to an important aspect of disease experience and impact.

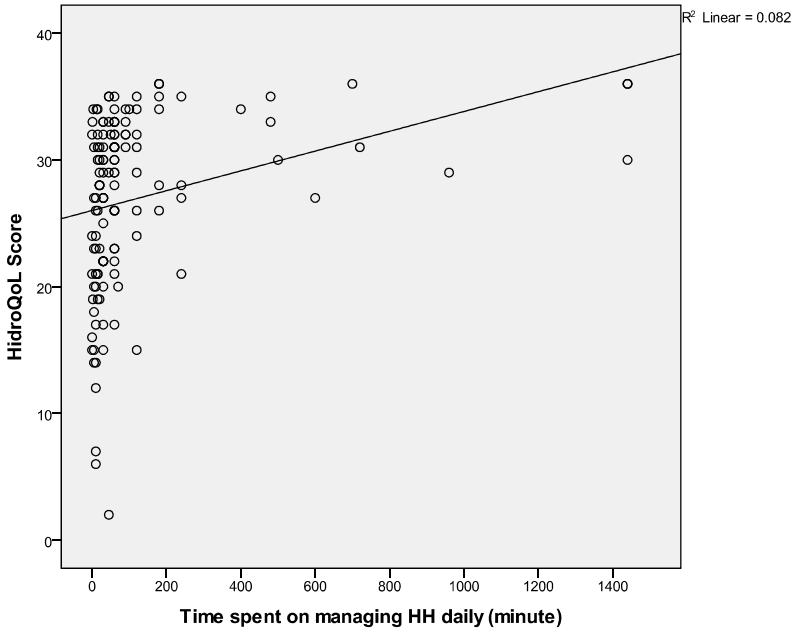
Table 8.22: Univariate regression analyses of HidroQoL score against daily time spent with hyperhidrosis.

M	Ind. Var	N	F(df)	p	R2	Adj R2	Dep Var	ß	SE	t	p	Beta
1	HidroQoL Score	127	11.23 (1, 125)	0.001	0.082	0.075	DMT	0.008	0.002	3.35	0.001	0.287
							C	26.008	0.657	39.57	0	

Note:

M: model; **Ind. Var.**: Independent Variable; **p:** p-value**; Adj. R2**: Adjusted r-squared; **Dep. Var**.: dependent variable; **ß:** Coefficient for dependent variable; **DMT**: Daily time spent dealing with hyperhidrosis; ***C*:** constant.

Figure 8.8: The relationship between the HidroQoL Score and daily time spent in managing the condition



Treatment history

The socio-demographic data collected from patients also included some information related to treatment history, particularly whether patient had been; treated for hyperhidrosis in the last 6 months; was currently receiving treatment; had received surgical therapy for hyperhidrosis previously; had received inter-dermal injection with Botox in the last six months. For each of the aspects patients either affirmed (yes) or rejected the notion (no). Scores of patients belonging to the two groups were compared along each aspect using the Mann-Whitney Test.

Treatment within last 6 months

In the US sample overall HidroQoL score (H-total) showed no statistically significant differences between those treated for hyperhidrosis in the last six months and those who were not ($p = 0.07$) (Median total score: Treated = 31; Not-treated = 28) (Table 8.23). The domain scores also indicated non-significant differences (H-DA, $p = 0.23$; H-PS, $p = 0.09$). In the UK sample, the contrary was observed, the overall score as well as scores for the two domains showed significant differences (total scale, $p = 0.01$, H-DA, $p = 0.04$, H-PS, $p < 0.001$) (Table 8.24). The median overall score was larger for those treated in the last 6 months (30), compared to those who had not been treated (23). Greater QoL impairment might have motivated patients to seek for treatment.

Table 8.23: Comparison of HidroQoL Scores by treatment history: US sample

		Median		MW test	
		Yes	*No*	*z*	*p*
Treated for HH in last 6-months	H-DA	10	9	1.208	0.23
	H-PS	20	19	1.694	0.09
	H-total	31	28	1.92	0.05
Currently receiving treatment for HH	H-DA	9	10	0.11	0.91
	H-PS	20	19	1.21	0.23
	H-total	30	28	0.90	0.37
Treated with inter-dermal Botox injection within the last 6 months	H-DA	11	10	1.15	0.25
	H-PS	20	19	1.19	0.23
	H-total	32	28	1.42	0.16
Have been treated with surgery	H-DA	12	9	2.57	0.01
	H-PS	20	19	1.38	0.17
	H-total	32	28	1.83	0.07

Current treatment

A comparison of those currently being treated and those getting no treatment showed no significant differences in the USA sample, both for the overall scale score ($p = 0.37$) and for the domain scores (H-PS, $p = 0.23$; H-DA, $p = 0.91$). However, the contrary was observed in the UK sample, differences of a statistical significance were observed for the scale score ($p < 0.01$) and also for the two domain scores (H-PS, $p = 0.02$; H-DA, $p < 0.001$).

Surgical intervention

Differences in the HidroQoL scores between patients who had received surgical therapy before and those who had not, were explored. The overall score and the psychosocial domain scores did not show statistically significant differences in the US sample. Significant differences were seen only on the impact on H-DA domain score ($p < 0.01$) only. In the UK sample, on the other hand, the overall score as well as the two domain scores showed non-significant differences.

Inter-dermal injection by Botox

In the US group, comparisons between patients previously treated with Botox injection and those not showed non-significant differences in the overall scale score; and the two domain scores. Similar findings were obtained in the UK group.

Table 8.24: Comparison of HidroQoL Scores by treatment history: UK Sample

		Median		MW test	
		Yes	*No*	*z*	*p*
Treated for HH in last 6-months	H-DA	11.00	9.00	3.02	0
	H-PS	20.00	15.00	2.05	0.04
	H-total	30	23	2.51	0.01
Currently receiving treatment for HH	H-DA	12.00	9.00	3.11	0
	H-PS	20.00	15.00	2.38	0.02
	H-total	31	23	2.75	0.01
Treated with inter-dermal Botox injection within the last 6 months	H-DA	12.00	9.00	1.9	0.06
	H-PS	18.50	16.50	0.88	0.38
	H-total	31	25	1.04	0.3
Have been treated with surgery	H-DA	8.50	9.50	-1.66	0.1
	H-PS	17.00	17.50	-0.66	0.51
	H-total	26	29	-1	0.32

Co-morbidities

Differences in the HidroQoL scores across patient with a co-morbidity and those without one were explored using the MW test. In both the USA and the UK samples, having or not having thyroid disorder; psychiatric or neurologic disorders; menopausal related complaints; diabetes and hypertension showed non-significant differences in the overall scale score and in the two domain scores (Table 8.25, Table 8.26). The *impact on daily life activities* domain score showed significant differences between participants reporting psychiatric or neurologic disorders and those without, in the USA group ($p = 0.03$).

As the diagnosis of primary hyperhidrosis involves ruling out the role of a secondary condition (Solish et al. 2008). it is expected that the measurement of HRQoL impairment resulting from hyperhidrosis should not be influenced by co-morbidities where they are present. Thus the current results are consistent with this notion. Finding significant differences would have meant that either the sample used in this study included patients with hyperhidrosis caused by other primary conditions; or that the measure was picking up something else other than the impacts of HH. Dysfunction of the sympathetic nervous system is known to play a role in hyperhidrosis (Hornberger et al. 2004). The link between psychological disorders such as social anxiety, and hyperhidrosis is quite strong, palmar sweating has been described as a symptom of social anxiety (Ruchinskas 2007), even though studies empirically investigating this matter have not found pathologic levels of psychological problems as social anxiety in HH patients (Weber et al. 2005).

Part II: Convergence and Divergence Validity

Further validation of the HidroQoL scores in assessing quality of life in hyperhidrosis involved developing and testing a number of hypothesis on how the scores of the HidroQoL relates to scores of other established measures of QoL based on the theoretical understanding of how the construct measured by the HidroQoL (QoL impairment in HH) related to the variable being measured by the external instrument. The descriptive score distribution of the each of the measures is presented in Table 8.27.

Table 8.25: Comparison of HidroQoL Scores by patient's co-morbidity (USA sample)

		Median		MW
		Comorbidity	*no-comorb.*	*p-value*
Thyroid disorder	H-DA	10	10	0.71
	H-PS	20	19	0.49
	H-Total	28	29	0.55
Psychiatric or neurologic disorders	H-DA	11	9	0.03
	H-PS	19	19	0.84
	H-Total	31	29	0.36
Menopausal related complaints	H-DA	11	10	0.24
	H-PS	19	19	0.57
	H-Total	31	29	0.33
Diabetes	H-DA	12	10	0.10
	H-PS	19	19	0.35
	H-Total	31	29	0.2
Hypertension	H-DA	11	10	0.50
	H-PS	19	19	0.97
	H-Total	28	29	0.93

Table 8.26: Comparison of HidroQoL Scores by patient's co-morbidity (UK sample)

		Median		MW	
		Comorbidity	*No comorbidity*	*Z-score*	*P-value*
Psychiatric or neurologic disorders	H-DA	10.5	9	1.13	0.26
	H-PS	17	17.5	0.1	0.92
	H-Total	27.5	28	0.63	0.53
Hypertension	H-DA	9.00	10.00	-1.3	0.19
	H-PS	16.00	18.00	-0.28	0.78
	H-Total	24	29	-0.64	0.52

Hypothesis 1: Impairment in hyperhidrosis-QoL has a positive association with disease severity: Patient's HDSS score is positively correlated with the HidroQoL score.

Testing this hypothesis involved assessing the degree and direction of association between the HDSS score and the HidroQoL's overall and domain scores. Spearman's rank sum correlation

analyses were performed between the two measures. The coefficient of the spearman's rank sum correlation showed significant association between the HDSS score and the HidroQoL overall scale score, in the US sample (rho = 0.653, $p < 0.01$) (Table 8.28). The HDSS score also showed correlation with the scores of the two domains (H-DA, rho = 0.655, $p < 0.01$; and H-PS, rho = 0.550, $p < 0.01$). The focus of the HDSS on both severity and on interference in daily life activities, places it closer to the content of the daily life activities domain than the psychosocial domain, explaining the small difference in the magnitude of the correlations. Similar analysis were carried out in the UK sample. The HDSS score was correlated with the HidroQoL total score (rho = 0.564, $p < 0.01$) as well as the two domain scores, daily life activities (rho = 0.518, $p < 0.01$) and psychosocial domain (rho = 0.551, $p < 0.01$) (Table 8.29).

Condition specific-QoL instruments given their attention on issues peculiar to a particular disease condition tend to have a greater connection to clinical outcomes (Salek 1998). In this study, the HidroQoL has demonstrated a strong association with a standard clinical measure in hyperhidrosis, the HDSS. Moreover, it is noteworthy that the strong correlation has been achieved despite the absence of items related to 'symptoms' in the HidroQOL; highlighting the strong relevance of the items as a reflection of impacts arising from the symptoms of hyperhidrosis. Not only is the initial set hypothesis confirmed but these findings also give some preliminary indications on the capabilities of the instrument to detect change in patients over time.

Hypothesis 2: The overall impact of HH on the patient's life is related to their hyperhidrosis-specific QoL

Hypothesis 2 was assessed by estimating a Spearman's rank correlation between HidroQoL scores and patient's GQ scores. The correlation coefficient showed an association between the GQ score and the HidroQoL overall score (rho = 0.610, $p < 0.01$). The two domains also showed a positive and strong correlation (*impact on daily life activities*, rho = 0.595, $p < 0.01$; *psychosocial aspect*, rho = 0.521, $p < 0.01$). Similar results were obtained from the UK sample (overall score, rho = 0.557, $p < 0.01$; *impact on daily life activities*, rho = 0.496, $p < 0.01$ and *psychosocial impact*, rho = 0.574, $p < 0.01$).

Table 8.27: Summary description of HidroQoL, DLQI, Skindex-17, EQ-5D scores.

	HidroQoL		DLQI		Skindex-17		EQ-5D									
							mobility		self-care		Usual-activities		Pain/discomfor		Anxiety/depressio	
	US	*UK*	*US*	*UK*	*US*	*UK*	*US*	*UK*	*US*	*UK*	*US*	*UK*	*US*	*UK*	*US*	*UK*
N	127	36	126	36	126	36	12	36	12	36	12	36	12	36	12	36
Minimum	2.0	1.0	.0	.0	1.0	.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Maximum	36.0	35.0	30.0	24.0	32.0	31.0	4.0	4.0	5.0	4.0	5.0	5.0	5.0	5.0	4.0	4.0
Range	34.0	34.0	30.0	24.0	31.0	31.0	3.0	3.0	4.0	3.0	4.0	4.0	4.0	4.0	3.0	3.0
Mean	27.0	25.6	10.4	9.7	13.9	15.8	1.4	1.7	1.2	1.3	1.8	2.2	1.7	2.1	2.0	2.2
SD	6.9	7.5	7.7	6.3	6.5	8.8	.7	1.1	.6	.7	1.0	1.2	1.0	1.4	1.0	1.0
Median	29.0	28.0	9.0	10.0	14.0	16.5	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	1.0	2.0
25th Percentile	22.0	21.5	4.0	4.0	12.0	9.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0	2.0
75th percentile	32.0	31.0	16.0	13.5	16.0	23.0	1.0	2.5	1.0	1.0	2.0	3.0	2.0	3.0	2.0	2.0
Skewness	-1.0	-1.3	.4	.2	.2	-.3	2.1	1.3	4.2	2.9	1.1	.7	1.3	.9	1.1	.9

The essence of the global question is to capture the overall impact of the disease on the patient's life. Presumably patients would reflect on those aspects of disease they consider to be affected when responding to this question. Thus, assessing the relationship between the score of the GQ provides an opportunity to see whether the HidroQoL is indeed useful as a measure of disease impact in hyperhidrosis. Beyond this, establishing that such a relationship exists implies that the content of the measure is appropriate and has the right emphasis for capturing quality of life impacts in hyperhidrosis. The current results support the construct and the content validity of the HidroQoL.

Hypothesis 3a: Hyperhidrosis-specific QoL is related to dermatology-specific QoL: Patient's HidroQoL score was positively correlated with the DLQI score

Hypothesis 3a was assessed by estimating a Spearman's rank sum correlation between the HidroQoL scores and the DLQI score. In addition, regression analysis was carried out. In the US sample, the correlation of DLQI score with the HidroQoL overall score was 0.572 ($p < 0.01$). The DLQI score had a correlation of 0.517 ($p < 0.01$) with the *impact on daily life activities* domain; and a correlation of 0.505 ($p < 0.01$) with the *psychosocial impact* domain.

Similar results were obtained from the UK sample. The DLQI total score correlated with the HidroQoL overall scale (rho = 0.562, $p < 0.01$) as well as with the two domain scores (*impact on daily life activities*, rho = 0.578, $p < 0.01$ and *psychosocial impact,* rho = 0.530, $p < 0.01$).

The regression analysis was carried out on the US sample only. The UK sample was considered too small for this. The model as a whole was statistically significance (F-statistic = 53.37, $p < 0.001$) indicating a relationship between the HidroQoL and the DLQI scores (Table 8.30). HidroQoL score explained 30% of variability in patients DLQI score. The beta coefficient showed that a 1 SD increase in HidroQoL score, resulted in an increase of 0.55 SD in the DLQI score. This demonstrates that hyperhidrosis-QoL was positively related to dermatology-QoL. This relationship was further illustrated in the scatter-plot between the two variables (**Figure 8.9**).

Hypothesis 3b: Patients HidroQoL scores were positively correlated with the Skindex-17 scores

Hypothesis 3b was tested by assessing the relationship between HidroQoL scores and the Skindex-17 scores using Spearman's rank sum correlation and using regression analysis.

Table 8.28: Multiple correlations between the HidroQoL scores and the Skindex, the DLQI, EQ-5D, general health, the HDSS, patients' WTP, and time spent in the daily management of the sweating (USA Sample)

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
hqdaily (1)	1.000															
hqpsy (2)	.582**	1.000														
hqtotal (3)	.802**	.943**	1.000													
Skpsy (4)	.436**	.573**	.584**	1.000												
Sksym (5)	.173	.289**	.276**	.196*	1.000											
Sktotal (6)	.415**	.536**	.551**	.856**	.588**	1.000										
Dtotal (7)	.517**	.505**	.572**	.758**	.342**	.764**	1.000									
Eqmobility(8)	.097	.011	.043	.063	.228*	.139	.196*	1.000								
Equsual (9)	.299**	.230**	.278**	.207*	.166	.242**	.325**	.496**	1.000							
Eqselfcare(10)	.044	.058	.040	.086	.085	.100	.162	.472**	.340**	1.000						
Eqpain (11)	.271**	.215*	.254**	.234**	.353**	.339**	.378**	.275**	.339**	.255**	1.000					
Eqanxiety(12)	.339**	.363**	.387**	.498**	.298**	.534**	.488**	.175*	.258**	.165	.301**	1.000				
gq (13)	.595**	.521**	.610**	.488**	.253**	.492**	.559**	.131	.303**	.053	.248**	.321**	1.000			
HDSS (14)	.655**	.550**	.653**	.447**	.201*	.445**	.405**	.066	.252**	-.099	.200*	.281**	.676**	1.000		
wtp (15)	.251**	.225*	.257**	.335**	.052	.276**	.301**	.043	.086	-.041	.053	.254**	.278**	.263**	1.000	
minute (16)	.569**	.372**	.474**	.388**	.275**	.434**	.488**	.102	.215*	.157	.295**	.361**	.482**	.386**	.185*	1.000

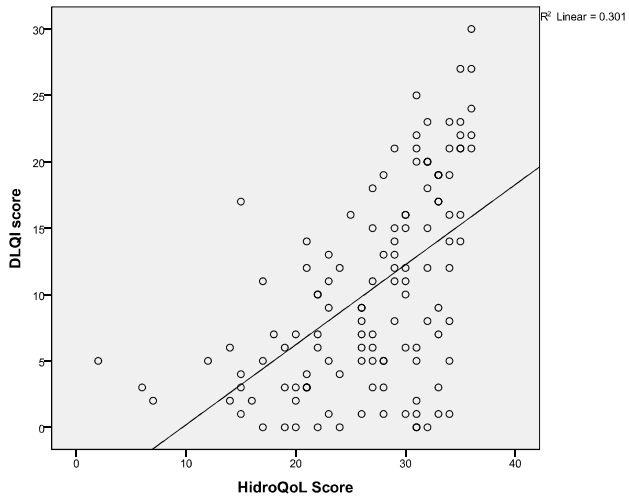
Table 8.29: Multiple correlations between the HidroQoL scores and the Skindex, the DLQI, EQ-5D, general health, the HDSS, patients' WTP, and time spent in the daily management of the sweating (UK Sample)

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
hqdaily (1)	1.000															
hqpsy (2)	.669**	1.000														
hqtotal (3)	.852**	.946**	1.000													
Skpsy (4)	.543**	.742**	.720**	1.000												
Sksym (5)	.378*	.381*	.376*	.407*	1.000											
Sktotal (6)	.559**	.719**	.702**	.938**	.680**	1.000										
Dtotal (7)	.578**	.530**	.562**	.648**	.622**	.754**	1.000									
Eqmobility(8)	-.059	.029	.004	.054	.088	.077	.191	1.000								
Eqselfcare(9)	-.090	.049	-.003	-.040	-.323	-.143	-.014	.430**	1.000							
Equsual (10)	.041	.135	.100	.321	.343*	.349*	.289	.497**	.345*	1.000						
Eqpain (11)	-.042	-.173	-.130	-.096	-.023	-.079	.107	.622**	.502**	.409*	1.000					
Eqanxiety(12)	.465**	.638**	.583**	.570**	.556**	.653**	.631**	.111	0.000	.338*	-.105	1.000				
gq (13)	.496**	.574**	.557**	.570**	.548**	.638**	.784**	.237	.114	.412*	.126	.633**	1.000			
HDSS (14)	.518**	.551**	.564**	.469**	.505**	.537**	.498**	.155	-.014	.324	.115	.598**	.764**	1.000		
wtp (15)	.117	.171	.134	.163	.316	.244	.196	-.168	-.189	.185	-.158	.367*	.303	.321	1.000	
minute (16)	.364*	.440**	.437**	.426**	.202	.399*	.477**	.096	.141	.233	.105	.439**	.532**	.410*	.120	1.000

Table 8.30: Results of univariate regression analyses of DLQI, Skindex, EQ-5D dimensions regressed on HidroQoL score.

M	Ind. Var	F(df)	p	R2	Adj R2	Dep. Var	B	SE	t	p	Beta
1	EQ-Anxiety/ depression	18.56 (1, 125)	0.00	0.13	0.12	H- Score	0.00	0.01	4.31	0.00	0.36
						Constant	0.65	0.32	2.01	0.05	
2	EQ-Mobility	0.42 (1, 125)	0.52	0.00	0.00	H- Score	0.01	0.01	0.65	0.52	0.06
						Constant	1.18	0.26	4.45	0.00	
3	EQ-Self-care	0.19 (1, 125)	0.66	0.00	0.01	H- Score	0.00	0.01	0.44	0.66	0.04
						Constant	1.09	0.21	5.24	0.00	
4	EQ-Pain/ discomfort	(1, 125) 8.65	0.00	0.07	0.06	H- Score	0.04	0.01	2.94	0.00	0.26
						Constant	0.71	0.35	2.05	0.04	
5	EQ-Usual-activities	12.17 (1, 125)	0.00	0.09	0.08	H- Score	0.04	0.01	3.49	0.00	0.30
						Constant	0.67	0.33	2.05	0.04	
6	Skindex-17 Score	35.41(1, 124)	0.00	0.22	0.26	H- Score	0.44	0.07	5.95	0.00	0.47
						Constant	1.95	2.06	0.94	0.35	
7	DLQI score	53.37 (1, 124)	0.00	0.30	0.30	H- Score	0.60	0.83	7.31	0.00	0.55
						Constant	-5.84	2.30	-2.54	0.01	

Figure 8.9: Scatter plot illustrating the relationship between DLQI score and the HidroQoL score



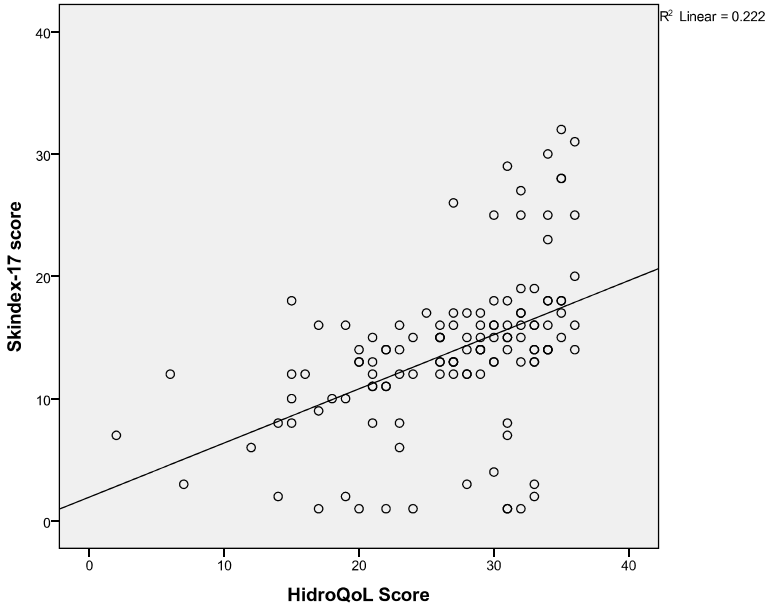
In the US sample, the Skindex total score correlated with HidroQoL total score (rho = 0.551, p < 0.01) (Table 8.28). The Skindex total score also showed correlation with the two HidroQoL domain scores (daily life activities, rho = 0.415, p < 0.01; psychosocial score, rho = 0.536, p < 0.01). Similar analyses in the UK sample showed slightly higher correlations: HidroQoL total score, rho = 0.702 (p < 0.01); daily life activities, rho = 0.559 (p < 0.01) and psychosocial score, rho = 0.719 (p < 0.01) (Table 8.29).

Analyses were also carried out between the Skindex-17's two domains scores and the HidroQoL's scores. In the US sample, the HidroQoL overall score correlated with the Skindex-17 psychosocial scale (rho = 0.584, p < 0.01) as well as the Skindex-17 symptom scale (rho = 0.276, p < 0.01). Correlations of comparable magnitude and direction were similarly observed in the UK sample (Skindex-daily life, rho = 0.720, p < 0.01; Skindex-symptom, rho = 0.376, p < 0.01). The lower correlations between the HidroQoL overall score and the Skindex-symptom, clarifies the focus of the HidroQoL (as a measure of QoL impact and not symptoms). On the other hand, the higher magnitude of the correlation between 'psychosocial scale' of the HidroQoL and the total Skindex score, serves to highlight the strength of the HidroQoL in addressing psychosocial impacts which largely underlie patients dermatology-QoL. Further analysis was carried out by regressing the skindex-17 overall score on HidroQoL overall score, using responses. This analysis was carried out for the US sample only, the UK sample was considered too small. The model as a whole was statistically significant (F-statistic = 35.41, p < 0.001). The HidroQoL score explained 22% of variability in the Skindex-17 scores, which was comparable to the results obtained with the DLQI. The inclusion of symptom related items might have been an influence (Table 8.30, model 6; Figure 8.10). Results on hypothesis 3 demonstrate that the HidroQoL captures aspects of QoL impairment which relate to patient's dermatology-QoL.

The relationship between the EQ-5D scores and the HidroQoL scores

The EQ-5D-5L, is a generic health status measure comprised of a five domain descriptive component and a VAS scale assessing overall health status (Brooks 1996). Each of the five domains are scored on a five point scale (from no-problems to extreme problems/inability to do function).

Figure 8.10: Scatter plot illustrating the relationship between HidroQoL score and Skindex-17 score



The EQ-5D was completed by patients taking part in this study alongside other measures. Two hypotheses were tested based on comparisons of the EQ – 5D domain scores and the HidroQoL scores.

Hypothesis 6a: The HidroQoL scores are not correlated with EQ-5Ds 'mobility' and 'self-care' domain scores.

Spearman's rank sum correlation analysis was used to assess the relationship between the HidroQoL scores and EQ-5Ds 'mobility' and 'self-care' domain scores. In the US sample, the scores for EQ-5Ds 'mobility' domain showed no correlation with the HidroQoL overall score; nor with the two HidroQoL domain scores ($p > 0.05$ in all instances). Similarly, the EQ-5Ds self-care domain score did not correlate with any of the HidroQoL scores, the total as well as the domain scores (Table 8.28).

The relationship between the HidroQoL scores and the EQ-5D was also explored in the UK sample. Scores of the HidroQoL (overall as well as domain scores) again showed no correlation with the EQ-5Ds "self-care" or the "mobility" domain scores (Table 8.29).

Univariate regression analyses of each of the two EQ 5Ds domains, "mobility" and "self-care", with the HidroQoL total score yielded non-significant models (F-statistic = 0.42, $p = 0.52$; F-

statistic = 0.19, p = 0.66) (Table 8.30). This shows an absence of any statistically significant relationship between the two EQ-5D domains and the HidroQoL scores. The hypothesis set initially is therefore confirmed: the HidroQoL's total and domain scores are not related to the 'mobility' and 'self-care' items supporting the divergence validity of the HidroQoL, as the two EQ-5D domains deal with themes that are unrelated to the impacts of HH.

Hypothesis 6b: The HidroQoL scores are correlated with EQ-5D-5L domains on 'usual activities', 'anxiety/depression' and 'pain or discomfort'.

Spearman's rank sum correlation analysis was used to assess the relationship between the HidroQoL scores and domain scores of the EQ-5D-5L related to 'usual activities', 'anxiety/depression' and 'pain or discomfort'. In the US sample the EQ-5D-5L score for 'usual activities' domain correlated with the HidroQoL scores (overall score, rho = 0.278, $p < 0.01$, *impact on daily life activities* domain, rho = 0.299, $p < 0.01$, *psychosocial impact* domain, rho = 0.230, $p < 0.01$). Significant correlations were also observed between the EQ-5Ds 'anxiety/depression' domain scores and HidroQoLs scores (HidroQoL total score: rho = 0.387, $p < 0.01$; impact on daily life activities domain: rho = 0.339, $p < 0.01$ and; *psychosocial impact* domain, rho = 0.363, $p < 0.01$). Similarly, the 'pain/discomfort' EQ-5D domain score correlated with the HidroQoL overall score (rho = 0.254, $p < 0.01$) as well as with the HidroQoL's two domain scores (*impact on daily life activities* domain, rho = 0.271, $p < 0.01$; *psychosocial impact* scale: rho = 0.215, $p < 0.01$). In the UK sample scores for EQ-5D-5L's 'usual activities' and 'pain/discomfort' were uncorrelated with HidroQoL's total score and the scores of the two domains. The 'anxiety/depression' domain showed correlation with the HidroQoL total score (rho = 0.583, $p < 0.01$), the HidroQoL daily life activities domain (rho = 0.465, $p < 0.01$) and HidroQoL psychosocial scale (rho = 0.638, $p < 0.01$). Further analyses employed univariate regression technique. Each of the three EQ-5Ds domain scores (usual activities, pain/discomfort and anxiety/depression) were regressed on HidroQoL total score. Only the regression model involving "anxiety/depression" was significant (F-statistic = 18.56, $p < 0.0001$); the HidroQoL explained 13 % of the variance in the EQ-5D "anxiety/depression" score.

Hypothesis 7: Greater impairment in quality of life is associated with more time spent in managing the condition

Living with a long term condition usually involves patients regularly taking treatment to address either symptoms or impacts of their condition, in addition to other measures to adapt to their condition. Both of these may be time consuming. Thus, the relationship between patient's quality of life (assessed by the HidroQoL score) and daily time spent in managing the condition (measured in minutes) was assessed using Spearman's rank correlation analysis.

In the US sample, the HidroQoL score showed moderate correlation with daily management time (overall scale score, rho = 0.474, $p > 0.01$, *impact on daily life activities* domain, rho = 0.569, $p < 0.01$; and the psychosocial domain score, rho = 0.372, $p < 0.01$). In the UK sample, the daily management time (DMT) showed correlation with the HidroQoL overall score (rho = 0.437, $p < 0.01$) as well as the two domain scores (impact on daily life activities, rho = 0.364, $p < 0.01$, and psychosocial domain score, rho = 0.44, $p < 0.01$).

This underscores the findings already observed on the link between global impact on patient's life and the HidroQoL scores, by demonstrating an equally strong correlation with a more specific aspect of the burden. Beyond this, the link between *daily time spent in caring* and impairment in hyperhidrosis has a relevance to the understanding of potential determinants of QoL, which is explored thoroughly in the next section.

Hypothesis 8: Greater HRQoL impairment is associated with higher WTP values.

Testing hypothesis 8 involved assessing how the HidroQoL scores relate to the patients WTP for a new cure. Spearman's rank sum correlation analysis was used for assessing the association between the HidroQoL scores and patient's WTP values. In the US sample, the HidroQoL scores had a weak correlation with WTP values (overall scale score, rho = 0.257, $p < 0.01$; *impact on daily life activities* domain, rho = 0.251, $p < 0.01$; and *psychosocial impact* domain, rho = 0.225, $p < 0.01$). Similar findings were obtained from the UK sample.

These findings partially confirm the null hypothesis, showing that impairment in HRQoL as reflected in HidroQoL scores is positively correlated with patient's WTP. On a broader basis, these findings offer insights into the nature of the relationship between hyperhidrosis-specific QoL and generic HRQoL. The two are weakly related. This was also noted earlier on the EQ-

5D-5L usual activities and pain scales. This underscores the uniqueness of hyperhidrosis-specific QOL and its role in understanding the impacts of disease. Nonetheless, in comparison with other tools for assessing generic HRQoL, WTP suffers a number of handicaps, the values obtained might be influenced by a range of factors, for example, the patient's ability to pay, cultural context, in this case in relation to the financing of health care services e.g. the proportion of out of pocket payments (Drummond et al. 2005).

Part III: Predictors Of HRQoL in Hyperhidrosis

Understanding the determinants of HRQoL in hyperhidrosis patients is of importance in both clinical and policy settings. In order to identify and address the health care needs of hyperhidrosis patients, factors that influence key patient outcomes offers opportunity for intervention during treatment or designing and planning of care. Multivariate regression analyses were carried out using stepwise and hierarchical regression techniques to identify factors influencing patient's HidroQoL score. The HidroQoL overall score was the dependent variable and independent variables used in the model included: the HDSS score, GQ score, DMT measured in minutes; body site affected; additional out of pocket expenditure on hyperhidrosis; treatment history (treatment within the last 6 months; currently receiving treatment; treated with surgery; treated with BTX-A); co-morbidities (thyroid disorder; psychiatric or neurologic disorders; diabetes; hypertension).

a) hierarchical regression analysis

In the hierarchical regression analysis, variables were sequentially added to the regression, one at a time, to facilitate the understanding of how much each contributed to explaining variance in patients HRQoL. The HDSS score (representing disease severity) explained the most variance in the HidroQoL score (38.74 %) (Table 8.31, Table 8.32). Other variables making significant contributions to explaining the HRQoL included: the GQ score (global impact) (7.9%); location of the hyperhidrosis (4.3%); patient's age (1.9%); co-morbidity: having thyroid disorders (1.8%). Overall the model explained 58% of the variance in the total score.

b) backward stepwise regression

In the backward stepwise regression, the regression model was estimated sequentially, first estimating a model with all variables; then estimating subsequent models by removing the least significant regressor if its significance level was ≥ 0.1 at each step, until there was no variable to

be excluded. The final model retained the following predictors: HDSS score, GQ score, co-morbidities: psychiatric or neurologic disorders, thyroid disorders; body site affected; age (Table 8.33). These predictors were jointly statistically significant in explaining the variability in the HRQoL of patients with HH, explaining 55.5% of the variation. However, only coefficients for the HDSS score, **GQ score**, and **age** were statistically significant:

- **The HidroQoL Score of patients with HDSS** score of 4, was 7.82 higher than those with HDSS score of 2.

- Patients with a GQ score of 3, 4, and 5 had respectively higher HidroQoL scores of 4.36, 6.84 and 7.44 than in patients with a GQ score of 2.

- HRQoL scores of patients with *palmar, feet and axillar* sweating were 3.53 lower than those with generalised sweating; those with *head or facial sweating* were 3.96 lower than those with generalised sweating.

Due to size of the UK sample, only the stepwise regression analysis was performed. Following estimation, only one predictor, the HDSS, remained in the model which was also significant. The model explained 33.7% of variance in QoL (Table 8.34). Model diagnostics analysis were carried out to see whether the OLS regression assumptions had been met. The residuals from the regression met normality assumptions and were homoskedastic (reflected equal variance) across the fitted values.

DISCUSSION

Apart from influencing the integrity of inferences, the lack of validity may have severe consequences for patient care, for example, misdiagnosis of patients, where a measure is used for screening purposes; or over-and-under treatment where an instrument is used for screening in patient management. This reflects the necessity of establishing the construct validity for measures of HRQoL. Thus, this chapter intended to provide evidence based on various tests, supporting the hypothesis that the HidroQoL accurately assesses HRQoL in patients with hyperhidrosis. The affirmation of the HidroQoL (response option a little or very much) exceeded 80% for most items reflecting on the relevance and importance of the content of the measure to hyperhidrosis patients.

Table 8.31: Predictors of HRQoL in hyperhidrosis (all variables included) based on the US sample (n = 127)

	B	SE	t	p-v	Beta
HDSS					
HDSS = 3	2.39	1.43	1.68	0.1	0.17
HDSS = 4	7.35	1.88	3.91	0	0.51
GQ Score					
GQ Score = 3	4.85	2.08	2.33	0.02	0.31
GQ Score = 4	7.49	2.26	3.32	0	0.52
GQ Score = 5	7.76	2.63	2.95	0	0.51
DMT (minute)	0	0	0.8	0.43	0.06
Location					
Palmar, feet & Axillar	-3.93	1.33	-2.95	0	-0.28
Palms and/or feet	-2.71	1.49	-1.82	0.07	-0.16
Head	-4.02	1.9	-2.12	0.04	-0.18
Axilla plus other	-3.41	2.15	-1.58	0.12	-0.12
Monthly Expenditure (£)	0	0	-0.42	0.67	-0.03
Age	-0.08	0.04	-1.96	0.05	-0.16
WTP					
WTP = £ 50 to 199	1.31	1.17	1.11	0.27	0.09
WTP = 200 or more	1.13	1.42	0.8	0.43	0.07
Female	-0.81	1.36	-0.6	0.55	-0.04
Treatment history					
Not treated within last 6 months	-0.95	1.2	-0.79	0.43	-0.07
Never received surgical therapy before	0.45	1.44	0.31	0.76	0.02
Never received Botox therapy before	-1.31	1.59	-0.82	0.41	-0.06
Not undergoing treatment currently	0.17	1.19	0.14	0.89	0.01
Co-morbidities reported absent:					
Thyroid disorder	-3.41	1.84	-1.85	0.07	-0.13
Psychiatric or neurologic disorders	-2.41	1.4	-1.72	0.09	-0.13
Menopausal related complaints	1.25	1.72	0.73	0.47	0.05
Diabetes	-0.39	2.02	-0.19	0.85	-0.02
Hypertension	0.01	1.44	0.01	1	0
Constant	28.14	4.95	5.69	0	.

