University of Alberta

Individualized Health Related Quality of Life Measures: their use in children and their psychometric properties

by

Sana Ishaque

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Masters of Science

in

Clinical Epidemiology

School of Public Health

©Sana Ishaque

Fall 2012
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Abstract

Health related quality of life (HRQL) has recently been recognized as an essential outcome of patient-centered care. Individualized HRQL (iHRQL) measures propose a patient-centered approach to HRQL measurement; these measures allow patients to nominate areas of life that are important and then score them from own perspective. The Measure Yourself Medical Outcome Profile (MYMOP) and four of its adaptations are adult iHRQL measuers; in this thesis they were critically appraised and a systematic review was conducted to identify any comparable pediatric measures. Reporting of the validation of MYMOP and its adaptations were inconsistent. Although pediatric iHRQL measures were identified, none of these measures met all currently recommended quality criteria for measurement properties. The available pediatric literature on iHRQL measures does not support their use in children without further validation work. In addition, reported HRQL definitions were heterogeneous, limiting their external validity.

Acknowledgments

I dedicate this thesis to all the children of the world.

I would like to thank my family and friends for their continuous support, patience, and understanding. To my mom, thank you for all the love, tenderness, and prayers that you make for me. I would specially like to thank my host family in Edmonton; without the support of Kamal family, this thesis would have been even more challenging.

I would like to thank Dr. Sunita Vohra as well as my other committee members Dr. Jeffrey A. Johnson, and Dr. Charlotte Paterson for their knowledge, support, and collaboration throughout my thesis. Their feedback was always timely, well-pointed, and meticulous.

Dr. Caroline Terwee, thank you for your support, for patiently replying to my lengthy emails, and for generously sharing your methodologic expertise. Thank you Fatima Al Sayah, for seconding the systematic review screening, and Tamara Dorfman for data extraction.

I would like to thank my supervisor for funding my thesis research.. I would also like to thank the members of her research team, especially her other graduate students. In particular I would like to thank Bradley C. Johnston, Larissa Shamseer, and Denise Adams for their support during my graduate studies.

Table of Content

List of Tables

List of Figures

List of Abbreviations

AQLQ: Asthma Quality of Life Questionnaire
COSMIN: COnsensus based Standards for the selection of Health Measurement INstrument
HRQL: Health related quality of life
iHRQL: Individualize health related quality of life
JAQQ: Juvenile Arthritis Quality of Life Questionnaire
MACTAR: McMaster-Toronto Arthritis Questionnaire
MYCaW: Measure Yourself Concern and Wellbeing
MYMOP: Measure Yourself Medical Outcome Profile
PAQLQ: Pediatric Asthma Quality of Life Questionnaire
PCORI: Patient Centered Outcomes Research Institute
PGI: Patient Generated Index
PSYCLOPS: Psychological Outcome Profiles
RG: Repertory Grid
SEIQoL-DW: Schedule for the Evaluation of Individual Quality of Life-Direct Weight

Chapter 1

1. Introduction

1.1. Emerging importance of patient centered research

There is growing recognition in healthcare that patients should be involved as partners in clinical care and research [1,2]. At an individual level, patients actively participate in the decision-making process about their own healthcare and are encouraged to take increased responsibility for managing their disease. On a broader scale, patients' feedback on healthcare services and provision is required to develop patient-centered healthcare policy and research, clinical practice guidelines, and patient information materials. A patient-centered health service performing these tasks should not only recognize the importance of patients' perspective and preferences but should also incorporate them into clinical practice and research. Patient-centered research, in its essence, is collaborative, inter-professional, and multidisciplinary [3,4]; its research agenda is informed by patients and captures outcomes that they consider (or perceive) are important[1,2].

A recent American initiative to encourage patient-centered research is the establishment of Patient Centered Outcomes Research Institute (PCORI)[1]. This Institute was established to bridge the identified gap [5,6] between healthcare research/practice and consumers' (or patients') interest. Health related quality of life (HRQL) is one the most important outcomes that are recognized by PCORI to be important to patients [7]. To our knowledge, at present there is no comparable Canadian initiative available.

1.2. Health related quality of life (HRQL) and HRQL measures

Health related quality of life (HRQL) measures have emerged as an essential part of

outcome assessment of patient-centered research [1]. HRQL is a multidimensional concept; it is the subjective perception of one's quality of life regarding health-related issues. There is no consensus on its definition. It may be defined as "*the physical, psychological, and social domains of health, seen as distinct areas that are influenced by a person's experience, belief, expectations, and perception"*[8]. Another definition is "*an individual's subjective perception of the impact of their health status, including disease and treatment, on physiological, psychological, and social functioning"*[9]. Both of these definitions are based on the World Health Organization's definition of *health* whereby *"health is a state of complete physical, mental, and social wellbeing and not merely the absence of disease or infirmity"* [10]. In the absence of a universal definition of HRQL, it is important that investigators define what they mean by the concept before opting to measure it.

The measurement of HRQL depends on the patient's subjective perception of the quality of his/her physiological, psychological, and social domains that are considered important by him/her[11].

Health related quality of life measures (or instruments) may be classified as generic or disease-specific [13-15]. Generic instruments can be used to measure HRQL of healthy as well as diseased populations; these measures allow for comparisons across different diseases, and between healthy and diseased populations[14,15]. Disease-specific measures, on the other hand, are used to assess HRQL of a population with a specific disease; they are good at picking small but important differences [14,15]. There is a clear mismatch between the number of disease-specifc individualized HRQL measures and the vast number of medical disease know and emerging. Generic individualized measures, when valid and reliable, can help fill this gap.

HRQL measures can also be classified as traditional or individualized measures[16,17]. Traditional or conventional measures of HRQL are comprised of a fixed predetermined set of questions for all patients (e.g. short form 36 or SF36).

Individualized (or patient-specific/patient-generated) measures do not have any preset questions/domains (e.g. Measure Yourself Medical Outcome Profile or MYMOP). Individualized HRQL (iHRQL) measures promote a measurement approach where patients can nominate and then score important aspects of their life from own perspective.

Each approach to HRQL measurement, i.e. traditional vs. individualized, has its advantages and disadvantages. Traditional measures produce convenient data that can be compared across patients; they can also be used for economic evaluation. However, their standard questions are generally based on some theoretical model, literature review, or adaptations of existing measures. They are static questionnaires, and they may not allow variation in individual perception, preferences and values of HRQL[18-20]. There is evidence that traditional measures may miss factors that are relevant/important to patients at a particular time [18-20], those aspects were picked by individualized measures; in such cases validity and responsiveness of traditional instruments may be questioned. In contrast, individualized measures can help in the provision of healthcare according to patients' needs, preferences, and values. HRQL can be evaluated in a patient-centered way by using these flexible questionnaires that allow individual variation is perception, values, and needs. Individualized measures are representative of relevant factors that may be important to a particular patient at a given time. However, the same measures have limitations, such as their inappropriateness for economic evaluations [21,22]. Individualized measures have also been criticized for their complicated weighting of scores[23], unwillingness of patients to share the most important area of their life[24], and change of the perceived *important area* of life for a particular patient [25]. No matter what the measurement approach might be, it is important to remember that the main purpose of HRQL measurement is to give voice to patients in their healthcare and research. Since both measurement approaches have strengths and weaknesses, no one can be preferred over another. Their use for outcome assessment depends on the purpose of application and potentially a combination of the two might be the preferred option.

To date there has been little research on use of individualized HRQL measures in children.

1.3. Individualized HRQL measures: Critical appraisal of MYMOP

There is increased recognition of individualized measures for their patient-centeredness, whether they are disease-specific or generic. For example, in rheumatology[17,20] and asthma[27-31], disease-specific individualized measures have been used for adults and children [16,17,20,27-31]. Examples of disease-specific individualized measures include: Asthma Quality of life Questionnaire (AQLQ)[32], Pediatric Asthma Quality of Life Questionnaire (PAQLQ)[27-29], Juvenile Arthritis Quality of Life Questionnaire (JAQQ), and McMaster-Toronto Arthritis Questionnaire (MACTAR)[17,20]. Examples of generic individualized measures include: Schedule for the evaluation of individualized quality of life (SEIQoL)[33], Patient Generated Index (PGI)[34], and Measure Yourself Medical Outcome Profile (MYMOP)[35,36].