F(24, 102) = 5.89; P < 0.001; R-squared = 0.58; Adj. R-squared = 0.48; Root MSE = 5

Note:1. The following groups were used as reference for comparing the influence of the dummy variables: HDSS = 4; GQ Score= 2; Generalised HH; WTP = £ 49 or less; Male; Co-morbidities confirmed; 2. B: coefficient; SE - standard error; t: t-statistic; p: p-value; Beta: beta coefficient

Table 8.32: Contribution of predictors included in 'all variables model' to explaining the variance in the HidroQoL Scores; with hierarchical inclusion of variables (USA sample)

Model fit statistics				**Impact of change on model**		
Model	*R-sq*	*F(df)*	*p-value*	*R2 Change*	*F(df) change*	*p-value*
1	0.39	38.741(2,124)	0			
2	0.46	20.918(5,121)	0	0.079	5.95(3,121)	0.001
3	0.47	17.491(6,120)	0	0.003	0.66(1,120)	0.42
4	0.51	12.052(10,116)	0	0.043	2.54(4,116)	0.043
5	0.51	10.883(11,115)	0	0	0.11(1,115)	0.737
6	0.53	10.659(12,114)	0	0.019	4.53(1,114)	0.036
7	0.54	9.395(14,112)	0	0.011	1.39(2,112)	0.255
8	0.54	8.699(15,111)	0	0	0.06(1,111)	0.809
9	0.54	8.200(16,110)	0	0.004	0.87(1,110)	0.354
10	0.55	7.683(17,109)	0	0.001	0.28(1,109)	0.6
11	0.55	7.277(18,108)	0	0.003	0.71(1,108)	0.401
12	0.55	6.840(19,107)	0	0	0.09(1,107)	0.764
13	0.57	6.926(20,106)	0	0.018	4.41(1,106)	0.038
14	0.58	6.862(21,105)	0	0.012	2.98(1,105)	0.087
15	0.58	6.546(22,104)	0	0.002	0.54(1,104)	0.464
16	0.58	6.206(23,103)	0	0	0.04(1,103)	0.845
17	0.58	5.889(24,102)	0	0	0.00(1,102)	0.995

	Variables included in model
Model 1	Constant (C); hdss (a)
Model 2	C; a; gq
Model 3	C; a; gq; m (m)
Model 4	C; a; gq; m; location (l);
Model 5	C; a; gq; m; l; expenditures (£)
Model 6	C; a; gq; m; l; £; age;
Model 7	C; a; gq; m; l; £; age; wtp
Model 8	C; a; gq; m; l; £; age; wtp; gender (g)
Model 9	C; a; gq; m; l; £; age; wtp; g; dm4
Model 10	C; a; gq; m; l; £; age; wtp; g; dm4; dm5
Model 11	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6;
Model 12	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6; dm7
Model 13	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6; dm7; dm13;
Model 14	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6; dm7; dm13; dm14
Model 15	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6; dm7; dm13; dm14; dm15
Model 16	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6; dm7; dm13; dm14; dm15; dm16
Model 17	C; a; gq; m; l; £; age; wtp; g; dm4; dm5; dm6; dm7 dm13; dm14; dm15; dm16; dm17

Table 8.33: Predictors of hyperhidrosis-QoL based on stepwise backward regression analysis (USA sample, N = 127)

	Coef.	*SE*	*t*	*p-value*	*[95% Int.]*	*Confidence*
HDSS = 3	2.71	1.34	2.03	0.05	0.06	5.36
HDSS = 4	7.89	1.71	4.62	0.00	4.51	11.27
GQ Score = 3	4.36	1.88	2.32	0.02	0.64	8.08
GQ Score = 4	6.84	1.97	3.48	0.00	2.95	10.74
GQ Score = 5	7.44	2.30	3.24	0.00	2.89	11.99
_Idm14_2						
Location:						
Palmar, feet & Axillar	-3.53	1.23	-2.86	0.01	-5.97	-1.08
Palms and/or feet	-2.70	1.41	-1.92	0.06	-5.49	0.08
Head	-3.96	1.74	-2.28	0.02	-7.41	-0.52
Axilla plus other	-3.48	2.01	-1.73	0.09	-7.46	0.51
Co-morbidities reported absent:						
Psychiatric or neurologic disorder	-2.28	1.29	-1.77	0.08	-4.83	0.28
Thyroid disorder	-3.24	1.71	-1.90	0.06	-6.62	0.14
age	-0.09	0.03	-2.80	0.01	-0.16	-0.03
Constant	28.23	3.09	9.12	0.00	22.10	34.36
F(12, 114)		11.88				
Prob > F		0				
R-squared		0.5557				
Adj R-squared		0.5089				
Root MSE		4.8676				

Table 8.34: Predictors of hyperhidrosis-QoL based on stepwise backward regression analysis (UK sample, N = 36)

	B	*SE*	*t*	*p-value*	*[95% CI]*	
HDSS = 3	6.61	2.63	2.51	0.017	1.26	11.97
HDSS = 4	10.52	2.58	4.07	0	5.27	15.77
Constant	19.64	1.90	10.33	0	15.77	23.50
F(2, 33)		8.38				
Prob > F		0.0011				
R-squared		0.3369				
Adj R-squared		0.2967				
Root MSE		6.3017				

Some items such as *my choice of clothing is affected* and *I feel embarrassed* showed extremely high endorsement rates, reflecting the high prevalence of these issues during interviews carried

out in the qualitative study reported in chapter 3. This shows that the items of the HidroQOL reflected issues of importance to patients. No responses were missing in both samples (UK & USA) used in the study, an advantage of electronic collection of data, which allowed use of a number of feature, such as prompts and reminders, to ensure all questions were responded to (Dillman 2006).

HidroQoL scores were compared across groups of patients on a number of various patient characteristics including gender, age, site of hyperhidrosis, co-morbidity, disease severity (HDSS score) and global impact to hyperhidrosis (GQ score). Significant differences in HidroQoL scores were observed across patients with differing sites of hyperhidrosis; but not across gender or age-groups. Previous work, though bearing little comparability with the current study, is still insightful. Amir et al. (Amir et al. 2000) observed lower quality of life in females than males, in an Israeli sample (N = 48); on the other hand Wolosker et al. (Wolosker et al. 2010) found no gender differences in the HRQoL of patients with Palmar HH awaiting surgery. This is similar to observations related to self-reported disease severity. While Kirimian-Teherani et al (Karimian-Teherani et al. 2009) found higher level of self-assessed severity in women; studies by Lear et al. (Lear et al. 2007), US and Canadian clinic samples, and Strutton et al. (Strutton et al. 2004), based on US households, could not find any gender differences. Thus while hyperhidrosis is most prevalent in the active working age group (Strutton et al. 2004, Lear et al. 2007), its impact is rather common across the board.

Differences in QoL based on the site of HH have been previously noted. Hamm et al. (2007) found greater QoL impairment in patients with axillary HH than in those with palmar HH, based on a German clinic sample. Patients with different sites of hyperhidrosis seem to differ in a number of important ways, with potential implications on how they experience the disease. For instance, palmer-plantar hyperhidrosis has onset in childhood or before puberty while axillary often starts during or post-puberty (Lear et al. 2007). Treatment pathways and their corresponding effectiveness differ by location of hyperhidrosis (Solish et al. 2007), leading to differences in the levels of control of the condition. Axillar hyperhidrosis for example has more treatment options which have demonstrated effectiveness, in comparison to craniofacial hyperhidrosis, which not only faces limited treatment options, but also has limited effectiveness information on the same. Additionally, the extent of visibility of the sweating is variable across different sites of hyperhidrosis; those with sweating involving the head/face or underarms may be more exposed than those with plantar hyperhidrosis (involving the feet).

For instruments used for discriminatory purposes, it is crucial to demonstrate the capability to distinguish between patients experiencing severe vs. those with milder forms of a condition, (Guyatt et al. 1992). Testing the HidroQoL for this capability was done by comparing scores from patients with differing levels of disease severity (HDSS score); also those with differing levels of overall impact of HH (GQ score). The HidroQoL scores showed significant differences in patient groups according to the HDSS score and the GQ score. Even more reassuring is that discriminative ability was also seen in all items, across both severity and overall impact. Capability to discriminate across different levels of disease severity is also an important early indication of the ability to detect important changes in patients condition (Fayers and Machin 2007). Taken together this suggests that the HidroQoL would be useful in the clinic, for the diagnosis as well as management of hyperhidrosis.

Further construct validation of the HidroQoL involved evaluating convergence and divergent validities. Convergence validity is demonstrated where an instrument correlates with other scales assessing a similar construct; on the other hand, discriminant validity is seen where an instrument does not correlate with instruments assessing un-related constructs (DeVellis, 2011). The HidroQoL score correlated with scores of the DLQI and Skindex-17, measures of dermatology-specific quality of life. The observed degree of relationship seemed consistent for measures assessing related but distinct constructs. Evidenced by higher correlations between the DLQI and the Skindex, which measure the same construct. Although therapeutic area instruments such as the DLQI offer better versatility and comparability across a therapeutic area as they deal with issues that have relevance across a group of related disease (Salek 1998), the potential for omissions of key items or irrelevant inclusions of items with respect to specific disease conditions may still be quite unavoidable. For instance items like ‘skin hurts’ or ‘skin condition bleeds’ in the Skindex-17 seem irrelevant in hyperhidrosis. The HidroQoL, on the other hand, as a disease specific instrument by definition and design contains content relevant only to hyperhidrosis. Moreover, the early development of the HidroQoL (See Chapter 3) actively involved hyperhidrosis patients, to ensure the measure’s content reflects themes of importance to them and; capture using the very expressions the patients use.

The lack of significant correlations between the HidroQoL scores and EQ-5D-5L’s ‘mobility’ and ‘self-care’ subscales demonstrates the divergent validity of the HidroQoL, as those domains are not relevant for HH patients let alone dermatology. On the other hand weak correlation with the other EQ-5D-5L items (usual activities, pain/discomfort and anxiety/depression) seems to

suggest that these aspects may have some relevance in HH, as expected, though convergence validity is not supported. It is encouraging that the magnitude of correlation between these three domains (aspects of generic QOL) and the HidroQoL, is smaller in comparison to that with Skindex and DLQI, measuring dermatology-specific QOL, a construct more closely related to hyperhidrosis-specific QOL.

The HidroQoL showed moderate-strong correlations with disease severity (HDSS score); overall global impact (GQ score); and daily time spent in dealing with the condition (measured in minutes) a further support of the convergence validity. Weak correlation was found between HidroQoL and patient's WTP. Although various studies have successfully demonstrated that WTP is indeed responsive in patients with skin disease, Müller and Augustin (Muller and Augustin 2013) found WTP to be non-responsive to change in disease severity; and uncorrelated to patient therapeutic benefit, in a sample of German outpatients with hyperhidrosis (n = 96). The relationship between WTP and self-reported disease severity and QoL has been explored in other skin conditions. WTP has shown significant correlations with the Psoriasis Disability Index ($r = 0.42$) in a sample of psoriasis patients (Finlay and Coles 1995), acne disability score ($r = 0.229$, $p < 0.005$) (Motley and Finlay 1989) and with the DLQI ($r = 0.249$) in patients with rosacea (Beikert et al. 2013). The current results suggests a limited usefulness of WTP in patients with hyperhidrosis.

Time spent on treatment has been previously used for studying the impacts of skin disease such as psoriasis and atopic dermatitis. Positive correlations have been confirmed between time spent on treatment and the Psoriasis Disability Index ($r = 0.37$) in psoriasis patients (Finlay and Coles 1995) and with IDQoL/DLQI (unclear whether CDLQI was used in this paper) in children with atopic dermatitis ($r = 0.31$) in (Jemec et al. 2006). This confirms a strong relationship between time spent on treatment and disease impacts on the patients. Moreover, (Holm and Jemec 2004) have reported on the test-retest reliability of time spent on treatment in atopic dermatitis patients. The results from the current study suggests that time spent on treatment represent an important measure of disease impacts in hyperhidrosis, consistent with observations in other skin conditions.

This study also looked into the predictors of HRQoL in hyperhidrosis patients, by exploring the determinants of the HidroQoL score. The majority of the variance in the patients was explained by the regression models (All-variables model: explained 58% of variance; backward stepwise regression explained 55%). Significant and important predictors included patient's disease severity; global life impact; site of hyperhidrosis; and age. This has not been explored previously in hyperhidrosis patients, nonetheless this is consistent with other skin conditions, for instance Psoriasis, where the most important predictors of HRQoL have been reported to be daily treatment time, disease severity; and patient benefit (Blome et al. 2010).

SUMMARY

- This chapter has reported on a validation study of the new HRQoL measure for hyperhidrosis, the HidroQoL, involving ***known-groups***, ***convergent*** and ***divergent validities***.
- The HidroQoL scores showed significant differences across patients with hyperhidrosis involving varying sites, but not across gender or age group.
- Patients with higher HDSS scores (level of disease severity) registered significantly higher HidroQoL scores.
- Patients with a higher GQ score (higher overall impact of disease on life) showed significantly higher HidroQoL scores.
- There was a significant difference in HidroQoL scores of patients who spent less than 1 hour; at least 1 hour but no more than 1.5 hours and; more than 1.5 hours in managing their condition daily.
- Scores of the daily life activities domain of the HidroQoL showed a significant difference w.r.t. to one aspect of treatment history (receiving surgery).
- There was no difference in HidroQoL scores with respect to co-morbidities, except for 'psychiatric or neurologic disorders'.
- Convergence validity of the HidroQoL was demonstrated through positive moderate correlations with the Skindex scores; the DLQI scores; Global impact score; disease severity; and daily time spent in dealing with condition.
- Divergent validity was demonstrated by lack of correlation between the HidroQoL Scores and the EQ-5D's 'mobility' and 'self-care' scale scores.

CHAPTER 9

Responsiveness and Interpretation of the Hyperhidrosis Quality of Life Index (HidroQoL) Scores

INTRODUCTION

Where an instrument is used in a longitudinal context, for example for monitoring the condition of individual patients over time, further psychometric attributes apart from internal consistency, reliability and construct validity may be required to ensure valid measurements. The measure must be capable of detecting important changes taking place in the patient's condition even if they are small, an attribute referred to as responsiveness (Guyatt et al. 1987). The assessment of responsiveness requires an external measure as a criterion for determining whether the patient's condition has changed, improved or worsened (Revicki et al. 2008). Previous clinical trial results, on differences between placebo and active treatment; or known distribution properties of the target patient population may also be useful as basis for assessing responsiveness.

Guyatt and colleagues (Guyatt et al. 1987) have argued that responsiveness is quite separate from validity and reliability, drawing the conclusion that an instrument can be responsive without being either valid or reliable. Demonstrating a measure's ability to capture score changes, without evidence that those changes are linked to the intended construct or the extent of measurement error in the change scores, would put the usefulness and relevance of such 'responsiveness' into question (Hays and Hadorn 1992). Considering construct validity as a "process by which empirical evidence from several validation procedures is assembled to support the inference that a particular instrument measures what it purports to measure" (Salek 1998), the evidence for responsiveness is one of such pieces of information, relating to the instrument's use in a longitudinal context. Establishing responsiveness requires demonstrating that the observed score changes reflect true changes in the concept being measured (longitudinal validity) and that such changes are not merely random variability (longitudinal reliability) (Terwee et al. 2003). Moreover, as measurement error increases, larger and larger magnitude of change might be required to demonstrate any treatment effect (Guyatt et al. 1987), reflecting an inverse relationship between reliability and responsiveness, similar to the observation on construct validity.

Establishing that an instrument produces reliable and valid measures and that it is responsive, is not sufficient to render it useful in routine clinical practice or in clinical research. Information facilitating assigning of easily understood meaning to an instrument's quantitative scores must also be available (Lohr 2002). Rather than just knowing whether patient scores have changed in a

statistically significant way, of relevance to patient management is whether a change in scores is clinically significant i.e. whether the change in scores is large enough to have an implication for patient care (Wyrwich et al. 2005). Such a cut-off change score is considered as the minimum clinically important difference (MCID) (Guyatt et al. 2002). Also, for a given absolute score, clinicians may want to know its implication on the patients' condition, whether it represents a mild, moderate or severe state of the patient's condition. In widely used dermatology QoL instruments such as the Skindex and DLQI both score categorisation and qualitative descriptors for each band have been provided (Hongbo et al. 2005; Prinsen et al. 2009). In addition, values for minimum clinically important difference (MCID) have been reported (Both et al. 2007; Basra et al. 2008). Nevertheless, efforts to ensure that QoL scores are interpretable are not limited to the above, results of clinical trials and statistical characteristics of samples where the measure was previously used and comparative data from non-diseased populations may be useful. Availability of any of such data is therefore considered among the important psychometric attributes of an instrument.

OBJECTIVES

Objectives of this study were to:

- Assess whether the HidroQoL is sensitive to change in patients whose condition had changed.
- Evaluate whether the HidroQoL was capable of discriminating between patients experiencing different levels of change in their sweating.
- Establish the MID value of the HidroQoL.
- Develop and propose scale banding for the interpretation of the HidroQoL scores.

METHODS

This study followed a prospective longitudinal study design with patient's assessed on two occasions, at baseline (assessment 1) and during a follow-up assessment (assessment 2) at least 21 days after initial assessment. There is no recommendation regarding the best interval between assessments in responsiveness studies. However, as a construct validation procedure, groups of patients expected to change should be identified a priori.

The study used the patient population used in the previous chapter.

Inclusion criteria:

- Patients with self-reported excessive sweating problems;

- Experiencing interference in their daily life (HDSS ≥ 2);
- Onset of hyperhidrosis in teenage years or early adulthood years;
- Aged 17 or above.

Exclusion criteria:

- Patients not experiencing any excessive sweating;
- Excessive sweating was not causing any interference in daily life (HDSS = 1);
- Onset of hyperhidrosis was after age of 30; and patient reporting a co-morbidity (hypertension, diabetes, PM hormonal disorders, and psychological disorders).
- Aged below 17.

Outcome measures

Apart from the HidroQoL, data were also collected on the patient's disease severity using the HDSS and on dermatology-specific QoL using the DLQI and the Skindex-17. In addition, two general questions were administered: an overall-impact global question (GQ), scored on a 5 point Likert scale; and a patient global assessment of change (PGA), also scored on a 5-point Likert scale. Further details on these measures are available in Chapter 2.

Procedures

After completion of the first assessment, patients received communication containing information on their follow-up assessment due in 21 days. On the 20th day following their initial assessment patients were sent an email containing the access details for their follow-up assessment (due in a day).

Data processing and analysis

All data analysis were performed using SPSS 20, STATA and MS Excel. Scale scores were tested for normality using Q-Q plots, histograms and the Shapiro-Wilk test. A paired t-test was used to assess whether the HidroQoL is capable of detecting any change in patient's condition from baseline (assessment 1) to follow-up (assessment 2). Significant score changes ($p < 0.05$) were expected in patient-groups that had minimally worsened or improved. Non-significant changes were expected for the no-change group ($p > 0.05$) was expected in the no-change group based on the anchor variables (PGA score and HDSS-change-score). The magnitude of change in the HidroQoL scores from baseline to follow-up assessment in each patient group was estimated using

the Effect sizes (ES) and Standard Response Mean (SRM). Cohen's Effect Size is estimated as the ratio of the mean score change (test 1 – test 2) to the standard deviation (SD) of baseline scores (Streiner and Norman 2008). SRM is given by the ratio of the mean score change (in each patient group) to its SD (Streiner and Norman 2008). In addition, the HidroQoL's efficiency at detecting change in patients was compared with that in of the DLQI and the Skindex using the relative efficiency index (REI). The REI indicates how much more or less valid and responsive an instrument is, relative to the comparator instrument (Ware et al. 1998) and is estimated as the ratio t-statistics of the target measure to the comparator (Fayers and Machin 2007). Oneway ANOVA was used to assess the HidroQoLs ability to discriminate among patients experiencing different levels of change from baseline to follow-up. F-statistics are calculated to test for differences in the mean changes in scale scores across groups (Ware et al. 1998).

A categorisation of the HidroQoL overall score was carried out by mapping scores of the HidroQoL to the GQ score; with the descriptors of GQ scores providing the descriptors for the bands. This involved first estimating the mean, mode and median of the GQ scores for each of the HidroQoL scores. The HidroQoL overall scores ranges associated with each mean GQ score was identified such that each score was allocated a unique band. To test the banding system, the level of agreement between the patients GQ scores and the banding (GQ predicted from the patients HidroQoL scores) was assessed using Intra-class correlations (ICC). In addition to assessing measurement error, the ICC also captures systematic bias among raters, thus measuring absolute agreement rather than consistency only (Streiner and Norman 2008). Its interpretation is similar to that of the kappa coefficient, where agreement is considered 'almost perfect' if at least 0.81, 'substantial' for 0.61-8, 'moderate' at 0.41- 6, 'fair' for .21-4. 'slight' for 0.00 – 0.20 and 'poor' for less than 0.00 (Landis and Koch 1977).

The MCID for the HidroQoL was estimated using three approaches:

1) Using an external anchor: the mean score change from baseline to follow-up in the minimally improved patient group (according the anchor) (Revicki et al. 2008).
2) Based on distribution of scores: One third standard deviation and one half standard deviation translate the effect sizes of 0.33 and 0.5, respectively, into the units of the scores being assessed (Yost and Eton 2005). Standard error of measurement (SEM) is calculated as (Standard Deviation $\sqrt{(1 - \text{reliability})}$) provides a relatively consistent measure of

precision across samples and settings (Wyrwich et al. 2005). One SEM is consistent with effect size of 0.5 when a measure's reliability is 0.75, and 0.33 when reliability is about 0.9 (Yost and Eton 2005).

3) Integrating anchor-based and distribution based methods: Upper bound of a 1-tail 95% CI for the mean score change in the patient group which did not change (de Vet et al. 2007).

Part I: Responsiveness Of The HidroQoL

Characteristics of the study participants

A total of 89 participants completed the HidroQoL questionnaire, on two occasions, 15 to 35 days apart. Fifty-four patients (60.7%) were from the USA and seventeen (19.1%), the UK (Table 9.1). The mean age of patients from the US was 37 (±12) with a range of 18 – 73, while those from the UK had a mean age of 38 (±13), ranging from 20 to 62. The largest age group was the 31 to 40 group in the both the USA (N = 19, 35.2%) and UK (N = 6, 35.3%) samples. Ninety-nine percent (N = 53) of patients from the US and one-hundred percent (N = 17) of those from the UK had previously sought for medical attention regarding their sweating (Table 9.2). Thirty five percent (N = 19) of the US patients, fifty nine percent (N = 10) of the UK sample, had been treated for their hyperhidrosis in the last 6 months, with a lower proportion being currently treated (USA, 29.6%, N = 16; UK, 47 %, N = 8). Oral systemic drugs (in pill-form) were the most common form of treatment (US, N = 7, 13%; UK, N = 4, 23%). The majority of participants did not have any co-morbidities, the most common were psychiatric and neurological disorders (US, N = 8, 14.8%; UK, N = 1, 5.9%).

Distribution of scale scores

The mean HidroQoL scores were 26.64 (± 7.14) and 25.08 (± 8.38), for baseline (test 1) and follow-up (test 2) assessments, and the range was 1 – 36 for both assessments (Table 9.3). The mode scores were 33 (test 1) and 32 (test 2), suggesting high levels of QoL impairment in both assessments. The DLQI mean scores were 10.13 (± 6.87) and 9.55 (± 6.96), during the first and second assessment; with ranges of 0 to 25 and 0 to 26, respectively. This reflects moderate to very large life impacts, lower cut-off for very large effect QoL effect for the DLQI is 11 (Hongbo et al. 2005). Most patients had low to moderate scores, as reflected in the inter-quartile range (5 – 16)

and (4 to 15) for both the first and second assessment, respectively, with no patients towards the upper extremity. The DLQI scores showed a slightly positive skew.

Table 9.1: Sociodemographic characteristics of the patients

Characteristic	USA Sample (n = 54)	UK Sample (n = 17)	Pooled Sample* (N = 89)
Gender, n (%)			
Male	9 (16.7%)	3 (17.6%)	18 (20.2%)
Female	45 (83.3%)	14 (82.4%)	71 (79.8%)
Age (years)			
Mean, SD	37, 12	38, 13	37, 12
Median	34	39	34
Mode	20	40	31
Range	18 – 73	20 – 62	16 - 73
Age (years), n (%)			
17 to 30	17 (31.5%)	5 (29.4%)	28 (31.9%)
31 to 40	19 (35.2%)	6 (35.3%)	33 (36.7%)
41 to 50	10 (18.5%)	3 (17.6%)	16 (17.8)
51 to 60	5 (9.3%)	1 (5.9%)	7 (7.8%)
> 60	3 (5.6%)	2 (11.8%)	5 (5.6%)
Body site involved, n (%)			
Generalised	11 (20.4%)	6 (35.3%)	20 (22.5%)
Palms, feet & axilla	21 (38.9%)	4 (23.5%)	28 (31.5%)
Palms and feet	8 (14.8%)	1 (5.9%)	16 (18%)
Head, Face	3 (5.6%)	2 (11.8%)	6 (18%)
Axilla	10 (18.5%)	2 (11.8%)	15 (16.9%)
Palms	1 (1.9%)	1 (5.9%)	3 (3.4%)
Trunk & rest of body	0 (0%)	1 (5.9%)	1 (1.1%)
Employment, n (%)			
Employed	38 (70.4%)	10 (58.8%)	61 (68.5%)
Unemployed	9 (16.7%)	3 (17.6%)	13 (14.6%)
Retired	3 (5.6%)	2 (11.8%)	5 (5.6%)
Full time student	4 (7.4%)	2 (11.8%)	10 (11.2%)
***Country of residence**			
USA			54 (60.7%)
UK			17 (19.1%)
Canada			8 (9%)
Australia			5 (5.6%)
Other			15 (5.6%)

Table 9.1 (continued)

	USA Sample (n = 54)	UK Sample (n = 17)	Pooled Sample (N = 89)
Disease severity (HDSS score), n			
1	0	0	0
2	28	17	51
3	63	32	120
4	51	24	89
Global impact (GQ score), n			
No effect	0	2	2
Small effect	9	5	18
Moderate effect	37	17	65
Large effect	53	33	106
Extremely large effect	43	16	69

Table 9.2: Patients' treatment history and disease characteristics

	USA Sample (n = 54)	UK Sample (n = 36)	Pooled Sample (n = 89)
Seen medical practitioner regarding hyperhidrosis, n (%)	53 (98.1 %)	17 (100%)	87 (97.8%)
Have received Surgical treatment, n (%)	4 (7.4%)	3 (17.6%)	12 (13.5%)
Received Botox within last 6 months, n (%)	7 (13%)	0 (0%)	8 (9%)
Received treatment within last 6 months, n (%)	19 (35.2%)	10 (58.8%)	36 (40%)
Currently receiving treatment*, n (%)	16 (29.6%)	8 (47.1%)	27 (30.3%)
Oral-systemic drugs (pill-form)	7 (13%)	4 (23.5%)	11 (12.4%)
Iontophoresis	5 (9.3%)	4 (23.5%)	10 (11.2%)
Aluminium Chloride Topical treatment	4 (7.4%)	0 (0%)	5 (5.6%)
Non-prescription/cosmetic preparations	5 (9.3%)	3 (17.6%)	8 (10%)
Other	4	1	6
Co-morbidities, n (%)			
Thyroid disorders	3 (5.1%)	0 (0%)	3 (3.4%)
Psychiatric of neurologic disorders	8 (14.8%)	1 (5.9%)	9 (10%)
Menopausal related complaints	4 (7.4%)	1 (5.9%)	6 (6.7%)
Diabetes	6 (11.1%)	0 (0%)	6 (6.7%)
Hypertension	4 (7.4%)	0 (0%)	5 (5.6%)
Other	9 (16.7%)	4 (23.5%)	14 (15.7%)

***** Some patients were on more than one of the listed treatments.

Table 9.3: Distribution of the scale scores, in pooled sample

	HidroQoL		DLQI		Skindex	
	test1	*test2*	*test1*	*test2*	*test1*	*test2*
Mean	26.62	25.08	**10.13**	**9.55**	17.27	16.42
SD Mean	7.14	8.38	6.87	6.96	8.26	8.44
SE Mean	0.76	0.89	0.73	0.74	0.88	0.89
95% CI	14, 35	9, 35	0, 22	0, 22	1, 27	0, 29
Range	1-36	1-36	0-25	0-26	0-31	0-30
IQR	22-33	18-32	5-16	4-15	14-14	12-22
Median	27	26	10	8	18	17
Mode	33	32	6	0	17	0
Skewness	-.914	-.770	.275	.424	-.690	-.423
SE skewness	.255	.255	.255	.255	.255	.255
Kurtosis	.775	-.055	-1.002	-.741	-.410	-.572
SE kurtosis	.506	.506	.506	.506	.506	.506
Ceiling, N (%)	2 (2.2%)	1 (1.1%)	-	-	-	-
Floor, N (%)	-	-	5 (5.6%)	10(11.2%)	4 (4.5%)	8 (8.9%)

In contrast, the score range used by the HidroQoL was wider and most patients were within the upper third of the scale. The distribution of the HidroQoL showed a negative skew, for both baseline and follow-up assessments. Kurtosis in all three scales shows minimal departure from that of a perfect bell-shape normal distribution (with kurtosis = 0).

Assessing the usefulness of the anchors

At baseline, the majority of USA patients (N = 22) had an HDSS score of 3, reflecting sweating that was barely tolerable and frequently interfered with their daily activities (Table 9.4). Among UK patients, there were equivalent numbers of patients with HDSS score of 3 and 4 at baseline, N = 6, for each. An HDSS score of 4 reflects intolerable sweating that always interferes with daily life activities. In the pooled sample, seventy-one percent of the patients (N = 64) had experienced no change in their condition, from baseline to follow-up, based on their HDSS-change-score (HDSS-cs = 0) (Table 9.5). Twenty-one percent of the patient (N = 19) had experienced a small improvement in their condition (HDSS-cs = -1) while only 6.7% (N = 6) had experienced a small deterioration (HDSS-cs = 1). No patient had experienced a major improvement or major deterioration i.e. an HDSS-cs of 2 or -2. Based on the second anchor, the patients global assessment (PGA), the proportion of patients reporting no change in their condition (PGA score = 3) was higher, seventy-four percent (N = 67) in the pooled sample (Table 5.1). Only eleven-percent of the patients (N = 10) reported a small improvement, while ten-

percent (N = 9) reported a slight deterioration. Two patients reported a major deterioration while one experienced a major improvement.

Table 9.4: Distribution of the HDSS score

Number of patients (N), %						
	USA		UK		Pooled Sample	
HDSS Score	*Baseline*	*Follow-up*	*Baseline*	*Follow-up*	*Baseline*	*Follow-up*
2	13 (24.1%)	19 (35.2%)	5 (29.4%)	6 (35.5%)	19 (21.3%)	27 (30.3%)
3	22 (40.7%)	21 (38.9%)	6 (35.3%)	6 (35.5%)	41 (46.1%)	38 (42.7%)
4	19 (35.2%)	14 (25.9%)	6 (35.3%)	5 (29.4%)	29 (32.6%)	24 (27%)

Table 9.5: Distribution of the anchors: HDSS change score and PGA score

Anchor	Score	USA	UK	Pooled Sample
HDSS score change	- 1	13 (24.1%)	4 (23.5%)	19 (21.3%)
	0	39 (72.2%)	11 (64.7%)	64 (71.1%)
	1	2 (3.7%)	2 (11.8%)	6 (6.7%)
PGA	1	1 (1.9%)	1 (5.9%)	2 (2.2%)
	2	4 (7.4%)	1 (5.9%)	9 (10%)
	3	41 (75.9%)	14 (82.4%)	67 (74.4%)
	4	8 (14.8%)	1 (5.9%)	10 (11.1%)
	5	0 (0%)	0 (0%)	1 (1.1%)

Table 9.6: Correlation of the HDSS and the PGA with the HidroQoL scores

	1	**2**	**3**	**4**	**5**	**6**	**7**
HDSS (baseline) - 1	1.000						
HDSS (follow-up) - 2	.765**	1.000					
HDSS change score) - 3	-.287**	.393**	1.000				
PGA - 4	.166	.032	-.166	1.000			
HidroQoL (baseline) - 5	**.607****	**.580****	**.001**	**.108**	1.000		
HidroQoL (follow-up) - 6	**.529****	**.605****	**.142**	**-.022**	.822**	1.000	
HidroQoL change score - 7	-.049	-.217*	**-.244***	**.132**	.023	-.501**	1.000

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

External measures used as anchors, for capturing the clinical change in the patient's condition need to be easy to understand and intuitive to interpret, must correlate with the target scale as basis for confidence that they measure the target construct (Guyatt et al. 2002). The HDSS change score (HDSS-cs) had a correlation of - 0.244 (p = 0.021) with the HidroQoL change score (HidroQoL-cs) (Table 9.6). The PGA, on the other hand, had a correlation of 0.132 (p = 0.217).The two anchors, the HDSS-cs and the PGA had a correlation of - 0.166 (p = 0.12) with each other. Further evaluation of the anchor measures, HDSS change score and the PGA involved cross-tabulation of patients under the two measures.

Seventy-five percent of the patients (n = 48) with HDSS-cs of 0 also rated themselves as not experiencing any change in their condition (PGA = 3) (Table 9.7) Eighty percent of patients reporting a slight improvement (n = 8), PGA = 4, had not registered any change on their HDSS-cs. On the other hand, eighty-four percent of patients with a small improvement in their condition i.e. HDSS-cs = -1, reported no change (PGA = 3). The observed lack of agreement between changes in patients' disease severity (HDSS-cs) and how patients perceived their condition (PGA) suggests that changes in perceived overall impact are unique from changes in disease severity.

Responsiveness of the HidroQoL

Patients were grouped according to the change in the HidroQoL between the two assessments, as follows: based on HDSS change score and the PGA score. Three groups were formed, patients not-experiencing any change (HDSS-change-score = 0; PGA = 3), patients minimally deteriorating (HDSS-change-score = 1; PGA = 2) and minimally improving (HDSS-change-score = -1; PGA = 3). The groups with major improvement (PGA = 5) and major deterioration were not considered.

Assessment of change in the HidroQoL Scores in patient groups

A paired t-test was carried out to assess the HidroQoL's sensitivity to change in each of the three patient groups: no-change, minimally worsened and minimally improved. First, the test was performed on the pooled sample, with the patients grouped based on the PGA.

Table 9.7: A comparison of the PGA against the HDSS change score in their comparison of patients.

		HDSS change score			Total
		-1	*0*	*1*	
PGA	1	0	1	1	2
	2	1	6	2	9
	3	16	48	3	67
	4	2	8	0	10
	5	0	1	0	1
Total		19	64	6	89

Note:
For *HDSS change score,* -1 is minimally improved, 0 is no-change, 1 is minimally worsened. For the PGA, 5 is sizeably improved, 4 is minimally improved, 3 is no-change, 2 is the minimally worsened and 1 is sizeably worsened.

The HidroQoL's - psychosocial domain (H-DA) score ($p < 0.05$) and the overall HidroQoL score ($p < 0.05$) significantly increased in the 'slight-improvement' group (Table 9.8). HidroQoL score changes in the 'slightly deteriorating' group were not statistically significant. Surprisingly, the 'no-change' group also showed a significant change in the domain as well as overall scores.

Similar analysis were carried out with patients grouped according to their HDSS change scores. Patients in the minimally improving group showed significant change in their HidroQoL-PS domain score ($p < 0.01$) and the total HidroQoL score ($p < 0.01$). The HidroQoL-DA domain score showed no significant changes in this group ($p = 0.08$). On the other hand, patients in the minimally worsening group did not change in a significant way ($p > 0.05$) in their total and domain HidroQoL scores. The magnitude of the mean change scores between the 'minimally improved' and the 'minimally deteriorating' were comparable (mean score change, 3.1±3.85 and -3±5.25 respectively) indicating some asymmetry. However, the 'no-change' group still showed unexpected change.

Relative efficiency of the HidroQoL at detecting change

Further, the HidroQoL's ability to detect change in the different patient groups was compared to that of the Skindex-17 and the DLQI using the relative efficiency index (REI).

Where patients were grouped based on the PGA, the REI for HidroQoL - DLQI in the minimally improved group was below 1, for the total HidroQoL score as well as for the two domains (HidroQoL-DA and HidroQoL-PS) (Table 9.9). The DLQI was more efficient at detecting change than the HidroQoL.

In contrast the HidroQoL seemed more efficient than the Skindex-17 in this patient group, REI > 2 for the HidroQoL total score as well as the domain scores. On the other hand, when patients were grouped based on HDSS-change-scores, the HidroQoL scale was more efficient than both the Skindex-17 and the DLQI in the minimally improved patient-group. Only the HidroQoL-DA domain score performed less efficiently than Skindex-17 (REI = 0.8), the total score and the HidroQoL-PS had REI greater than 1.2 with Skindex. These results were mirrored in the group worsening.

Table 9.8: Sensitivity of the HidroQoL scores in patients experiencing 'no change', 'slight improvement' and 'slight deterioration' based on paired t-test, in the pooled sample

Anchor	Patient-group	Score	Mean	SD	SE mean	95% CI		t	df	p-value
						Lower	*Upper*			
PGA	No change (n=67)	HidroQoL-DA	0.63	2.33	0.29	0.06	1.2	2.2	66	0.03
		HidroQoL-PS	0.97	2.93	0.36	0.25	1.69	2.71	66	0.01
		HidroQoL	1.6	4.5	0.55	0.5	2.69	2.91	66	0
	Minimally worsened (n= 9)	HIdroQoL-DA	-0.44	1.88	0.63	-1.89	1	-0.71	8	0.5
		HidroQoL-PS	1.33	3.61	1.2	-1.44	4.1	1.11	8	0.3
		HidroQoL	0.89	4.94	1.65	-2.91	4.68	0.54	8	0.6
	Minimally improved (n=10)	HIdroQoL-DA	1.5	2.46	0.78	-0.26	3.26	1.93	9	0.09
		HidroQoL-PS	1.3	1.77	0.56	0.04	2.56	2.33	9	0.05
		HidroQoL	2.8	3.85	1.22	0.04	5.56	2.3	9	0.05
HDSS	No-change (n = 64)	HidroQOL-DA	0.64	2.24	0.28	0.08	1.2	2.29	63	0.03
		HidroQoL-PS	0.94	2.96	0.37	0.2	1.68	2.53	63	0.01
		HidroQoL	1.58	4.49	0.56	0.46	2.7	2.82	63	0.01
	Minimally worsened (n = 6)	HidroQOL-DA	-1.5	1.64	0.67	-3.22	0.22	-2.24	5	0.08
		HidroQoL-PS	-1.5	3.83	1.57	-5.52	2.52	-0.96	5	0.38
		HidroQoL	-3	5.25	2.14	-2.51	8.51	1.4	5	0.22
	Minimally improved (n = 20)	HidroQOL-DA	1.05	2.44	0.55	-0.09	2.19	1.93	19	0.07
		HidroQoL-PS	2.05	2.48	0.55	0.89	3.21	3.7	19	0
		HidroQoL	3.1	3.85	0.86	1.3	4.9	3.6	19	0

Note: HidroQoL-DA is HidroQoL daily activities domain, HidroQoL-PS is HidroQoL psychosocial impact domain

Table 9.9: A comparison of the HidroQoL's ability to detect change with that of the DLQI and Skindex in the pooled sample

	Patient-group	**PGA**			**HDSS-change-score**		
		H-total	*H-DA*	*H-PS*	*H-total*	*H-DA*	*H-PS*
DLQI	Minimally improved	0.8	0.7	0.8	2.3	1.2	2.3
	Minimally worsened	-0.5	0.7	-1.0	**5.4**	**8.6**	**3.7**
Skindex-17	Minimally improved	2.5	2.1	2.6	1.4	0.8	1.5
	Minimally worsened	1.9	-2.5	3.8	**2.7**	**4.4**	**1.9**

Note: H-total, HidroQoL total Score; H-DA, HidroQoL impact on daily life activities domain score; H-PS, HidroQoL psychosocial impact domain;

Magnitude of change HRQoL using standard response mean and effect size

In order to estimate the magnitude of change from baseline (assessment 1) to follow-up (assessment 2) standard response mean (SRM) and effect size (ES) coefficients were calculated. Where the PGA was used as an anchor, minimally improving patients showed a small to moderate effect size on the total HidroQoL score (ES = 0.37) (Table 9.10). Patients in the 'no change' group had a small effect size on the total HidroQoL score (ES = 0.22). A comparable pattern was observed with regard to the SRM, although these figures tended to be higher. Similar analyses were carried out based on HDSS anchored patient groups. The HidroQoL showed a moderate effect size (ES = 0.47) in the minimally improving group. The DLQI and the Skindex, on the other hand, had small effect sizes (DLQI, ES = 0.21; Skindex, ES = 0.25). Although for the HidroQoL the ES for the slightly improving group remains largely the same regardless of the choice of anchor, the change in ES for the DLQI and Skindex-17 seems dramatic.

Testing for the HidroQoL's responsiveness: USA sample

Responsiveness of the HidroQoL was also assessed in the USA and the UK samples separately. The minimally improving group, showed non-significant difference in the two HidroQoL domains scores as well as the overall scale score scores (Table 9.11). The 'no change' group on the other hand showed significant differences between the baseline and the follow-up assessment. Nonetheless, the mean HidroQoL overall score change was greater in the minimally improving group (2.63± 4.14) than in the no change group (1.83±5.05).

Similar analyses were carried out using the HDSS change score as an anchor. The minimally improving group showed a significant difference in the HidroQoL overall score and the psychosocial domain score but not in the daily life activities domain score. On the other hand, the no-change group showed significant differences in the HidroQoL total score and in the daily activities domain score but not in the psychosocial score.

Comparing sensitivity to change among measures (DLQI, Skindex, HidroQoL).

The responsiveness of the HidroQoL was compared with that in other measure, the Skindex-17 and the DLQI using the relative efficiency index in patients from the USA sample. The DLQI was more efficient at detecting changes (REI < 1) where the minimally improving group was based on the PGA (Table 9.12). The HidroQoL showed greater efficiency at detecting change in patients, when the HDSS change score was used as an anchor. REI was greater than two for the overall scale and psychosocial domain score, except for the daily activities domain. These results are highly comparable to results obtained from the pooled sample.

Magnitude of change in HRQoL using standard response mean and effect size

The magnitude of change from baseline (assessment 1) to follow-up (assessment 2) was also analysed at the subgroup level for the USA. It was not possible to carryout similar analysis on the UK sample due to the size of the sample.

According to the PGA anchor, the group experiencing minimal improvement had a small-to-moderate effect size (ES) for the HidroQoL score (ES: HidroQoL total score, 0.33, daily life activities domain, 0.61 and psychosocial domain, 0.2) (Table 9.13). This is only slightly smaller than the observed figure for the pooled sample. ES in the no change group was similar (ES = 0.3). This is consistent with the observation on the paired t-test where the no-change group showed highly significant differences. Standard Response Mean was also slightly smaller than figures observed in the pooled sample (SRM: HidroQoL total, 0.63, daily life activities domain, 0.54, Psychosocial domain, 0.69). Where the HDSS was used as the anchor, the ES in the minimally improving group was higher than in the pooled sample (ES: HidroQoL overall score, 0.52, H-DA score, 0.44, H-PS score, 0.47).

Table 9.10: Estimating the responsiveness of the HidroQoL based on standardised response mean and effect size with patient groups (pooled sample)

Anchor	Patient group	Scores	Test 1		Test2		Test 1 - Test 2		SRM			ES		
			Mean	SD	Mean	SD	Mean	SD	Est.	95% CI		Est.	95% CI	
										Lower	*Upper*		*Lower*	*Upper*
PGA	**No change**	HidroQoL-DA	8.87	2.95	8.24	3.21	0.63	2.33	0.27	0.02	0.51	0.21	0.02	0.41
		HidroQoL-PS	17.84	4.85	16.87	5.28	0.97	2.93	0.33	0.09	0.57	0.2	0.05	0.35
		HidroQoL	26.7	7.34	25.1	8.13	1.6	4.5	0.36	0.11	0.6	0.22	0.07	0.37
	Minimally worsening	HidroQoL_DA	8	2.24	8.44	3.28	-0.44	1.88	-0.24	-1.01	0.53	-0.2	-0.84	0.45
		HidroQoL-PS	16.22	3.77	14.89	6.47	1.33	3.61	0.37	-0.4	1.14	0.35	-0.38	1.09
		HidroQoL	24.22	5.47	23.33	9.62	0.89	4.94	0.18	-0.59	0.95	0.16	-0.53	0.86
	Minimally improving	HidroQoL-DA	9.9	2.38	8.4	3.84	1.5	2.46	0.61	-0.11	1.32	0.63	-0.11	1.37
		HidroQoL-PS	17.3	5.48	16	6.15	1.3	1.77	0.74	0.02	1.45	0.24	0.01	0.47
		HidroQoL	27.2	7.57	24.4	9.89	2.8	3.85	0.73	0.01	1.44	0.37	0.01	0.73
HDSS	**No change**	HidroQoL-DA	8.91	2.87	8.27	3.15	0.64	2.24	0.29	0.04	0.54	0.22	0.03	0.42
		HidroQoL-PS	17.52	5.08	16.58	5.64	0.94	2.96	0.32	0.07	0.57	0.18	0.04	0.33
		HidroQoL	26.42	7.52	24.84	8.49	1.58	4.49	0.35	0.1	0.6	0.21	0.06	0.36
	Minimally worsening	HidroQoL-DA	9.67	1.86	11.17	1.6	-1.5	1.64	-0.91	-1.96	0.14	-0.81	-1.73	0.12
		HidroQoL-PS	18	3.63	19.5	3.73	-1.5	3.83	-0.39	-1.44	0.66	-0.41	-1.52	0.69
		HidroQoL	27.67	4.97	30.67	5.13	-3	5.25	-0.57	-1.62	0.48	-0.6	-1.71	0.51
	Minimally improving	HidroQoL-DA	8.75	2.9	7.7	3.47	1.05	2.44	0.43	-0.04	0.9	0.36	-0.03	0.76
		HidroQoL-PS	17.9	4.24	15.85	5.54	2.05	2.48	0.83	0.36	1.29	0.48	0.21	0.76
		HidroQoL	26.65	6.58	23.55	8.66	3.1	3.85	0.8	0.34	1.27	0.47	0.2	0.75

Table 9.11: Sensitivity of the HidroQoL scores (USA sample)

Anchor	Patient-group	Instrument	Mean	SD	SE mean	95% CI		t	df	p
						Lower	*Upper*			
PGA	Minimally improving (n = 8)	HidroQoL	2.63	4.14	1.46	-0.84	6.09	1.79	7	0.12
		H-DA	1.5	2.78	0.98	-0.82	3.82	1.53	7	0.17
		H-PS	1.13	1.64	0.58	-0.25	2.5	1.94	7	0.09
	No-change (N = 41)	HidroQoL	1.83	5.05	0.79	0.23	3.43	2.32	40	0.03
		H-DA	0.71	2.18	0.34	0.02	1.4	2.08	40	0.04
		H-PS	1.12	3.46	0.54	0.03	2.21	2.08	40	0.04
HDSS	Minimally improving (n = 13)	HidroQoL	3.08	4.01	1.11	0.65	5.5	2.77	12	0.02
		HidroQoL-DA	1.08	2.56	0.71	-0.47	2.63	1.51	12	0.16
		HidroQoL-PS	2	2.45	0.68	0.52	3.48	2.94	12	0.01
	No-change (n = 39)	HidroQoL	1.8	5.04	0.81	0.16	3.43	2.22	38	0.03
		HidroQoL-DA	0.72	2.14	0.34	0.02	1.41	2.1	38	0.04
		HidroQoL-PS	1.08	3.5	0.56	-0.06	2.21	1.92	38	0.06

Note: The minimally worsening group was excluded from analysis due to the small sample size

Table 9.12: Relative efficiency Index of HidroQoL with DLQI and Skindex-17 in detecting change in the minimally improving group.

Patient-group	PGA			HDSS-change-score		
	HidroQoL	*HidroQoL-DA*	*HidroQoL-PS*	*HidroQoL*	*HidroQoL-DA*	*HidroQoL-PS*
DLQI	0.77	0.66	0.83	4.62	2.52	4.9
Skindex-17	2.89	2.47	3.13	2.10	1.14	2.23

The ability of the HidroQoL to discriminate patients experiencing different levels of change

The HidroQoL's ability to discriminate across groups of patients differing in various characteristics such as level of disease severity, amount of daily time spent in managing the condition, has been previously demonstrated (See Chapter 7). In longitudinal HRQoL measurement, it is important for an instrument to be capable of discriminating between patients experiencing different levels of change over time (Stratford and Riddle 2005). A oneway ANOVA test of the change scores across patient groups showed non-significant differences in the overall score as well as domain scores of the HidroQoL (overall score, p = .63; H-PS, p = 0.9; H-DA, p = 0.192) (Table 9.14).

Table 9.13: Estimating the responsiveness of the HidroQoL based on standardised response mean and effect size (USA sample)

Anchor	Patient group	Scores	Test 1		Test2		Test 1 - Test 2		SRM			ES		
			Mean	SD	Mean	SD	Mean	SD	Est.	95% CI		Est.	95% CI	
										Lower	*Upper*		*Lower*	*Upper*
PGA	No change (n = 41)	HidroQoL	27.59	6.00	25.76	7.82	1.83	5.05	0.36	0.05	0.68	0.30	**0.04**	**0.57**
		HidroQoL-DA	9.17	2.54	8.46	3.13	.707	2.18	0.32	0.01	0.64	0.28	**0.01**	**0.55**
		HidroQoL-PS	18.41	4.14	17.29	5.02	1.12	3.46	0.32	0.01	0.64	0.27	**0.01**	**0.53**
	Minimally improving (n = 8)	HidroQoL	26.75	7.89	24.13	10.23	2.63	4.14	0.63	-0.20	1.47	0.33	-0.11	0.77
		HidroQoL-DA	9.88	2.48	8.38	4.14	1.50	2.78	0.54	-0.3	1.38	0.61	-0.33	1.54
		HidroQoL-PS	16.88	5.74	15.75	6.16	1.13	1.64	0.69	-0.15	1.52	0.20	-0.04	0.43
HDSS	No change (n = 39)	HidroQoL	27.36	6.45	25.56	8.24	1.795	5.04	0.36	0.03	0.68	0.28	0.02	0.53
		HidroQoL-DA	9.31	2.61	8.59	3.17	.718	2.14	0.34	0.01	0.66	0.28	0.01	0.54
		HidroQoL-PS	18.05	4.46	16.97	5.37	1.08	3.5	0.31	-0.02	0.63	0.24	-0.01	0.50
	Minimally improving (n = 13)	HidroQoL	26.15	5.93	23.08	8.26	3.08	4.01	0.77	0.16	1.37	0.52	0.11	0.93
		HidroQoL-DA	8.54	2.47	7.46	3.38	1.08	2.57	0.42	-0.18	1.02	0.44	-0.19	1.06
		HidroQoL-PS	17.62	4.29	15.62	5.21	2.00	2.45	0.82	0.21	1.42	0.47	0.12	0.81

Similar analyses were carried out utilising the HDSS change score (HDSS-cs) as an anchor. Significant differences were obtained on the overall score and the psychosocial domain score (overall HidroQoL score, $p = 0.026$; psychosocial domain score, $p = 0.035$; impact on daily activities, $p = 0.05$) showed non-significant differences.

Analyses were also carried out for the USA and the UK samples, separately. In both samples, there were no statistically significant differences in the HidroQoL change scores among patients who were minimally improving, minimally worsening and experiencing no change, based on the two anchors, PGA and HDSS-cs.

The relationship between change scores from the HidroQoL, the DLQI and the Skindex

The ability of the HidroQoL to assess valid change i.e. longitudinal validity was assessed using Spearman's rank sum correlation. The HidroQoL change score showed a correlation of 0.254 with the HDSS change score, 0.263 with the DLQI change score and 0.203 with the Skindex-17 (Table 9.15). Correlations varied across the samples (UK, USA and pooled samples). The correlation of HidroQoL scores with the DLQI and the HDSS was highest in the USA sample, while that with the Skindex-17 was the highest in the pooled group. The correlations involving the HidroQoL total scores was higher than that involving the domains. On the other hand, the daily life activities domain showed lower correlation than the psychosocial domain or the total, for all instruments.

Part II: Development of A Banding System Using An Anchor-Based Approach

The analyses reported in this section to establish the scale banding for the HidroQoL used the sample from the reliability study reported in Chapter 7. A total of 260 participants completed the HidroQoL and the global question (GQ) on overall impact of hyperhidrosis on patient's life. One hundred and forty two patients (54.6%) were from the USA and seventy three (28.1%) were from the UK. A detailed description of the socio-demographic characteristics of this patient population are described in Chapter 7.

Table 9.14: Comparison of amount of change in patients between those with 'slight deterioration', 'no-change' and patients with 'slight improvement': Pooled sample

Anchor			Sum of Squares	df	Mean Square	F	Sig.
PGA	HidroQoL	Between Groups	18.61	2	9.306	.465	.630
		Within Groups	1662.61	83	20.031		
		Total	1681.22	85			
	HidroQoL-DA	Between Groups	17.94	2	8.972	1.683	.192
		Within Groups	442.39	83	5.330		
		Total	460.34	85			
	HidroQoL-PS	Between Groups	1.77	2	.887	.105	.900
		Within Groups	700.04	83	8.434		
		Total	701.81	85			
HDSS	HidroQoL	Between Groups	171.98	2	85.99	4.433	**0.015**
		Within Groups	1687.41	87	19.396		
		Total	1859.39	89			
	HidroQoL-DA	Between Groups	30.60	2	15.302	3.004	**.055**
		Within Groups	443.18	87	5.094		
		Total	473.79	89			
	HidroQoL-PS	Between Groups	59.76	2	29.878	3.493	**.035**
		Within Groups	744.20	87	8.554		
		Total	803.96	89			

Table 9.15: Correlation of the HidroQoL score with the HDSS, DLQI and Skindex-17

	Sample	a	b	c	d	e	f
DLQIcs (a)	*Pooled*	1					
	US	1					
	UK	1					
HDSScs (b)	*Pooled*	.133	1				
	US	.144	1				
	UK	.227	1				
HidroQoL-DA (c)	*Pooled*	.198	.205	1			
	US	.273*	.211	1			
	UK	.003	.065	1			
HidroQoL-PS (d)	*Pooled*	.249*	.230*	.478**	1		
	US	.214	.281*	.561**	1		
	UK	.371	-.030	.308	1		
HidroQoL (e)	*Pooled*	.263*	.254*	.821**	.894**	1	
	US	.267	.285*	.829**	.928**	1	
	UK	.174	.035	.898**	.696**	1	
Skindex (f)	*Pooled*	.522**	.192	.228*	.133	.203	1
	US	.479**	.256	.272*	.073	.172	1
	UK	.579*	.069	-.028	.276	.106	1

Note: **. Correlation is significant at the 0.01 level (2-tailed); *. Correlation is significant at the 0.05 level (2-tailed); Pooled, N = 89; US, N = 54; UK, N = 17

Distribution of the scale scores (HidroQoL, HDSS and GQ)

In the pooled sample, forty six percent of the patients (n = 120) had an HDSS score of 3, indicating that their sweating was barely tolerable and frequently interfered with their daily activities (Figure 9.1). Forty-one percent of the patients (n = 106) had a GQ score of 3, indicating that the largest number of patients were experiencing a large effect on their lives (Figure 9.2). The mean HidroQoL total score was 27 (±6.8) with a range of 1 to 36 in the pooled sample (Table 9.16). The 5-95th percentile shows that most patients had scores in the upper half of the scale range, indicating a positively skewed score distribution (Figure 9.3, Figure 9.4). This is reflected in the mode score of 31, suggesting high QoL impairment. Four percent of pooled sample (n = 11) achieved maximum score while the minimum score was achieved by none. This indicates that there were no ceiling or floor effects associated with the scale. The distribution of the HidroQoL scores in this sample shows a skew to the left. Kurtosis reflects a slightly tighter curve with steeper tails reflecting minor departure from that of a perfect bell-shape as would be expected for a normal distribution. This was only of particular concern for the UK patient population (Kurtosis = 1.53).

Figure 9.1: Distribution of the Hyperhidrosis Disease Severity Score (HDSS)

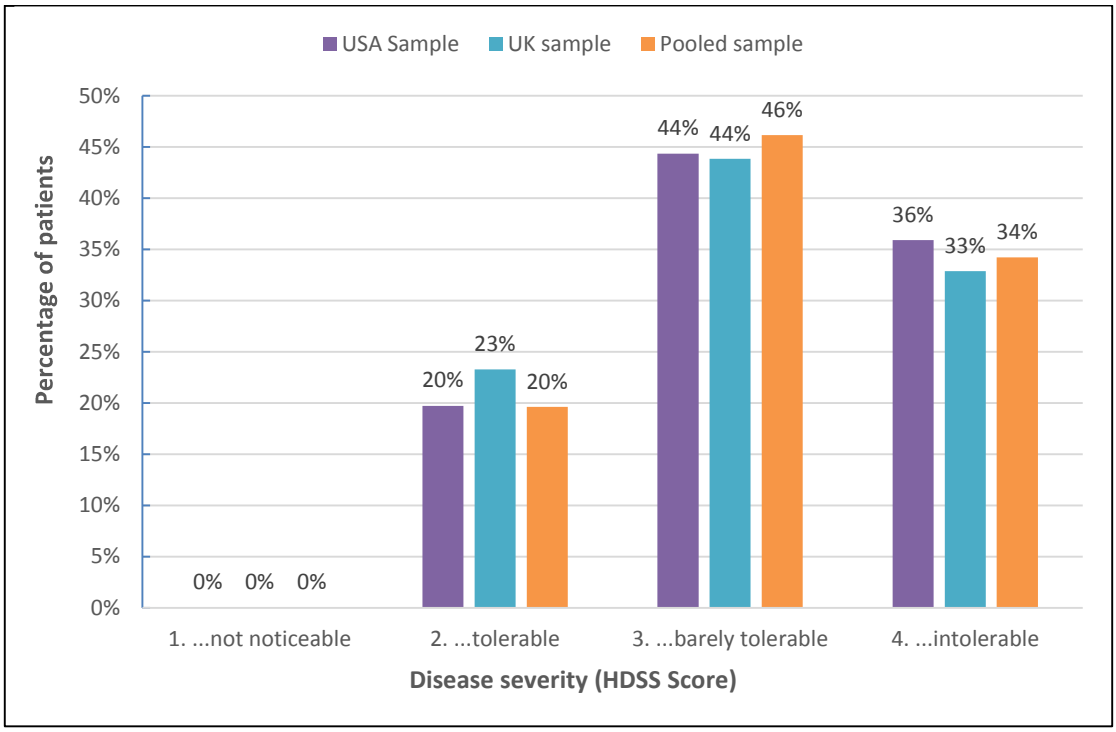


Figure 9.2: Distribution of GQ Score

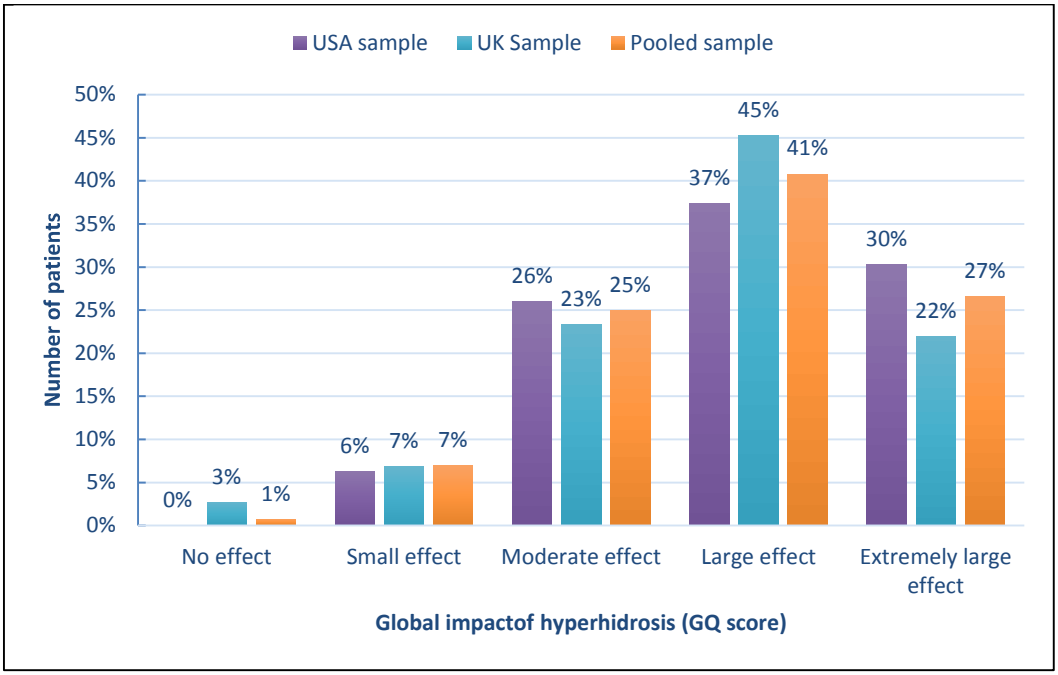


Table 9.16: Distribution of the HidroQoL scores

	Sample		
	USA	*UK*	*Pooled*
N	142	73	260
Mean	27	26	27
Std. Deviation	6.76	6.98	6.76
Std. Error of Mean	.57	.82	.42
Range	2_36	1_36	1_36
Median	29	28	29
Mode	31	29	30
IQR	23 - 32	22 - 32	22 - 32
05-95%	15 - 35	16-36	15 - 35
Skewness Stat.	-1.08	-1.00	-0.95
Skewness SE	0.2	0.3	0.2
Kurtosis Stat.	1.09	1.53	0.88
Kurtosis SE	0.40	0.56	0.30
Ceiling, N (%)	5 (3.5%)	5 (6.8%)	11 (4.2%)
Floor, N (%)	-	-	-

Scale banding for the HidroQoL

Ultimately, an understanding as to what a certain score represents to the patient or clinician, is essential for the application of HRQoL measures in clinical decision-making (Terwee et al. 2007).

Figure 9.3: Distribution of the HidroQoL total scores using box and whisker plot (USA, N = 142; UK, N = 73)

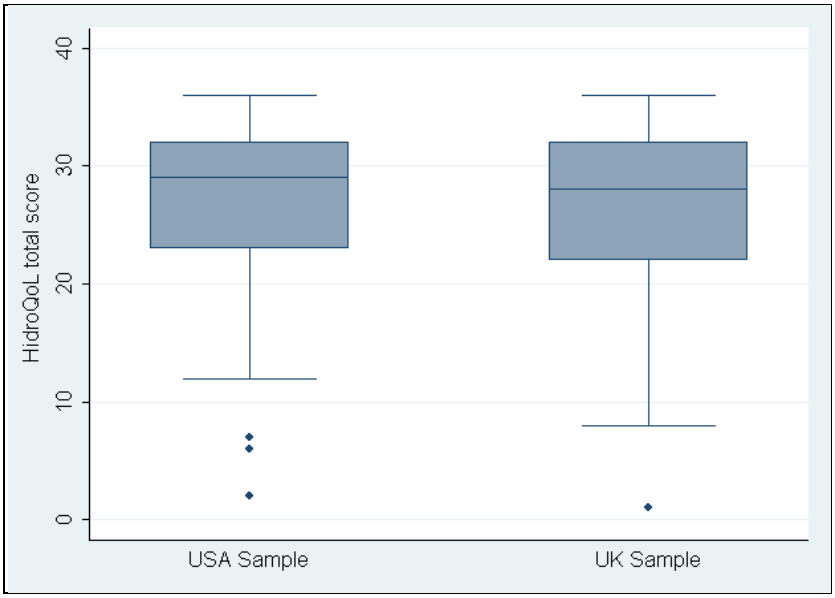
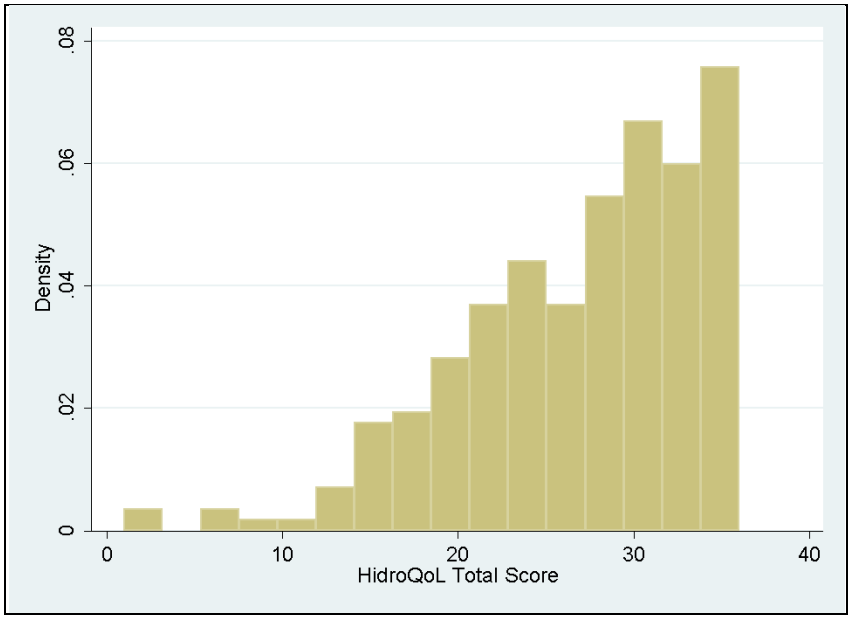


Figure 9.4: Distribution of the HidroQoL total score using a Histogram (pooled sample, N = 260)



Not only does such information enhance the practicality of the instrument, especially in routine clinical practice, but the process of establishing interpretability offers an opportunity for further insight into how the measure and its scores relates to other key measures of disease impact. Thus a categorisation of the HidroQoL scores into bands, related to levels of overall impact experienced by patients was explored. A simple mapping of HRQoL scores to the GQ scores was carried out. The frequency, mean, mode and median GQ score was calculated, for each HidroQoL score. Based on the integer-value of the mean GQ score, each HidroQoL was classified under one GQ score.

HidroQoL-score ranges related to each of the five GQ scores were identified, providing the scale-banding cut-offs. Each HidroQoL score was classified into a single band, corresponding to a level of overall HRQoL impact from Hyperhidrosis: 'no effect at all' (GQ = 0), 'a small effect' (GQ = 1), 'moderate effect' (GQ = 2), 'large effect' (GQ = 3), 'extremely large effect' (GQ = 4). Each band corresponded to a single GQ score (Aawar, 2011). Although previous chapters demonstrated the invariance of the factorial structure (Chapter 5) and item calibration (Chapter 6) of the HidroQoL across countries (UK and USA), separate banding systems were explored for the UK and the USA.

HidroQoL Scale banding for the USA sample

The HidroQoL scores of patients from the USA are presented in Table 9.17 and Figure 9.5. For each score, the distribution of GQ scores is presented including, the mean; its integer-value (mean2); median; and mode. A simple sorting of the HidroQoL score - GQ score table according to the GQ score (integer-value of the mean), allowed the identification of score ranges associated with each GQ score. This provided score bands representing different levels of overall impact of disease on patients' life. A number of HidroQoL scores might have been classified in either of adjacent bands, reflecting the discrepancy between mode, mean and median GQ score. For example, scores 15 and 19 could have been included in band one or band two, scores 20 and 22 could have been classified under band two or band three. Similarly scores 30, 31 and 33 could have been included in band three or band four. This means that a number of alternative bands were feasible (Table 9.18). The level of agreement between the banding (representing predicted level of overall impact based on patient's score) and the GQ score (actual level of impairment) was estimated for each set using ICC. In addition, the accuracy of the banding in classifying patients across their GQ scores (actual level of overall impact) was considered. The highest ICC (0.726)

was seen on the banding 0 to 1, 2 to 11, 12 to 22, 23 to 32, 33 to 36 (set 2 in Table 9.18). This banding also provided the most accurate classification of patients across GQ scores (level of overall impact of condition)

Table 9.17: Frequency, mean, mode and median of GQ scores for each HidroQoL score (USA Sample).

HidroQoL	Number of patients					GQ Scores				Patient-total
	GQ =0	*GQ =1*	*GQ =2*	*GQ =3*	*GQ =4*	*mean*	*mean2*	*median*	*mode*	
0						NA	NA	NA	NA	
1						NA	NA	NA	NA	
2		1				1	**1**	1	1	1
3						NA	**NA**	NA	NA	
4						NA	**NA**	NA	NA	
5						NA	**NA**	NA	NA	
6		1				1	**1**	1	1	1
7			1			2	**2**	2	2	1
8						NA	**NA**	NA	NA	
9						NA	**NA**	NA	NA	
10						NA	**NA**	NA	NA	
11						NA	**NA**	NA	NA	
12			1			2	**2**	2	2	1
13						NA	**NA**	NA	NA	
14			2			2	**2**	2	2	2
15		2	1	1		1.75	**2**	1.5	1	4
16			1			2.00	**2**	2	2	1
17		1	2	1		2.00	**2**	2	2	4
18			1			2.00	**2**	2	2	1
19		1	1	1		2.00	**2**	2	1,2,3	3
20			2	2		2.50	**3**	2.5	2,3	4
21			4	2		2.33	**2**	2	2	6
22			3	3		2.50	**3**	2.5	2,3	6
23			2	3		2.60	**3**	3	3	5
24		1		2		2.33	**2**	3	3	3
25				1	1	3.50	**4**	3.5	3,4	2
26			2	5	1	2.88	**3**	3	3	8
27		1	2	2	3	2.88	**3**	3	4	8
28			1	5	1	3.00	**3**	3	3	7
29			1	4	3	3.25	**3**	3	3	8
30			3	2	4	3.11	**3**	3	4	9
31			3	4	5	3.17	**3**	3	4	12
32			2	8	2	3.00	**3**	3	3	12
33		1		4	5	3.30	**3**	3.5	4	10
34			2	2	7	3.45	**3**	4	4	11
35				1	6	3.86	**4**	4	4	7
36					5	4.00	**4**	4	4	5
Patient-total	0	9	37	53	43					142

Notes: NA, not applicable, no patient with that reported score

Figure 9.5: Relationship between the HidroQoL score and the mean, median and mode of the GQ score for USA patients

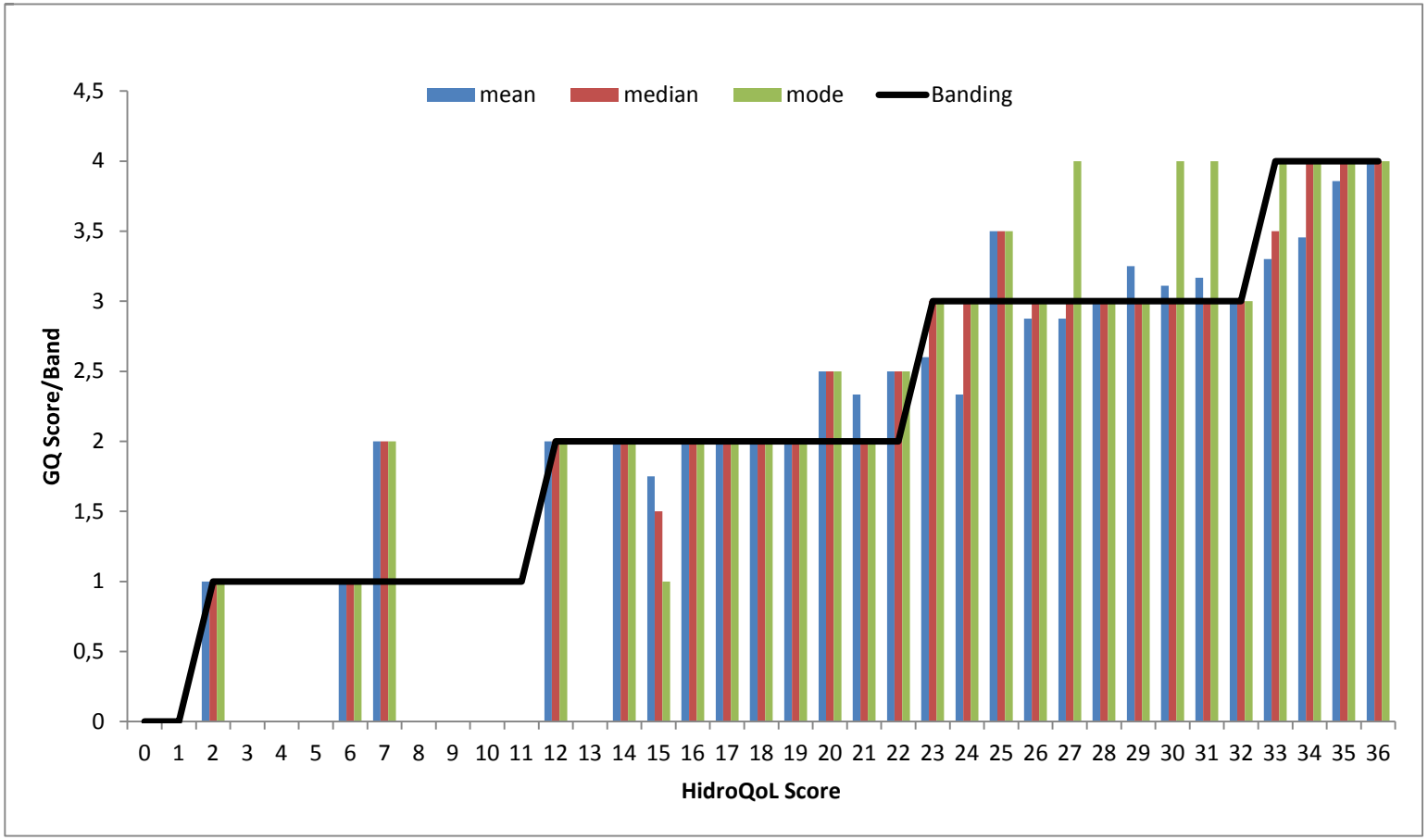


Table 9.18: Alternative HidroQoL score banding for the USA sample

Set	Range of HidroQoL scores					ICC
	Band 0	*Band 1*	*Band 2*	*Band 3*	*Band 4*	
1	0 to 1	2 to 6	7 to 24	25 to 34	35 to 36	0.712
2	0 to 1	2 to 11	12 to 22	23 to 32	33 to 36	0.726
3	0 to 1	2 to 6	7 to 22	23 to 29	30 to 36	0.687
4	0 to 1	2 to 11	12 to 22	23 to 29	30 to 36	0.689
5	0 to 1	2 to 6	7 to 21	22 to 33	34 to 36	0.718
6	0 to 1	2 to 6	7 to 21	22 to 32	33 to 36	0.715
7	0 to 1	2 to 10	11 to 20	21 to 30	31 to 36	0.669

HidroQoL Scale banding for the UK sample

Similar analysis were carried out on the U.K. sample. For each score of the HidroQoL, from 0 to 36, the number of patients with each score and the distribution of their GQ scores (mean, median, mode) was calculated (Table 9.19, Figure 9.6). A number of HidroQoL scores could have potentially fitted in either of adjacent bands. For example, score 15 could have fitted band one, two or three; score 23 could have fitted band two, three or four; score 31 could have fitted band three or four; score 33 could have been in band two or three. This entails multiple ways of categorising the scores into bands of varying overall life impact (Table 9.20). For each possible banding, the level of agreement with actual overall life impact (GQ score) was estimated. The level of agreement ranged from 0.652 to 0.726. Further, the accuracy of each alternative banding in classifying patients to their actual level of overall impact was also analysed. The banding 0 to 1, 2 to 10, 11 to 23, 24 to 33, 34 to 36 (0.726) (set 3) had the highest level of agreement with actual level of overall life impact experienced by patient. This banding also offered the best accuracy in classifying patients. This, therefore, was the recommended categorisation of the HidroQoL scores in patients from the UK

Universal banding system for the HidroQoL

Considering the intended universal use of the HidroQoL across multiple countries, a scale banding system that would apply across different cultures would be advantageous in a number of ways in supporting international clinical trials; and would facilitate comparison of results across countries.

The first step in exploring a universal banding scale was to assess the level of agreement between the two banding systems, calibrated on the UK and USA samples, respectively.