Individualized measures such as the PGI, SEIQoL-DW, AQLQ, and Repertory Grid (RG), have been criticized for their use in clinical trials [16]. Concerns included a lack of standardization required for making group comparisons for estimating population effect. Additional concerns included missing evidence of sound psychometric properties (PGI, RG), length of and complexity of weighting procedure (SEIQoL)[16]. Potential limitations of this review included unclear methods, lack of use of a formal risk of bias tool, and narrow search strategy [16]. There is a need to further assess the applicability of individualized measures in clinical trials.

To address some of the identified weaknesses of existing generic individualized measures, a new simplified generic iHRQL measure was developed, called MYMOP (Measure Yourself Medical Outcome Profile)[37,38]. Due to its simple format and patient-centeredness, MYMOP has been used as the basis for additional

outcome measures in specific populations, such as spiritual healing for chronic pain, physiotherapy, acupuncture for tinnitus [35,39-42].

There is a need to assess HRQL in a patient-centered way in a variety of diseases and individualized measures might assist in this. In particular we are interested to see how MYMOP and its adaptations were developed, and if they have sound psychometric properties. We also opt to determine if there is any validated generic individualized HRQL measures for children, or if current MYMOP is valid for pediatric use.

1.4. Why individualized outcome measures matter for pediatric population

Current pediatric research is criticized for the absence of high quality studies as compared to adults [43,44]. Generally this distinction is made between the difference in the number of clinical trials and observational studies. However, the problem is not only fewer clinical trials being done, it also includes the quality of the pediatric trials (and observational studies) that are conducted. There are numerous challenges to perform high quality pediatric research [45,46], and one of them is valid measurement of relevant outcomes. Regardless of study design, a research study is not credible in the absence of valid and reliable measurement of appropriate outcomes [47].

In April 2009, an international group of pediatric researchers gathered under the umbrella of Standards for Research in Child Health (STaRChild)[45]. The group acknowledged heterogeneity of outcome measurement in pediatric clinical trials [45,46,48] and identified provision of valid measurement of relevant outcomes as one of the priority areas in pediatric research.

Outcome measurement for any study is a two-step process. The first step is to choose an appropriate outcome for a given condition. The second step is the selection of valid and reliable measurement tools to capture the chosen outcome(s).

In a systematic review, Sinha et al, have pointed the neglected state of outcome selection in pediatric trials[46]. The use of insufficiently validated outcome measures has also been identified as a major problem in some areas of pediatric research [48].

Given the emerging importance of patient-centered outcomes, and iHRQL, this thesis was developed to evaluate if a valid generic pediatric iHRQL measure exists.

1.5. Assessing risk of bias in HRQL measures

Risk of bias assessment is recommended for the evaluation of measurement properties for validation studies[49]. The COSMIN checklist was developed by an international Delphi study in which consensus was reached on definitions of measurement properties; and standards were set for satisfactory design and statistical analysis of a study on measurement properties of health related patient reported outcomes[50]. The checklist consists of 10 boxes, and a three-step process to assess the methodological quality of studies evaluating the measurement properties of health status measurement instruments[50-52] . In the first step, COSMIN taxonomy and definitions are used to determine which measurement properties were evaluated in each included article. The corresponding COSMIN boxes are marked for each article. In step2 corresponding COSMIN boxes are completed for each article. In step3 each item is scored on a 4-point rating scale i.e. “poor”, “fair”, “good” or “excellent”. An overall score for the methodological quality of a study is determined by taking the lowest rating of any of the items in a box. A study on a measurement property is rated as having excellent quality if all relevant items on COSMIN are marked adequate. A study is rated as having ‘good’ methodological quality if some aspects of a measurement property are not reported, but one can deduce that these issues would have been adequate. A study is rated as having ‘fair’ quality if there are minor flaws in methods or statistical analysis. A study is rated as having ‘poor’ quality if there are major flaws in methods or statistical analysis[53]. Sample size also affects the rating of a study for any particular measurement property. Details on

4-point rating of COSMIN checklist can be found on www.cosmin.nl [53].

2. Thesis Objective

The need to measure HRQL in clinical trials and observational studies is widely recognized [54]; the intent of this thesis is to identify valid, reliable, and patient-centered measures to capture this essential outcome for infants, children, and youth. More specifically, the objective of this thesis was to identify if a pediatric individualized health related quality of life (HRQL) measure exists, such as a pediatric version of MYMOP. For this, a critical review of MYMOP and its adaptations' qualitative attributes and psychometric properties was performed. Secondly, a systematic review was performed to identify available generic and disease-specific individualized HRQL measures used in children, and to assess reporting of psychometric properties and HRQL as primary study outcome. Thirdly, a comprehensive risk of bias analysis was performed on the studies reporting on development and/or validation of identified generic individualized measures. This risk of bias analysis was performed to evaluate the robustness of current generic individualized HRQL measures to ascertain if they were adequate to recommend these measures for use in pediatric research. Table 1-1 summarizes the objectives of this thesis.

3. Specific objectives and thesis outline

Chapter 2: Critical appraisal of MYMOP and its adaptations: Psychological Outcomes Profiles (PSYCLOPS), Measure Yourself Concern and Wellbeing (MYCaW), MYMOP Pictorial, and Chinese MYMOP using Terwee criteria

Chapter 3 is a systematic review of six databases to identify individualized HRQL measures used in children (0-18 years).

Chapter 4 is the application of COnsensus based Standards for the selection of Health

Measurement INstrument (COSMIN) and Terwee quality criteria to perform risk of bias analysis of three generic iHRQL measures identified in chapter 3.

Chapter 5 presents overall future research and clinical implications of this thesis.

Table 1-1 Thesis objectives

Specific objective	Methods employed
1. To identify if a pediatric individualized health related quality of life (HRQL) measure exists, such as a pediatric version of MYMOP	A critical review of MYMOP and its adaptations' qualitative attributes and psychometric properties was performed using Terwee quality criteria[55]. COSMIN[50] was used to analyze the translation procedure of C-MYMOP
2. To identify available generic and disease-specific individualized HRQL measures used in children, and to assess reporting of psychometric properties and HRQL as primary study outcome	A systematic review was performed.
3. To assess the methodological quality of identified generic iHRQL measures for children	A comprehensive risk of bias analysis was performed on the studies reporting on development and/or validation of identified generic individualized measures. COSMIN[50] and Terwee criteria[55] were used for this appraisal.

4. References

1. Admins. Patient-centered outcomes research institute amends draft research agenda in response to public comment *Patient-Centered Outcomes Research Institute* 2012; .

2. Kjeken I, Ziegler C, Skrolsvik J, Bagge J, Smedslund G, Tovik A, Dagfinrud HS, Petersson IF, and Hagen KB. How to develop patient-centered research: Some perspectives based on surveys among people with rheumatic diseases in scandinavia. *Phys Ther* 2010; 90: 450-460.

3. Noble J. Patient-centered research: Through the looking glass in search of a paradigm. *Journal of General Internal Medicine* 1989; 4: 555-557.

4. Nickman NA, Haak SW, and Kim J. Use of clinical simulation centers in health professions schools for patient-centered research. *Simulation in Healthcare: The Journal of The Society for Medical Simulation* 2010; 5: 295-302.

5. Tallon D, Chard J, and Dieppe P. Relation between agendas of the research community and the research consumer. *Lancet* 2000; 355: 2037-2040.

6. Wolfe F, Pincus T. Listening to the patient: A practical guide to self-report questionnaires in clinical care. *Arthritis & Rheumatism* 1999; 42: 1797-1808.

7. Siteupdate. PCORI outlines a patient-centered vision for research and methodology in JAMA theme issue on comparative effectiveness research *Patient-Centered Outcomes Research Institute* 2012; .

8. Testa MA, Simnson DC. Assessment of quality-of-life outcome. N Engl J Med 1996;334: 835-40.

9. Matza LS, Swensen AR, Flood EM, Secnik K, and Leidy NK. Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value Health* 2004; 7: 79-92.

10. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

11. Slevin ML, Plant H, Lynch D, Drinkwater J, and Gregory WM. Who should measure quality of life, the doctor or the patient?. *Br J Cancer* 1988; 57: 109-112.

12. Ashworth M, Evans C, and Clement S. Measuring psychological outcomes after cognitive behaviour therapy in primary care: A comparison between a new patient-generated measure "pSYCHLOPS" psychological outcome profiles and "hADS" hospital anxiety and depression scale. *J Ment Health* 2009; 18: 169-177.

13. Guyatt GH, Feeny DH, and Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622-629.

14. Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics* 1999; 16: 605-625.

15. Ravens-Sieberer U, Erhart M, Wille N, Wetzel R, Nickel J, and Bullinger M. Generic health-related quality-of-life assessment in children and adolescents: Methodological considerations. *Pharmacoeconomics* 2006; 24: 1199-1220.

16. Patel KK, Veenstra DL, and Patrick DL. A review of selected patient-generated outcome measures and their application in clinical trials. *Value Health* 2003; 6: 595-603.

17. Jolles BM, Buchbinder R, and Beaton DE. A study compared nine patient-specific indices for musculoskeletal disorders. *J Clin Epidemiol* 2005; 58: 791-801.

18. Bowling A. *Measuring disease: a review of disease-specific quality of life measurement scales.* Buckingham, ENGLAND: Open University Press. 1995: 374 p.

19. Bowling A. What things are important in people's lives? A survey of the public's judgements to inform scales of health related quality of life. *Soc Sci Med* 1995; 41: 1447-1462.

20. Clinch J, Tugwell P, Wells G, and Shea B. Individualized functional priority approach to the assessment of health related quality of life in rheumatology. *J Rheumatol* 2001; 28: 445-451.

21. Paterson C, Britten N. In pursuit of patient-centred outcomes: A qualitative evaluation of the 'measure yourself medical outcome profile'. *J Health Serv Res Policy* 2000; 5: 27-36.

22. Paterson C. Measuring outcomes in primary care: A patient generated measure, MYMOP, compared with the SF-36 health survey. *BMJ* 1996; 312: 1016-1020.

23. Macduff C, Russell E. The problem of measuring change in individual health-related quality of life by postal questionnaire: Use of the patient-generated index in a disabled population. *Quality of Life Research* 1998; 7: 761-769.

24. Vachon ML, Kristjanson L, and Higginson I. Psychosocial issues in palliative care: The patient, the family, and the process and outcome of care. *Journal of Pain & Symptom Management* 1995; 10: 142-150.

25. Higginson I, Priest P, and McCarthy M. Are bereaved family members a valid proxy for a patient's assessment of dying?. *Soc Sci Med* 1994; 38: 553-557.

26. Jolles BM, Buchbinder R, and Beaton DE. A study compared nine patient-specific indices for musculoskeletal disorders. *J Clin Epidemiol* 2005; 58: 791-801.

27. Al-Akour N, Khader YS. Quality of life in jordanian children with asthma. *Int J Nurs Pract* 2008; 14: 418-426.

28. Basaran S, Guler-Uysal F, Ergen N, Seydaoglu G, Bingol-Karakoc G, and Ufuk Altintas D. Effects of physical exercise on quality of life, exercise capacity and pulmonary function in children with asthma. *J Rehabil Med* 2006; 38: 130-135.

29. Garcia-Marcos L, Carvajal Uruena I, Escribano Montaner A, Fernandez Benitez M, Garcia De La Rubia S, Tauler Toro E, Perez Fernandez V, and Barcina Sanchez C. Seasons and other factors affecting the quality of life of asthmatic children. *Journal of Investigational Allergology and Clinical Immunology* 2007; 17: 249-256.

30. Trzcieniecka-Green A, Bargiel-Matusiewicz K, and Wilczynska-Kwiatek A. Quality of life and activity of children suffering from bronchial asthma. *Eur J Med Res* 2009; 14: 7.

31. Ziora D, Madaj A, Wieckowska E, Ziora K, and Kozielski J. Correlation of spirometric parameters taken at a single examination with the quality of life in children with stable asthma. *Journal of Physiology and Pharmacology* 2007; 58: 801-809.

32. Qoltech - measurement of health-related quality of life & asthma control Retrieved 7/20/2012, 2012, from http://www.qoltech.co.uk/aqlq.html .

33. Wagner J. Acceptability of the schedule for the evaluation of individual quality of life-direct weight (SEIQoL-DW) in youth with type 1 diabetes. *Qual Life Res* 2004; 13: 1279-1285.

34. Ruta DA, Garratt AM, Leng M, Russell IT, and MacDonald LM. A new approach to the measurement of quality of life. the patient-generated index. *Med Care* 1994; 32: 1109-1126.

35. Abbot NC, Harkness EF, Stevinson C, Marshall FP, Conn DA, and Ernst E. Spiritual healing as a therapy for chronic pain: A randomized, clinical trial. *Pain* 2001; 91: 79-89.

36. Jenkinson C. MYMOP, a patient generated measure of outcomes. research into outcomes has moved away from symptom based assessments. *BMJ* 1996; 313: 626.

37. MYMOP - intro Retrieved 7/20/2012, 2012, from http://sites.pcmd.ac.uk/mymop/ .

38. Hyland ME, Lewith GT, and Wheeler P. Do existing psychologic scales measure the nonspecific benefit associated with CAM treatment? *J Altern Complement Med* 2008; 14: 185-189.

39. Bikker AP, Mercer SW, and Reilly D. A pilot prospective study on the consultation and relational empathy, patient enablement, and health changes over 12 months in patients going to the glasgow homoeopathic hospital. *J Altern Complement Med* 2005; 11: 591-600.

40. Kim T-, Kang JW, Kim KH, Kang K-, Shin M-, Jung S-, Kim A-, Jung H-, Lee S-, Choi J-, and Choi S-. Acupuncture for dry eye: A multicentre randomised controlled trial with active comparison intervention (artificial tear drop) using a mixed method approach protocol. *Trials* 2010; 11: .

41. Jackson A, MacPherson H, and Hahn S. Acupuncture for tinnitus: A series of six n=1 controlled trials. *Complement Ther Med* 2006; 14: 39-46.

42. Hill S, Eckett MJH, Paterson C, and Harkness EF. A pilot study to evaluate the effects of floatation spa treatment on patients with osteoarthritis. *Complement Ther Med* 1999; 7: 235-238.

43. Martinez-Castaldi C, Silverstein M, and Bauchner H. Child versus adult research: The gap in high-quality study design. *Pediatrics* 2008; 122: 52-57.

44. DeMauro SB, Giaccone A, Kirpalani H, and Schmidt B. Quality of reporting of neonatal and infant trials in high-impact journals. *Pediatrics* 2011; 128: e639-44.

45. Hartling L, Wittmeier KD, Caldwell PH, van der Lee JH, Klassen TP, Craig JC, Offringa M, and StaR Child Health g. StaR child health: Developing evidence-based guidance for the design, conduct, and reporting of pediatric trials. *Clinical Pharmacology & Therapeutics* 2011; 90: 727-731.

46. Sinha I, Jones L, Smyth RL, and Williamson PR. A systematic review of studies that aim to determine which outcomes to measure in clinical trials in children. *PLoS Medicine / Public Library of Science* 2008; 5: e96.

47. Tugwell P, Boers M. OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: Introduction. *J Rheumatol* 1993; 20: 528-530.

48. Johnston BC, Shamseer L, da Costa BR, Tsuyuki RT, and Vohra S. Measurement issues in trials of pediatric acute diarrheal diseases: A systematic review. *Pediatrics* 2010; 126: e222-31.

49. Terwee CB, Mokkink LB, Knol DL,Ostelo RW,Bouter LM, - de VH. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2011; 10: .

50. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HCW. The COSMIN study reached international consensus on taxonomy,

terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737-745.

51. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, and de Vet HC. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.

52. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HC. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international delphi study. *Qual Life Res* 2010; 19: 539-549.

53. Cosmin | cosmin Retrieved 7/19/2012, 2012, from http://cosmin.nl/ .

54. Admins. PCORI adopts definition of patient-centered outcomes research *Patient-Centered Outcomes Research Institute* 2012; .

55. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, and de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.

Chapter 2
Individualized Health Related Quality Of Life Instrument Measure Yourself Medical Outcome Profile (MYMOP) and Its Adaptations: A Critical Appraisal

Sana Ishaque[1, 2]
Jeffrey A. Johnson[1]
Sunita Vohra[1, 2, 3]

[1] School of Public Health
University of Alberta, Edmonton, Alberta, Canada
[2] CARE Program for Integrative Health & Healing, University of Alberta, Edmonton, Alberta, Canada
[3] Faculty of Medicine and Dentistry and School of Public Health, University of Alberta

Acknowledgements: The authors are grateful to Charlotte Paterson for her helpful comments. Jeffrey Johnson is a Senior Scholar with AIHS and is a Centennial Professor at the University of Alberta. Sunita Vohra receives salary support from Alberta Innovates Health Solutions (AIHS) as a Health Scholar.

Corresponding author:
Sunita Vohra MD FRCPC MSc
Professor, Faculty of Medicine and Dentistry and School of Public Health, University of Alberta

Email: svohra@ualberta.ca

Word count (excluding abstract, keywords, tables, and references): 4082

1. Abstract

Background: Health related quality of life (HRQL) is increasingly recognized for its importance in health research. In the evaluation of individualized therapies, it may be helpful to apply individualized measures of HRQL. Due to the usability of the Measure Yourself Medical Outcome Profile (MYMOP), a validated individualized HRQL measure, several adaptations have been developed.

Objective: This review was conducted to identify adaptations of MYMOP, and evaluate the psychometric properties of MYMOP and its adaptations.

Methods: Adaptations were identified using MYMOP website and personal communication, supplemented by a SCOPUS search. Bibliographies of included studies were hand-searched. Modified Terwee and COSMIN criteria were used to evaluate the psychometric properties.

Results: Sixteen studies were included in this review. Adaptations were developed to evaluate individualized therapies in cancer, psychiatry, acupuncture, and the Chinese population. The included measures were MYMOP, Measure Yourself Concern and Wellbeing, Psychological Outcome Profiles (PSYCHLOPS), MYMOP-pictorial (MYMOP-P), and Chinese MYMOP (C-MYMOP). The quality of the psychometric properties varied; none of the included measures met all currently recommended quality criteria for psychometric properties.

Conclusion:

Current literature provides evidence that MYMOP and its adaptations offer individualized assessment for individualized therapies, such as those typically offered in complementary, alternative, or integrative medicine. Further validation work is recommended.

Key words: Health related quality of life, patient-generated, individualized, patient-centered, domain specific

2. Abbreviations

C-MYMOP: Chinese MYMOP

CAM: Complementary and Alternative Medicine

CORE-OM: Clinical Outcomes Routine Evaluation – Outcome Measure

COSMIN: Consensus based Standards for the selection of health status Measurement INstruments

DA: Depression Alliance

EQ-5D: EuroQol Group health status index 5-Dimensions

HADS: Hospital Anxiety Depression Scale

HRQL: Health Related Quality of Life

ICC: Intraclass correlation coefficient

MOS-6A: Medical outcome study 6-items general health survey

MYCaW: Measure Yourself Concerns and Wellbeing

MYMOP: Measure Yourself Medical Outcome Profile

MYMOP-Pictorial: MYMOP-P

PRIMHE: Primary Care Mental Health Education

PSYCHLOPS: Psychological Outcome Profiles

QoL: Quality of Life

SD: Standard Deviation

SEIQoL: Schedule for the Evaluation of Individual Quality of Life-Direct Weighting,

PGI: Patient Generated Index

SF-36: Medical Outcomes Study 36-item Short-Form Health Survey

3. Introduction

3.1 Quality of Life

Health related quality of life (HRQL) has grown in its importance as an essential outcome for patient-centered research [1]. Advances in medical research have resulted in prolonged survival for those with chronic diseases, making the patient's experience vital to assessment of therapeutic effectiveness. Arguably, effective therapies not only alleviate the patient's signs or symptoms, but also make a significant difference in the patient's HRQL.

3.2 Defining Health Related Quality of Life (HRQL)

HRQL is the subjective perception of an individual's quality of life regarding the health related issues. As defined by the World Health Organization, health is a state of complete physical, mental, and social wellbeing and not merely the absence of disease or infirmity [2]. Interchangeable terms often used to refer to the domain of *health* include health status, functional status, quality of life, and wellbeing. Quality of life includes everything that is part of one's life, whereas HRQL represent those parts of quality of life that directly relate to an individual's health. There are numerous definitions of HRQL such as "a person or group's perceived physical and mental health over time"[3]. Matza et al[4], defines HRQL for adults as "an individual's subjective perception of the impact of their health status, including disease and treatment, on physiological, psychological,

and social functioning." The construct provides insight into patient-directed goals and values.

3.3 Individualized Health Related Quality of Life Measures

HRQL is a multidimensional concept and its measurement is not straightforward. Traditional HRQL measures consist of predetermined set of questions usually developed by experts, with or without input from patient groups (e.g. focus groups). This list of questions, organized in various domains, is intended to capture important aspects of patients' lives. However, the predetermined questions may not be relevant to particular patients at different stages of their disease [5]. Moreover, what may be important for one patient may not have similar value for another patient [6]. Each patient may seek help with varying goals, values, and treatment effect for any particular intervention. Personalized approaches to treatment also enhance potential differences between patients, such as those preferred in complementary and alternative medicine (CAM) [6].

Traditional HRQL measures, with their predetermined domains, provide standardized measurement for all respondents; they are beneficial for between and within group comparisons. However, these measures may not represent all health domains valued by each individual patient [7-9]. Some researchers have expressed concern on the patient-centeredness of current HRQL measures in practice [5]. In contrast, individualized measures allow patients to nominate and score the important aspect of their lives that contribute most to their overall quality of life. Examples of individualized measures

include: Schedule for the Evaluation of Individual Quality of Life-Direct Weighting (SEIQoL-DW) [10-12], Patient Generated Index (PGI) [13], and Juniper's measures of asthma-related HRQL for adults and children [14,15]. Paterson *et al*'s Measure Yourself Medical Outcome Profile (MYMOP) [16] furthered the concept of these measures. MYMOP has been used by our pediatric CAM clinic to assess outcomes in children; we found that it is important to know if the questionnaire is validated for pediatric use.

The purpose of this paper is to critically appraise the psychometric properties of MYMOP[17] and its adaptations.

4. Methods

4.1 Search Method

We conducted a SCOPUS search for articles' titles, abstracts, and keywords. The names of the adaptations were identified using the MYMOP website and personal communication with instruments developers. The search terms included MYMOP, MYCaW, PSYCHLOPS, MYMOP-P, and C-MYMOP. We searched the reference lists of the included articles to identify any additional relevant articles. In addition to this, we scanned each instrument's primary website to identify additional articles. Finally, we reviewed the abstracts to identify studies conducting formal psychometric evaluation, or qualitative evidence collection to validate the instruments of interest.

4.2 Quality Assessment

We evaluated the results of measurement/psychometric properties for each measure, identified, using a modified version of the Terwee criteria [18][19]. There are three domains of psychometric properties: reliability, validity, and responsiveness. Reliability domain is further subdivided into: internal consistency, reliability, and measurement error. Validity domain is also subdivided into: content validity, construct validity, and criterion validity. The possible overall rating for a measurement property is "positive" (+), "indeterminate" (?), "negative" (-), or "no information available" (0) (Appendix 1).