Table 9.19: Frequency, mean, mode and median of GQ scores for each HidroQoL score (UK patients)

H-Score	Number of patients					GQ Scores				Patient-total
	GQ = 0	*GQ = 1*	*GQ = 2*	*GQ =3*	*GQ = 4*	*mean*	*Mean2*	*median*	*mode*	
1	1					0	0	0	0	1
2						NA	NA	NA	NA	
3						NA	NA	NA	NA	
4						NA	NA	NA	NA	
5						NA	NA	NA	NA	
6						NA	NA	NA	NA	
7						NA	NA	NA	NA	
8		1				1	1	1	1	1
9						NA	NA	NA	NA	
10						NA	NA	NA	NA	
11			1			2	2	2	2	1
12						NA	NA	NA	NA	
13						NA	NA	NA	NA	
14						NA	NA	NA	NA	
15						NA	NA	NA	NA	
16		1	1	1		2	2	2	1,2,3	3
17			2			2	2	2	2	2
18						NA	NA	NA	NA	
19				1		3	3	3	3	1
20	1	1		2	1	2.20	2	3	3	5
21		1				1.00	1	1	1	1
22			3	2		2.40	2	2	2	5
23			1	1	1	3.00	3	3	2,3,4	3
24			2	3		2.60	3	3	3	5
25		1	1	2		2.25	2	2.5	3	4
26				1		3.00	3	3	3	1
27				1		3.00	3	3	3	1
28				3	2	3.40	3	3	3	5
29			1	3	2	3.17	3	3	3	6
30			1	2	1	3.00	3	3	3	4
31			1	2	2	3.20	3	3	3,4	5
32			1	4	1	3.00	3	3	3	6
33			2	2		2.50	3	2.5	2,3	4
34						NA	NA	NA	NA	
35				1	3	3.75	4	4	4	4
36				2	3	3.60	4	4	4	5
Patient-total	2	5	17	33	16					73

Notes: H-Score, HidroQoL score; mean2, integer-value of the mean for GQ score

Figure 9.6: Relationship between the HidroQoL score and the mean, median and mode of the GQ score for UK patients

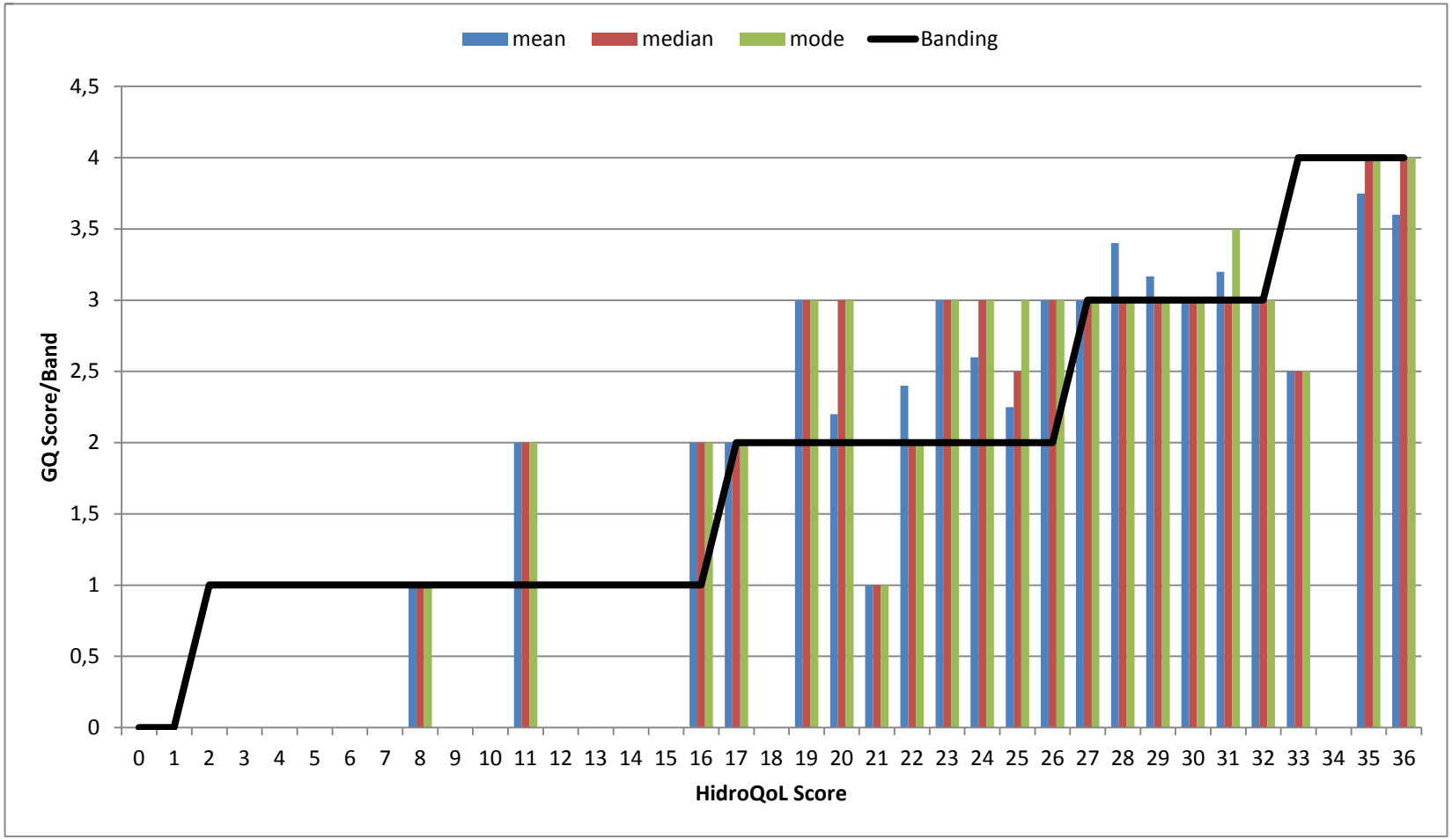


Table 9.20: Alternative HidroQoL score banding based on the UK sample

Set	Range of HidroQoL scores					ICC
	Band 0	*Band 1*	*Band 2*	*Band 3*	*Band 4*	
1	0 to 1	2 to 10	11 to 22	23 to 30	31 to 36	0.692
2	0 to 1	2 to 16	17 to 23	24 to 32	33 to 36	0.691
3	0 to 1	2 to 10	11 to 23	24 to 33	34 to 36	0.726
4	0 to 1	2 to 11	12 to 22	23 to 32	33 to 36	0.704
5	0 to 1	2 to 10	11 to 18	19 to 34	35 to 36	0.670
6	0 to 1	2 to 10	11 to 20	21 to 30	31 to 36	0.652

A near perfect agreement was observed between set 2 (US calibrated) and set 3 (UK calibrated) in both the US sample (ICC = 0.947) and the UK sample (ICC = 0.951). This justified the development of a universal banding system for the HidroQoL, for potential application in the UK, the USA and beyond. Following the approach used in the UK/USA banding systems, for each HidroQoL score, from 0 to 36, the corresponding number of patients and the distribution of GQ scores (mean, median and mode) were estimated (Table 9.21, Figure 9.7). A simple sorting of the HidroQoL score – GQ score table based on the GQ score, was applied to identify the ranges of the HidroQoL score corresponding to each level of overall impact. This provided the banding of the HidroQoL scores, with each score allocated into a single band. There were a few HidroQoL scores where adjacent bands were equally appropriate. For example HidroQoL scores 16, 19 and 20 could have been included in band 2 or band 3. HidroQoL scores 27, 31 and 34 could have been appropriately classified under band three or band four. Thus, there was more than one possible banding-system (Table 9.22). For each banding system the agreement between bands (reflecting a prediction of patients level of impact based on patients score) and patients' GQ scores (representing actual level overall impact) was estimated using ICC. This ranged from 0.633 (set 7) to 0.712 (set 2). Further consideration was made of the accuracy with which the bands classified patients into their actual level of overall QoL impairment (GQ score). These analyses are presented in the next section. The banding system 0 – 1, 2-10, 11-22, 23-32, 33-36 (ICC = 0.679) provided the most accurate classification of patient to their true level of overall impact, and was therefore the recommended banding.

Table 9.21: Frequency, mean, mode and median of GQ scores for each HidroQoL score (pooled sample)

H-score	Number of patients					GQ score				Patient-Total
	GQ=0	*GQ=1*	*GQ=2*	*GQ=3*	*GQ =4*	*mean*	*mean2*	*median*	*mode*	
0						NA	NA	NA	NA	
1	1					0	0	0	0	1
2		1				1	1	1	1	1
3						NA	NA	NA	NA	
4						NA	NA	NA	NA	
5						NA	NA	NA	NA	
6		1				1	1	1	1	1
7			1			2	2	2	2	1
8		1				1	1	1	1	1
9						NA	NA	NA	NA	
10						NA	NA	NA	NA	
11			1			2	2	2	2	1
12			1			2	2	2	2	1
13						NA	NA	NA	NA	
14			3			2	2	2	2	3
15		3	1	1		1.60	2	1	1	5
16		1	2	2		2.20	2	2	2,3	5
17		1	4	3		2.25	2	2	2,3	8
18			3			2.00	2	2	2	3
19		1	1	2		2.25	2	2.5	3	4
20	1	1	5	4	1	2.25	2	2	2,3	12
21		1	5	2		2.13	2	2	2	8
22		1	6	5	1	2.46	2	2	2	13
23			4	5	1	2.70	3	3	3	10
24		1	3	5		2.44	2	3	3	9
25		1	1	3	1	2.67	3	3	3	6
26			2	6	2	3.00	3	3	3	10
27		1	2	4	4	3.00	3	3	3,4	11
28			2	10	4	3.13	3	3	3	16
29			2	8	5	3.20	3	3	3	15
30			4	9	6	3.11	3	3	3	19
31			4	7	8	3.21	3	3	4	19
32			3	13	3	3.00	3	3	3	19
33		2	2	6	5	2.93	3	3	3	15
34		1	3	5	8	3.18	3	3	4	17
35				4	11	3.73	4	4	4	15
36				2	9	3.82	4	4	4	11
Patient-Total	2	18	65	106	69					260

Figure 9.7: Relationship between the HidroQoL score and the mean, median and mode of the GQ score for pooled sample

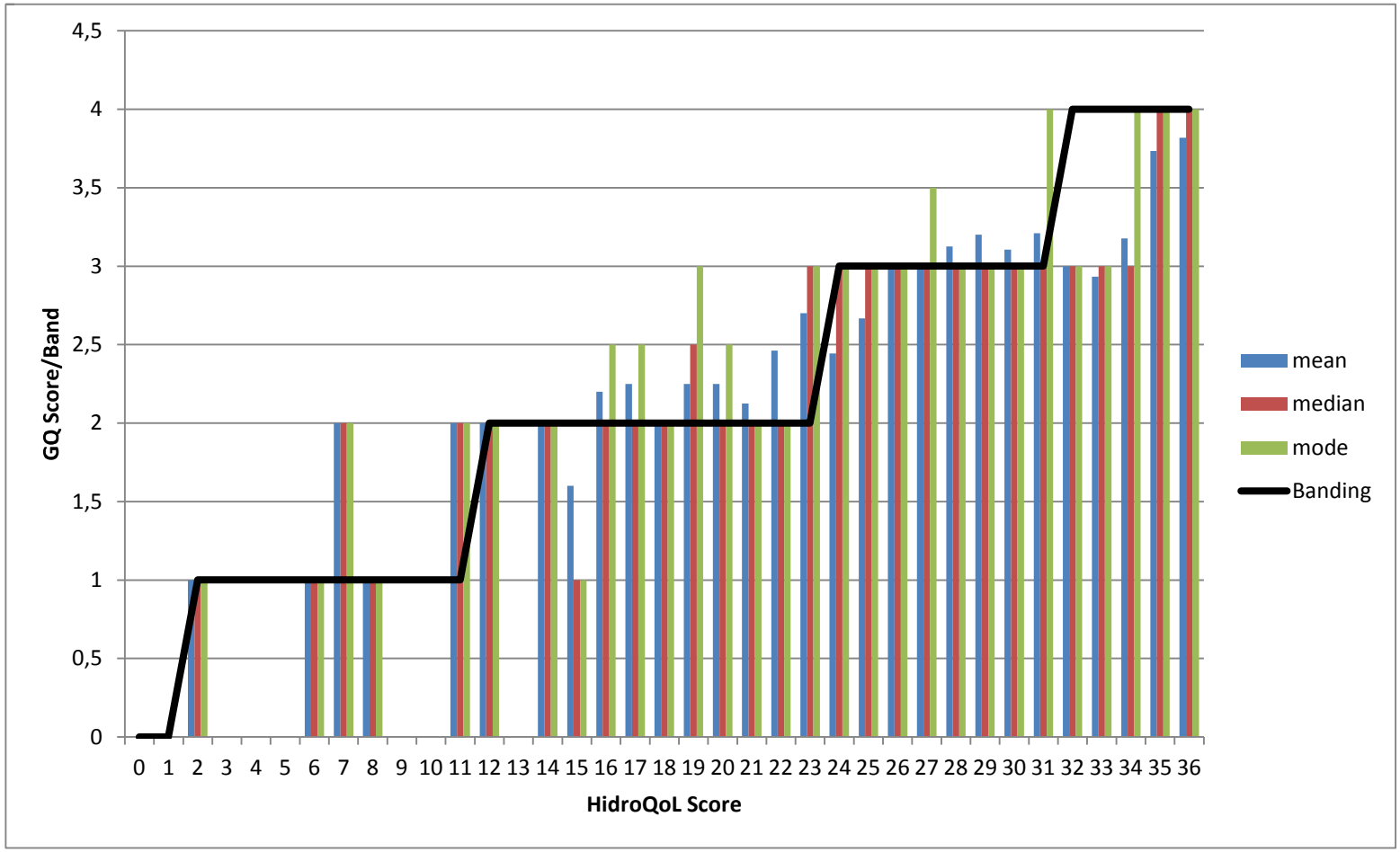


Table 9.22: Alternative HidroQoL score banding based on the pooled sample

	Range of HidroQoL scores					ICC
Set	*Band 0*	*Band 1*	*Band 2*	*Band 3*	*Band 4*	
1	0 to 1	2 to 8	9 to 22	23 to 30	31 to 36	0.666
2	0 to 1	2 to 15	16 to 22	23 to 33	34 to 36	0.712
3	0 to 1	2 to 16	17 to 23	24 to 31	30 to 36	0.677
4	0 to 1	2 to 10	11 to 22	23 to 32	33 to 36	0.679
5	0 to 1	2 to 6	7 to 25	26 to 33	34 to 36	0.699
6	0 to 1	2 to 10	11 to 22	23 to 34	35 to 36	0.700
7	0 to 1	2 to 10	11 to 20	21 to 30	31 to 36	0.633

Assessing accuracy of banding system (pooled sample)

The accuracy of the various banding systems for the pooled sample (presented in Table 9.22) was assessed using the cross-tabulation of each banding (score predicted level of overall impact) against GQ scores (actual level of overall impact). For accurate classification, the majority of patients in each band, must have the corresponding GQ score. On the other hand, for each GQ score, the majority of patients must be captured by the corresponding band. This simple notion embodies key criteria in optimal identification of test cut-offs, that of maximising sensitivity and specificity. For example, situations where the majority/largest group falling in a band are not the corresponding GQ score indicate high false positive rate i.e. poor sensitivity. Conversely, situations where a band does not 'capture' the largest number of patients in a GQ score, reflect a high false negative rate, i.e. poor specificity.

GQ-score – banding cross tabulation

The scale banding 0-1, 2-8, 9-22, 23-30, 31-36 (set 1) showed optimal classification of patients across the bands (Table 9.23). For each band, patients with the corresponding GQ score were the majority. In band 0, GQ = 0 had the highest number of patients. In band 1, GQ = 1 had the highest number of patients.

Table 9.23: Distribution of GQ scores for proposed scale banding (set 1)

		GQ Score					
		0	*1*	*2*	*3*	*4*	Total
Band	0	1	0	0	0	0	1
	1	0	3	1	0	0	4
	2	1	9	32	19	2	63
	3	0	3	20	50	23	96
	4	0	3	12	37	44	96
	Total	2	18	65	106	69	260

The same applies to bands two, three and four. On the other hand, across the GQ scores, the corresponding band had the most patients only for GQ score two, three and four. Overall, this banding system classifies a disproportionately high number of patients as experiencing extremely high impact (band 4), while allocating a disproportionately smaller than expected number of patients into the 'small effect' bracket (band 1).

The distribution of GQ scores across the banding scale 0 to 1, 2 to 15, 16 to 22, 23 to 33, 34 to 36 (Set 2) is presented in Table 9.24. Bands zero, two, three, four, showed the optimal distribution of GQ scores i.e. the respective GQ scores (zero, two, three, four) had the largest number of patients. In band one, patients with GQ score two were in majority (n = 7). The distribution of bands within each GQ score was as expected only for GQ score three, where most patients were in band three (n = 29). This banding system had most patients classified under band three (indicating large effect).

The tabulation of the banding 0-1, 2-16, 17-23, 24-31, 32-36 (set 3), against GQ scores is presented in Table 9.25. In bands zero, two, three and four, the associated GQ scores showed the highest number of patients. This patterns was disturbed only in band one where GQ score 2 showed the highest number of patients. On the other hand, within each GQ score, the corresponding bands had the highest number of patients, except for GQ score zero. Bands zero and two had a single patient each, in band zero. Overall this offered a balanced distribution of patients across bands comparable to distribution of GQ scores.

Table 9.24: Distribution of GQ scores for proposed scale banding (set 2)

		GQ					Total
		0	*1*	*2*	*3*	*4*	
	0	1	0	0	0	0	1
	1	0	6	7	1	0	14
Band	2	1	6	26	18	2	53
	3	0	5	29	76	39	149
	4	0	1	3	11	28	43
Total		2	18	65	106	69	260

Table 9.25: Distribution of GQ scores for proposed scale banding (set 3)

		GQ Score					Total
		0	1	2	3	4	
	0	1	0	0	0	0	1
	1	0	7	9	3	0	19
Band	2	1	5	28	21	3	58
	3	0	3	20	52	30	105
	4	0	3	8	30	36	77
Total		2	18	65	106	69	260

The relationship between GQ scores and banding system 0 to 1, 2 to 10, 11 to 22, 23 to 32, 33 to 36 (set 4) is shown in Table 9.26. All bands showed expected pattern i.e. the matching GQ score had the highest number of patients. Similarly, an analysis of the spread of bands within each GQ score showed expected pattern for GQ scores one, two, three, four. In GQ score 0 band zero and band two had a single patient each. This banding system offers a well-balanced and spread of patients across the different bands reflecting the spread in patients GQ score.

The distribution of GQ scores across the banding 0 – 1, 2 – 6, 7 – 25, 26 – 33, 34 – 36 (Set 5) is presented in Table 9.27. This banding scale showed a proportional spread of GQ scores across the bands, such that in each band, the corresponding GQ score had the highest number of patients. In contrast, the vertical pattern of the matrix was not as expected for GQ scores zero, one and four. For patients with GQ score 0 there were equal numbers falling in bands zero (n = 1) and two (N =

1), as previously noted. The highest number of patients with GQ score one was classified into band two, while among those with GQ score four, band three was the most prevalent.

Table 9.28 shows the distribution of GQ score across the banding 0 to 1, 2 to 10, 11 to 22, 23 to 34, 35 to 36 (Set 6). The spread of GQ scores across the bands was proportional i.e. for each band the highest number of patients showed the matching GQ score. Further the distribution of bands for each GQ score was also assessed. Only for GQ scores one and three showed the highest number of patients in the matching bands (band one and band three). GQ scores zero and three showed equal numbers of patients in bands zero and band two; and in bands two and three, respectively. Overall this banding resulted in a disproportionally large number of patients placed in band three at the expense of all other bands.

Table 9.26: Distribution of GQ scores for proposed scale banding (set 4)

		GQ					Total
		0	*1*	*2*	*3*	*4*	
Bands	0	1	0	0	0	0	1
	1	0	3	2	0	0	5
	2	1	9	35	24	3	72
	3	0	3	20	52	30	105
	4	0	3	8	30	36	77
Total		2	18	65	106	69	260

Table 9.27: Distribution of GQ scores for proposed scale banding (set 5)

		GQ					Total
		0	*1*	*2*	*3*	*4*	
Bands	.00	1	0	0	0	0	1
	1.00	0	2	0	0	0	2
	2.00	1	12	41	32	4	90
	3.00	0	3	21	63	37	124
	4.00	0	1	3	11	28	43
Total		2	18	65	106	69	260

Table 9.29 presents the distribution of GQ scores across the banding system 0 – 1, 2 – 10, 11 to 20, 21 to 30, 31 - 36 (set 7). In each band, the corresponding GQ score was most prevalent. This pattern however was not portrayed in the distribution of bands within each GQ score. Bands zero and two had a single patient each in GQ score zero, band two had the highest number of patients in GQ score one. GQ two had most patients allocated in band three. Although this was the most practical banding it led to a higher proportion of patients being classified in band three and four disproportionately more than the distribution of GQ scores.

The distribution of GQ scores for the banding 0 – 1, 2 – 10, 11 – 22, 23 - 32, 33 – 36 (set 8) is presented in Table 9.30

.

Table 9.28: Distribution of GQ scores for proposed scale banding (set 6)

		GQ					Total
		0	1	2	3	4	
PB6	.00	1	0	0	0	0	1
	1.00	0	3	1	0	0	4
	2.00	1	9	32	19	2	63
	3.00	0	6	32	81	47	166
	4.00	0	0	0	6	20	26
Total		2	18	65	106	69	260

Table 9.29: Distribution of GQ scores for proposed scale banding (set 7)

		GQ					Total
		0	1	2	3	4	
Bands	.00	1	0	0	0	0	1
	1.00	0	3	1	0	0	4
	2.00	1	7	21	12	1	42
	3.00	0	5	31	57	24	117
	4.00	0	3	12	37	44	96
Total		2	18	65	106	69	260

Table 9.30: Distribution of GQ scores for proposed scale banding (set 8)

		GQ					Total
		0	*1*	*2*	*3*	*4*	
Bands	.00	1	0	0	0	0	1
	1.00	0	3	1	0	0	4
	2.00	1	9	32	19	2	63
	3.00	0	3	27	70	34	134
	4.00	0	3	5	17	33	58
Total		2	18	65	106	69	260

The spread of GQ scores for each band appears balanced, the corresponding GQ score had the highest number of patient. Furthermore, the vertical pattern of the matrix was assessed for the distribution of bands in each GQ score. The band corresponding to the GQ score showed the highest number of patients only for GQ score two (n = 32) and GQ score three (n = 70).

Return Operating Characteristic Curve Analysis

An analysis using a Receiver Operating Characteristic curve was carried out to assess the accuracy of each HidroQoL score, as a cut-off for different levels of overall disease impact (based on GQ scores). This method is used for assessing the internal validity of diagnostic tests, by plotting sensitivity against 1- specificity for each score.

Sensitivity captures the true positive rate, the probability of detecting cases i.e. correct diagnosis of those with a condition (Kumar and Indrayan 2011). On the other hand, specificity measures the false positive rate, probability of correctly identification of non-cases (Kumar and Indrayan 2011). As there are five GQ scores, to carry out the ROC analysis, these were regrouped in multiple ways:

- Grouping 1: GQ score 0 vs. GQ scores 1- 4
- Grouping 2: GQ score 0 -1 vs. GQ scores 2 – 4
- Grouping 3: GQ score 0 – 2 vs. GQ scores 3 – 4
- Grouping 4: GQ score 0 – 3 vs. GQ score 4

Separate ROC curves were estimated for each groupings, to identify the most optimal HidroQoL score cut-offs for classifying patients within each grouping. ROC analysis for grouping 1 was not carried out due to the number of patients with GQ score 0 (n = 2). The ‘proportion of correctly

classified' was used to identify optimal HidroQoL score cut-off. In addition, the score with the highest Youden Index (specificity + sensitivity – 1) is considered to offer the highest discrimination.

Cut-off between GQ score 0 – 1 and GQ scores 2 – 4

According to the area under the curve (AUC) for the ROC curve for grouping 1 (Figure 9.8), the HidroQoL scores provided a better classification of patients between GQ score 0 – 1 and GQ scores 2 – 4 than a random guessing (Area under ROC curve = 0.78) (Table 9.31). This means the scores of the HidroQoL were able to classify patients between those experiencing no or small impact and those experiencing a greater overall impact. HidroQoL score cut-offs >= 7 and > = 11 offered the highest accuracy in classifying patients (percent correctly classified = 93.5% for both scores) (Table 3.2). However, a cut-off score of >= 26 was considered most discriminating (Youden Index = 47.9).

Figure 9.8: Receiver operating characteristic curve for classifying patients between GQ 0-1 and GQ 2 – 4 using the HidroQoL

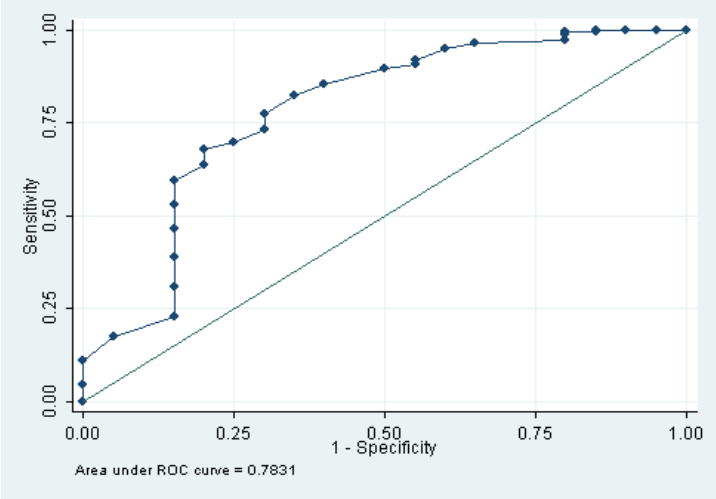


Table 9.31: Area under curve for the ROCs for each grouping

	AUC (CI)	Std. Err.
Grouping 2 cut-off	0.78 (0.66, 0.91)	0.06
Grouping 3 cut-off	0.78 (0.72, 0.83)	0.03
Grouping 4 cut-off	0.79 (0.73, 0.84)	0.03

ROC for grouping 3, GQ score 0 – 2 vs. GQ scores 3 – 4

The ROC for the HidroQOL for grouping 3 (GQ scores 0 - 2 vs. GQ score 3 – 4) shows a statistically significant classification (AUC = 0.78) (Figure 9.9). The cut-off score >23 showed the highest proportion of accurately classified patients (77.3%) although its discriminatory power was not the highest (Youden Index = 43.3%) (Table 9.32). A cut-off score of >=26 showed the greatest discriminatory power although its accuracy was not superior (Youden Index = 46.5%, percent correctly classified = 75.4%).

ROC for grouping 4, GQ score 0 – 3 vs. GQ score 4

The HidroQoL provided a statistically significant discrimination of patients according to their GQ score (GQ scores 0 – 3 vs. GQ score 4) (Area under ROC = 0.79, Figure 9.10). The highest proportion of patients was correctly classified at cut-off >= 35 although this did not have the highest discriminatory power (Youden Index = 25.9%). The highest discriminatory power was seen on cut-off ≥ 27 (Youden Index = 42.1%, proportion of patients correctly classified = 61.5%).

Part III: Establishing MCID for The HidroQoL

The relationship of the HidroQoL with HDSS change score and the PGA was explored in the previous section, a small-to-moderate correlation (r = -0.244, p = 0.021) was reported for the HDSS-cs, while that for the PGA was small (r = 0.142, p = 0.186). The relationship between the HidroQoL and the PGA raises questions over its validity as measure of change in hyperhidrosis-QoL, threatening its relevance as an anchor. Thus for estimation of the MID, in this study, only the HDSS-cs was used as an anchor.

Anchor-based approach

The estimation of the MID based on the anchor approach involved first grouping patients according their HDSS-cs, score of -1, as slightly improved, score of 0, as experiencing no-change, score of 1, as slightly deteriorating. The mean score change in the slightly improving group provides the MID estimate (Crosby et al. 2003).

Table 9.32: Operating characteristics of the HidroQoL score cut-offs in classifying patients according to their GQ score.

H-Score	Grouping 1				Grouping 2				Grouping 3			
	Sn	*Sp*	*Class.*	*y*	*Sn*	*Sp*	*Class.*	*y*	*Sn*	*Sp*	*Class.*	*y*
(>= 1)	100%	0%	92.3%	0.0%	100%	0%	67.3%	0.0%	100%	0%	26.5%	0.0%
(>= 2)	100%	5%	92.7%	5.0%	100%	1%	67.7%	1.2%	100%	1%	26.9%	0.5%
(>= 6)	100%	10%	93.1%	10.0%	100%	2%	68.1%	2.4%	100%	1%	27.3%	1.1%
(>= 7)	**100%**	**15%**	**93.5%**	**15.0%**	100%	4%	68.5%	3.5%	100%	2%	27.7%	1.6%
(>= 8)	100%	15%	93.1%	14.6%	100%	5%	68.9%	4.7%	100%	2%	28.1%	2.1%
(>= 11)	**100%**	**20%**	**93.5%**	**19.6%**	100%	6%	69.2%	5.9%	100%	3%	28.5%	2.6%
(>= 12)	99%	20%	93.1%	19.2%	100%	7%	69.6%	7.1%	100%	3%	28.9%	3.1%
(>= 14)	99%	20%	92.7%	18.8%	100%	8%	70.0%	8.2%	100%	4%	29.2%	3.7%
(>= 15)	98%	20%	91.5%	17.5%	100%	12%	71.2%	11.8%	100%	5%	30.4%	5.2%
(>= 16)	97%	35%	91.9%	31.7%	99%	16%	72.3%	15.9%	100%	8%	32.3%	7.9%
(>= 17)	95%	40%	90.8%	35.0%	98%	20%	72.7%	18.3%	100%	10%	34.2%	10.5%
(>= 18)	92%	45%	88.5%	37.1%	97%	26%	73.5%	22.5%	100%	15%	37.3%	14.7%
(>= 19)	91%	45%	87.3%	35.8%	97%	29%	74.6%	26.0%	100%	16%	38.5%	16.2%
(>= 20)	90%	50%	86.5%	39.6%	95%	32%	74.6%	27.2%	100%	18%	40.0%	18.3%
(>= 21)	85%	60%	83.5%	45.4%	93%	40%	75.4%	32.6%	99%	24%	43.9%	22.6%
(>= 22)	83%	65%	81.2%	47.5%	91%	47%	76.9%	38.5%	99%	28%	46.9%	26.8%
(>= 23)	78%	70%	76.9%	47.5%	**88%**	**55%**	**77.3%**	**43.3%**	97%	35%	51.2%	31.7%
(>= 24)	73%	70%	73.1%	43.3%	85%	60%	76.5%	44.6%	96%	39%	54.2%	34.9%
(>= 25)	70%	75%	70.4%	45.0%	82%	65%	76.2%	46.4%	96%	44%	57.7%	39.6%
(>= 26)	**68%**	**80%**	**68.9%**	**47.9%**	**79%**	**67%**	**75.4%**	**46.5%**	94%	47%	59.2%	40.8%
(>= 27)	64%	80%	65.0%	43.8%	75%	69%	73.1%	44.3%	**91%**	**51%**	**61.5%**	**42.1%**
(>= 28)	60%	85%	61.5%	44.6%	70%	73%	71.2%	43.2%	86%	54%	62.7%	40.0%
(>= 29)	53%	85%	55.4%	37.9%	62%	75%	66.5%	37.6%	80%	61%	65.8%	40.4%
(>= 30)	47%	85%	49.6%	31.7%	55%	78%	62.3%	32.5%	72%	66%	67.7%	38.4%
(>= 31)	39%	85%	42.3%	23.8%	46%	82%	58.1%	28.6%	64%	73%	70.4%	36.5%
(>= 32)	31%	85%	35.0%	15.8%	38%	87%	53.9%	24.8%	52%	79%	71.5%	30.7%
(>= 33)	23%	85%	27.7%	7.9%	29%	91%	48.9%	19.2%	48%	87%	76.5%	34.7%
(>= 34)	18%	95%	23.5%	12.5%	22%	95%	46.2%	17.6%	41%	92%	78.5%	32.7%
(>= 35)	11%	100%	17.7%	10.8%	15%	100%	42.7%	14.9%	**29%**	**97%**	**78.9%**	**25.9%**
(>= 36)	5%	100%	11.9%	4.6%	6%	100%	36.9%	6.3%	13%	99%	76.2%	12.0%
(> 36)	0%	100%	7.7%	0.0%	0%	100%	32.7%	0.0%	0%	100%	73.5%	0.0%

Figure 9.9: Receiver operating characteristic curve for classification of patients between GQ 0-2 and GQ 3 – 4 using the HidroQoL

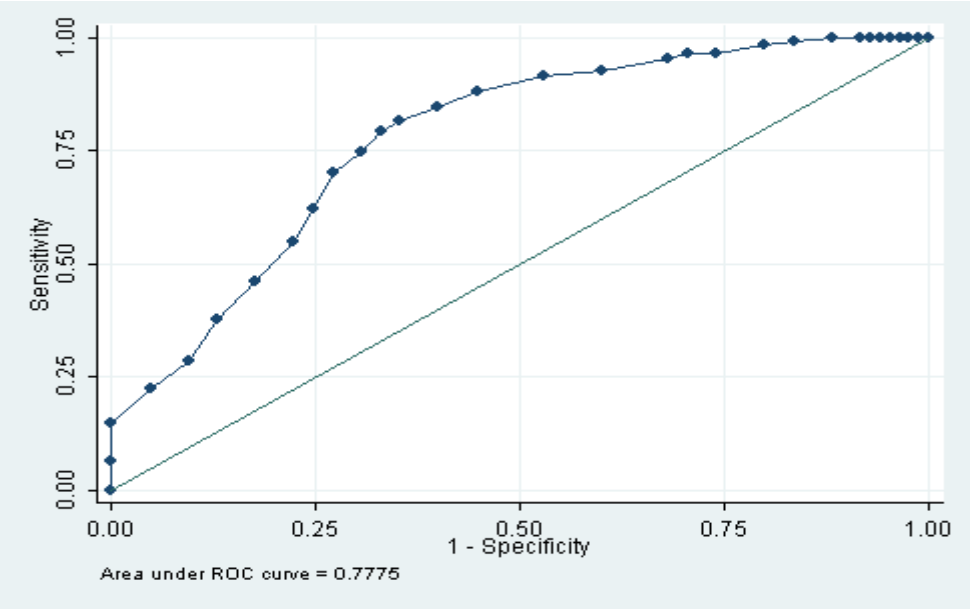


Figure 9.10: Receiver operating characteristic curve for classification of patients between GQ score 0-3 and GQ 4 using the HidroQoL

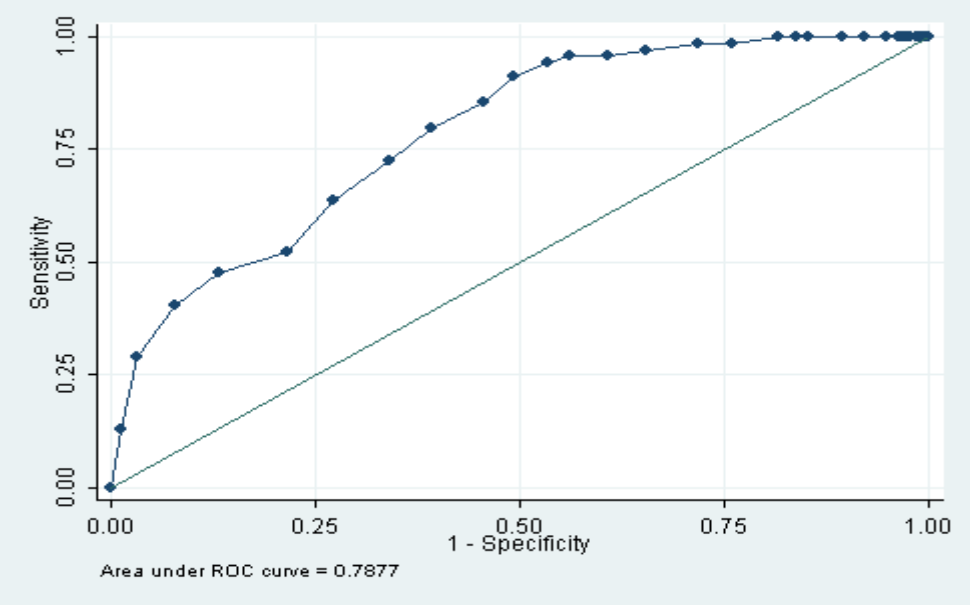


Table 9.33: Mean HidroQoL score change in the 'slightly improving' patient group as an estimate of the MID

Site affected	Sample	N	Mean (Test 1-Test 2)	SD Mean	SE Mean	95% CI of Mean	
All types	Pooled	19	**2.84**	3.78	0.87	1.02	4.66
	US	13	**3.08**	4.01	1.11	0.65	5.50
Localised	Pooled	16	**2.63**	3.67	0.92	0.67	4.58
	US	11	**2.45**	3.91	1.18	-0.17	5.08
Axillary	pooled	14	**2.93**	4.08	1.09	0.57	5.29
	US	10	**3.10**	4.56	1.44	-0.16	6.36

Integrated approach

The MID for the HidroQoL was also estimated by integrating the anchor-based and distribution-based methods i.e. using statistical characteristics of patient groups defined based on the external anchor. The upper bound for 95% CI of the mean HidroQoL-cs of the group that had not changed was estimated as a measure of MCID (Table 9.34) (de Vet et al. 2007). Including all patients in the pooled sample (all types of hyperhidrosis) gave an MID estimate of 2.5, while the corresponding USA and UK estimates were 2.5 and 3.55, respectively. Patients with generalised hyperhidrosis showed MID of 2 in the pooled sample and 1.81 in the US sample. For localised hyperhidrosis, patients from the UK had the most conservative MID estimate, 5.17, while the pooled sample had a MID estimate of 2.94. Due to a small sample size the MID estimate in some patient sub-groups of the UK-sample (generalised and axillary hyperhidrosis) would not be calculated

Distribution-based approach

A third approach used in establishing cut-offs for important change utilised the statistical characteristics of the sample of baseline patient responses (N = 64). Specifically, the standard deviation (1/2 SD and 1/3SD) and standard error of measurement were estimated. The pooled sample, including all types of hyperhidrosis showed a SEM of 2.14 and ½ SD of 3.39. These compared to the figures observed for the U.S (SEM = 2.16, ½ SD = 3.41) and the UK (SEM = 2.22, ½ SD = 3.51). The mean HidroQoL baseline scores of patients generalised hyperhidrosis from the US showed a ½ SD of 2.6 and SEM of 1.65 while a similar group from the UK had ½ SD of 4.12 and SEM of 2.6. For the U.S. patient population with localised hyperhidrosis, SEM

Table 9.34: Upper-bound of 1 tailed 95% CI for the mean HidroQoL-cs in the 'no-change' patient group as an estimate of the MCID

Site affected	Sample	N	Mean (Test1-est2)	SD Mean	SE Mean	Mean 95% CI		mean + 1.645*SE
						Lower	*Upper*	*1-tail, 95% CI*
All types	pooled	64	1.58	4.49	0.56	0.46	2.70	**2.50**
	US	39	1.79	5.04	0.81	0.16	3.43	**3.12**
	UK	11	1.91	3.30	1.00	-0.31	4.13	**3.55**
Generalised	Pooled	15	0.93	2.55	0.66	-0.48	2.34	**2.02**
	US	8	0.13	2.90	1.03	-2.30	2.55	**1.81**
	UK	5	1.40	2.07	0.93	-1.17	3.97	**2.93**
Localised	Pooled	49	1.78	4.93	0.70	0.36	3.19	**2.94**
	US	31	2.23	5.41	0.97	0.24	4.21	**3.83**
	UK	6	2.33	4.23	1.73	-2.10	6.77	**5.17**
Axillar	pooled	20	2.35	5.45	1.22	-0.20	4.90	**4.36**
	US	12	3.25	5.96	1.72	-0.53	7.03	**6.08**
	UK	5	3.00	3.39	1.52	-1.21	7.21	**5.49**
Palmo-plantar	Pooled	45	1.38	3.99	0.59	0.18	2.58	**2.36**
	US	28	1.32	4.50	0.85	-0.42	3.07	**2.72**
	UK	8	1.63	3.02	1.07	-0.90	4.15	**3.38**

of 2.30 and ½ SD of 3.42 were observed. In the comparable group in the UK, SEM of 1.99 and ½ SD were observed. Estimates for patients with axillary and palmo-plantar hyperhidrosis in the US and the pooled samples were very similar to those observed for localised hyperhidrosis. In patients with axillary hyperhidrosis from the UK sample the estimate for ½ SD was 2.3 while that for SEM was 1.45. On the other hand, the UK patients with palmo-plantar hyperhidrosis showed a ½ SD of 3.29 and SEM of 2.08

DISCUSSION

There is consensus regarding the impairment on patient's QoL resulting from skin disease (Finlay and Ryan 1996). Numerous skin-specific and disease-specific questionnaires have, therefore, been developed and validated for assessing QoL impairment. The formal measurement of HRQoL using standardised instruments, however, is yet to be fully integrated into the dermatology clinic (Finlay 2011). This state of affairs might be attributable to a number of issues, institutional (the organisation and delivery of care within the clinic), behavioural (physicians' views and outright inertia) as well as technical (quality and appropriateness of QoL questionnaires in routine clinical practice) (Lohr and Zebrack 2009).