The important psychometric properties depend on the purpose of the measurement instrument [20,21]. Measurement tools are generally applied for three purposes: evaluation, discrimination, and prediction. For an evaluative instrument (e.g. used in clinical trials), psychometric properties considered most important are test-retest reliability and responsiveness [22]. For discriminative purpose, the important measurement properties are internal consistency reliability, and cross-sectional construct validity [23,24]. HRQL measures can be used for prediction of future health, survival (e.g. in life threatening conditions such as cancer)[25], resource allocation, and health policy[26]. Predictive measures are intended to provide some independent prognosis of future clinical event, or to provide an estimate of HRQL score using a brief version of a longer original measure.

To evaluate the translation process of C-MYMOP we applied item # 4 to 11 of Consensus based Standards for the selection of health status Measurement INstruments (COSMIN) Box G [27] that evaluates the quality of translation procedure.

5. Results

The search yielded 63 unique studies. We retrieved 24 articles in full text after screening the "title," "abstract," and "keywords." We finally included 16 studies, evaluating four original and one translated questionnaire [16,28-42]. Table 1 presents the general characteristics of these studies. It is notable that 10 of 16 (63%) were applied in populations taking CAM. In addition to MYMOP, we identified three adaptations and one translated version: MYCaW**,** PSYCHLOPS, C-MYMOP, and MYMOP-P (Table 2). The new questionnaires were adapted for evaluation of therapies in cancer [41,42], psychiatry [29,36], the Chinese population [33], and acupuncture [35]. Table 2, Table 3, and Table 4 respectively presents the description of the included measures, their psychometric properties, and translation procedure.

5.1 Measure Yourself Medical Outcome Profile (MYMOP)

MYMOP is a problem specific, individualized measure that was developed in a primary care setting (Table 2) [16]. Each patient is asked to report on two symptoms that bother them the most, one activity limited by the reported symptoms, and general well-being. After the pilot study, a brief medication questionnaire was added to the scale [40].

However, medication questions are not scored and thus do not contribute to the final MYMOP score for a patient [43]. A single score can be calculated by taking the average of item scores, but can only be interpreted in the presence of individual item scores. An important aspect of MYMOP that makes it unique is the absence of predetermined domains or items. The scores are based on severity of these issues over the previous week. For meaningful comparison, the items chosen must remain unchanged between the first and the subsequent completion of the questionnaire.

5.1.1 Quality Assessment of MYMOP

We did not identify any studies evaluating measurement error, floor, ceiling effect, and interpretability of the MYMOP. Three studies assessed content validity (Table 3). The first of these gives clear description of the measurement aim and information on the target population [16]. The second study, reports on the content validity [39] gathered from patients' views about MYMOP's ability to measure outcomes that are important to patients. This study compared the qualitative interview data for 20 interviewees to their corresponding quantitative MYMOP score [39]. Incorporation of participants and practitioners' views resulted in the development of the second version called "MYMOP 2," which is the current version in use. The third study exploring content validity [38], involved interviewing 23 new patients of eight acupuncturists in the UK. They used three qualitative analytical techniques: focus groups, in-depth interviews, and cognitive interview. The issues identified about MYMOP2 were floor effect, inability of patients

to score symptoms of episodic conditions, and inaccurate measurement of medication change. No revisions of MYMOP2 were performed based on the study results[38].

Two studies that assessed construct validity examined the correlation between "perceived change in condition" and MYMOP scores [16,40]. Both studies confirmed the MYMOP scores correlated with the perceived change in condition. Similar results were observed for the correlation of clinical-outcome assessed by physicians and MYMOP scores [40]. Another parameter used to evaluate construct validity was comparison of MYMOP scores in individuals with acute conditions and in those with chronic conditions; it was hypothesized that changes in MYMOP score would correlate well with changes in acute conditions (<4 weeks) rather than chronic conditions (>4 weeks). This correlation was confirmed in this study [16]. In addition, expected correlations of MYMOP and SF-36 scores were also reported [16].

Responsiveness of MYMOP was determined by gradient change in score at repeat applications across perceived changes by clincians [16] and by patients [16,40]. Standardized response mean, and index of responsiveness were also reported as evidence for responsiveness [16,40]. A t-test was conducted to compare the scores of patients who described themselves as a "little better" to "about the same, [40]" and gradient change in scores at two and four weeks was determined [16]. The authors applied the SF-36, MOS-6A, and EQ-5D, simultaneously to the study population, but did not report correlation coefficients for changes [16,40].

5.2 Measure Yourself Concerns and Wellbeing (MYCaW)

MYCaW [37,41,44] was adapted from MYMOP to evaluate cancer support services, especially for patients undergoing integrative treatment (Table 2). Like MYMOP, it allows patients to define and measure their two most important concerns and general wellbeing on a seven-point ordinal scale; higher score signifies poorer health [41]. MYCaW also has pictorial faces, and the wording added at the each end of the seven point scale:“not bothering me at all=0,” “bothers me greatly=6” [45]. There are two MYCaW versions, for self-administrated and face-to-face interview. Each version has initial and follow-up forms. The questionnaire consists of three scored domains, two of which are individualized. The follow up form includes two open-ended questions: "other things affecting your health" and "reflecting on your time with (service name) what were the most important aspects for you? [41]." MYCaW provides quantitative (mean change in score and SD), and qualitative data.

5.2.1 Quality Assessment of MYCaW

To our knowledge content validity is the only reported psychometric property of MYCaW. Adaptation and validation of MYCaW started in 2002 [44](Table 3). Initial draft was discussed with experts and patient-representatives resulting in subsequent revision to the layout and wording of the draft instrument [44][41]. A later study defined minimal important change for the interpretation of scores as 0.5, 1, and 1.5 as minimal,

moderate, and large changes, respectively [37]. One of the advantages of MYCaW is its ability to capture range of qualitative information at individual patient's level [41]. There have been substantial efforts to provide a frame of analysis for the rich qualitative information gathered by the questionnaire [42,46]. Three questions of MYCaW were qualitatively analysed: "concerns and problems" question on the first form; and "other things affecting your health," and "what has been most importat for you?" of the follow up form. Sample of 782, 407, and 588 patients reported on "concerns and problems," "other things affecting your health," and "what has been important for you?" repectively. Their responses were organized into categories; focus group of five women validated the categories for appropriateness and acceptability.

5.3 Psychological Outcome Profiles (PSYCHLOPS)

PSYCHLOPS is an individualized mental health outcome measure [28]. Similar to MYMOP, PSYCHLOPS measures the score of unique issue(s) for an individual (Table 2).

PSYCHLOPS is a one-page questionnaire, the reverse side of the questionnaire is completed by a therapist [47]. On the reverse side, it has instructions for scoring, an open text box where therapist can note their comments, patient identification, and therapist's scoring of patients' overall health after therapy. The patient questionnaire consists of three domains: problems, function, and wellbeing. There are three different versions of the questionnaire: pre-therapy, during-therapy, and post-therapy. Four questions are

common to each version. The initial two questions ask patients to identify and measure their most bothersome problems, the third identifies and measures one function limited due to the identified problem(s), and fourth is about general wellbeing over the last week. A fifth question in the during-therapy version identifies any new problem that arises amidst therapy. A sixth question on the post-therapy version asks the patients to score how they feel compared to the start of therapy. PSYCHLOPS does not assign a score to every question. The questions related to Problems, Functioning and Wellbeing have six-point (0-5) scales, where higher score signify worse outcomes. The "individually identified" items from the initial form are transferred to the subsequent versions for patient to rescores them. This process provide changes in score from pre to post therapy[47].

5.3.1 Quality Assessment of PSYCHLOPS

A group of clinical psychologists, counseling psychologists, psychotherapists, counselors, general practitioners, and academic mental health researchers interested in mental health started adaptation of PSYCHLOPS in 2004(Table 3) [28].

To address content validity, the developers consulted patient representatives, and three expert groups. The initial draft was piloted to 30 patients [28]; and it was revised as required [28]. In 2005 (Table 2), Ashworth et al gathered information about the feasibility, validity, and usefulness of PSYCHLOPS from experts [31]. Internal consistency was determined via Cronbach's alpha, and the values were within acceptable range [29,32,36].