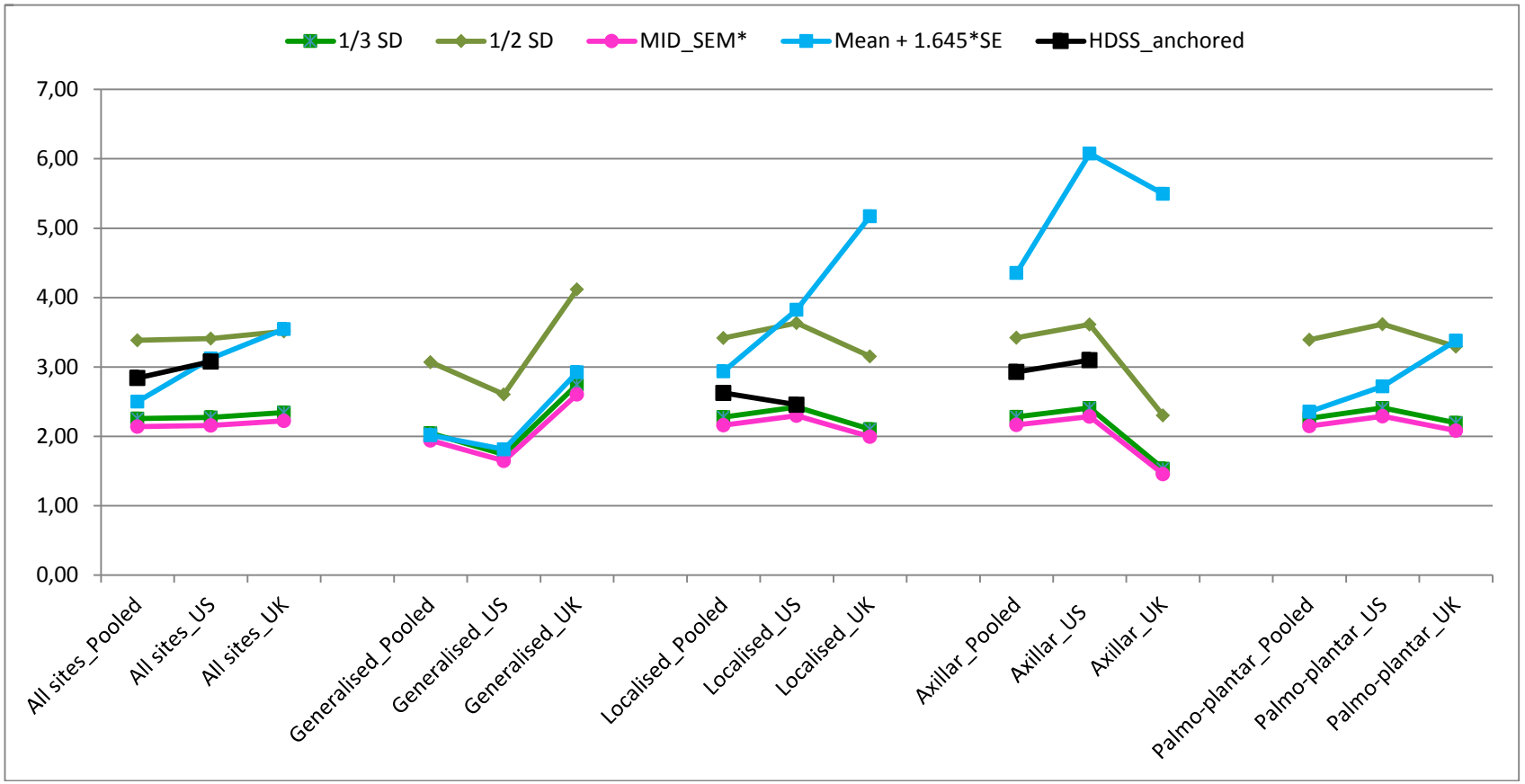
Table 9.35: Measures of precision of the HidroQoL, Standard deviation and Standard Error of Measurement (SEM) as MCID estimates.

Site of hyperhidrosis	Sample	N	Mean	SD	SE	1/3 SD	1/2 SD	SEM*
All sites	**Pooled**	234	27.14	6.77	0.44	**2.26**	**3.39**	**2.14**
	US	137	27.15	6.82	0.58	**2.27**	**3.41**	**2.16**
	UK	56	27.11	7.03	0.94	**2.34**	**3.51**	**2.22**
Generalised-hyperhidrosis	**Pooled**	79	29.30	6.13	0.69	**2.04**	**3.07**	**1.94**
	US	53	29.62	5.21	0.72	**1.74**	**2.60**	**1.65**
	UK	20	28.20	8.24	1.84	**2.75**	**4.12**	**2.60**
Localised-hyperhidrosis	**Pooled**	155	26.03	6.83	0.55	**2.28**	**3.42**	**2.16**
	US	84	25.58	7.27	0.79	**2.42**	**3.63**	**2.30**
	UK	36	26.50	6.30	1.05	**2.10**	**3.15**	**1.99**
Axillar	**Pooled**	71	26.80	6.84	0.81	**2.28**	**3.42**	**2.16**
	US	46	26.11	7.22	1.07	**2.41**	**3.61**	**2.28**
	UK	10	29.60	4.60	1.45	**1.53**	**2.30**	**1.45**
Palmo-plantar	**Pooled**	113	25.58	6.79	0.64	**2.26**	**3.39**	**2.15**
	US	64	25.38	7.23	0.90	**2.41**	**3.62**	**2.29**
	UK	21	25.05	6.58	1.44	**2.19**	**3.29**	**2.08**

Reliability (internal consistency) = 0.9

* SEM = SD * SQRT (1 - Reliability)

Figure 9.11:Estimates for MCID for the HidroQoL across different patient sub-populations and reflecting multiple analytical approaches



Technical issues directly pertains to psychometric attributes of a measure overall and those that are particular for clinical integration of measures e.g. practicality, applicability and interpretability. Instruments used to measure HRQoL in the clinic, for example, in monitoring the patient's condition over time, must have demonstrated the ability to capture important changes in the patient's condition over time regardless of how small they are (Guyatt et al. 1987). Furthermore, additional information may be required for the conversion of the abstract scores into clinically meaningful values that can be incorporated into clinical decision-making (Sampogna and Abeni 2011). The current study, therefore, assessed whether the HidroQoL is capable of capturing change in the patient's condition, even when such change is small. An additional aim was to provide information to facilitate the interpretation of its scores particularly providing estimates for minimal clinically important differences (MCID) and a banding-system for the scores.

In this study, patient's HRQoL was assessed on two occasions, at baseline and on a follow-up visit, two to four weeks apart. Change in the patient's condition was assessed using two external measures. The patient's global rating of change (PGA), where patients rated their perception of the amount of change they experienced from baseline to follow-up, and the HDSS. The validity, reliability and responsiveness of the HDSS has been previously demonstrated (Kowalski et al. 2004; Lowe et al. 2007). A 1-point change is comparable to a 50% change in sweating while a 2-point change is associated with a 80% change in the level of sweating (Solish et al. 2007). Responsiveness was analysed using multiple approaches including paired t-tests, change magnitude coefficients (Effect size, Standard Response Mean); longitudinal discriminant validity based of change scores assessed using ANOVA; and spearman's correlations. This permitted a more refined definition of the change construct, not only evaluating the capability to detect change in patients, but also the capability to differentiate between patients experiencing different levels of change (Stratford et al. 1996).

The results of the paired t-test in the pooled, USA and UK samples, shows that the HidroQOL was sensitive to changes in the patient condition for those who improved, both when change in patients condition was defined using the PGA or using the HDSS. The HidroQoL was more efficient at detecting these changes, relative to the DLQI and the Skindex-17, where change in patient's condition was defined based on the HDSS. Nonetheless, where the PGA was used in defining the patient's condition, the performance of the HidroQoL was comparable to, though slightly worse

than, that of the DLQI, and still more efficient than that of the Skindex-17. The HidroQOL discriminated between patients experiencing different levels of change. The small-to-moderate effect sizes obtained in the minimal improvement group are as expected and demonstrate that the HidroQoL can detect change consistent with true changes in the patient's conditions.

A major concern when applying responsiveness results is whether the capability of a measure to detect change is underpinned by longitudinal validity i.e. that the observed changes in a scale reflect true-changes in the underlying construct (Terwee et al. 2003). The HidroQoL change scores correlated with change scores for the DLQI, Skindex and the HDSS. In addition, previous validation studies (reported in CH 3 – CH 7) reported content validity, dimensional structure and inter-temporal stability of the HidroQoL. These results suggest that the evaluative use of the HidroQoL scores is supported.

The holy-grail to the use of HRQoL in routine clinical practice is the ability of being able to interpret scores: an understanding of what changes in the score from one visit to the next actually means in clinical terms and how the interpretation of change in those scores influences treatment decision-making (Salek and Kamudoni 2013). There are no rules regarding what information serves this purpose, as long as there is a rationale for its validity and clinical relevance. Examples include, reference scores from research studies with similar patients or from the general or healthy populations; comparison with previous scores from patients; scale banding for different levels of impairment; and MCID cut-offs (Snyder et al. 2012). These may be established in a cross-sectional or longitudinal set-up, although the latter is especially suitable for defining 'important' score change, MID (Crosby et al. 2003). In this study, a scale banding-system for different levels of impairment associated with HH and MCID cut-offs were established.

To develop the scale banding-system, patients' responses on the HidroQoL were mapped to a global question (GQ) on patients' perception of the overall impairment on their life resulting from hyperhidrosis. This provided score-bands corresponding to the levels of the GQ question (no effect, small effect, moderate effect, large effect and very large effect). An alternative to this approach is to create score categories based on distribution characteristics of the sample using techniques such as mixture modelling. Although such an approach may seem robust (i.e. offering an efficient means to determining the appropriate number of categories and cut-offs based on model-fit) (Nijsten et al. 2009), the meaning attached to score categories remains arbitrary given

that the assigned interpretation is not underpinned by any patient or clinical input. The main advantage of the approach followed in this study, therefore, was that the qualitative meaning to the score ranges was rooted in patient's judgement.

The banding-system 0 – 1, 2 – 10, 11-22, 23 – 32 and 33 – 36 ($r_s = 0.679$) was proposed. This provided the most accurate classification of patients according to their actual level of impairment i.e. the majority of the patients were placed in the band that would have been predicted by their GQ score (Aawar 2011). There were a number of patients with either a low or high GQ score for their score-range (or the reverse, a high/low score-range for their GQ score). This is possibly a consequence of the particular anchor question used (Prinsen et al. 2011), which might have been interpreted differently by different patients given its generic nature. On the other hand, the questionnaire enabled patients to reflect on aspects of their life in greater detail, picking up issues missed on a generic question. Thus it was possible for a patient showing minimal impairment on the generic question to end up showing greater impairment (and similarly, the opposite was equally possible). Ultimately, this is an inherent feature of any categorisation system; and the degree to which misclassifications are minimised provides a measure of the rigor of any categorisation system.

The definition of the magnitude of important change score is central to the application of HRQoL instruments because, rather than simply knowing that a given change in the scores is beyond that which may be attributable to chance (i.e. is statistically significant), clinical decisions are based on the clinical significance of those changes. Therefore, the MCID score for the HidroQoL was explored, using anchor-based approach, distribution-based (standard deviation – SD; Standard error of measurement - SEM) of the baseline assessment and an integration of the two. Applying both anchor-based and distribution-based methods provided confidence for the proposed estimate. The two approaches provided answers to different questions, namely: what value of the minimal change-score that would be considered important ?; and 'is the instrument capable of detecting such a value ?'(De Vet et al. 2006). Therefore estimates of 1-SEM or ½ SD, representing the Minimum Detectable Change (MDC) score (De Vet et al. 2006).

Across all patient sub-groups the 1-SEM and 1/3 SD, was smaller than anchor-based MID or the one based on integrated approaches. As previously noted, the 1/3SD closely reflected 1-SEM

estimate as expected in a measure with reliability of 0.9 (Yost and Eton 2005). This ranged from 1.65 (USA, generalised hyperhidrosis) to 2.3 (USA, localised hyperhidrosis). Using the anchored approach MID ranged from 2.45 (USA, localised hyperhidrosis) to 3.1 (U.S.A axillary). On the other hand, MID for the the integrated approach ranged from 1.81 (U.S.A, generalised) to 6.08 (U.S.A, axillary). Patients with generalised hyperhidrosis showed slightly lower MID estimates in comparison to those with localised hyperhidrosis, on both the anchor-based and distribution-based estimate. Due to the small size of some patient groups (UK palmo-plantar patients, minimally deteriorating patients, minimally improving UK patients, generalised hyperhidrosis) anchor-based MCID estimates were not established. Taking into account practicality considerations and the available evidence, an MCID of 3 was proposed for the HidroQoL.

SUMMARY

- This study has developed and proposed a banding system for interpreting the scores of the HidroQoL scale.
- Separate banding-scales were proposed for generic and localised hyperhidrosis, however, a strong level of agreement between the two banding systems supported use of a common banding across the different countries.
- An anchor-based approach was used based on a global question (GQ) on the overall HRQoL impact of hyperhidrosis on patient's life, ensuring that the proposed banding is patient-centred and not arbitrary.
- In addition to facilitating the integration patient-centred care into routine clinical practice through greater engagement with the patient on issues affecting them, the banding system suggested may be an important tool in minimising decision uncertainty faced by clinicians especially in assessing their patient's condition.
- The MID of the HidroQoL has been established using anchor and integrated approaches.

CHAPTER 10

General Discussion

In disorders related to the skin, the subjective experience of patients, particularly in relation to impairment in QoL are regarded as a vital sign of disease activity (Chren 2005). First, the severity of most chronic skin conditions is mainly linked to the impact on social life, patient discomfort and psychosocial functional limitations (Grob 2007). Second, 'hard' endpoints such as clinical measures or symptoms like redness or lesion size may be challenging to measure and interpret (Grob 2007). In hyperhidrosis, for example, gravimetric measurement is used for determining severity of disease based on the amount of sweat produced. Not only is the procedure cumbersome for routine clinical practice, but also the amount of sweat produced has shown great intra and inter-individual variability(Wörle et al. 2007). Moreover, the threshold amount of sweating for the diagnosis of hyperhidrosis is unclear (Hund et al. 2002).

Apart from concerns over symptoms, patients may have greater worries regarding their diminishing QoL which might be the main driver in seeking medical attention. The ability of PROs such as HRQoL to take into account a broader spectrum of impacts experienced by the patient beyond just symptoms is not only useful in reducing uncertainty associated with clinical decision making but provides a better framework for evaluating the risks and benefits from therapy.

Given the foregoing, there has been a rapidly growing role of humanistic outcomes in recent years not only in patient management but also in assessment of efficacy of therapies, in health services-audit and epidemiological studies. This suggests that the way in which QoL information is gathered, processed, interpreted, presented and utilised needs to transition from being an art to a science (Finlay 2011). However, the development of a unified conceptual framework for QoL and its measurement has often been hampered by the subjective nature of QoL, including the fact that it is an abstract concept which is associated with personal perception, beliefs and values (Spilker and Revicki 1996). For example, QoL may change over time, even without a corresponding change in the patient's underlying condition, depicting 'response shift'. This makes ensuring objectivity in the measurement process quintessential. Fisher (2000) has discussed two main elements of objectivity in measurement, methodological and social. The former requires that results from measurement are not dependent on the object being measured. On the other hand, social objectivity reflects consensus and agreement on standards and their interpretation as applied to the units of measurement. In the context of QoL, one can translate this as attaining a unified conceptual framework on QoL, guidelines on its measurement and its interpretation. While 'objective

measurement' may be considered idealistic and unattainable in practice, it provides a benchmark and standard for rigorous instruments. This issue is taken up later.

One of the major objectives of this research, therefore, included to evaluate instruments used in assessing QoL in hyperhidrosis. Various types of instruments have been applied in hyperhidrosis, generic, skin-specific and disease-specific. Not all instruments identified measured the concept QoL as claimed, suggesting a wider issue relating to the lack of clarity, consensus and transparency in the definition of the conceptual framework for hyperhidrosis-QoL. Among the generic and the skin-specific measures, only the IIRS, the DLQI and the PBI had been validated in hyperhidrosis patients. Furthermore, the majority of instruments lacked patient's input in their content development. This has implications for content validity and applicability, reflected in poor coverage of core issues or lack of appropriate emphasis. Such instruments may particularly be difficult to use in the clinic (Higginson and Carr 2001). A promising measure, the HHIQ was not developed for use in routine clinical practice, lacked applicability and interpretability. The need for a clear and transparent conceptual framework for QoL in hyperhidrosis; and a new disease-specific instrument for its measurement was made clear through the review. The new instrument would aim to be useful in clinical research, but also adapted and refined for routine clinical practice. Thus it would need to be short, have a simple scoring procedure, psychometrically sound and responsive in individual patient evaluation, with data facilitating the interpretation of scale scores provided.

The position of patients as experts on their condition makes their account of how they experience their disease a rich and important source of information. A literature search of studies investigating the impacts of hyperhidrosis using qualitative research methods was carried out. Only one study recruiting females only was found. The rest of the studies were based on quantitative methods which may not provide information on the inner perceptions, values and beliefs of patients (e.g. underlying patient's self-image, their health needs and priorities). This indicated a need for a qualitative investigation into the experiences of patients living with hyperhidrosis, especially how their life is affected. This would offer a unique opportunity to understand phenomena from the eyes and voice of the subjects, capturing inner thought processes as well as the context influences such as cultural and social norms and beliefs (Bowling 2009). Most crucially, this recognises that long-term QoL outcomes in patients is subject to many influences besides therapy including coping

strategies and accommodation of the disease, internalisation of negative social stereotypes, patient's level of self-esteem; importance attached to appearance (Greenhalgh 2009). Moreover, during the development of a new patient reported outcome measure qualitative investigation into the experiences of patients provides a means for their involvement in content development. The data gathered from patients, particularly, how their condition affects them is useful in defining the conceptual framework of the instrument. The actual phrasing used by patients to describe their condition may also be useful in formulating the items, ensuring that the content is not only relevant but also appropriate, comprehensible and interpretable by the target patient population. Documentation of the evidence demonstrating the link between items in the measure and the impacts experienced by patients is particularly important for instruments used to assess therapeutic benefit in clinical studies submitted to the FDA (Rothman et al. 2009).

In this study, data collection used a triangulation of focus groups, interviews and survey ensuring a balance in the strengths and weakness associated with each method. The interviews conducted were semi-structured, starting with an open question allowing the patient to recall all their experiences and to narrate them according to their perception and prioritisation, ensuring that the final data was authentic. A similar approach was taken during focus group discussions where patients were invited to share their experiences and talk about how their lives had been affected. The interference of the interviewer to the patients' description was minimal and limited to prompts for more clarity. Their phrasing was also such that the patients would not be influenced to provide a particular answer. If anything, the passive listening ear might have encouraged the patients to be candid in their explanation allowing them to share more. Patients initiated discussion on a number of issues e.g. about new upcoming treatments and underlying causes of hyperhidrosis. . Overall the patients were enthusiastic about the discussions/interviews and saw this as an opportunity to contribute towards the general good of all other patients with hyperhidrosis. Participants to the focus groups were grateful for the opportunity to discuss their condition with other sufferers.

Although the selection of patients in qualitative research need not necessarily be probabilistic, this ought to be at least purposive, in order to achieve a sample that reflects all the key characteristics of the target patient population. The overriding consideration is whether the sample is 'adequate' to supply all information needed for a comprehensive analysis (Yardley 2008). Achieving 'data saturation' helps to demonstrate this. In this study data saturation was demonstrated; data collection

continued even after saturation was reached, and no new additional themes emerged. This ensured the thoroughness and depth in the information collected.

The QoL issues reported by patients reflected strong social underpinnings. The most frequent emotional distresses such as anxiety were associated with patient's uncertainty about when the sweating would start and how other people would react to it. Patients often felt embarrassed particularly because they thought others had noticed their sweating. The importance of anxiety in hyperhidrosis is reflected in earlier theories on the condition, which considered it to be primarily a psychological condition (e.g. social anxiety) (Ruchinskas 2007). This is understandable considering that palmar hyperhidrosis is a key symptom of pathological social anxiety.

Patients have reported physical discomforts related to hyperhidrosis, for example having drenched clothes and the related unpleasant body odour. Patients with plantar hyperhidrosis (affecting the feet) reported discomfort associated constant wet feet, often leading to bad feet odour and to athletes feet in a few. The condition also had an impact on patient's work-life and career choices. Patients suffered reduced productivity due to challenges with using computers, smart screens or working with paper document. Relating with colleagues or clients was taxing. Patients had their own 'little rituals' just to keep dry or to avoid their sweating from being noticeable, which required extra effort, work and time e.g. showering more than once a day, changing clothes more than once a day, carrying a towel everywhere. Patients with hyperhidrosis spend 15 to 60 minutes in managing the symptoms of the condition and more than 50% change their clothes more than twice a day (Hamm et al. 2006). Comparable experiences have been reported in other skin diseases. In psoriasis, daily time needed for treatment was found to be the strongest predictor of HRQoL, in a cross-sectional study involving 1210 patients in 130 dermatology practices in Germany (Blome et al. 2010). The reported impacts did not seem to diminish with age of patient, suggesting that the amount of accommodation taking place might be minimal

The impacts of hyperhidrosis cross-cut multiple areas of life with a common linkage to patient's social life. Amir (Amir et al. 2000) using regression analysis showed that impairment in social functioning alone explained 81% of subjective suffering in hyperhidrosis patients, based on 48 Israeli patients attending a dermatology clinic. This emphasises the need for tools that comprehensively address such multidimensionality in disease-impacts and in particular e.g. assessing HRQoL to understand the broad impacts of the condition on multiple aspect of the

patient's life, simultaneously. Modern society's emphasis on healthy skin as part of sexy-perfect-body image; the high visibility of skin; and the importance of skin to self-identity provide some explanation for the high impairment in QoL suffered by patients with skin disease (Beltraminelli and Itin 2008). For hyperhidrosis, patients also deal with the fact that sweating in and of itself is associated with lack of hygiene. The results of the qualitative study provided a rich source of material for the development of a conceptual framework for QoL in hyperhidrosis and a new instrument for its measurement. A clear and structured process was followed in the development of the first version of the new instrument from the identified QoL issues, to ensure not only the appropriate coverage and emphasis in its content but also technically quality.

Clarity regarding the internal structure of an instrument not only reflects the rigor of the conceptual framework and its translation into measurement, but also provides the rationale for combining the items into domain or overall scale scores (Lohr 2002). The developmental version of the HidroQoL underwent field testing in the target population (comprised of patients from the U.S. and the UK) to test its internal structure as well as the relevance and acceptability of the content. This also facilitated the revision of the instrument, eliminating items not contributing to measurement and retaining those such a contribution. The study population used had self-assessed hyperhidrosis and the majority (85%) had previously seen a doctor for their condition.

In order to perform item reduction using the classical test theory approach, the subjects were divided into two groups. The first, comprising of patients from the USA, was used for the exploratory factor analysis (EFA), and the second group, made up of patients from the UK, was used for confirmatory factor analysis (CFA). Prior to the EFA, redundant items were removed based on results of correlation analysis, consideration of content coverage and importance of the issue to patients. For example the item *my self-confidence is affected* and my *self-esteem is affected* were highly correlated. The *self-confidence item* was more prevalent during the qualitative study, making it the preferred item. EFA was then carried out on the remaining 36 items to explore the number of factors underlying the HidroQoL, as well as to assess the role of the items to measurement. Items not meeting criteria were sequentially removed, with 21 items retained. These items showed optimal fit to both a single-factor as well as a two-factor structure. Although the single factor solution was based on parallel analysis (considered a more robust factor extraction approach), the two factor solution, offered more insight into the nature of hyperhidrosis impacts.

Moreover, the two factors were interpretable as *impact on daily life activities* and *psychosocial impact*.

Although EFA is informative as an exploratory tool, confirming hypothesised number of factors can only be undertaken using confirmatory factor analysis. As such, the CFA carried out on the UK subsample tested the single and two factor solutions observed from the EFA. Both the single factor and two factor solutions showed optimal fit. Still, the latter showed better fit, which might be due to the inclusion of more parameters than the single factor solution.

Rasch analysis was carried out on the 36-item HidroQoL (following the removal of multi-collinear items). Performing the analysis on this version of the HidroQoL ensured comparability with the EFA; and guaranteed that the first stage of item reduction still incorporated qualitative consideration. Items showing poor fit to the Rasch model (RM) were identified and removed. This was done sequentially, one item at a time, taking into account impacts on content validity, impairment continuum covered by the scale and impact on the reliability. This provided thorough insights into the contribution of each item to the conceptual definition of the target construct. The Rasch analysis allowed the conceptualisation of hyperhidrosis-QoL as a construct relevant to all types of hyperhidrosis. Therefore the realised conformity to the RM demonstrated the unidimensionality of the HidroQoL.

In as much as recommendation on the most optimal response categorisation for QoL instrument exist, and suggest seven (plus/minus 2) (Streiner and Norman), such guidance must consider the target population and concept being measured. The consequences of an inappropriate categorisation are costly both in terms of measurement efficiency and time (response burden). Following the Rasch model the number of response categories was reduced from 5 to 3, such that the responses for items were ‘no, not at all’, ‘a little’ and ‘very much’. Based on the RM, this new response categorisation minimised ambiguity.

The item calibration of the HidroQoL on the RM was cross-validated on a fresh sample, comprised of patients from the UK. A comparison of item hierarchies showed that the majority of items retained their level of difficulty, five items showed a shift in their item difficulty locations. For example, the item *I worry about people’s reactions* was slightly difficult for the UK group, while

the item *I feel frustrated* seemed easier. Taken literally, this may mean UK patients cared less about other people's reactions compared to the US patients, while they easily got frustrated in comparison to their US counterparts. An alternative approach to invariance compared the item calibrations from the two samples using a scatter plot and quality control lines, accounting for measurement error. This showed that all items except one, were invariant within measurement error. The observed differences in item difficulties in some of the items might have been a result of the small size of the UK sample. The results from the analysis taking into account measurement error supports this.

Using two alternative approaches in establishing dimensionality was important, not only as a means for cross-validating results from the two approaches, but also because the two methods provide slightly different perspectives on the same issues. In the RA, for example, all hyperhidrosis-site-specific item showed poor fit suggesting that they were not assessing the same Rasch latent variable (hyperhidrosis-QoL). During the EFA these items all belonged to a single factor. The EFA, however, would not indicate whether this factor was part and parcel of a broad QoL construct relevant for all forms of hyperhidrosis or not.

Although the Rasch analysis and EFA produced slightly different instruments, eleven items were common. The major difference was in items assessing the psychosocial impact domain. One reason for this might be the fact that the Rasch model assesses whether an item is used consistently, in line with Rasch probabilistic condition i.e. whether patients with greater impairment have a higher probability of a higher score than those with a lower impairment (Tennant et al. 2004). Furthermore, the RM conceptualises the latent variable as a linear metric measuring the latent variable/construct from a low to high severity level; with items placed hierarchically on the metric according to their level of difficulty (Pallant and Tennant 2007). In contrast, the FA linear model does not accommodate the latent variable's severity dimension; it makes no consideration of item difficulty; and thus lacks the capability to deal with item redundancy. Since FA assesses items based on shared covariation, those where this is low may be penalised despite their contribution to overall scale for example, the item *My sex life is affected*.

The final version of the HidroQoL utilised the taxonomy from the EFA to provide two sub-scales, impact on daily life activities and psychosocial impact, in addition to the overall scale. The choice

of items was based on the Rasch analysis, in order to simultaneously take into account the entire continuum of impairment in HRQoL and realise a unidimensional construct. Three items were added on the fifteen selected based on RA optimisation, *my physical activities are affected*, *I feel embarrassed* and *my choice of clothing is affected*. The first two, were included in the FA reduced instrument. Although the RM showed some response dependence between the item *my physical activities are affected* and *my hobbies are affected,* the two items represent separate and mutually exclusive concepts. The items *I feel embarrassed* and *my choice of clothing is affected* emerged as the most prevalent themes during qualitative research, thus their omission might have negatively impacted content validity and applicability of the instrument. Thus the process of selecting items for the final version of the HidroQoL and the development of a measurement model, explicitly addressed the friction between the qualitative and quantitative methods as well as between the classical test theory and modern test theories, applied in this study. The most statistically viable measurement model was implemented, but not at the neglect priorities of patients.

The internal structure of the construct of hyperhidrosis-specific Quality of Life has been previously explored. Kuo et al. (Kuo et al. 2004) using EFA, identified five domains including: functional, psychological, social, affective and physical function explaining 69% of QoL, with the functional domain explaining most of the variance (42%). The contents of all domains, except for the 'physical domain', are covered by the HidroQoL. Amir et al. (2000), on the other hand, included six domains (functional, social, inter-personal, emotional-self, emotional-other and conditions) in their conceptualisation of hyperhidrosis-QoL. Results of a regression analysis they carried out showed subjective suffering from hyperhidrosis to be explainable by three factors (social, interpersonal and emotional-other); with the social domain accounting for 80% of the variation. The studies by Kuo et al. and Amir et al. included items site-specific questions e.g. items relevant for palmar or plantar hyperhidrosis only. The findings from the current study have shown that special considerations might be required for including such items if the intention is to measure HRQoL across patients with hyperhidrosis of different sites.

The final phase of the study involved establishing the psychometric properties of the final HidroQoL version, testing its reliability, construct validity and responsiveness. In addition information facilitating interpretation of scores was generated. Patients with all types of hyperhidrosis, according to body site affected, were included. This is particularly important given

that the HidroQoL is intended for use in all forms of hyperhidrosis. Moreover, the different sites of hyperhidrosis tend to have slightly differing prognosis and impact. In this research the level of impact reported by patients with generalised hyperhidrosis was the highest while those experiencing sweating of the hands and feet were the least affected.

Reliability was tested by assessing the internal consistency and test-retest reliability of the domain scores as well as the full scale score. In view of the influence of sample distribution characteristics, for example heterogeneity of sample, on reliability, separate analyses were carried out for patient from UK and USA patients, in addition to analysis in a pooled sample. The obtained results supported homogeneity of the domain overall scale score and the two domains (*impact on daily life activities impact* and *psychosocial impact domains*). These results demonstrate the clarity with which the construct being assessed by the HidroQoL has been defined.

Test-retest reliability particularly relevant for measures used in a longitudinal context. Its assessment needs to take into account a number of issues which may confound reproducibility. The period of time between baseline and follow-up should be close enough to ensure that the underlying construct does not change, but not too close to avoid carry-over effects from initial assessment (Streiner and Norman 2008). A period of three to fourteen days has been recommend, with a one week offering a good balance (Salek and Luscombe 1992; Streiner and Norman 2008). The choice of the correlation coefficient has to be appropriate, the ICC fits the purpose, as it also captures systematic bias. In this study, time interval from baseline to follow-up ranged from 5 to 25 days. The ICC was the choice coefficient, the results showed strong test-retest reliability for the individual item scores, the domain scores; and the overall score. This provides confidence that the HidroQoL can be used for evaluative purposes as change taking place in the patient's life would not be obscured by measurement or systematic errors.

The validity of the construct underlying the HidroQoL (hyperhidrosis-specific QoL) was thoroughly evaluated by testing various a priori hypothesis, on how the HidroQoL scores relates to other measures of constructs related to the hyperhidrosis-specific QoL. A moderate to strong correlation was obtained between scores of the HidroQoL and scores of the DLQI and the Skindex-17. This confirmed the apriori hypothesis that hyperhidrosis-QoL is related to skin-specific QoL.

Disease-specific scales are expected to reflect issues of particular concern to patients with a given condition(Guyatt et al. 1993). Therefore a moderate to strong relationship was expected between the scores of HidroQoL and other measures of disease severity and impact in hyperhidrosis. The HidroQoL showed expected relationship with level of disease severity (HDSS score), overall impact of hyperhidrosis on patient's life and the amount of time spent in managing the sweating daily. The results observed suggests that the content of the HidroQoL: addresses those features of QoL which are linked to level of disease severity, is representative of those issues pertinent to hyperhidrosis patients such as overall life impact and the daily amount of time spent in managing hyperhidrosis. This suggests that impairment in patient's QoL, as measured by the HidroQoL, would be a plausible indicator of disease activity in hyperhidrosis, providing the justification for its use as a primary endpoint in clinical trials.

Application of a measure for evaluative purposes presupposes that it has an ability to detect clinically meaningful changes (Guyatt et al.). The HDSS score change and a retrospective patient's global change assessment were used as anchors to determine the degree to which the patient's condition had changed. Baseline to follow-up HidroQoL score changes were as expected for patients whose condition had minimally improved; but not in the 'no change' or 'minimally worsening' groups. The HidroQoL was capable of discriminating between patients only where the HDSS change scores was used as an anchor. Although this may be regarded as a sign of poor specificity, it is, nonetheless, equally plausible that the anchors were not offering the best discrimination among patients. The moderate correlation observed between the change scores of the HidroQoL and those of the Skindex-17 and DLQI offered confidence that score change were valid. This evidence demonstrates that the HidroQoL has the capability to detect important changes even if they were small.

Interpretation of QoL scores may require different types of information, addressing the various applications that a measure might be subjected to. Clinicians want to know what a given magnitude of score tells them about how their patient is doing. For example consideration of whether a score change seen in patients from one visit to the next is clinically significant, may require estimates of minimal clinically important difference (MCID). To facilitate the interpretation of the HidroQoL, MCID cut-off scores and a scale banding system have been proposed. Two anchors were used, the HDSS, for establishing the MCID; and the GQ, for the development of a scale score categorisation.

Both anchors were easily understood by the subjects, interpretable and showed adequate correlation with the target measure (Guyatt et al. 2002). The HDSS scale, apart from being widely validated and used in hyperhidrosis, is used in routine clinical practice. A 1-point change represents a 50% reduction in sweat production while a 2-point change relates to an 80% change in the amount of sweating (Solish et al.). A 2-point improvement on the HDSS is also used in establishing the MCID cut-off values for the DLQI (Kowalski 2007). On the other hand, the general question (GQ) used in this study is an adaptation of similar question used in establishing the scale banding for the DLQI; and the Renal Quality of Life Profile (Hongbo et al. 2005; Aawar 2011)

Caution is needed in applying MCID cut-off scores. Estimates based on minimally improving patients might differ from those based on minimally worsening patients; for example a larger magnitude of change might be needed for patients to feel that their condition has deteriorated (Testa 2000). On the other hand, patients with high baseline scores (i.e. high level of impairment) are likely to show greater improvement than those with low baseline scores (low level of impairment). Considering the high baseline level of impairment in this sample, MCID estimate might have been lower in a patient group with less impairment. The use of multiple approaches in the calculation of MCID (e.g. the integrated approach based on the 'no change' patients) provided a means to control for such biases.

Being able to interpret QoL score has practical and conceptual implications on the application of QoL scales, whether in routine clinical practice, in clinical research or in health policy decision-making. The proposed banding may facilitate screening and diagnosis of hyperhidrosis. Its simplicity avails a means for capturing the subjective experience of the patient into the consultation. The additional information provided by the banding may also alert the clinician to the severity of QoL impairment, which may influence treatment strategy. The scale score categorisation may offer a useful common language for describing hyperhidrosis, which may aid in minimising decision uncertainty on the part of the clinicians, aligning some practice variations in management and handling of hyperhidrosis patients. All in all the interpretability information will provide a bridge between the scores and appropriate actions to be taken. These issues are expounded upon below.

First, the HidroQoL might be applied in routine clinical practice as a screening tool. Considering that diagnosis of hyperhidrosis also relies on the degree of impairment experienced by patients in their day-to-day life. A one-off completion of the HidroQoL suffices for this purpose. Patients may have to complete the HidroQoL prior to their consultation, either at home or while waiting for their consultation in the clinic. Their scores may then be made available to the clinician during consultation alongside other records. The provided scale banding may help clinicians in determining magnitude of impact.

The HidroQoL may also be useful in detailing the specific functional areas patients might be experiencing problems. Despite the known high prevalence of psychosocial problems among dermatology patients, these tend to go unnoticed (Picardi et al.). Identifying problem areas may encourage clinicians to discuss the highlighted issues or to refer patients to other health professionals for psychiatric support or counselling.

The HidroQoL may be useful in patient management as a tool for monitoring the patient's condition over time. Information provided through the instrument would be used alongside other pieces of information to determine whether a treatment strategy was working, allowing for precisely planned treatment strategies (Hahn et al. 2007). For example, the MCID cut-off values, can be applied in deciding whether an observed change score necessitates a review of treatment strategy. Patients would have to complete the HidroQoL prior to or during their visit to the clinic, in order for the information to available during consultation. The data would have to be systematically stored, to facilitate longitudinal comparisons on later visits.

Bringing patients to the centre of the process of care has characterised recent reforms of the NHS in the UK and in other health care systems, for example in Sweden and the U.S. A key component of patient centred care is to empower patients to self-manage their condition. This enables patients to play a more active and central role for example in monitoring their symptoms and QoL; complying to treatment; and in decision-making on risk and benefit assessment of treatments (Frost et al. 2007a). In this context the HidroQoL may allow patients to voice their concerns, priorities and needs bringing them to the clinical agenda (Higginson and Carr 2001). This may be in the form of discussions on particular issues, collaborative setting of treatment goals or choice of therapeutic management strategy (Marshall et al. 2006). This may require the patient completing

the HidroQoL and interpreting score results on their own before visiting the clinic, to self-assess their disease severity, form their own priorities and to bring such information to the consultation. The simplicity of the HidroQoL scoring and the scale banding provided would facilitate such a process.

Use of PROM to advance patient-centred care, is of course contingent upon the instrument chosen having optimal applicability, adequate coverage and emphasis on the issues of most relevance to patients. Using a measure whose content is of low priority to the patients may achieve the contrary, aggravate feelings by the patient that their needs are not being met (Lohr and Zebrack 2009). Evidence provided in this research, indicates the patient rooting of the HidroQoL's content, guaranteeing that its items are representative of the views, experiences and priorities of the patients with the phraseology used reflecting language used by patients. Patients seek to learn more about their disease condition especially regarding its causes, prognosis, impact on their QoL, available treatments and their related effectiveness, the impact of treatment on QoL (Brundage et al. 2005). In this context, results from clinical trials applying the HidroQoL scores as an endpoint may also serve wider patient education objectives. Such data may facilitate the understanding of the how various therapies may affect patient's QOL, as they reflect those outcomes patients might be more familiar with and care about most.

Symptoms or disease severity consideration alone may fail to capture the full therapeutic benefit and risks to be considered in choice of therapies given the known adverse events and complications associated with the majority of hyperhidrosis treatments, Botox, Iontophoresis and ETS surgery. The approach taken in the HidroQoL in measuring hyperhidrosis QoL, may offer a more comprehensive framework.

The comprehensive development and validation of the HidroQoL makes it potentially important as a PROM for hyperhidrosis within the National Patient Reported Outcome Programme (Black 2013), if the program is extended to cover dermatology. In particular, the HidroQoL scores may be used as a performance indicator in service contracts or in evaluating performance of providers for hyperhidrosis treatment. Quality of care from different service providers and outcomes from different interventions may be compared across the entire NHS, which might be useful in decision-making related to commissioning of services, choice of provider or interventions to be covered

(Devlin and Appleby 2010). Further, the possibility of adding up items at the domain and overall scale level suggests that the HidroQoL might also be useful as an outcome measure in disease registries on hyperhidrosis.

The dilemma on how to measure health not only complicates efforts at identifying population health needs but also presents challenges in the provision of care that is of good quality, effective, accessible and satisfactory to the patients (Dalgard and Finlay 2006). For example, the accuracy and efficiency in the measurement of quality of life impairment is central to decision-making in both resource allocation and clinical management settings. The HidroQoL may, in this regard, be used to investigate disease burden from hyperhidrosis in the wider population. This may facilitate the monitoring of health disparities across regions, based on outcome/indicator of most relevance to patients. In addition the HidroQoL may facilitate economic evaluation of interventions in hyperhidrosis, for example, by using responder definitions that are based on the HidroQoL's composite scores are used or where Quality adjusted life year (QALY) using the HidroQoL scores by mapping of the HidroQoL score to the scores of a preference based measures like EQ-5D.

The presentation and communication of data from QoL instruments has an impact on how this information is interpreted and used. For the HidroQoL, the scores for the domains and the overall scale can be calculated by simple summation of individual items. For cross sectional or one-off use of the HidroQoL, for example in patient-screening, patient score may be compared against the scale banding provided. Patient's absolute score may be presented as a point on a cascading bar (with different colours reflecting the different levels of impairment). For evaluative use, involving longitudinally collected scores, patient-scores from different assessments spread over time might be compared. A line graph of the mean scores over time can be used for presentation. Patients have shown a preference for this format over others (such as stacked graphs, text data, or side by side bar graphs) (Frost et al. 2007a). Although the individual items provide insights on the specific areas patients may be experiencing impairment, the current findings do not support their application in hypothesis testing.

The limited access and challenge associated with obtaining permission to use many PRO instruments is a reality many researchers know only too well. Additionally, clear, transparent and easy to use instruction on how to use, interpret and present information collected from the scales

are often also missing for many instruments. In order to facilitate the use of the HidroQoL, documentation on all the necessary information pertaining to the known psychometric information, target population in which the HidroQoL might be applied, scoring system of the HidroQoL, will be provided through a User's Manual developed for the measure. A special website will be developed through which the User's Manual and the instrument will be made available for download. This will also host the web-version of the instrument.

The HidroQoL was developed for assessing hyperhidrosis-specific QoL in clinical research and in routine clinical practice. The different applications demand slightly different qualities. In a clinical research situation where analyses are at the group level, the greater availability of expertise and resources for data collection and analysis may permit a lengthier and complex questionnaire. On the other hand, in routine clinical practice resources might be limited, making other considerations such as suitability, appropriateness and acceptability, interpretability at the individual patient level, responsiveness to change at the individual level (Higginsons and Carr, 2001). Furthermore, such instruments must also emphasise on issues that patients consider relevant and that are most likely to be influenced by therapy. The HidroQoL, with 18 items, is short and fits on a single sheet of paper. The web-version fits a single screen shot. This avoids the risk that some questions might be left unattended. The fact that patients do not feel overwhelmed by the number of sheets may contribute towards the quality of answers obtained. Furthermore, the organisation of the items and responses was done such that subjects flow naturally through the questionnaire, from left to right, down the instrument. The actual item stems were short, not exceeding 7 words except for 2 items; furthermore they were expressed in the first person. The response categorisation is simplistic, with the number of options and their descriptors, highly unambiguous. Moreover all items use a common categorisation.

During field testing patients highlighted the ease of completion of the HidroQoL. Even with the field-testing version patients considered the completion time to be acceptable. Furthermore, another strong advantage attributable to patient involvement in the early development of the measure, was the relevance of the items to patients. Involvement of patient population from multiple countries enhanced the universality of the instrument, avoiding cultural colloquialisms. This suggests high translatability of the HidroQoL.

A number of disease-specific QoL instruments for hyperhidrosis are available, including the HDSS, the HHIQ, the Hyperhidrosis Scale (HS), the Hyperhidrosis Questionnaire (HQ), and the Hyperhidrosis Quality of Life Questionnaire. Although the HDSS and the HS have been used as HRQoL instruments, they assess the level of disease severity and interference in daily life activities caused by hyperhidrosis (Keller et al. 2001). The former has one item assessing impairment in daily activities (Solish et al. 2005) while the latter has 15 items assessing distress with a range of daily activities (Keller et al. 2001); issues related to the social and psychological burden of hyperhidrosis are not included in either scales. The validity, test-retest reliability and responsiveness of the HDSS have been demonstrated (Kowalski et al. 2004; Solish et al. 2005), while for the HS internal consistency, sensitivity and specificity for the HS scale has demonstrated (Keller et al. 2001; Keller et al. 2009).

The HQ assesses disease-specific QoL in hyperhidrosis covering four domains (functional domain, psychological domain, social domain, affective domain, physical domain). The content of the physical domain seems more related to side effects of surgical treatment than hyperhidrosis per se. In addition, a few items included in this measure seem to be relevant only to a sub-population sub-population of patients such as those with palmar or plantar hyperhidrosis. While construct validity and internal consistency was reported, test-retest reliability and responsiveness have not been assessed (Kuo et al. 2004). Another measure, the HLQLQ, disease-specific QoL using four domains (functional/social domain, personal domain, emotional-self or others, sweating under special circumstances). The use of this instrument has largely been restricted to patients with surgical treatment. Although the application studies (de Campos et al. 2003; Ambrogi et al. 2009) report sensitivity results (in terms of t-test of before and after surgery) this is hardly interpretable without clear demonstration of construct validity or reliability. In comparison to the domain ‘sweating under different circumstances’ (7 items) the content addressing emotional impacts is rather narrow (2 items). The issue relating to hyperhidrosis-type specific items was also observed here i.e. the construct was not defined in a way that the content would be relevant in the target population otherwise the fact that some items were type-specific was not addressed in the scoring system. The definition of the construct *disease-specific QoL in hyperhidrosis* offered in the existing measures, therefore, seems to be at a level not accommodating all types of hyperhidrosis.

The most promising instrument, the HHIQ, has reported the involvement of hyperhidrosis patients in its early development (Teale et al. 2002). Its test-retest reliability, construct validity and responsiveness have been reported (Naumann et al. 2002; Hamm et al. 2006). Nonetheless, the HHIQ does not cover key emotional issues in hyperhidrosis such as impacts on 'self-image' or 'embarrassment'. Evidence on its scoring procedures is not available. In addition, the baseline questionnaire includes 42 items, reflecting the purported use of the instrument in clinical trials rather than in routine clinical practice.

The key issues in the current measures identified through the review may be summarised as follows. First, the definition of the construct 'disease-specific QoL' reflect inadequate conceptual frameworks and a lack of patient involvement, as the appropriate balance and emphasis is lacking in the content. Internal structure was assessed only for a single instrument (the HQ). Most of the instruments reviewed lacked some key psychometric information. Thus, the HidroQoL fits into a space none of the current disease specific measures cover, starting with its conceptualisation, its qualitative development process and the validation of the final instrument. Its conceptualisation is based on patient experiences, and a literature review, its development combined multiple studies to provide various psychometric information, and finally the end-product (the instrument) is relevant to all forms of hyperhidrosis (with special attention paid to achieve this). Therefore the HidroQoL is applicable and practical enough for use in both routine clinical practice and in research settings.

Apart from involving patients to obtain input for a new measure, the use of a patient population from online social networking sites means that the views and issues that underpin the HidroQoL represent the experience of a hard-to-reach patient population often overlooked during typical clinic-based research. Although (Cinà and Clase 1999) used a population from an email discussion panel to validate the IIRS; during the current study these shaped the actual design of the questionnaire.

A key aspect of the current research was the use of patient populations from online social networking communities throughout all phases. As a novelty, a number of issues are still outstanding. Online social networking sites like the rest of the world wide web are not exempt from falsification of information and identity theft (Bilge et al. 2009). Of greatest interest was whether participants were indeed who they claimed to be or whether they indeed suffered from the

condition under study (Redmond 2010). This is not only of interest when considering the validity of data collected, but even more importantly the suitability of the participants for a given study. While it is counter intuitive that a falsified identity will be used to engage in social networking sites medical research, it is still necessary to put in place some validity checks. Firstly, additional avenues/channels which can be used to reach the participants e.g. telephone or physical address can be sought. In this study patients who responded to the online open questionnaire were asked if they would be willing to participate in an interview and, if so, to provide a telephone. Another alternative is to ask patients to supply medical records related to the condition under study(Lenert and Kaplan 2000). Considering that this requires extra effort some participants are likely to be put off. Finally, modern test theory models such as the Rasch Model hold a lot of promise for identifying participants with response patterns that are unexpected or out of the norm as a result of guessing, carelessness. Such patterns might be likely for persons using other people's identity to complete the instrument (Lenert and Kaplan 2000).

Ethical considerations represent the final hurdle online social networks must overcome to achieve a wider proliferation as a research tool. Of particular importance are issues related to patient consent, users' privacy expectations, confidentiality and data anonymisation (Zimmer 2010). The Helsinki declaration requires that informed consent be obtained from human subjects for their participation in medical research. The main challenge for social networking sites such as Facebook is that it may be possible to access personal data of users and their friends without their explicit acknowledgement (Redmond 2010). Use of such data for clinical or health outcomes research would be considered unethical. Even where the data were to be willingly provided by user through their use of various "Applications or platforms", such data can only be ethically used for medical research if the user granted specific consent for that purpose. The ethical standard for SNS research is to ensure that expectations and intent of users in relation to their personal data are understood and respected (Zimmer 2010). While it is understandable that collecting patient consent using traditional methods, pen and paper, may actually be challenging for SNS based research, other alternatives for achieving the same are available and widely accepted (Lenert and Kaplan 2000). Informed consent can be obtained online by providing a web-version of traditional consent forms, where an electronic signature is used. In this study we applied both the traditional written consent and we also made use of an electronic signature.

Availability of a rigorously developed PRO measure does not guarantee automatic application. Translating the 'basic science' of QoL measurement into application in routine clinical practice, clinical research or in health system development is at the mercy of numerous factors, practical/logistical, behavioural and methodological, affecting the actual measurement, analysis of data, or how the data is actually employed. Also, various stakeholders, sources of information (the patient), analysts (researchers/nurses/clinicians) and end-users of information may influence the process.

The introduction of systematic QoL data collection has resource and organisational implications, which might not be available or possible in the context of the clinic. For example, restricted budgets would imply such resources would have to be drawn from elsewhere unless time spent collecting, analysing and using QoL information would be reimbursed (Greenhalgh 2009). Moreover, the initial introduction of QoL may be associated substantial fixed costs, for example the need for staff-training, changes in the clinic flow (Fung and Hays 2008). For computer based/electronic systems, new equipment such as computers and service support may be required.

Further challenges may relate to the choice of instrument and means of data collection (instrument administration). The instrument chosen must not only be psychometrically sound, but must be suitable for intended use i.e. application in the routine clinical practice requires necessary adaptation. The majority of HRQoL instruments have not been developed for use in routine clinical practice (Higginson and Carr 2001). For example, generic instruments may seem inappropriate for clinical trials or use in routine clinical practice in dermatology. Clinical feasibility has to be reflected not only in the response and administration burden but in key psychometric properties, i.e. whether validity in the clinic population or individual patient analyses is established. A fundamental issue also relates to the ease of attaching meaning to the scores, PROs may be irrelevant without the ability of being able to interpret scores in terms of what change in scores from one visit to the next may mean (Salek and Kamudoni 2013).

The role of producers and end-users of QoL information in the process of measuring health outcomes cannot be ignored. The views, beliefs, practices of the patients and clinicians seems central in this regard. Clinicians may be not be interesting/willing to use QoL instruments, where they are unconvinced of the benefit of routine QoL measurement to patient management or where they view the understanding of psychosocial or daily life impact of disease as irrelevant to the

clinic context (Greenhalgh and Meadows 1999). Even in situations where there is an interest in understanding the QoL impairment experienced by the patients, they may still be unwilling to measure QoL using an instrument where they mistrust or question the credibility of existing QoL instruments. Furthermore, having collected QoL information based on validated instruments clinicians may still be at loss as to the most appropriate response and action. Thus, clear decision aids on the most appropriate action on the patient's QoL may be necessary. There is currently a paucity of such add-ons to PROMs.

In spite the tendency for clinicians to trust their own *adhoc* assessment of patients QoL impact, they are not able to always accurately predict this (Basra and Shahrukh), especially for patients experiencing either extremely low or high level of impairment. Furthermore, fears that integrating PROs may be excessively costly or require more time are often contrary to current evidence. The thinking that subjectivity of PRO information implies a lack of reliability may also be unfounded. There is evidence that the level of precision and predictive ability of PRO is comparable to that of clinical variables (Hahn et al. 2007). These issues require deliberate effort at providing information that might help to allay some of the fears, for example as part of general clinician education or through change management processes (these are discussed later on).

Considering the resources implications of implementing PROs, whether in terms of fixed costs associated with initial set up, the training of staff, the reformulation of flow in the clinic, a clear demonstration of the added value of routine measurement is a prerequisite. Routine QoL measurement should not only make sense from a theoretical or conceptual point of view, but should also lead to tangible benefits on the process and outcomes of care. This would be key in establishing the case for PRO measurement to clinicians or other consumers of PRO information.

The patient, being the source of the PRO information, cannot be left out of the equation. Their motivation has an influence on the quality of the information collected and their enthusiasm may provide a push-factor to the physicians, broadening what is possible and feasible with PROMs. At the moment, QoL-discussions within consultation tend to be initiated by the clinician (Davies et al. 2008), which highlights potential for greater involvement and changing role of the patient. Patient education emphasising patient self-efficacy and their participation in treatment decision-

making may present PROMs as a means to these aims (Luckett et al. 2009), facilitating a shift in how patients view routine PROM measurement (Valderas et al. 2008).

The appropriateness of PRO instruments for routine PRO measurement in the clinic has been highlighted. The need for the harmonisation of guidelines for the development and validation of disease-specific dermatology questionnaires has also been highlighted. Two areas are particularly important when it comes to the development of disease specific instruments, one relates to ensuring that measures are rooted in the experience of patients and thus retain relevance and applicability to the target patient population. This has a bearing on not only the level of motivation patients will have in completing the questionnaire, but also how much clinicians will deem the measure useful as comprehensively capturing the unique disease impacts. Indeed, these elements may favour disease specific questionnaires over generic instruments.

Another consideration affecting the use of a measure, especially in the clinic, is whether the scales can be added together to form composite scores, an aspect of practicality (Lohr 2002). This should be evidence-based, showing that items indeed tap into a common construct. Some investment into the development and assessment of a scale's measurement model is required. Supporting the internal structure of the instrument is not only relevant for justifying the use of composite score, but forms a key part of the construct validity of the measure. Also, this evidence is used in subsequent validity test, for example, identifying an external measure which measures the same construct as a given instrument.

A prerequisite to the application of QoL-questionnaires, once reliability, validity and responsiveness are adequately demonstrated is the ability to attach qualitative meaning in terms that are relevant for patient management. Already much development has been undertaken regarding this issue in assessing QoL in dermatology. However, interpretation should not end at identifying the MCID, as is the current practice. The connection between such cut-off and specific clinical decisions or action should be established (Testa 2000). This means estimating chance of QoL improvement or worsening for alternative treatments, in populations taking into account broader range of factors such as health-care resources and costs in addition to benefit (Testa 2000). Apart from reporting what the cut-off estimate for MCID is, a further analysis may then indicate that, for example 23 % of patient receiving AC and 16% who received Botox would have remained stable (change of less than 2.15) had they received surgical treatment. Essentially using the MCID to explain the risks-benefits associated with each treatment. The second level information provides

connection between QoL impairment and clinical decisions or use of resources. This might offer more assistance in a decision-making context. For example would a 10 % risk of worsening QoL be worth the extra hospital days or resource expenditure.

A focus on methods to facilitate clinical interpretation of scale scores and the development of guidelines for the development of disease-specific QoL measures were singled out as being key to the next generation of dermatology QoL in the special issue of *Dermatology Clinic* journal on QoL measurement, published in 2012 (DeLong and Chen 2012). Earlier recommendations of the *International Dermato-Epidemiology Association (IDEA)* from 2008 Nottingham meeting raised similar issues, emphasising the need for streamlining HRQoL instruments given the current mushrooming of disease-specific measures, to ensure consensus on methods and to facilitate identification of optimal instruments for measuring disease-specific QoL in various settings in dermatological diseases (Chen 2012; DeLong and Chen 2012). Apart from ensuring that instruments have rigorously tested psychometric properties, instrument development guidelines would usher in clarity in the conceptual framework of QoL in dermatology at large. This may also encourage transparency in measurement, crucial to achieving truly objective measurement.

Finally, as with all new innovations, taking QOL measurement from bench to application in the clinic has to be carefully managed. How such a process is managed has a bearing on the uptake and acceptance of QOL measurement. Without careful consideration there could be a backlash from the physicians, especially if they perceive QoL as invading their professional judgement or autonomy. Such sentiments have been reported especially clinicians perceiving a push from the research community (Greenhalgh 2009). This may require paying attention to several issues. First, the engagement of clinicians or clinical researchers in the process, ensuring their active ownership of the change process, would mean that they are active in planning and implementing the necessary changes to research design or clinic flow, identifying possible limitations together with instrument developers and other stakeholders. In this context clinicians and researchers work together to develop action plans to address barriers identified by physicians and other health professionals. This would increase the chance that QOL is useful in the clinics, in addition to reinforcing change in practice, which may not be possible through clinician training alone.

Implications of the study

- Given the mushrooming number of disease-specific QoL instruments in dermatology, a thorough evaluation of their attributes in urgently needed, to provide guidance on the optimal choice (DeLong and Chen 2012). Not only would this facilitate comparison of clinical trial results, but might create greater consensus regarding methods of assessing disease-specific QoL in various disease areas, which might smoothen the process of integrating QoL assessment into the clinic. The review of literature and measures used in assessing HRQoL in hyperhidrosis undertaken in this research, serves this purpose, within hyperhidrosis. This means researchers and clinicians now have a resource to aid their choice of the most appropriate end-point in hyperhidrosis, while being made aware of the limitations to the data obtained.

- The qualitative research undertaken as part of this research has provided deep insights into the experiences of patients with hyperhidrosis, particularly the impairment in their QoL. Although previous accounts used quantitative approaches, the qualitative methods applied in this study means that the perceptions and beliefs of the patients, were captured using the patient's own words.

  Patients revealed perceptions that the general public is mostly unaware of hyperhidrosis. One patient contrasted hyperhidrosis to diabetes, with the later receiving much sympathy from the public in comparison with hyperhidrosis while another suggested that hyperhidrosis was not 'sexy' enough for researchers either. Also reported were frustrations with how general practitioners were ignorant about the condition. Thus, patients expressed much excitement at the ongoing research study and were quiet enthusiastic about the opportunity to make a contribution to the cause of hyperhidrosis. Most often patients inquired about the future direction of the study and whether they would be able to take part. For most patients, this was the first time they ever discussed the experience of living with the condition, reflecting on how their life had been affected. Most patients mentioned just carrying on with life and not giving their condition much thought. From that perspective, this study accorded the patients a real opportunity to reflect on how their life might have been affected and to talk about those experiences. Patients taking part in the focus group discussion found the experience to be beneficial as some patients had never met anyone

with their condition before. Therefore, not only did the study serve the research objectives but the actual process of data gathering also had positive side effects on the patients.

- The evidence generated in this research provided clarity regarding the definition of the concept 'hyperhidrosis disease-specific QoL' and its constituents 'psychosocial impact' and 'daily life activities impact'. This is not only relevant for future theoretical development to the understanding of hyperhidrosis, but also provides greater reason for dermatologists to pay close attention to these consequences in their management of patients. An even greater imperative is that the means to accurately measure these constructs has been provided.
- This research has provided a framework for the development of a new generation of QoL measures, utilising a mix of qualitative and quantitative methods ensuring that the resultant measure was efficient, applicable and content valid. The advantages of using factor analysis in parallel to Rasch analysis rather in sequence, were demonstrated. The balanced application of these methods avoided over-reliance on statistical approaches while addressing potential frictions that arose
- Considering that this is the first hyperhidrosis QoL instrument to undergo a thorough development and evaluation, the HidroQoL has potential to profoundly affect patient management in hyperhidrosis.
- The confidence in the scores of the HidroQoL and their interpretation provides a new opportunity to incorporate HRQoL considerations in hyperhidrosis in patient management, for self-efficacy, by epidemiological studies of disease burden and in health-care-service planning. For example the MCID score cut-off values may facilitate the examination of the efficacy of various treatment strategies by allowing the estimation and valuation of population risk and benefit of deteriorating QoL, according to the clinical significant cut-offs.
- The findings from this research offer an important rationale and basis for revisiting treatment guidelines in hyperhidrosis for example redefining treatment goals in terms of HRQoL

- The use of patient populations from online social networking community might be a particularly useful approach for disease conditions with a low prevalence rate, a situation characterising most orphan diseases, where it is challenging to accumulate sufficient

patient numbers locally around a single study centre. This study has provided demonstrated the feasibility of such an approach. The current study also suggests that a virtual hyperhidrosis disease registry established in collaboration with the existing patient support groups might be feasible.

Through the use of the online SNS it was possible to recruit patients under varying treatment modalities, most critically including those under self-treatment (taking OTC medications) as well as those not on any treatment. Patients who had not yet sought for medical attention for their condition despite experiencing QoL impairment, were included.

Limitations

- Data collection in all phases of this research were undertaken using electronic means, including telephone, videoconferencing, internet, with no clinic visits involved. One idea during the design phases was to request patients to provide records that would demonstrate that they had the diagnosis of hyperhidrosis in the form of a prescription receipt for hyperhidrosis medication or historical medical records. However, such an approach was not implemented as it was seen to impose an excessive burden on the study participants, which might have scared away potential participants. Thus it is possible that this study included subject who might not have had hyperhidrosis. Nonetheless, given that nearly all patients had seen a clinician regarding hyperhidrosis, with almost half treated over the last 6 months for their condition, helped to reduce the number of subjects with symptoms of hyperhidrosis but did not have the actual diagnosis. Moreover, only a small proportion of patients reported having co-morbidities known to cause excessive sweating.
- The use of the internet for both the recruitment and the data collection processes means it is not possible to calculate response rates. The only statistic showing how well the questionnaire was responded to is the completion rate, among those who registered how many completed the assessments.
- The reliability and responsiveness studies involved longitudinal data collection. Ideally all patients completing the baseline assessment should have completed the follow-up assessments as well. This was not the case, even after a follow-up email a large minority still did not complete the assessments. Nonetheless, the number of patients responding was sufficient to perform all planned analyses.

- The use of an internet based patient population raises other concerns. For example, Langenbruch et al. (Langenbruch et al. 2010) in a study using an internet-based psoriasis patient population to collect data on treatment and therapeutic benefit contrasted their results with those from of comparable clinic-based study. Although the clinical attributes seemed similar, patients in the online study showed greater dissatisfaction with treatment and less patient defined benefit. This may reflect a form of self-selection, where patients who are more impaired may engage more in online health-sites and in research than less impaired patients. On the other hand, similar self-selection can also be seen in clinic populations. Patients who seek for medical attention tend to have greater impairment than non-seekers, similarly those volunteering for research participation may differ from non-participants (Bland 1995).
- A more serious source and form of selection bias, however, is the need for computer literacy and access to internet, required to fall into our study population. This may pose a threat to extending the generalisability of results to those without computer literacy and without internet access. This limitation however may be irrelevant. With current figures on internet usage at 82% (UK) and 77.2% (USA) (World Bank, 2013), it is those without the internet and without computers who might not be representative of the majority of patients.
- In as much as patients reported whether they were on active treatment or not, patients did not receive an 'active' therapy in this research, thus responsiveness could not be measured based on hypothesis relating to therapy of known treatment effect.
- The development and validation of a new PRO instrument is a lengthy, tedious and resources intensive task, with whole teams dedicated to such effort, for instance the EUROQOL group or the EORTC group, to mention but a few, with budgets of up to half a million dollars. This illustrates the stringent conditions under which the HidroQoL was developed, in terms of time, finances and human resources. There is no doubt that the amount and quality of data collected, might have been enhanced without the said constraints. Nonetheless, the thorough and systematic process followed ensured rigor of the new instrument in spite of such limitations

Future work

- Although the patient population used in this research, recruited through the online social networks, included sub-groups representing key disease characteristics, a study to test the psychometric properties of the HidroQoL in a patient population recruited from the clinic

is planned.

- The responsiveness of the HidroQoL was firmly established utilising multiple analytical approaches. Nonetheless, evaluating responsiveness in a clinical trial by assessing treatment-effect of a therapy of known efficacy, might further provide confidence regarding the use of the measure in clinical trial situations.
- Given the controversy surrounding the long-term 'effectiveness' of various treatments in hyperhidrosis and on the other hand the restriction of the NHS to cover two cycles of BOTOX, the development of a PROM disease registry for hyperhidrosis would be valuable as a resource for assessing the real-life long-term benefits of the various available treatments.
- This research provided evidence on the psychometric properties of a web-version of the HidroQoL. Building on this, a comprehensive website containing the HidroQoL, where patients can self-assess their condition, obtain a total score as well as its interpretation on a scale banding, is planned. Such a website can integrate other valuable information which patients seek such as the causes of hyperhidrosis or available treatments and their efficacy (included on the basis of QoL).
- Given the tight financial climate in Europe as well as the UK, health-care budgets are under high pressure as never before. This suggests an ever growing role of bodies such as NICE, which assess pharmaceuticals in terms of their value, as a basis for reimbursement decisions. To facilitate provision of pharmacoeconomic information on hyperhidrosis therapies, a future study could perform a mapping of the HidroQoL score to a preference based measure such as the EQ-5D, in order to develop a conversion algorithm that might be used to easily translate information from clinical trials using the HidroQoL in future, and utilise such information in pharmacoeconomic evaluation studies.
- The current research has developed and validated an English version of the HidroQoL. Studies to perform the cultural adaptation of the instrument into German and Portuguese are planned. With the efforts made to ensure universality and translatability of the measure, this process should be less complex.
- The HidroQoL w-s designed and developed for use in routine clinical practice, with the simple structure, the number of items and the scoring system intended to support this. In order to establish whether this intention was achieved it would be useful to assess the HidroQoL's feasibility in the clinic. Evaluating how clinicians utilise the instrument, its

impact on the consultation and in patient management. Such a study can also study whether ultimately the use of the HidroQoL makes any differences in outcomes realised by the patient.

- The banding scale proposed in this study needs to be confirmed in a larger patient population. Information on how the banding may vary across for example cultural groups might also provide a form of conceptual invariance test. Similarly, further studies may be needed to confirm the MID in other patient populations.
- An initial understanding evaluating the impact of the measure might focus on for example, how training general practitioners in the use of the measure affected the treatment decision-making and patient satisfaction with the consultation encounter.

REFERENCES

Aawar, N. 2011. *Evaluation of patient compliance with renal replacement therapy and its impact on patient reported outcomes.* Cardiff University.

Abeni, D. et al. 2002. Further evidence of the validity and reliability of the Skindex-29: an Italian study on 2,242 dermatological outpatients. *Dermatology* 204(1), pp. 43-49.

Abeni, D. et al. 2001. Lo Skindex-29, un nuovo strumento per la valutazione della qualita della vita in dermatologia. Costruzione e validazione della versione italiana. *Giornale Italiano di Dermatologia e Venereologia* 136, pp. 73-76.

Abram, C. 2012. *Facebook For Dummies*. Hoboken (NJ): John Wiley & Sons.

Ak, M. et al. 2013. The evaluation of primary idiopathic focal hyperhidrosis patients in terms of alexithymia. *Journal of health psychology* 18(5), pp. 704-710.

Altman, D. et al. 2000. *Statistics with Confidence: Confidence intervals and statistical guidelines*. Bristol: BMJ Books.

Ambrogi, V. et al. 2009. Bilateral Thoracoscopic T2 to T3 Sympathectomy Versus Botulinum Injection in Palmar Hyperhidrosis. *The Annals of Thoracic Surgery* 88(1), pp. 238-245.

Amini, M. et al. 2008. Patient's estimation of efficacy of various hyperhidrosis treatments in a dermatological clinic. *Acta Derm Venereol* 88(4), pp. 356-362.

Amir, M. et al. 2000. Impairment in quality of life among patients seeking surgery for hyperhidrosis (excessive sweating): Preliminary results. *Israel Journal of Psychiatry and Related Sciences* 37(1), pp. 25-31.

Andrich, D. et al. 2012a. Quantifying Local, Response Dependence Between Two Polytomous Items Using the Rasch Model. *Applied Psychological Measurement* 36(4), pp. 309-324.

Andrich, D. et al. 2012b. *Interpreting RUMM 2030 Analysis: Part I Dichotomous Data*. Perth: RUMM Laboratory.

Andrich, D. et al. 2012c. *Interpreting RUMM 2030 Analysis: Part II polytomous data*. Perth: RUMM Laboratory.

Atkins, J. L. and Butler, P. E. M. 2002. Hyperhidrosis: a review of current management. *Plastic and reconstructive surgery* 110(1), pp. 222-228.

Augustin, M. et al. 2010. Quality of life evaluation in wounds: validation of the Freiburg Life Quality Assessment-wound module, a disease-specific instrument*. *International Wound Journal* 7(6), pp. 493-501.

Augustin, M. et al. 2004a. Validation of a comprehensive Freiburg Life Quality Assessment (FLQA) core questionnaire and development of a threshold system. *Eur J Dermatol* 14(2), pp. 107-113.

Augustin, M. et al. 2009. The patient benefit index: a novel approach in patient-defined outcomes measurement for skin diseases. *Arch Dermatol Res* 301(8), pp. 561-571.

Augustin, M. et al. 2008. Development and validation of a new instrument for the assessment of patient-defined benefit in the treatment of acne. *Journal der Deutschen Dermatologischen Gesellschaft* 6(2), pp. 113-120.

Augustin, M. et al. 2004b. German Adaptation of the Skindex-29 Questionnaire on Quality of Life in Dermatology: Validation and Clinical Results. *Dermatology* 209(1), pp. 14-20.

Augustin, M. et al. 2000. Validation and Clinical Results of the FLQA-d, a Quality of Life Questionnaire for Patients with Chronic Skin Disease. *Dermatology and Psychosomatics* 1(1), pp. 12-17.

Baez, A. et al. 2007. [Frey syndrome secondary to submaxillectomy and botulinic treatment]. *Medicina (B Aires)* 67(5), pp. 478-480.

Baghaei, P. 2008a. Local Dependency and Rasch Measures. *Rasch Measurement Transactions* 21(3), pp. 1105 - 1106.

Baghaei, P. 2008b. The Rasch model as a construct validation tool. *Rasch Measurement Transactions* 22(1), pp. 1145-1146.

Basra, M. et al. 2006. The Family Dermatology Life Quality Index: measuring the secondary impact of skin disease. *British Journal of Dermatology* 156(3), pp. 528-538.

Basra, M. K. A. et al. 2008. The Dermatology Life Quality Index 1994–2007: a comprehensive review of validation data and clinical results. *British Journal of Dermatology* 159(5), pp. 997-1035.

Basra, M. K. A. and Shahrukh, M. 2009. Burden of skin diseases. *Expert Review of Pharmacoeconomics and Outcomes Research* 9(3), pp. 271-283.

Beikert, F. C. et al. 2013. Willingness to pay and quality of life in patients with rosacea. *J Eur Acad Dermatol Venereol* 27(6), pp. 734-738.

Beltraminelli, H. and Itin, P. 2008. Skin and psyche – From the surface to the depth of the inner world. *Journal der Deutschen Dermatologischen Gesellschaft* 6(1), pp. 8-14.

Bhatti, Z. U. et al. 2011. Major life changing decisions and cumulative life course impairment. *Journal of the European Academy of Dermatology and Venereology* 25(2), pp. 245-246.

Bieling, P. et al. 2001. Factor Structure of the Illness Intrusiveness Rating Scale in Patients Diagnosed with Anxiety Disorders. *Journal of Psychopathology and Behavioral Assessment* 23(4), pp. 223-230.

Bilge, L. et al. eds. 2009. *All your contacts are belong to us: automated identity theft attacks on social networks*. 18th International world wide web conference. Madrid, 20 - 24 April.

Black, N. 2013. Patient reported outcome measures could help transform healthcare. *BMJ* 346(f137).

Bland, M. 1995. *An introduction to medical statistics*. Oxford: Oxford University Press.

Blome, C. et al. 2009a. Measuring patient-relevant benefits in pruritus treatment: development and validation of a specific outcomes tool. *British Journal of Dermatology* 161(5), pp. 1143-1148.

Blome, C. et al. 2009b. Measurement of patient-relevant benefits in the treatment of chronic hand eczema – a novel approach. *Contact Dermatitis* 61(1), pp. 39-45.

Blome, C. et al. 2010. Time needed for treatment is the major predictor of quality of life in psoriasis. *Dermatology* 221(2), pp. 154-159.

Bloom, J. 1984. Response to J.E. Ware Jr., Methodology in behavioral and psychosocial cancer research. Conceptualizing disease impact and treatment outcomes. *Cancer* 53 (Suppl.), p. 3.

Bond, T. 2004. Validity and assessment: a Rasch measurement perspective. *Metodologia de las Ciencias del Comportamiento* 5, pp. 179-194.

Bond, T. G. and Fox, C. M. 2007. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah (NJ): Lawrence Erlbaum Associates Publishers.

Both, H. et al. 2007. Critical review of generic and dermatology-specific health-related quality of life instruments. *J Invest Dermatol* 127, pp. 2726-2739.

Bottomley, A. et al. 2009. Patient-reported outcomes: assessment and current perspectives of the guidelines of the Food and Drug Administration and the reflection paper of the European Medicines Agency. *European Journal of Cancer* 45(3), pp. 347-353.

Bowling, A. 2001. *Measuring disease: a review of disease-specific quality of life measurement scales*. 2nd ed. Buckingham: Open University Press.

Bowling, A. 2005. *Measuring health : a review of quality of life measurement scales*. 3rd ed. Berkshire: Open University Press.

Bowling, A. 2009. *Research methods in health: investigating health and health services*. 3rd ed. ed. Berkshire: Open University Press.

Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3(2), pp. 77-101.

Brazier, J.-. et al. 1992. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ: British Medical Journal* 305(6846), p. 160.

Brod, M. et al. 2009. Qualitative research and content validity: developing best practices based on science and experience. *Quality of Life Research* 18(9), pp. 1263-1278.

Brooks, R. 1996. EuroQol: the current state of play. *Health Policy* 37(1), pp. 53-72.

Brown, B. et al. 2005. Using the general practice research database to esetimate the incidence, prevalence and management of hyperhidrosis in the UK. *Value in Health* 8(6), p. A143.

Brown, T. A. 2006. *Confirmatory Factor Analysis for Applied Research*. New York (NY): Guilford Press.