In terms of construct validity, PSYCHLOPS has moderate to strong correlation with Clinical Outcomes Routine Evaluation-Outcome Measure (CORE-OM) [32] and Hospital Anxiety Depression Scale (HADS) [19]. Responsiveness was defined as "sensitivity to change" and was measured by effect size [29,32]. Interpretability was assessed by mean and SD of pre and post-therapy scores [29,32]. Test-retest reliability was reported as intraclass correlation coefficients (ICC) between baseline and retest as 0.70, 0.68, 0.69, and 0.79 for problems domain, activity that was hard-to-do, wellbeing, and overall score respectively [36]. The study participants for reliability assessment were healthy individuals and stable during the interim period.

In 2007, Ashworth analyzed if the preset items on CORE-OM identify the individualized PSYCHLOPS responses [30]. There were 611 individual responses on PSYCHLOPS and the responses were categorized into 8 themes and 61 sub-themes. Of 61 sub-themes, 27 (44%) were not mapped to preset questions of CORE-OM. Of 215 clients, 128 (60%) reported at least one response that could not be mapped to CORE-OM.

5.4 MYMOP-Pictorial

MYMOP-P was developed, by Anthony Day, to audit patient outcomes after acupuncture (Table 2) [34,35]. During the study, the author found that patients who were "elderly," "having low confidence in completing forms," "low literacy," or "mother tongue not English" were not able to fill MYMOP2 properly. To solve this issue MYMOP-P was

developed. MYMOP-P has six point scale (0-5) that range from “as good as it could be” to “as bad as it could be.” Each response option has a “face” that corresponds to the current state of patient, and patients are asked to choose one face in order to score their reported issue. The author did not explain the method of questionnaire adaptation any further, it is not clear if any patient representatives were involved. To our knowledge, no formal evaluations of the instrument’s measurement properties are reported yet.

5.5 Chinese Version of MYMOP (C-MYMOP)

Chung et al have performed translation and validation of MYMOP2 to Chinese language [33]. The translation process of C-MYMOP was excellent on all parameters except description of the sample (Table 4). The *recall-period* was increased from 7 days to two weeks on the follow-up form. They inferred content validity from the current MYMOP version (Table 3). In order to assess construct validity, *a priori* hypotheses were reported, considering both direction and magnitude of expected relationships. Interpretability was estimated by minimal important difference, and minimal detectable change values for all domains except activity domain; information on the mean change in the score of four perceived global change (patient-perceived change) groups was also reported. The authors indicated that the distribution of mean change scores met the expected increment with the perceived global change, implying that the C-MYMOP was responsive to such change.

6. Discussion

We reviewed the format, content and evidence of measurement properties for MYMOP[16,38-40], three adaptations[28-37,41,42], and one translation in this review[33]. Of these measures, PSYCHLOPS was the most frequently evaluated[28-32,36], and therefore had the greatest evidence of its psychometric properties, including test-retest and internal consistency reliability. To our knowledge, MYCaW[37,41,44] and MYMOP-P[34,35] are the least psychometrically tested questionnaires; their only reported measurement property is content validity. We found the quality of translation of C-MYMOP[33] to be excellent generally; the C-MYMOP[33] also had positive rating for construct validity, responsiveness, and interpretability.

Content validity was the most widely reported measurement property[16,28,31,33-35,37-39,44]. Of five measures, three had positive[16,28,31,37,44], and two had indeterminate rating[33-35,37,38] for content validity. Construct validity was the second commonly tested measurement property[16,29,32,33]; it was reported for all measures except MYMOP-P and MYCaW. Evidence on construct validation was limited in terms of reporting *a priori* hypotheses regarding expected correlations. Modern day reporting standards for assessment of construct validity [20,21,27] suggest that *a priori* hypotheses regarding the strength and direction of the correlation also be specified. Given our results, future validation studies should consider developing and reporting a priori hypotheses for construct validity evaluation.

Criterion validtiy was reported for three measures in five studies [12, 19, 21,22,40]; however, we find that all claims of criterion validity were actually supportive of construct validity under the current definitions suggested by COSMIN taxonomy [21]. We find it difficult to see an instrument as a "gold standard," unless a short version of a questionnaire was tested against its long version [20,21,27]. We therefore evaluated these claims as we would evaluate construct validity. Our approach did not affect the grading of the evidence. For future researchers we recommend to avoid reporting such evaluations as criterion validity, unless it involves testing a short version of a questionnaire against a long version (gold standard). Further, assessment against SF-36 may be considered assessment of construct validity, not criterion validity, since some would argue that SF-36 is not a universally accepted "gold standard". Evidence on structual validity, and internal consistency reliability are not relevant to the included measures. Structural validity is relevant for measures based on item response theory [20,21,27]; internal consistency reliability is required for questionnaires with predetermined multidimensional domains. Internal consistency was reported for PSYCHLOPS, however it is not relevant to individualized measures because there are no predetermined domains to these measures. If internal consistency is not reported it would not affect the intended use of these measures.

Of five studies reporting on responsiveness [12, 19, 21, 22, 40], two [19,22] assessed responsiveness by effect sizes. We were unable to evaluate this evidence because the reported statistic did not meet the COSMIN and modified Terwee criteria for evaluation of responsiveness; both studies [19, 22] were published before the Terwee criteria was

developed. As of the recent criteria for psychometric properties we would like to see further evaluation of responsiveness.

Another limitation of the included studies is the imprecise use of terminology to define psychometric properties. This finding is not unique to these studies; Mokkink et al [48] reported similar finding in a study of quality assessment of systematic reviews of measurement properties. Of note, international concensus on taxonomy of measurement properties is a recent development in the field of psychometrics [21].

Unlike a systematic review, study inclusion, data abstraction, and quality assessment were not independently duplicated. We acknowledge that lack of independent duplication can be a source of error to a review; however single data extraction does not result in any difference in the effect estimates for many outcomes [49]. Moreover, to strengthen our critical appraisal, we chose objective checklist criteria to evaluate the quality of measurement properties, enhancing the reproducibility of our results. Although we only included studies published in English, a Chinese tool was identified in the review, demonstrating the sensitivity of our search method.

Critical appraisal is essential to evaluate medical research; it helps identify methodological strengths and limitations. Critical appraisal can be done using checklist or score based scales. For our review, we considered appraisal tools/articles such as Criteria by the Scientific Advisory Committee of the Medical Outcomes Trust (MOT) [50], Evaluating the Measurement of Patient-Reported Outcomes (EMPRO) [51], and Terwee

[18, 19]and COSMIN criteria [27]. The MOT criteria provide a list of items that instrument developers should have considered to ascertain optimal properties of their tool. However, MOT does not provide guidance on how the reported evidence should be classified if any of the listed items are absent. The EMPRO criteria have an integral scoring system, the weighting of which is not clearly described nor explicitly justified with empiric data [51]. We used Terwee and COSMIN criteria because the COSMIN checklist was developed through a consensus-based Delphi study and has empirical evidence supporting its measurement properties [27]; the Terwee criteria have been developed to be consistent with the COSMIN checklist [18, 19]. We preferred to use a checklist rather than a summary score because a summary score does not provide specific details on methodological strengths or limitations. A checklist approach is also preferred by the Cochrane Collaboration, based on empirical evidence that the summary scores of quality assessment tools can be problematic [52, 53, 54]. As such, Cochrane has moved from the popular use of a score-based quality assessment tool [55], to the new descriptive checklist assessment, the Risk of Bias tool.

Traditional and individualized HRQL instruments provide two different approaches to HRQL measurement. Traditional measures with their standardized questions can be used for group comparisons (e.g. clinical trials). Some traditional measures (e.g., utility or preference-based) are also suited for economic evaluations. In comparison, individualized instruments can be beneficial in primary care settings or N-of-1 trials where each patient is usually the main focus. Individualized HRQL measures can help both patient and

healthcare provider to tailor healthcare according to the patients' needs, values, and preferences.