Browne, M. and Cudeck, R. 1992. Alternative ways of assessing model fit. *SMR* 21, pp. 230 - 258.

Brundage, M. et al. 2005. Communicating quality of life information to cancer patients: A study of six presentation formats. *Journal of Clinical Oncology* 23(28), pp. 6949-6956.

Burke, M. J. and Dunlap, W. P. 2002. Estimating interrater agreement with the average deviation index: A user's guide. *Organizational Research Methods* 5(2), pp. 159-172.

Byrne, B. 2011. *Structural Equation Modeling With Mplus: Basic Concepts, Applications, and Programming*. New York (NY): Routledge.

Calman, K. C. 1984. Quality of life in cancer patients--an hypothesis. *J Med Ethics* 10(3), pp. 124-127.

Cetindag, I. B. et al. 2008. Long-term Results and Quality-of-Life Measures in the Management of Hyperhidrosis. *Thoracic Surgery Clinics* 18(2), pp. 217-222.

Chang, K. Y. et al. 2011. Application of the rasch model to develop a simplified version of a multiattribute utility measurement on attitude toward labor epidural analgesia. *Anesthesia and Analgesia* 113(6), pp. 1444-1449.

Chen, S. C. 2012. Health-related quality of life in dermatology: introduction and overview. *Dermatologic Clinics* 30(2), pp. 205-208.

Chren, M.-M. et al. 1996. Skindex, a Quality-of-Life Measure for Patients with Skin Disease: Reliability, Validity, and Responsiveness. *J Investig Dermatol* 107(5), pp. 707-713.

Chren, M. M. 2005. Measurement of vital signs for skin diseases. *Journal of Investigative Dermatology* 125(4), pp. viii-ix.

Chren, M. M. et al. 1997. Convergent and discriminant validity of a generic and a disease-specific instrument to measure quality of life in patients with skin disease. *Journal of Investigative Dermatology* 108(1), pp. 103-107.

Chren, M. M. et al. 2001. Measurement properties of Skindex-16: a brief quality-of-life measure for patients with skin diseases. *J Cutan Med Surg* 5(2), pp. 105-110.

Cina, C. S. et al. 2007. Endoscopic thoracic sympathectomy for hyperhidrosis: Technique and results. *J Minim Access Surg* 3(4), pp. 132-140.

Cinà, C. S. and Clase, C. M. 1999. The Illness Intrusiveness Rating Scale: A measure of severity in individuals with hyperhidrosis. *Quality of Life Research* 8(8), pp. 693-698.

Cina, C. S. et al. 2006. Effect of endoscopic sympathectomy for hyperhidrosis on quality of life using the illness intrusiveness rating scale. *Minerva Chir* 61(3), pp. 231-239.

Connor, K. M. et al. 2006. Botulinum toxin treatment of social anxiety disorder with hyperhidrosis: a placebo-controlled double-blind trial. *J Clin Psychiatry* 67(1), pp. 30-36.

Coons, S. J. et al. 2009. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO good research practices task force report. *Value in Health* 12(4), pp. 419-429.

Coons, S. J. and Kaplan, R. M. 1992. Assessing health-related quality of life: application to drug therapy. *Clin Ther* 14(6), pp. 850-858; discussion 849.

Coste, J. et al. 1997. Methodological approaches to shortening composite measurement scales. *J Clin Epidemiol* 50, pp. 247 - 252.

Costello, A. and Osborne, J. 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* 10, pp. 173-178.

Crosby, R. D. et al. 2003. Defining clinically meaningful change in health-related quality of life. *Journal of Clinical Epidemiology* 56(5), pp. 395-407.

Cummins, R. A. 1997. Assessing quality of life. *Quality of life for people with disabilities: Models, research and practice* 2, pp. 116-150.

Cummins, R. A. et al. 1998. *Quality of Life Definition and Terminology: A Discussion Document from the International Society for Quality of Life Studies*. Blacksburg (VA): The International Society for Quality-of-Life Studies.

Dalal, A. A. et al. 2011. Evaluating patient-reported outcome measurement comparability between paper and alternate versions, using the lung function questionnaire as an example. *Value in Health* 14(5), pp. 712-720.

Dalgard, F. and Finlay, A. Y. 2006. Need and greed in dermatology: disease and outcome measures in dermatological healthcare planning. *Dermatology* 213(4), pp. 279-283.

Davidson, W. L. 2003. *Assessment of nutrition intervention for patients with unresectable pancreatic cancer in a fish oil supplement trial - does it.* Queensland University of Technology.

Davies, E. et al. 2008. Does ad hoc quality-of-life discussion in inflammatory skin disease consultations reflect standardized patient-reported outcomes? *Clinical and Experimental Dermatology* 33(1), pp. 16-21.

de Campos, J. et al. 2003. Quality of life, before and after thoracic sympathectomy: Report on 378 operated patients. *Annals of Thoracic Surgery* 76(3), pp. 886-891.

de Vet, H. C. et al. 2007. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Quality of life research* 16(1), pp. 131-142.

De Vet, H. C. et al. 2006. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes* 4(1), p. 54.

DeLong, L. K. and Chen, S. C. 2012. Future Directions in Dermatology Quality of Life Measures. *Dermatologic clinics* 30(2).

DeMars, C. 2010. *Item Response Theory*. Oxford: Oxford University Press.

DeVellis, R. F. 2011. *Scale Development: Theory and Applications*. Los Angeles (CA): Sage Publications.

Devins, G. et al. 2001. Structure of lifestyle disruptions in chronic disease: a confirmatory factor analysis of the Illness Intrusiveness Ratings Scale. *Medical Care* 39(10), pp. 1097-1104.

Devins, G. M. 2010. Using the Illness Intrusiveness Ratings Scale to understand health-related quality of life in chronic disease. *Journal of psychosomatic research* 68(6), pp. 591-602.

Devins, G. M. et al. 1983. The emotional impact of end-stage renal disease: importance of patients' perception of intrusiveness and control. *Int J Psychiatry Med* 13(4), pp. 327-343.

Devlin, N. J. and Appleby, J. 2010. *Getting the most out of PROMS*. London: The Kings Fund, Office of Health Economics.

Dillman, D. A. 2006. *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Hoboken (NJ): John Wiley & Sons.

Doward, L. C. and McKenna, S. P. 2004. Defining patient-reported outcomes. *Value in Health* 7, pp. S4-S8.

Drummond, M. et al. 2005. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.

Ebrahim, S. 1995. Clinical and public health perspectives and applications of health-related quality of life measurement. *Social science & medicine* 41(10), pp. 1383-1394.

Edelen, M. O. et al. 2006. Identification of differential item functioning using item response theory and the likelihood-based model comparison approach: application to the Mini-Mental State Examination. *Medical care* 44(11), pp. S134-S142.

Epstein, R. 2000. Responsiveness in quality-of-life assessment: nomenclature, determinants, and clinical applications. *Medical care* 38(9 Suppl), p.91.

Fairclough, D. L. 2010. *Design and analysis of quality of life studies in clinical trials*. 2nd ed. ed. Boca Raton, Florida (USA): Chapman & Hall.

Fayers, P. and Machin, D. 2007. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes*. 2nd ed. West Sussex: John Wiley & Sons.

Ferrans, C. 2005. Definitions and conceptual models of quality of life. In: Lipscomb, J. et al. eds. *Outcomes Assessment in Cancer: Measures, Methods and Applications*. Cambridge: Cambridge University Press.

Finlay, A. 1998. Quality of life assessments in dermatology. *Semin Cutan Med Surg* 17, pp. 291 - 296.

Finlay, A. and Coles, E. 1995. The effect of severe psoriasis on the quality of life of 369 patients. *British Journal of Dermatology* 132(2), pp. 236-244.

Finlay, A. and Khan, G. 1994. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol* 19, pp. 210 - 216.

Finlay, A. Y. 2011. Practice Gaps. Dermatologists should better integrate quality-of-life measures to inform and improve clinical decision making: comment on "The impact of pruritus on quality of life". *Archives of dermatology* 147(10), p. 1157.

Finlay, A. Y. and Ryan, T. J. 1996. Disability and handicap in dermatology. *International Journal of Dermatology* 35(5), pp. 305-311.

Fisher Jr, W. P. 2000. Objectivity in psychosocial measurement: what, why, how. *Journal of Outcome Measurement* 4(2), p. 527.

Fitzpatrick, R. 1996. Alternative approaches to the assessment of health-related quality of life. In: Offer, A. ed. *In Pursuit of the Quality of Life*. Oxford: Oxford University Press.

Flanagan, J. C. 1978. A research approach to improving our quality of life. *American Psychologist* 33(2), pp. 138-147.

Flanagan, K. H. et al. 2008. Botulinum toxin type a versus topical 20% aluminum chloride for the treatment of moderate to severe primary focal axillary hyperhidrosis. *J Drugs Dermatol* 7(3), pp. 221-227.

Frost, M. H. et al. 2007a. Applying quality-of-life data formally and systematically into clinical practice. *Mayo Clinic Proceedings* 82(10), pp. 1214-1228.

Frost, M. H. et al. 2007b. What is Sufficient Evidence for the Reliability and Validity of Patient-Reported Outcome Measures? *Value in Health* 10(Suppl. 2), pp. S94-S105.

Fung, C. H. and Hays, R. D. 2008. Prospects and challenges in using patient-reported outcomes in clinical practice. *Qual Life Res* 17(10), pp. 1297-1302.

Gill, P. et al. 2008. Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal* 204(6), pp. 291-295.

Gill, T. M. and Feinstein, A. R. 1994. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 272(8), pp. 619-626.

Glogau, R. G. 2001. Treatment of palmar hyperhidrosis with botulinum toxin. *Seminars in Cutaneous Medicine and Surgery* 20(2), pp. 101-108.

Gnanasakthy, A. et al. 2013. Potential of patient-reported outcomes as nonprimary endpoints in clinical trials. *Health Qual Life Outcomes* 11, p. 83.

Goh, C. 2007. Aluminum chloride hexahydrate versus palmar hyperhidrosis. *International journal of dermatology* 29(5), pp. 368-370.

Golics, C. et al. 2009. Adolescents with skin disease have specific quality of life issues. *Dermatology* 218(4), pp. 357-366.

Gorecki, C. et al. 2011. Applying mixed methods to pretest the Pressure Ulcer Quality of Life (PU-QOL) instrument. *Quality of Life Research*, pp. 1-11.

Gorsuch, R. L. 1997. Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment* 68(3), pp. 532-560.

Greenhalgh, J. 2009. The applications of PROs in clinical practice: what are they, do they work, and why? *Qual Life Res* 18(1), pp. 115-123.

Greenhalgh, J. et al. 2005. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Soc Sci Med* 60(4), pp. 833-843.

Greenhalgh, J. and Meadows, K. 1999. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of Evaluation in Clinical Practice* 5(4), pp. 401-416.

Grob, J. et al. 2005. Comparative study of the impact of chronic urticaria, psoriasis and atopic dermatitis on the quality of life. *British Journal of Dermatology* 152(2), pp. 289-295.

Grob, J. J. 2007. Why are quality of life instruments not recognized as reference measures in therapeutic trials of chronic skin disorders? *Journal of Investigative Dermatology* 127(10), pp. 2299-2301.

Group, T. W. 1948. *Constitution of the World Health Organisation*. Geneva, Switzerland: World Health Organisation, pp. 29-43.

Group, T. W. 1995. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 41(10), pp. 1403-1409.

Grozdev, I. et al. 2012. Physical and mental impact of psoriasis severity as measured by the compact Short Form-12 Health Survey (SF-12) quality of life tool. *J Invest Dermatol* 132(4), pp. 1111-1116.

Gujarati, D. N. 2003. *Basic econometrics*. New York (NY): McGraw Hill.

Guyatt, G. et al. 1987. Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of chronic diseases* 40(2), pp. 171-178.

Guyatt, G. H. et al. 1993. Measuring health-related quality of life. *Annals of Internal Medicine* 118(8), pp. 622-629.

Guyatt, G. H. et al. 1992. Measuring health status: what are the necessary measurement properties? *Journal of clinical epidemiology* 45(12), pp. 1341-1345.

Guyatt, G. H. et al. 2002. Methods to explain the clinical significance of health status measures. *Mayo Clinical Proceedings* 77(4), pp. 371-383.

Gwet, K. 2002. Computing inter-rater reliability with the SAS system. *Stat Methods Inter-rater Reliability Assess* 3, pp. 1-16.

Hagquist, C. et al. 2009. Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies* 46(3), pp. 380-393.

Hahn, E. A. et al. 2007. Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clin Proc* 82(10), pp. 1244-1254.

Haider, A. and Solish, N. 2005. Focal hyperhidrosis: diagnosis and management. *Canadian Medical Association Journal* 172(1), pp. 69-75.

Halioua, B. et al. 2000. Quality of life in dermatology. *International Journal of Dermatology* 39(11), pp. 801-806.

Hambleton, R. K. and Rogers, H. J. 1991. Advances in criterion-referenced measurement. In: Hambleton, R.K. ed. *Advances in educational and psychological testing: Theory and Applications*. New York (NY): Kluwer Academic/Plenum Publishers, pp. 3 - 43.

Hamm, H. et al. 2006. Primary focal hyperhidrosis: disease characteristics and functional impairment. *Dermatology* 212(4), pp. 343-353.

Hattie, J. 1985. Methodology review: assessing unidimensionality of tests and ltenls. *Applied Psychological Measurement* 9(2), pp. 139-164.

Haynes, S. N. et al. 1995. Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment* 7(3), p. 238.

Hays, R. and Hadorn, D. 1992. Responsiveness to change: an aspect of validity, not a separate dimension. *Quality of Life Research* 1(1), pp. 73-75.

Hays, R. D. et al. 2005. Application of Structural Equation Modeling to Health Outcomes Research. *Evaluation & the Health Professions* 28(3), pp. 295-309.

Heckmann, M. et al. 2001. Botulinum toxin A for axillary hyperhidrosis (excessive sweating). *N Engl J Med* 344(7), pp. 488-493.

Henteleff, H. J. and Kalavrouziotis, D. 2008. Evidence-Based Review of the Surgical Management of Hyperhidrosis. *Thoracic Surgery Clinics* 18(2), pp. 209-216.

Herdman, M. et al. 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 20(10), pp. 1727-1736.

Hermansen, S. E. et al. 2002. Patients' and doctors' assessment of skin disease handicap. *Clin Exp Dermatol* 27(3), pp. 249-250.

Higginson, I. J. and Carr, A. J. 2001. Measuring quality of life: Using quality of life measures in the clinical setting. *BMJ* 322(7297), p. 1297.

Hill, S. and Glade, D. 2012. Hyperhidrosis - a review of hyperhidrosis and its current management. *Clinical Dermatology* 6(2), pp. 19 - 24.

Holgado–Tello, F. P. et al. 2010. Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity* 44(1), pp. 153-166.

Holm, E. A. and Jemec, G. B. 2004. Time spent on treatment of atopic dermatitis: a new method of measuring pediatric morbidity? *Pediatr Dermatol* 21(6), pp. 623-627.

Hongbo, Y. et al. 2005. Translating the Science of Quality of Life into Practice: What Do Dermatology Life Quality Index Scores Mean&quest. *Journal of Investigative dermatology* 125(4), pp. 659-664.

Hornberger, J. et al. 2004. Recognition, diagnosis, and treatment of primary focal hyperhidrosis. *J Am Acad Dermatol* 51(2), pp. 274-286.

Hund, M. et al. 2002. Definition of axillary hyperhidrosis by gravimetric assessment. *Archives of Dermatology* 138(4), p. 539.

Hunt, S. et al. 1985. Measuring health status: a new tool for clinicians and epidemiologists. *The Journal of the Royal College of General Practitioners* 35(273), p. 185.

Hunt, S. M. and McKenna, S. P. 1993. Measuring patients' views of their health. SF 36 misses the mark. *BMJ* 307(6896), p. 125.

Jakobsson, U. et al. 2011. Construct validity of the SF-12 in three different samples. *Journal of Evaluation in Clinical Practice*.

Jemec, G. B. et al. 2006. Time spent on treatment (TSOT). An independent assessment of disease severity in atopic dermatitis. *Acta Dermatovenerol Alp Panonica Adriat* 15(3), pp. 119-124.

Jemec, G. B. and Wulf, H. C. 1996. Patient-physician consensus on quality of life in dermatology. *Clin Exp Dermatol* 21(3), pp. 177-179.

Jenkinson, C. and Layte, R. 1997. Development and testing of the UK SF-12 (short form health survey). *J Health Serv Res Policy* 2(1), pp. 14-18.

Jobling, R. and Naldi, L. 2006. Assessing the impact of psoriasis and the relevance of qualitative research. *Journal of Investigative Dermatology* 126(7), pp. 1438-1440.

Johnson, C. et al. 2011. The European Organization for Research and Treatment of Cancer: guidelines for developing questionnaire modules. Fourth Edition. *Quality of Life Research*. Brussels: EORTC Quality of Life Group.

Jowett, S. and Ryan, T. 1985. Skin disease and handicap: An analysis of the impact of skin conditions. *Social Science & Medicine* 20(4), pp. 425-429.

Juniper, E. F. et al. 1997. Clinical impact versus factor analysis for quality of life questionnaire construction. *Journal of Clinical Epidemiology* 50(3), pp. 233-238.

Kalkan, M. T. et al. 1998. The measurement of sweat intensity using a new technique. *Turkish Journal Of Medical Sciences* 28, pp. 515-518.

Karimian-Teherani, D. et al. 2009. New epidemiological aspects of patients with severe hyperhidrosis presenting for sympathetic surgery. *Journal of the European Academy of Dermatology and Venereology* 23(6), pp. 651-655.

Keller, S. et al. 2009. Diagnosis of palmar hyperhidrosis via questionnaire without physical examination. *Clinical Autonomic Research* 19(3), pp. 175-181.

Keller, S., M et al. 2001. A novel scale for assessing quality of life following bilateral endoscopic thoracic sympathectomy for palmer and plantar hyperhidrosis. In: *4th International Symposium on Sympathetic Surgery.* Tampere, Finland.

Kerr, C. et al. 2010. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. *Expert review of pharmacoeconomics & outcomes research* 10(3), pp. 269-281.

Kim, D. et al. 2005. Factor structure of the Korean version of illness intrusiveness rating scale: cross-cultural implications. *J Korean Med Sci* 20(2), pp. 302-306.

Kini, S. P. and DeLong, L. K. 2012. Overview of Health Status Quality-of-Life Measures. *Dermatologic Clinics* 30(2), pp. 209-221.

Kline, P. 1994. *An Easy Guide to Factor Analysis*. London: Taylor & Francis Group.

Kowalski, J. 2007. Minimal important difference (MID) of the Dermatology Life Quality Index in patients with axillary and palmar hyperhidrosis. *Journal of the American Academy of Dermatology* 56(2), p. AB52.

Kowalski, J. et al. 2005. Cost effectiveness and budget impact of botulinum toxin type A (BoNTA) treatment for severe primary axillary hyperhidrosis inadequately managed with topical agents. In: *American Academy of Dermatology Summer Meeting.* Chicago, Illinois, USA, July 20 - 24.

Kowalski, J. W. et al. 2004. Validity and reliability of the hyperhidrosis disease severity scale (HDSS). In: *62nd Annual Meeting of the American Academy of Dermatology* Washington D.C., 6 - 10 February. Journal of the American Academy of Dermatology, p. A.198.

Krueger, R. A. 1994. *Focus Groups: Practical Guide for Applied Research*. London: Sage Publications Ltd.

Kumar, R. and Indrayan, A. 2011. Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics* 48(4), pp. 277-287.

Kuo, C. H. et al. 2004. Developing an instrument to measure quality of life of patients with hyperhidrosis. *The journal of nursing research : JNR* 12(1), pp. 21-30.

Lackey, N. R. et al. 2003. *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. London: SAGE Publications Ltd.

Landis, J. R. and Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, pp. 159-174.

Langenbruch, A. K. et al. 2010. Internet-supported gathering of treatment data and patient benefits in psoriasis. *Journal of the European Academy of Dermatology and Venereology* 24(5), pp. 541-547.

Lasch, K. E. et al. 2010. PRO development: rigorous qualitative research as the crucial foundation. *Quality of Life Research* 19(8), pp. 1087-1096.

Lear, W. et al. 2007. An epidemiological study of hyperhidrosis.*Dermatol Surg*. Vol. 33. United States, pp. S69-75.

Leidy, N. K. et al. 1999. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value in Health* 2(2), pp. 113-127.

Lenert, L. and Kaplan, R. M. 2000. Validity and interpretation of preference-based measures of health-related quality of life. *Medical Care* 38(9 SUPPL. 2).

Leung, A. K. C. et al. 1999. Hyperhidrosis. *International Journal of Dermatology* 38(8), pp. 561-567.

Liang, M. H. 2000. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. *Medical care* 38(9), p. II.

Linacre, J. 2009. *A User's Guide to WINSTEPS/MINISTEPS*. Chicago (IL): Winsteps.com.

Linacre, J. M. 1998. Detecting multidimensionality: which residual data-type works best? *Journal of Outcome Measurement* 2, pp. 266-283.

Linacre, J. M. 1999. Investigating rating scale category utility. *Journal of Outcome Measurement* 3(2), p. 103.

Lipsey, M. W. 1990. *Design sensitivity: Statistical power for experimental research*. London: SAGE Publication Ltd.

Lohr, K. N. 2002. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research* 11(3), pp. 193-205.

Lohr, K. N. and Zebrack, B. J. 2009. Using patient-reported outcomes in clinical practice: challenges and opportunities. *Quality of Life Research* 18(1), pp. 99-107.

Lowe, N. et al. 2004. The place of botulinum toxin type A in the treatment of focal hyperhidrosis.*Br J Dermatol*. Vol. 151. England, pp. 1115-1122.

Lowe, N. J. et al. 2007. Botulinum toxin type A in the treatment of primary axillary hyperhidrosis: a 52-week multicenter double-blind, randomized, placebo-controlled study of efficacy and safety. *J Am Acad Dermatol* 56(4), pp. 604-611.

Luckett, T. et al. 2009. Improving patient outcomes through the routine use of patient-reported data in cancer clinics: future directions. *Psychooncology* 18(11), pp. 1129-1138.

Lynn, M. R. 1986. Determination and quantification of content validity. *Nursing research*.

MacCallum, R. C. et al. 1996. Power analysis and determination of sample size for covariance structure modeling. *Psychological methods* 1(2), p. 130.

MacKeigan, L. D. and Pathak, D. S. 1992. Overview of health-related quality-of-life measures. *Am J Hosp Pharm* 49(9), pp. 2236-2245.

Magasi, S. et al. 2012. Content validity of patient-reported outcome measures: perspectives from a PROMIS meeting. *Quality of Life Research* 21(5), pp. 739-746.

Magin, P. et al. 2009. The psychological sequelae of psoriasis: results of a qualitative study. *Psychology, Health & Medicine* 14(2), pp. 150-161.

Marshall, S. et al. 2006. Impact of patient-reported outcome measures on routine practice: a structured review. *Journal of evaluation in clinical practice* 12(5), pp. 559-568.

Maslow, A. H. 1973. *The Farther Reaches of Human Nature*. New York (NY): Penguin Books.

Masters, G. N. 1982. A Rasch model for partial credit scoring. *Psychometrika* 47(2), pp. 149-174.

Masters, N. G. and Wright, D. B. 1997. The Partial Credit Model. In: Linden, W. and Hambleton, R. eds. *Handbook of Modern Item Response Theory*. New York (NY): Springer Verlag New York, pp. 101-121.

McDowell, I. 2006. *Measuring health : a guide to rating scales and questionnaires*. 3rd ed. Oxford: Oxford University Press.

Meredith, W. and Teresi, J. A. 2006. An essay on measurement and factorial invariance. *Medical care* 44(11), pp. S69-S77.

Messick, S. 1988. The once and future issues of validity: Assessing the meaning and consequences of measurement. *Test validity* 33, p. 45.

Messick, S. 1995. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist* 50(9), p. 741.

Moons, P. et al. 2006. Critique on the conceptualisation of quality of life: a review and evaluation of different conceptual approaches. *International journal of nursing studies* 43(7), pp. 891-901.

Motley, R. and Finlay, A. 1989. How much disability is caused by acne? *Clinical and experimental dermatology* 14(3), pp. 194-198.

Mount, B. M. and Scott, J. F. 1983. Whither hospice evaluation. *J Chronic Dis* 36(11), pp. 731-736.

Muller, C. and Augustin, M. 2013. Willingness-to-pay and patient-defined benefits in the treatment of hyperhidrosis: results from the first German health services research study in hyperhidrosis. *Br J Dermatol* 168(2), pp. 448-450.

Muller, C. et al. 2012. Efficacy and safety of methantheline bromide (Vagantin((R))) in axillary and palmar hyperhidrosis: results from a multicenter, randomized, placebo-controlled trial. *J Eur Acad Dermatol Venereol* 27(10), pp. 1278 - 1284.

Mullin, P. A. et al. 2000. Applying cognitive design principles to formatting HRQOL instruments. *Quality of Life Research* 9(1), pp. 13-27.

Munro, B. H. 2005. *Statistical Methods for Health Care Research*. 5th ed. Philadelphia (PA): Lippincott Williams & Wilkins.

Muthén, L. K. and Muthén, B. O. 1998 - 2010. Mplus User's Guide. Sixth Edition. Los Angeles (CA): Muthén & Muthén.

Naumann, M. et al. 2003. Botulinum toxin type A is a safe and effective treatment for axillary hyperhidrosis over 16 months: a prospective study. *Archives of Dermatology* 139(6), p. 731.

Naumann, M. K. et al. 2002. Effect of botulinum toxin type A on quality of life measures in patients with excessive axillary sweating: a randomized controlled trial. *British Journal of Dermatology* 147(6), pp. 1218-1226.

Nelson, P. A. et al. 2012. Recognition of need in health care consultations: a qualitative study of people with psoriasis. *British Journal of Dermatology* 168(2), pp. 354-361.

Neumayer, C. et al. 2005. Effect of Endoscopic Thoracic Sympathetic Block on Plantar Hyperhidrosis. *Arch Surg* 140(7), pp. 676-680.

Neumayer, C. et al. 2004. Limited endoscopic thoracic sympathetic block for hyperhidrosis of the upper limb: reduction of compensatory sweating by clipping T4. *Surg Endosc* 18(1), pp. 152-156.

Nijsten, T. 2012. Dermatology Life Quality Index: time to move forward. *Journal of Investigative Dermatology* 132(1), pp. 11-13.

Nijsten, T. et al. 2007. Cross-Cultural Inequivalence of Dermatology-Specific Health-Related Quality of Life Instruments in Psoriasis Patients. *J Invest Dermatol* 127(10), pp. 2315-2322.

Nijsten, T. et al. 2009. Categorization of Skindex-29 scores using mixture analysis. *Dermatology* 218(2), pp. 151-154.

Nijsten, T. et al. 2006a. Refinement and reduction of the Impact of Psoriasis Questionnaire: Classical Test Theory vs. Rasch analysis. *Br J Dermatol* 154, pp. 692 - 700.

Nijsten, T. E. C. et al. 2006b. Testing and Reducing Skindex-29 Using Rasch Analysis: Skindex-17. *J Invest Dermatol* 126(6), pp. 1244-1250.

Norman, G. R. et al. 2003. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical Care* 41(5), p. 582.

Norman, G. R. and Streiner, D. L. 2007. *Biostatistics: The Bare Essentials*. 3rd ed. Shelton (CT): PMPH B.C. Decker.

Norquist, J. M. et al. 2011. Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. *Quality of Life Research*, pp. 1-8.

Nunnally, J. C. and Bernstein, I. H. 1994. *Psychometric theory*. London: McGraw-Hill.

Ohhashi, T. et al. 1998. Human perspiration measurement. *Physiol Meas* 19(4), pp. 449-461.

Ozer, Z. C. et al. 2009. Confirmatory and exploratory factor analysis of the caregiver quality of life index-cancer with Turkish samples. *Quality of Life Research* 18(7), pp. 913-921.

Padilla, G. et al. 1996. Quality of Life - Cancer. In: Spilker, B. ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, USA: Lippincott-Raven, pp. 301-308.

Pallant, J. F. and Tennant, A. 2007. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology* 46(1), pp. 1-18.

Panhofer, P. et al. 2005. A survey and validation guide for health-related quality-of-life status in surgical treatment of hyperhidrosis. *European Surgery* 37(3), pp. 143-152.

Park, E. J. et al. 2010. An epidemiological study of hyperhidrosis patients visiting the Ajou University Hospital hyperhidrosis center in Korea. *J Korean Med Sci* 25(5), pp. 772-775.

Patrick, D. L. et al. 2011a. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value in Health* 14(8), pp. 967-977.

Patrick, D. L. et al. 2011b. Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 - Assessing respondent understanding. *Value in Health* 14(8), pp. 978-988.

Patrick, D. L. et al. 2007. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value in Health* 10, pp. S125-S137.

Patrick, D. L. and Chiang, Y. 2000. Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. *Medical Care* 38(9), p. II.

Patrick, D. L. and Deyo, R. A. 1989. Generic and disease-specific measures in assessing health status and quality of life. *Medical care*, pp. 217-232.

Phillips, D. 2006. *Quality of Life: Concept, Policy and Practice*. Oxon: Routledge.

Picardi, A. et al. 2000. Psychiatric morbidity in dermatological outpatients: an issue to be recognized. *British Journal of dermatology* 143(5), pp. 983-991.

Polit, D. F. and Beck, C. T. 2006. The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health* 29(5), pp. 489-497.

Pope, C. and Mays, N. 2008. *Qualitative Research in Health Care*. 3rd ed. Oxford: Wiley-Blackwell.

Powell, V. and Powell, C. 1987. Quality of life measurement. *Journal of Medical Ethics* 13(4), p. 222.

Prieto, L. et al. 2003. Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes* 1(1), p. 27.

Prieto, L. et al. 1998. Rasch measurement for reducing the items of the Nottingham Health Profile. *J Outcome Meas* 2(4), pp. 285-301.

Prinsen, C. et al. 2013. Measurement of health-related quality of life in dermatological research and practice: outcome of the EADV Taskforce on Quality of Life. *JEADV* 27(10), pp. 1195-1203.

Prinsen, C. A. C. et al. 2011. Interpretation of Skindex-29 scores: cutoffs for mild, moderate, and severe impairment of health-related quality of life. *Journal of Investigative Dermatology* 131(9), pp. 1945-1947.

Prinsen, C. A. C. et al. 2009. Health-related quality of life assessment in dermatology: interpretation of Skindex-29 scores using patient-based anchors. *Journal of Investigative Dermatology* 130(5), pp. 1318-1322.

Prinsen, C. A. C. et al. 2010. Health-Related Quality of Life Assessment in Dermatology: Interpretation of Skindex-29 Scores Using Patient-Based Anchors. *J Invest Dermatol* 130(5), pp. 1318-1322.

Quality-Metric. 2013. *Short-Form Survey 36 (SF-36)* [Online]. Quality Metric. Available at: www.sf-36.org [Accessed: 12 July 2013].

Rabin, R. et al. 2011. EQ-5D-5L User guide: Basic information on how to use the EQ-5D-5L instrument. Brussels: EuroQoL Group.

Ramos, R. et al. 2006. An assessment of anxiety in patients with primary hyperhidrosis before and after endoscopic thoracic sympathicolysis. *European Journal of Cardio-Thoracic Surgery* 30(2), pp. 228-231.

Rani, Z. et al. 2005. Quality of life issues and new benchmarks in the assessment of skin diseases. *Journal of Pakistan Association of Dermatologists* 15, pp. 339-344.

Raphael, D. et al. 1996. Quality of life indicators and health: current status and emerging conceptions. *Social Indicators Research* 39(1), pp. 65-88.

Rapley, M. 2003. *Quality of Life Research: A Critical Introduction*. London: SAGE Publications Ltd.

Redmond, F. 2010. Social Networking Sites: Evaluating and Investigating their use in Academic Research. In: *International Conference of Education, Research and Innovation 2010.* Madrid, 15 - 17 November.

Reeve, B. B. et al. 2007. Applying item response theory to enhance health outcomes assessment. *Quality of Life Research* 16, pp. 1-3.

Reeve, B. B. and Mâsse, L. C. 2004. Item Response Theory Modeling for Questionnaire Evaluation. In: Presser, S. et al. eds. *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken (NJ): John Wiley & Sons, Inc., pp. 247-273.

Reise, S. P. and Haviland, M. G. 2005. Item response theory and the measurement of clinical change. *Journal of personality assessment* 84(3), pp. 228-238.

Reise, S. P. and Henson, J. M. 2003. A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *Journal of Personality Assessment* 81(2), pp. 93-103.

Reise, S. P. et al. 2000. Factor analysis and scale revision. *Psychological Assessment* 12(3), p. 287.

Renzi, C. et al. 2001. Factors associated with patient satisfaction with care among dermatological outpatients. *British Journal of Dermatology* 145(4), pp. 617-623.

Revicki, D. et al. 2008. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology* 61(2), pp. 102-109.

Ritter, P. et al. 2004. Internet versus mailed questionnaires: a randomized comparison. *Journal of Medical Internet Research* 6(3).

Ro, K. M. et al. 2002. Palmar hyperhidrosis: evidence of genetic transmission. *J Vasc Surg* 35(2), pp. 382-386.

Robson, C. 2011. *Real World Research*. West Sussex: Wiley.

Rothman, M. et al. 2009. Use of existing patient-reported outcome (PRO) instruments and their modification: The ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. *Value in Health* 12(8), pp. 1075-1083.

Rothman, M. L. et al. 2007. Patient-Reported Outcomes: Conceptual Issues. *Value in Health* 10, pp. S66-S75.

Ruchinskas, R. 2007. Hyperhidrosis and anxiety: chicken or egg? *Dermatology* 214(3), pp. 195-196.

Ruchinskas, R. A. et al. 2002. The relationship of psychopathology and hyperhidrosis. *Br J Dermatol* 147(4), pp. 733-735.

Russo, P. A. J. et al. 2004. Psychiatric morbidity in psoriasis: A review. *Australasian Journal of Dermatology* 45(3), pp. 155-161.

Saldana, J. 2009. *The Coding Manual for Qualitative Researchers*. London: SAGE Publications Ltd.

Salek, M. and Kamudoni, P. 2013. Quality of life measurement in dermatology consultation: impact on patient reported outcomes. *Giornale italiano di dermatologia e venereologia: organo ufficiale, Societa italiana di dermatologia e sifilografia* 148(3), pp. 263-275.

Salek, M. and Luscombe, D. 1992. Health-related quality of life assessment: A review. *Journal of Drug Development* 5(3), pp. 137-153.

Salek, S. 1998. *Compendium of quality of life instruments*. Haslemere: Euromed Communications.

Salek, S. et al. 2007. The practical reality of using a patient-reported outcome measure in a routine dermatology clinic. *Dermatology* 215(4), pp. 315-319.

Sampogna, F. and Abeni, D. 2011. Interpretation of Skindex-29 scores. *J Invest Dermatol* 131(9), pp. 1790-1792.

Sayeed, R. A. et al. 1998. Quality of life after transthoracic endoscopic sympathectomy for upper limb hyperhidrosis. *Eur J Surg Suppl* (580), pp. 39-42.

Schäfer, I. et al. 2012. Prevalence and burden of hyperhidrosis in the adult population. *DMW-Deutsche Medizinische Wochenschrift* 137(S 03), p. A292.

Schipper, H. et al. 1990. Definitions and conceptual issues. In: Spilker, B. ed. *Quality of Life Assessments in Clinical Trials*. New York: Raven Press, pp. 11 - 23.

Schmidt, J. et al. 2006. Endoscopic thoracic sympathectomy for severe hyperhidrosis: impact of restrictive denervation on compensatory sweating. *Ann Thorac Surg* 81(3), pp. 1048-1055.

Schmitt, J. and Ford, D. E. 2007. Understanding the relationship between objective disease severity, psoriatic symptoms, illness-related stress, health-related quality of life and depressive symptoms in patients with psoriasis - a structural equations modeling approach. *Gen Hosp Psychiatry* 29(2), pp. 134-140.

Schmitt, T. A. 2011. Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment* 29(4), pp. 304-321.

Schneier, F. R. et al. 2012. Social anxiety and functional impairment in patients seeking surgical evaluation for hyperhidrosis. *Compr Psychiatry* 53(8), pp. 1181-1186.

Shea, T. et al. 2009. Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC Psychiatry* 9(1), p. 21.

Shikiar, R. et al. 2006. The validity and responsiveness of three quality of life measures in the assessment of psoriasis patients: results of a phase II study. *Health and Quality of Life Outcomes* 4(1), p. 71.

Smith, A. B. et al. 2008. Rasch fit statistics and sample size considerations for polytomous data. *BMC Medical Research Methodology* 8(1), p. 33.

Smith, E. 2002. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 3(2), pp. 205 - 231.

Snyder, C. F. et al. 2012. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Quality of Life Research* 21(8), pp. 1305-1314.

Solish, N. 2006. Assessing hyperhidrosis disease severity and impact on quality of life. *Cutis* 77(Suppl. 5), pp. 17-27.

Solish, N. et al. 2005. Prospective open-label study of botulinum toxin type A in patients with axillary hyperhidrosis: effects on functional impairment and quality of life. *Dermatol Surg* 31(4), pp. 405-413.

Solish, N. et al. 2007. A comprehensive approach to the recognition, diagnosis, and severity-based treatment of focal hyperhidrosis: recommendations of the Canadian Hyperhidrosis Advisory Committee. *Dermatol Surg* 33(8), pp. 908-923.

Solish, N. et al. 2008. Evaluating the Patient Presenting with Hyperhidrosis. *Thoracic Surgery Clinics* 18(2), pp. 133-140.

Spilker, B. and Revicki, D. 1996. Taxonomy of Quality of Life. *Quality of Life and Pharmacoeconomics in Clinical Trails*, pp. 25 - 32.

Stewart, A. and Ware, J. E. 1992. *Measuring functioning and well-being*. Durham (NC): The Rand Cooperation.

Stratford, P. W. et al. 1996. Health status measures: strategies and analytic methods for assessing change scores. *Physical Therapy* 76(10), pp. 1109-1123.

Stratford, P. W. and Riddle, D. L. 2005. Assessing sensitivity to change: choosing the appropriate change coefficient. *Health Qual Life Outcomes* 3, p. 23.

Streiner, D. L. and Norman, G. R. 2008. *Health Measurement Scales : A practical guide to their development and use*. Oxford: Oxford University Press.

Streker, M. et al. 2012. Hyperhidrosis plantaris – a randomized, half-side trial for efficacy and safety of an antiperspirant containing different concentrations of aluminium chloride. *JDDG: Journal der Deutschen Dermatologischen Gesellschaft* 10(2), pp. 115-119.

Strutton, D. R. et al. 2004. US prevalence of hyperhidrosis and impact on individuals with axillary hyperhidrosis: results from a national survey. *J Am Acad Dermatol* 51(2), pp. 241-248.

Swartling, C. et al. 2001. Botulinum A toxin improves life quality in severe primary focal hyperhidrosis.*Eur J Neurol*. Vol. 8. England, pp. 247-252.

Tabolli, S. et al. 2011. Assessment of the health status of 2,499 dermatological outpatients using the 12-item Medical Outcomes Study Short Form (SF-12) questionnaire. *Br J Dermatol* 165(6), pp. 1190-1196.

Tan, S. R. and Solish, N. 2002. Long-term efficacy and quality of life in the treatment of focal hyperhidrosis with botulinum toxin A. *Dermatol Surg* 28(6), pp. 495-499.

Teale, C. et al. 2002. Development, validity, and reliability of the Hyperhidrosis Impact Questionnaire (HHIQ). *Quality of Life Research* 11(7), pp. 702-702.

Tennant, A. and Conaghan, P. G. 2007. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research* 57(8), pp. 1358-1362.

Tennant, A. et al. 2004. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 7 Suppl 1, pp. S22 - 26.

Tennant, A. and Pallant, J. 2006. Unidimensionality Matters. *Rasch Measurement Transactions* 20(1), pp. 1048 - 1051.

Tennant, A. and Pallant, J. 2007. DIF matters: A practical approach to test if Differential Item Functioning makes a difference. *Rasch Measurement Transactions* 20(4), pp. 1082-1084.

Teresi, J. A. 2006. Different approaches to differential item functioning in health applications: Advantages, disadvantages and some neglected topics. *Medical Care* 44(11), pp. S152-S170.

Teresi, J. A. and Fleishman, J. A. 2007. Differential item functioning and health assessment. *Quality of Life Research* 16, pp. 33-42.

Terwee, C. et al. 2003. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Quality of Life Research* 12(4), pp. 349-362.

Terwee, C. B. et al. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology* 60(1), pp. 34-42.

Tesio, L. 2003. Measuring behaviours and perceptions: rasch analysis as a tool for rehabilitation research. *J Rehabil Med* 35, pp. 105 - 115.

Testa, M. A. 2000. Interpretation of quality-of-life outcomes: issues that affect magnitude and meaning. *Medical care* 38(9), pp. 166 -174.

The-EuroQoL-Group. 1990. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy* 16, pp. 199 - 208.

Thomas, M. et al. 2006. Panel discussions among hyperhidrosis patients to assess lifestyle limitations, treatment regimes, and compensating behaviors. In: *64th Annual Meeting of the American Academy of Dermatology.* San Fransisco, California, USA, March 3 - 7.

Thornicroft, G. and Slade, M. 2000. Are routine outcome measures feasible in mental health? *Quality in Health Care* 9(2), pp. 84-84.

Tristan, A. 2006. An Adjustment for Sample Size in DIF Analysis. *Rasch Measurement Transactions* 20(3), pp. 1070-1071.

Tu, Y. R. et al. 2007. Epidemiological survey of primary palmar hyperhidrosis in adolescent in Fuzhou of People's Republic of China. *Eur J Cardiothorac Surg* 31(4), pp. 737-739.

Tweet, M. S. et al. 2011. Spontaneous coronary artery dissection: a disease-specific, social networking community–initiated study. 86(9), pp. 845-850.

Twiss, J. et al. 2011. Can We Rely on the Dermatology Life Quality Index as a Measure of the Impact of Psoriasis or Atopic Dermatitis. *Journal of Investigative Dermatology* 132(1), pp. 76-84.

US-F.D.A. 2009. *US Food & Drug Administration guidance for industry-patient reported outcome measures: Use in medical product development to support labelling claims* [Online]. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf [Accessed: 15 January 2012].

Valderas, J. M. et al. 2008. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 17(2), pp. 179-193.

Visser, M. R. M. et al. 2005. Methods to detect response shift in quality of life data: a convergent validity study. *Quality of Life Research* 14(3), pp. 629-639.

Vorkamp, T. et al. 2010. Hyperhidrosis: Evolving concepts and a comprehensive review. *The surgeon: Journal of the Royal Colleges of Surgeons of Edinburgh and Ireland* 8(5), pp. 287-292.

Ware, J. 2003. Conceptualisation and measurement of Health Related Quality of Life: Comments on an Evolving Field. *Arch Phys Med Rehabil* 84(Suppl 2), pp. S43 - 51.

Ware, J. et al. 1998. The responsiveness of disease-specific and generic health measures to changes in the severity of asthma among adults. *Quality of Life Research* 7(3), pp. 235-244.

Ware, J. et al. 1994. *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston (MA): The Health Institute, New England Medical Centre.

Ware, J. et al. 1996. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 34, pp. 220 - 233.

Ware, J. et al. 1993. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: The Health Institute, New England Medical Center.

Ware, J. E. J. 1984. Methodology in behavioral and psychosocial cancer research. Conceptualizing disease impact and treatment outcomes. *Cancer* 53(10 Suppl), pp. 2316-2326.

Weber, A. et al. 2005. Psychosocial aspects of patients with focal hyperhidrosis. Marked reduction of social phobia, anxiety and depression and increased quality of life after treatment with botulinum toxin A. *Br J Dermatol* 152(2), pp. 342-345.

Whalley, D. et al. 2004. A new instrument for assessing quality of life in atopic dermatitis: international development of the Quality of Life Index for Atopic Dermatitis (QoLIAD). *British Journal of Dermatology* 150(2), pp. 274-283.

Whittemore, R. et al. 2001. Validity in qualitative research. *Qualitative Health Research* 11(4), pp. 522-537.

Williams, B. et al. 2010. Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care (JEPHC)* 8(3), pp. 1-13.

Wilson, I. B. and Cleary, P. D. 1995. Linking clinical variables with health-related quality of life. *JAMA* 273(1), pp. 59-65.

Wohl, Y. et al. 2007. Screening for common dermatologic disorders amongst Israeli adolescents. *Int J Dermatol* 46(10), pp. 1046-1049.

Wolfe, E. W. and Smith, E. V., Jr. 2007a. Instrument development tools and activities for measure validation using Rasch models: part I - instrument development tools. *J Appl Meas* 8(1), pp. 97-123.

Wolfe, E. W. and Smith, E. V., Jr. 2007b. Instrument development tools and activities for measure validation using Rasch models: part II--validation activities. *J Appl Meas* 8(2), pp. 204-234.

Wolosker, N. et al. 2010. Is gender a predictive factor for satisfaction among patients undergoing sympathectomy to treat palmar hyperhidrosis? *Clinics (Sao Paulo)* 65(6), pp. 583-586.

Wörle, B. et al. 2007. Definition and treatment of primary hyperhidrosis. *JDDG: Journal der Deutschen Dermatologischen Gesellschaft* 5(7), pp. 625-628.

Wright, B. D. and Masters, G. N. 1982. *Rating Scale Analysis*. Chicago (IL): Mesa Press.

Wynd, C. A. et al. 2003. Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research* 25(5), pp. 508-518.

Wyrwich, K. W. et al. 2005. Estimating clinically significant differences in quality of life outcomes. *Quality of Life Research* 14(2), pp. 285-295.

Yardley, L. 2008. Demonstrating validity in qualitative psychology. *Qualitative psychology: A practical guide to research methods* 2, pp. 235-251.

Yost, K. J. and Eton, D. T. 2005. Combining distribution-and anchor-based approaches to determine minimally important differences. *Evaluation & the Health Professions* 28(2), pp. 172-191.

Young, O. et al. 2003. Evaluation of the impact of transthoracic endoscopic sympathectomy on patients with palmar hyperhydrosis. *Eur J Vasc Endovasc Surg* 26(6), pp. 673-676.

Zimmer, M. 2010. "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology* 12(4), pp. 313-325.

PUBLICATIONS AND PRESENTATIONS

Conference-Oral presentations

Kamudoni, P., Mueller, B, Salek, S. Validation of the final version of the Hyperhidrosis Quality of Life Index (HidroQoL). *International Society for Quality of Life Research (ISOQOL) 19th Annual Conference, October 2013. Miami, USA*

Conference-Poster presentations

Kamudoni, P., Mueller, B, Mueller, C., Salek, S. Hyperhidrosis greatly influences patients quality of life: qualitative study using online social networks, Poster presentation, *21st Congress European Academy of Dermatology and Veneology, September 2012. Prague, Czech Republic*

Kamudoni, P, Salek, MS, Mueller, B. and Mueller, C. The Hyperhidrosis Quality of Life Index (Hidroqol©): a novel patient reported outcome measure in hyperhidrosis. *42nd Annual Meeting of the European Society for Dermatological Research, September 2012. Venice, Italy.*

Kamudoni, P., Mueller, B, Mueller, C., Salek, S. Qualitative development and content validation of a new patient reported outcome measure in hyperhidrosis, the Hyperhidrosis Quality of Life Index (HidroQoL). *International Society for Quality of Life Research (ISOQOL) 19th Annual Conference, October 2012. Budapest, Hungary.*

Journal articles

Salek, S.S., Kamudoni, P. Quality of life measurement in dermatology consultation: impact on patient reported outcomes. *Giornale Italiano di Dermatologia e Venereologia*. May 2013.

Abstracts

Kamudoni, P, Salek, MS, Mueller, B. and Mueller, C.. Hyperhidrosis Quality of Life Index-Hidroqol©: a novel patient reported outcome measure in hyperhidrosis. *Journal of Investigative Dermatology* 2012; 132 (S2): S73.

Kamudoni, P., Mueller, B, Mueller, C., Salek, S. Qualitative development and content validation of a new patient reported outcome measure in hyperhidrosis, the Hyperhidrosis Quality of Life Index (HidroQoL). Quality of Life Research 2012; 21 (Supplement 1) (A3089).

APPENDICES

Appendix I – Study protocol

Appendix II – Approval from University Hospital Greifswald Ethics Committee

Appendix III – Qualitative study: Patient information sheet and consent form

Appendix IV – Qualitative study: Topic guide for Interviews

Appendix V – Qualitative study: screenshot for online survey

Appendix VI – Content validation questionnaire

Appendix VII – Copy of the EQ – 5D

Appendix VIII – Copy of the DLQI

Appendix IX – Copy of the Skindex-17

Appendix I

Conceptualisation, development and validation of a bi-lingual novel disease-specific hyperhidrosis quality of life questionnaire for use in routine clinical practice

STUDY PROTOCOL

Submitted to the University Hospital Greifswald Ethics Committee

Prepared by
Paul Kamudoni

July 2011

1. Background and Rationale

Hyperhidrosis is a skin-disorder characterised by excessive sweating beyond thermo-regulation requirements of the body and it affects 1-2% of the population. The impact of hyperhidrosis on the quality of life of its sufferers has been previously studied. Patients have reported suffering emotional distress, considerable impairment in relation to performance of work-related as well as household tasks, and dysfunctional social life (Hamm et al. 2006; Solish et al. 2006, Strutton et al., 2004). Various questionnaires were utilised in these studies including those specific for hyperhidrosis (HDSS, HHIQ and HS), those for all skin-disorders, measuring dermatology specific quality of life (e.g. DLQI) and more generic questionnaire measuring health related quality of life in general (e.g. SF-12). These studies underscore the importance or relevance of understanding the quality of life of hyperhidrosis patients, as an important component of any efforts at improving the care and treatment of hyperhidrosis patients. Considering the patients' subjective experience with the disease and their personal account of how it affects them also reflects the WHO's concept of health, where health is defined holistically beyond clinical measures of disease (See Review Paper for Ref.).

Most hyperhidrosis specific instruments are appropriate and validated for sub-groups of hyperhidrosis patients, mostly based on location of the disease or the treatment strategy received (e.g. Hyperhidrosis Scale). The promising few (e.g. *Hyperhidrosis Impact Questionnaire*) seem more fitting for research settings rather than daily routine clinical practice. As such a gap still exists for a disease specific tool, generic enough to cover all types of hyperhidrosis (according to body location), concise enough for use in routine clinical practice. While previous studies have reported negative impact of hyperhidrosis on quality of life, an in-depth investigation of quality of life issues of hyperhidrosis patients using a triangulation of various qualitative methods has not been done. Additionally a comparison across various sub-groups (between UK and German patients) has also not been done.

A novel disease specific questionnaire covering all forms of hyperhidrosis will enhance the understanding of the impact of hyperhidrosis on the life of patients. We expect that this will assist in improving the diagnosis and clinical management of hyperhidrosis. Additionally an in-depth understanding of the impact of hyperhidrosis on the patients' quality of life would reveal the treatment needs of hyperhidrosis patients, particularly pointing towards additional care or services that these patients require. The potential implication is a modification of current treatment guidelines for hyperhidrosis based on the new evidence that this study would create.

2. Research Question and Objectives

The aim of this study is to develop and validate a disease specific questionnaire for measuring the impact of hyperhidrosis on the quality of life of patients. Additional secondary objectives will also be achieved, this includes to investigate the impact of hyperhidrosis on the daily lives of patients; and to conduct sub-group comparisons of the impact of hyperhidrosis on the quality of life of patients across cultures (in particular between U.K and Germany).

3. Study design and methodology

This study will utilise a variety of methodologies combining qualitative and quantitative approaches. Document research (review of literature), semi-structured interviews and focus groups will be used in the first part of the study. The latter steps will include Quantitative approaches - survey questionnaire. The study will involve patients with clinically diagnosed hyperhidrosis, recruited from Germany and from the UK Multiple steps will be followed, including conceptualisation, qualitative data collection, and qualitative development of questionnaire, pre-testing, and validation reflecting the complex process of developing new questionnaires. Each step is elaborated on below:

3.1. Step 1: Conceptualisation

Conceptualization will involve extensive review of the literature on the impact of hyperhidrosis on the health related quality of life of patients, including reviewing questionnaires previously used in studies evaluating quality of life of hyperhidrosis patients. The main output of this stage is a conceptual framework of quality of life issues in hyperhidrosis patients.

3.2. Step 2: Qualitative data collection

A triangulation of qualitative data collection approaches will be applied for this phase combining the use of the internet, telephone and postal mail. Online text based focus groups will be organised via secure, password protected online discussion platform designed for this purpose. Interviews will be held via telephone or through online Instant Messaging facilities (Skype & Facebook). Additionally, postal and online surveys (containing open questions) will also be used.

At least 70 hyperhidrosis patients will be sampled for this phase. We anticipate this phase to take 8 weeks. Given that the goal in qualitative analysis is to gain deep insight in an issue rather than make statistical inference, there is no recommendable sample size or formulae for deriving one. Nevertheless, a general rule of thumb is to continue sampling until a 'saturation point' is reached i.e. where no additional themes and issues are emerging from additional subjects (Bowling 2009, p.410). The proposed sample size is based on studies addressing similar research question.

In consideration of the clear differences between online patient populations and patients accessing care in the clinic (Langenbruch et al. 2009) a purposive sampling strategy will be employed, in order to obtain patients via various channels. This is a deliberate strategy to minimise bias in the collection of issues. Thus patients will be recruited via online patient support groups and forums as well as through the dermatology practices or clinics (primary care and secondary care setting) through doctors.

This process will take place in parallel, for the U.K and German patients.

3.3. Step 3: Qualitative development of questionnaire

This phase will involve a structured content analysis of the transcripts of the focus group and interview as well as the data collected via the open questions of postal survey information with the aim of identifying major themes and issues emerging from both the UK and German patients. This process will again take place in parallel, for the English and German patient groups. Two panels, one English-speaking and the other for German speaking, comprised of dermatologists and patients will be asked to review the themes and issues identified. The first drafts of the English and German versions of the questionnaire will be prepared based on this review.

3.4. Step 4: Pre-testing

The first draft of the questionnaire will be pretested in a sample of 30 patients. Purposive sampling strategy will be used to ensure that representativeness based on type of hyperhidrosis is maintained, in both the German and English patient groups. Patients will firstly be asked to complete the draft questionnaire, followed then by an interview on their understanding of the questions as well as their evaluation of other aspects of the questionnaire e.g. comprehensiveness, relevance of questions, flow of questions.

Patients in Germany will be recruited via primary care physicians and will complete the paper-and-pencil version of the questionnaire while the interviews will be conducted via telephone. For UK patients the same process will be completed using electronically sent questionnaires (E-Mail) and telephone for the interviews.

3.5. Step 5: Field-testing

This will mainly focus on establishing the questionnaire's construct validity, reliability and sensitivity. The questionnaire will be given to at least 250 patients recruited through the dermatologists or general practitioners; and through online-patient support groups.

Construct Validity

Construct validity is one of the core psychometric properties that a valid questionnaire must posses. It reflects the theoretical relationships of items to each other and to the hypothetical scale (Basra et al. 2007). This will be assessed using factor analysis. A purposive sample of 400 patients will be drawn, including German (n=200) and English (n=200) patients. Same channels used to recruit patients in the qualitative data collection phase will be applied in this phase also.

Convergence Validity

Besides the new questionnaire, patients will also be asked to complete a second questionnaire – the DLQI (this instrument was also developed by Cardiff University – approval from license holder is yet to be obtained. We expect that the patients' skin quality of life to be related to the quality of life based on the disease specific measure. The DLQI score should therefore be correlated with the score of the new measure.

Reliability

A questionnaire's ability to produce consistent and reproducible results reflects its reliability properties. Intra-class correlations among items within and with other scales will be used in assessing internal consistency of the new questionnaire. Data already collected for construct validity assessment will be utilised in establishing this. Assessment of reliability over time, test-retest reliability, will be done by asking the patients sampled for field testing to complete the new questionnaire again 7 – 10 days after initial completion.

Sensitivity

The ability of a questionnaire to detect existing differences between individual or groups of patients (Fayers 2007, p.101) reflects its sensitivity. Such a property is very useful especially in discriminative instruments intended for use in diagnosing patients i.e. identification of clinically relevant differences (Ibid). This will be assessed by measuring the standard response means based on the data already provided by the patients for construct validation. Disease-free subjects, German (n = 50) as well as English (n = 50), will be recruited to act as a control group, in further analyses of the instruments sensitivity.

4. Patient Selection

Participants for steps 2 to 5 of this study must fulfil the following inclusion criteria;

- Must have a medically confirmed diagnosis of hyperhidrosis except for the control group recruited in the sensitivity done in step 5;
- Be seeking for treatment for the hyperhidrosis.
- Aged 16 and above.
- Able to understand and read German (in the case of the German patients).
- Able to understand and read English (in the case of UK patients)
- Capable of giving informed consent.

5. Ethical Considerations

As this study does not involve any interventions, the risk to which the study participants will be exposed is minimal. Nevertheless considerations have been given to the studies research processes and how possible effects on patients can be minimised.

- Qualitative data collection will involve discussing the impact of hyperhidrosis on the patient's everyday life. The participant's consent will be obtained before the start of interview/focus group. We envisage that personal and embarrassing issues may arise in the process. Should a participant feel uncomfortable to proceed with the discussion, they

will be free to withdraw their consent, which will lead to a prompt termination of interview/focus group and deletion of information collected as per participants' wishes.

- Written consent will also be sought from patients responding to both the survey administered as part of qualitative data collection and the questionnaires completed in both the pre- and field testing. Patients will be given the opportunity to ask questions via contacts provided on the information sheet. The set of documents sent to the patients both via post and electronic will include an information sheet and a consent form which the patients will be asked to enclose together with their questionnaires in their reply.
- While the study does not confer direct immediate benefits to the participants, we believe this research will lead to better management and treatment of hyperhidrosis patients, thus enhancing their quality of life.
- Handling of the personal data will be in line with both Germany and UK personal data and privacy laws. All data received through electronic means will be stored in a password protected server/database and will be only accessible to the research team. Similarly all data collected in paper and pencil format will be stored in a secure storage location, only accessible to the research team. All data from patients will be confidential and thus will not be shared outside the research team. The data collected will be held for a period of 6-12 months after the completion of the project.

6. Communication of results

- Results of the structured content analysis will be presented at international conferences – and also published in a peer reviewed journal under the title: "comparison of quality of life issues between German and UK patients."
- The major themes/issues identified from the structured content analysis will form the input for the questions in the new questionnaire (English and German versions).
- The results from the pre-test and field-testing phase will be presented at international conferences, and will also lead to publications of the various aspects of the psychometric properties of both the English-version and the German version of the questionnaire.

Appendix II

Universitätsmedizin
GREIFSWALD

Universitätsmedizin Greifswald · Fleischmann Strasse 8 · D-17475 Greifswald

Herrn P. Kamudoni
RIEMSER Arzneimittel AG
An der Wiek 7

D-17493 Greifswald-Insel Riems

EINGANG
04. AUG. 2011
erledigt

Ethikkommission

GESCHÄFTSSTELLE
Universitätsmedizin Greifswald
Ethikkommission
Institut für Pharmakologie
Friedrich-Löffler Strasse 23 d
D-17487 Greifswald

BEARBEITER
Frau Dr. K. Kindermann

DATUM: 28.07.2011

Votum der Ethikkommission

Titel der Studie:	Konzeptualisierung, Entwicklung und Validierung eines neuartigen, bilingualen, krankheitsspezifischen Lebensqualitätsfragebogens zum routinemäßigen Einsatz in der Hyperhidrosetherapie (RIQOL-H)
Antrag vom:	22.07.2011
Eingegangen am:	22.07.2011
Reg.-Nr.:	BB 84/11

Sehr geehrter Herr Kamudoni,

die Ethikkommission der Universitätsmedizin Greifswald hat die zum o.g. Versuchsplan eingereichten Unterlagen in ihrer Sitzung am 26.07.2011 geprüft.

Die Kommission stellte mehrheitlich fest, dass gegen die Durchführung der Studie keine ethischen und rechtlichen Bedenken bestehen und befürwortet deshalb das Vorhaben.

Die Mitglieder der Kommission wünschen Ihnen viel Erfolg bei der Durchführung des Vorhabens.

Mit freundlichen Grüßen

Prof. Dr. W. Siegmund
Vorsitzender der Ethikkommission

Telefon:
03834 – 86 5644

Telefax:
03834 – 86 5631

e-mail:
ethik@uni-greifswald.de

Internet:
www.medizin.uni-greifswald.de

Bankverbindung:
Deutsche Bundesbank Rostock
Konto-Nr. 130 015 30
BLZ 130 000 00

USt ID: DE137584813

Zur Begutachtung haben der Kommission vorgelegen:

- Antrag an die Ethikkommission vom 22.07.2011
- Datenschutzerklärung vom 22.07.2011
- Kostenübernahmeerklärung vom 22.07.2011
- Prüfplan
- Patienteninformation, Einwilligungs- und Datenschutzerklärung
- Patientenfragebogen
- Lebenslauf P. Kamudoni

Der Ethikkommission gehören an:

reguläre Mitglieder	*ständige Stellvertreter*
Prof. Dr. M. Lerch Klinik für Innere Medizin A	*Prof. Dr. J. Mayerle ** *Klinik für Innere Medizin A*
Prof. Dr. R. Biffar Zentrum für Zahn-, Mund- und Kieferheilkunde	*Dr. I. Polzer* *Zentrum für Zahn-,* *Mund- und Kieferheilkunde*
Prof. Dr. U. Runge Klinik und Poliklinik für Neurologie	*Prof. Dr. A. Hamm* *Institut für Psychologie*
OA Dr. M. Gründling * Klinik für Anästhesiologie und Intensivmedizin	*OÄ Dr. S. Friesecke* *Klinik für Innere Medizin B*
Prof. Dr. W. Siegmund * Institut für Pharmakologie	*Dr. Dr .G. Engel* *Universitätsapotheke*
Prof. Dr. Th. Kohlmann * Institut für Community Medicine	*Prof. Dr. W. Hoffmann* *Institut für Community Medicine*
PD Dr. B. Bockholdt * Institut für Rechtsmedizin	*Prof. Dr. W. Joecks ** *Fakultät für Rechts- und Staatswissenschaft*
Prof. Dr. H.-W. Eckert Fakultät für Rechts- und Staatswissenschaft	*Prof. Dr. J. Lege und Prof. Dr. C. D. Classen* *Fakultät für Rechts- und Staatswissenschaft*
Prof. Dr. H. Assel Theologische Fakultät	*Prof. Dr. K. Ott* *FB Biologie*
Prof. Dr. H. Lauffer Klinik und Poliklinik für Kindermedizin	*PD Dr. R. Bruns* *Klinik und Poliklinik für Kindermedizin*
Prof. Dr. R. Sudick Klinik und Poliklinik für Frauenheilkunde Klinikum Neubrandenburg	*Prof. Dr. M. Zygmunt* *Klinik und Poliklinik für Frauenheilkunde und* *Geburtshilfe*
PD Dr. H.-C. Schober * Klinik für Innere Medizin Klinikum Südstadt Rostock	*PD Dr. R. Möllmann* *niedergelassener Internist, Greifswald*
Katharina Schade, Medizinstudentin	*Rahel Österreicher-Lutz, Medizinstudentin **

* bei der Sitzung am 26.07.2011 anwesend

Appendix III

Informed Consent Form for Hyperhidrosis patients being invited to participate in a focus group discussion/interview on the impact of hyperhidrosis on your daily life.

09 August 2011

This Informed Consent Form has two parts:

- Information sheet (to give you the background to this study)
- Certificate of Consent (to show your acceptance to take part in this study).

Part I: Information Sheet

Introduction

Hyperhidrosis, excessive sweating beyond physiological and environmental requirements of the body is known to have a number of unpleasant effects on the patients. Previous research has observed its impact on the patients' everyday life and well-being. Assessing the impact of hyperhidrosis on daily life of patients is at the heart of the process of diagnosis of hyperhidrosis

Purpose of research

We would like to develop an easy to use hyperhidrosis-specific instrument for evaluating the impact of hyperhidrosis on quality of life that can be applicable across all severity of the disease and appropriate in evaluating all treatment strategies. To do this we are interested in learning from your experiences with hyperhidrosis. We would like to know the different ways in which hyperhidrosis affects your daily life and quality of life.

Type of Research Intervention

This research will involve your participation in an focus group discussion/interviews.

Voluntary Participation

Be aware that your participation in this research is entirely voluntary.

Procedure

If you accept the invitation to participate in this research you will be asked to take part in an online discussion with other persons also experiencing problems with excessive sweating. The online discussion will require a few minutes daily to read the discussion board to take note of the discussion topic, to respond to the issue introduced by the moderator and also respond the posts made by other respondents. Once the discussion is concluded you will be asked to verify a summary list of main themes/issues discussed (this may happen at about seven days after the discussions are concluded).

Duration

The online discussion will run for a period of 14 days.

Confidentiality

The experiences you share will be treated anonymously. Your personal information will not be shared outside the research team. Your information will only be identified by an ID number instead of your name. To ensure this, right from the beginning your will choose a user name which you can be identified with in all the discussions, to follow your contributions.

Use of Results

The knowledge gained from the discussion will be used as input into a new questionnaire/instrument. Our summary of the issues from the discussions will first be shared with you before the final instrument is developed and shared with the public.

Part II: Certificate of Consent

I have read the foregoing information in part I. I have had the opportunity to ask questions where I had my doubts. All issues and questions I raised have been addressed to my satisfaction. I consent voluntarily to be a participant in this study.

Name: _____

Signature: _____

Date: _____

I confirm that the participant was given an opportunity to ask questions about the study, and all their questions and concerns have been addressed and to the best of my ability. I confirm that the participant was not coerced into giving consent, and the consent has been given freely and voluntarily.

A copy of this ICF has been provided to the participant.

Name of Researcher: Paul Kamudoni, MSc.

Signature: _____

Date: _____

Appendix IV

Topic guide for focus group discussion/interviews

9 July 2011

- Tell me a bit about yourself e.g. how long you have had hyperhidrosis and any background information in relation to the disease.
- Can you tell me about your experience of the first time you got diagnosed with hyperhidrosis i.e. How difficult was it for your physician/doctor to diagnose you with hyperhidrosis?
- What was it about the sweating that made you seek for medical attention i.e. what specific factors did you consider in judging that your sweating is problematic [what issues does the patient consider in evaluating the severity of his/her sweating]
- Which areas of your life have been most affected by hyperhidrosis?
  [probes based on previous studies]
  - Treatment history & Level of satisfaction with past treatments.
  - Physical ailment.
  - Functional e.g. writing, manual work
  - Limitation on daily life/ everyday life e.g. at home, dressing.
  - Social life
  - Psychological well-being,
  - Personal domain or affection – hugging, intimate touching
  - Therapy
  - Satisfaction with life
  - Impact on employment and productivity
  - Special circumstances e.g. writing exams,
  - Special effects particular/specific to sweating of hands, feet or underarms or any other special area.

- Can you describe how each of the areas you have mentioned has been affected?
- Which usual daily activities are you uncomfortable to do due to your hyperhidrosis?
- What aggravates your hyperhidrosis?
- How satisfied are you with your life? [am not interested in general satisfaction with life of patient per se]
- Can you share with me about the treatments you have taken for your hyperhidrosis
  [Interest here is to get the patient to talk about how they feel about the treatments they have taken]
  - What are your views and feelings on the currently available treatments for hyperhidrosis?
- How are you currently managing your excessive sweating i.e. treatment or any efforts to control its effects?
- How have you adapted your life to living with hyperhidrosis? [this may reflect a combination of limitations in functioning and other psychological limitation]

Appendix VI

Content validation Questionnaire

(Only first three pages included)

20 January 2012

Name : _____

Thank you for agreeing to take part in questionnaire feedback process as part of content validation.

Each item on the questionnaire needs to be assessed for language clarity, completeness, and relevance and scaling.

The following definitions are provided to ensure standardisation so that each person has the same understanding of the criteria.

Please rate each of the items on the following:

A. **Language clarity**: The sentences and wording should be clear, understandable, straightforward and simple. Phrases and wording should be unambiguous and jargon free and should be understood by someone with a reading ability of 12 years.
B. **Completeness**: The sentences should be complete, not broken and should end appropriately.
C. **Relevance**: Each statement should be relevant to the subject area of the target population
D. **Scaling**: Scaling refers to the scoring system, with the five response options. Panel members should rate whether the response options fit the statement or not.

Statement 1: My choice of clothing is affected					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	

Statement 2: My choice of footwear is affected					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 3: My holiday is affected					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 4: I have difficulties gripping objects					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 5: I have difficulties handling money					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 6: I have difficulties with physical contact with others					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	

Statement 7: My hobbies are affected					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 8: I avoid speaking with groups of people					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 9: My physical activities are affected					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	
Statement 10: My outdoor activities are affected					
	Strongly agree	**Agree**	**Agree**	**Strongly disagre**	
Language Clarity	☐	☐	☐	☐	
Completeness	☐	☐	☐	☐	
Relevance	☐	☐	☐	☐	
Scaling	☐	☐	☐	☐	

Exit this survey

Living with Hyperhidrosis

Dear Sir/Madam,

I am inviting you to take part in this study investigating the impact of hyperhidrosis on the the daily lives of hyherhidrosis patients, being undertaken by the Welsh School of Pharmay (Paul Kamudoni, M.Sc.). Our aim is to create a user-friendly disease specific quality of life measure that can reduce the physician-patient communication gap for hyperhidrosis patients, thereby enhancing the diagnosis and treatment of hyperhidrosis. Your input, as a hyperhidrosis patients is crucial to make this achievable.

We ask you to respond to the questions contained in this surveys, being as detailed as you can be where necessary. Further to that we invite you to a telephone or internet based interview for a more indepth discussion on the same topic. If you have any questions please send me an email on kamudoniP@cardiff.ac.uk. Your participation in this study is voluntary.

I would like to assure you that the information you share with us will be treated with utmost confidentiality, in line with existing data protection and privacy regulations in the U.K. Moreover, your data will not be shared outside the research team. The information that you provide will not be directly attributed to you in the publication of our findings. The data will be stored in a password protected computer, only accessible to the research team, for a period of up to12 months, after when it will be deleted.

If you consent to the above, please proceed to the questions.

Sincerely,

Paul Kamudoni

School of Pharmacy (Cardiff Uni.)
Redwood Building,
King Edward VII Avenue,
Cardiff CF10 3XF,
Wales, U.K.

Submit

Exit this survey

Living with Hyperhidrosis

1. What is your nationality; date of birth ? (*mm/dd/yyyy*)

2. What is your gender ?

male

female

3. How long have you suffered from hyperhidrosis ?

4. Which body-site does your hyperhidrosis affect ?

5. How was your hyperhidrosis diagnosed?

6. Which areas of your daily life have been most affected by

hyperhidrosis ?
(please give at least 5 areas)

7. Please describe in detail how each of the areas you mentioned has been affected?

8. What activities are you uncomfortable doing because of your hyperhidrosis?

9. How have you adapted to living with hyperhidrosis?

10. Would you like to further discuss your experiences through an interview ?
If yes, please indicate your contacts and preferable date and time.

E-mail/ Tel.

Date & time

Prev Submit

Consent Form

Before you proceed to the questionnaire, please take a moment to complete this consent form.

Click here to read the hyperhidrosis study patient information sheet (pdf) before you give consent

*1. I confirm that I have read the patient information sheet and that I understand the purpose and nature of this study.

*2. I have had the opportunity to ask questions about all aspects of the study. All issues I raised have been addressed to my satisfaction.

*3. I understand that the study will handle all information from me with confidentiality, any publication resulting from the study shall not reveal my name or any other personally identifying information

I accept | I decline

CONTACT

For any enquiry about the survey, please contact Paul Kamudoni

Kamudonip@cardiff.ac.uk

www.cardiff.ac.uk/phrmy/cser

Please read the patient information page here

Appendix XI

(English version for the UK)

Under each heading, please tick the ONE box that best describes your health TODAY

MOBILITY

I have no problems in walking about ❑
I have slight problems in walking about ❑
I have moderate problems in walking about ❑
I have severe problems in walking about ❑
I am unable to walk about ❑

SELF-CARE

I have no problems washing or dressing myself ❑
I have slight problems washing or dressing myself ❑
I have moderate problems washing or dressing myself ❑
I have severe problems washing or dressing myself ❑
I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

I have no problems doing my usual activities ❑
I have slight problems doing my usual activities ❑
I have moderate problems doing my usual activities ❑
I have severe problems doing my usual activities ❑
I am unable to do my usual activities ❑

PAIN / DISCOMFORT

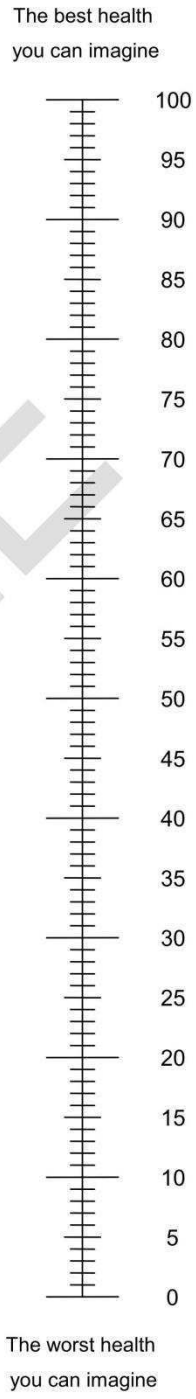
I have no pain or discomfort ❑
I have slight pain or discomfort ❑
I have moderate pain or discomfort ❑
I have severe pain or discomfort ❑
I have extreme pain or discomfort ❑

ANXIETY / DEPRESSION

I am not anxious or depressed ❑
I am slightly anxious or depressed ❑
I am moderately anxious or depressed ❑
I am severely anxious or depressed ❑
I am extremely anxious or depressed ❑

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



Appendix XII

DERMATOLOGY LIFE QUALITY INDEX

DLQI

Hospital No: Date:
Name: Score:
Address: Diagnosis:

The aim of this questionnaire is to measure how much your skin problem has affected your life OVER THE LAST WEEK. Please tick ☑ one box for each question.

1.	Over the last week, how **itchy**, **sore**, **painful** or **stinging** has your skin been?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	
2.	Over the last week, how **embarrassed** or **self conscious** have you been because of your skin?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	
3.	Over the last week, how much has your skin interfered with you going **shopping** or looking after your **home** or **garden**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
4.	Over the last week, how much has your skin influenced the **clothes** you wear?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
5.	Over the last week, how much has your skin affected any **social** or **leisure** activities?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
6.	Over the last week, how much has your skin made it difficult for you to do any **sport**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
7.	Over the last week, has your skin prevented you from **working** or **studying**?	Yes No	❒ ❒	Not relevant ❒
	If "No", over the last week how much has your skin been a problem at **work** or **studying**?	A lot A little Not at all	❒ ❒ ❒	
8.	Over the last week, how much has your skin created problems with your **partner** or any of your **close friends** or **relatives**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
9.	Over the last week, how much has your skin caused any **sexual difficulties**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
10.	Over the last week, how much of a problem has the **treatment** for your skin been, for example by making your home messy, or by taking up time?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒

Please check you have answered EVERY question. Thank you.

Appendix XIII

No. item Skindex-17	Description of the item **In the last 4 or 1 weeks, how frequent have you experienced the following…..**	Scoring†
1*	My skin hurts	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
2+	My skin condition makes it hard to work or do hobbies	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
3+	My skin condition affects my social life	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
4+	My skin condition makes me feel depressed	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
5+	I tend to stay at home because of my skin condition	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
6*	My skin itches	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
7+	My skin condition affects how close I can be with those I love	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
8+	I tend to do things by myself because of my skin condition	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
9*	Water bothers my skin condition (bathing, washing hands)	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
10+	My skin condition makes showing affection difficult	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
13*	My skin is irritated	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
11+	I am embarrassed by my skin condition	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
12+	I am frustrated by my skin condition	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
14+	My skin condition affects my desire to be with people	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
15+	I am humiliated by my skin condition	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
16*	My skin condition bleeds	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time
17+	My skin condition interferes with my sex life	[ ] Never [ ] Rarely [ ] Sometimes [ ] Often [ ] All the time