There are limitations to both measurement approaches. Standardized format of traditional measures do not allow respondents to measure any areas of HRQL that are important to them, but are not already included in the questionnaire. Scores on individualized measures represent measurement of unique patient issues, and thus these scores cannot be used for discriminative purposes and economic evaluations. Also, since the patient nominates the individualized domains or symptoms that are important to them, changes in the importance of certain domains or symptoms over time may limit the feasibility of these measures in providing meaningful evaluations of interventions. Some may argue that this would limit measurement properties of individualized measures. However, we believe that this does not affect their validity. Change in nominated symptoms for evaluation can inform clinicians that the patients' experience and priorities have changed over time; and hence clinicians now need to focus on other aspect of life that are important to patient. Overall, both measurement approaches are complementary to each other; they can be used alone or in combination depending on the purpose for their use.

Furthermore, assessing HRQL offers the opportunity to improve physician-patient communication and achieve better psychological outcomes [56,57]. Given the multiple demands put on the health care system and the time constraints faced by health care providers, individualized measures that are short, straightforward and quick to administer may help integrate routine HRQL assessment in clinical settings. MYMOP and its

adaptations offer a set of brief and easy-to-complete questionnaires for a variety of clinical conditions. These measures offer an advantage in conditions when patients report varied concerns (e.g. cancer, CAM use). MYMOP has been criticized for being symptom specific [58,59], however the recent development by Patient Reported Outcomes Information System (PROMIS) encourages the use of *domain-specific* rather than *disease-specific* measures[60]. Researchers at PROMIS state that the experience of *fatigue, headache, nausea, sleep problems, and etc.* are less likely to be influenced by the mere presence or absence of a disease. MYMOP was developed primarily to overcome the diagnostic differences in different disciplines of health care (within CAM streams, and CAM to conventional medicine). MYMOP (and its adaptations) being *generic domain (patient selected) specific* measure can be used to overcome issue of variability in outcome measurement in clinical trials.

Current literature provides evidence that MYMOP and its adaptations are *symptom specific* measures that can be used to measure variation in patient-complaints regardless of their diagnosis. These measures can be a starting point for *domain specific* measurement of symptoms like pain, nausea, anxiety, etc. Given that validation is an iterative/ongoing process and considerable efforts have been put to develop and achieve sound psychometrics of these measures, we would recommend researchers to further the validation of MYMOP and its adaptations before considering to develop a new measure.

Table 2-1: Study characteristics

Instrument	Study & Country	Study design	Sample size and age (range)	Target Population (diagnosis/underlying condition)	Practice setting/ context	Psychometric Property (ies) involved/evaluated	"Gold Standard" or Comparator	Application (time)
MYMOP	Paterson et al 1996, the UK[16]	Longitudinal	N= 265, 218 general practice patients, 47 CAM patients (2-84 years; mean & SD: 47 +- 17.6); 109 completed the f/u	Evaluation; patients receiving complementary med	Primary care (GPs and CAM providers)	Content validity, Construct validity (hypothesis testing), Responsiveness,	SF- 36	0, 2, 4, 8, and 16 week
	Paterson et al 2000, Scotland[40]	Longitudinal	N= 81 (32.4-82.8 years)	Evaluation; chronic bronchitis with acute exacerbation	General practice (Glasgow, Scotland)	Responsiveness, construct validity (hypothesis testing)	MOS-6A, EQ-5D	0, and after 1 week within completion of treatment
	Paterson et al 2000, the UK [39]	Longitudinal	176 patients completed MYMOP, 20 interviews; age: 16-86 years	Evaluation	New patients of 12 complimentary practitioners Somerset UK	Qualitative analysis to assess content validity	Semi-structured interviews	MYMOP was completed for a minimum of twice and maximum of nine times over 9 months
	Paterson et al 2004, the UK [38]	Longitudinal	23 patients: 64 interviews; age: 26-83 years	Evaluation of primary care acupuncture patients	New patients of acupuncturists	Qualitative analysis to assess content validity	NA	Thrice to 18 patients, twice to 4 patients

MYCaW	Peace et al 2002, the UK [44]	Longitudinal	N= 157 (18-86 yrs; peak 50-59 yrs)	Evaluation of cancer patients and their caregivers	Multidisciplinary clinic (Complementary & Alternative med)	Content validity	None	Time period not defined (at first visit & time of completion of treatment)
	Cooke 2000, the UK [37]	Dissertation:	N= 100 (81 patient + 19 carers), 61% completed f/u at 3rd wk, 40% at 3rd month; 18-70+	Evaluation of cancer patients (or various primary and secondary cancers) taking CAM along with conventional medicine	Multidisciplinary cancer support	Content validity	None	0, 3 weeks, & 3 months
	Paterson et al 2007 [41]	Publication based on findings of Peace et al 2002 & Cooke 2000	Cavendish Center: n= 254 (21-84 yrs); Bristol center: n= 267	NA	NA	NA	NA	NA
PSYCHLOPS	Ashworth et al, 2004, the UK [28]	Group consultations re adaption	Not reported	Primary care based psychotherapy patients	NA	Content validity	NA	NA
	Ashworth 2005, the UK[31]	Survey	Four primary care mental health practitioners; age: not reported	Primary care patients undergoing talk/psychotherapy	Primary care	Content validity	CORE-OM	NA
	Ashworth 2005, -the UK [32]	Longitudinal	N= 235 completed pre-therapy questionnaires; n= 110 post-therapy; age for whom complete set available for	Patients entering psychotherapy in primary care	Primary care	Internal consistency (cronbach's alpha), construct validity (hypothesis testing), interpretability, responsiveness[32]	CORE-OM	Pre and post therapy (no time interval specified)

			analysis: 15- 64 years					
	Ashworth 2007, the UK [30]	Cross sectional	N= 215; 16-64 year (<18 included if not full time students)	Patients entering to primary care mental health	Primary care	Qualitative analysis performing comparison of PSYCHLOPS responses to CORE-OM (content validity)	CORE-OM	NA
	Ashworth 2009, the UK [29]	Longitudinal	N= 336, complete data available for n = 114; 17-75 years	Primary care psychological therapy patients	Primary care clinical psychologists performing talking therapy based on CBT model	Internal consistency (cronbach's alpha), construct validity (hypothesis testing), interpretability, responsiveness [29]	HADS	Pre and post therapy (no time interval specified)
	Evans C 2010, the UK [36]	Longitudinal	N= 73 (1^{st} time responders), n= 56 (completed both 1^{st} and 2^{nd} rounds); age not reported	Students from three institutes from London	Non clinical sample	Internal consistency, reliability	None	0, 1 -2 weeks later
MYMOP-Pictorial (MYMOP-P)	Anthony Day 2004, the UK [35]	Audit of acupuncture	N= 62 initial form, 55 f/u completed (23-80 years)	Patients undergoing acupuncture	Varying complaints	Content validity	None	Not specified
	Anthony Day 2004, the UK [34]	Article explaining MYMOP-P	n/a	n/a	n/a	n/a	n/a	n/a

C-MYMOP	Chung et al 2010, China [33]	Longitudinal study explaining development, piloting, and psychometric evaluation	Pilot n= 28, 539 at baseline, 343 at 2 weeks, 272 at 4 weeks (18-79 years)	Patients taking Chinese medicine	Varying complaints	Construct validity (hypothesis testing, cross cultural validity), interpretability, Responsiveness	Hong Kong Chinese version of SF-36	Baseline, 2 week, 4 week

Table 2-2: Description of included measures

	Definitions					
Included measure (s)		**MYMOP**[16,38-40]	**MYCaW** [37,41,44]	**PSYCHLOPS** [28-32,36]	**MYMOP Pictorial** [34]	**C-MYMOP**[33]
Version(s) available		Two: initial form and follow up form	Two: face to face, and self completion version; each version has initial and f/u forms	Three: pre-therapy, during therapy, post-therapy	One	One
Construct	Description of what the questionnaire intends to measure	Physical symptoms, activity, wellbeing	Concern/problems, wellbeing	Mental health outcomes	Same as MYMOP	Same as MYMOP
Domain	A domain or dimension refers to the area of behavior that we are trying to measure.	Symptom(s), and activity as chosen by patient; wellbeing	Concern(s) chosen by patient; wellbeing	Problems, and function chosen by patient; wellbeing	Same as MYMOP	Same as MYMOP
Setting (e.g. Clinical, general population, epidemiological study)	In what setting the measurement was made?	Clinical trial, primary care (Paterson 1996); Clinical trial, general practice (Paterson 2000)	Observational, multidisciplinary clinic CAM (Peace 2002); questionnaire development, multidisciplinary clinic CAM	Clinical and general population	Clinical (acupuncture patients	Clinical (patients undergoing complementary and alternative med)
Recall Period	What is the recall period to which the questionnaire refers	Last week (last 7 days)	Current concerns are asked	Range from under one month to over five year	Same as MYMOP	Same as MYMOP for initial form, 2 weeks for f/u forms
Purpose (evaluative, discriminative, or both)	Purpose of questionnaire	Evaluative	Evaluation	Evaluation of primary care mental health patients undergoing psychotherapy	Evaluation	Evaluation

Target Population: diagnosis, age	For the kind of people questionnaire was originally developed	N= 265 (2-84 years; mean & SD: 47 +- 17.6); 109 completed the f/u patients receiving CAM (P 1996); N= 81 (32.4-82.8 years) patients of chronic bronchitis with acute exacerbation (p 2000)	N= 157 (18-86 yrs; peak 50-59 yrs), cancer patients (peace et al 2002); Cavendish Center: n= 254 (21-84 yrs), Bristol center: n= 267, cancer patients (Paterson 2006)	Primary care psychotherapy sessions' patient (s)	Acupuncture patients suffering from various conditions	Heterogeneous population taking complementary and alternative medicine, Chinese med and allopathic
Mode of administration (e.g. self, interview, proxy administered if proxy administered: name of proxy e.g. parent or health care provider)		Self administered	Self and interview administered	Self administered	Self administered	Not specified
# of items	Number of questions in a questionnaire	Eight on the initial form, four scored four un-scored; six on the f/u form four of which are scored	Three scored items (of these two are patient generated) and two open ended questions	Six items on pre-therapy and post-therapy questionnaires: three preset and three patient/individual specific; five items on during therapy: two preset and four individualized	Same as MYMOP	Same as MYMOP
# and type of response options	Scale type	0-6 point scale (Seven points)	0-6 point scale with smiley face adjacent to "0," and sad face adjacent to "6"	0-5 point scale	Six point faces scale	Seven point scale

Time to complete	Average time to complete the questionnaire	Not specified	Not specified	Not specified	Not specified	Not specified
Full copy available (if available full copy will be attached as an appendix)		Yes	Yes	Yes	Yes	No
Instructions (not described, clearly described, unclear)		Clear description on filling and scoring are available	Completion and scoring methods clearly described	Clearly described	Not reported	Not reported if available or not
Country (related to cross cultural validity)		The UK	The UK	The UK	The UK	China
Translation/ cultural adaptations available		Chinese language	None	None	None	None
Generic of specific (disease of population specific)		Problem specific (Individualized)	Problem specific	Specific- Condition specific	Problem specific	Problem specific

South Thames Primary Care Research Network = STaRNeT; Depression Alliance (DA); Primary Care Mental Health Education (PRIMHE)

Table 2-3: Summary of the assessment of measurement properties (based on Quality Criteria by Terwee et al, 2007[18,19])

	Validity			Reliability					
Questionnaires	**Content Validity**	**Construct Validity:, HT, CC**	**Criterion validity**	**Internal consistency**	**Measurement error**	**Reliability**	**Responsiveness**	**Floor or ceiling effect**	**Interpretability**
MYMOP	+ [16,39]; - [38]	+ [16], ? [16,40]	0	0	0	0	+ [16,40][17,40]	0	0
MYCaW	+ [37,44]	0	0	0	0	0	0	0	? [44]
PSYCHLOPS	+ [28][31]	? [29,32]	0	? [29,32,36], not relevant to individualized measures	0	+ [36]	? [29,32], the information does not meet Terwee criteria standards	0	? [29,32]
MYMOP-P	?[34,35]	0	0	0	0	0	0	0	0
C-MYMOP	?[33]	CC: +; HT: +[33]	0	0	0	0	+ [33]	0	+ [33]

HT= hypothesis testing, CC = cross cultural adaptation
Rating: + = positive, ? = indeterminate, - = poor (negative), 0 = no information available

Table 2-4: Methodological criteria for the translation process and cross-cultural validation (MYMOP Chinese)

COSMIN box-G (item # 4-11) with 4-point scale[48]					
Methodological criteria	Excellent	Good	Fair	Poor	Acquired rating of C-MYMOP [33]
Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described?	Both source language and target language described			Source language NOT known	Excellent
Was the expertise of the people involved in the translation process adequately described? e.g. expertise in the disease(s) involved, in the construct to be measured, or in both languages	Expertise of the translators described with respect to disease, construct, and language	Expertise of the translators with respect to disease or construct poor or not described	Expertise of the translators with respect to language not described		Excellent
Did the translators work independently from each other?	Translators worked independent	Assumable that the translators worked independent	Unclear whether translators worked independent	Translators worked NOT independent	Excellent
Were items translated forward and backward?	Multiple forward and multiple backward translations	Multiple forward translations but one backward translation	One forward and one backward translation	Only a forward translation	Excellent
Was there an adequate description of how differences between the original and translated versions were resolved?	Adequate description of how differences between translators were	Poorly or NOT described how differences between translators			Excellent

	resolved	were resolved			
Was the translation reviewed by a committee (e.g. original developers)?	Translation reviewed by a committee (involving other people than the translators, e.g. the original developers)	Translation NOT reviewed by (such) a committee			Excellent
Was the HR-PRO instrument pre-tested (e.g. cognitive interviews) to check interpretation, cultural relevance of the translation, and ease of comprehension?	Translated instrument pre-tested in the target population	Translated instrument pre-tested, but unclear if this was done in the target population	Translated instrument pre-tested, but NOT in the target population	Translated instrument NOT pre-tested	Excellent
Was the sample used in the pre-test adequately described?	Sample used in the pre-test adequately described		Sample used in the pre-test NOT (adequately) described		Fair

Appendix 2-1

Quality criteria for measurement properties (Based on Terwee et al[18,19,21])

Property (definitions are based on COSMIN taxonomy)		Rating	Quality Criteria
Reliability: The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions	Internal consistency: The degree of the interrelatedness among the items	+ ? -	Sub)scale uni-dimensional AND Cronbach's alpha(s) ≥ 0.70 Dimensionality not known OR Cronbach's alpha not determined (Sub)scale not uni-dimensional OR Cronbach's alpha(s) < 0.70
	Measurement error: The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured	+ ? -	MIC > SDC OR MIC outside the LOA MIC not defined MIC ≤ SDC OR MIC equals or inside LOA
	Reliability: The proportion of the total variance in the measurements which is due to 'true' differences between patients	+ ? -	ICC/weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80 Neither ICC/weighted Kappa, nor Pearson's r determined ICC/weighted Kappa < 0.70 OR Pearson's r < 0.80
Validity: The degree to which an HR-PRO instrument measures the construct(s) it purports to measure	Content validity: The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured	+ ? The target population considers items in	The target population considers all items in the questionnaire to be relevant AND considers the questionnaire to be complete No target population involvement

			-	the questionnaire to be irrelevant OR considers the questionnaire to be incomplete
	Construct validity: The degree to which the scores of an HR-PRO instrument are consistent with hypotheses	Cross-cultural: The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument	+ ? -	Original factor structure confirmed OR no important DIF Confirmation original factor structure AND DIF not mentioned Original factor structure not confirmed OR important DIF
		Structural: The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured	+ ? -	Factors should explain at least 50% of the variance Explained variance not mentioned Factors explain < 50% of the variance

		Hypothesis testing: Idem construct validity	+ ? -	(Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses) AND correlation with related constructs is higher than with unrelated constructs; Solely correlations determined with unrelated constructs; Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR correlation with related constructs is lower than with unrelated constructs
	Criterion validity: The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'		+ ? - 0	Convincing arguments that gold standard is ''gold'' AND correlation with gold standard >0.70; No convincing arguments that gold standard is ''gold'' OR doubtful design or method; Correlation with gold standard <0.70, despite adequate design and method; No information found on criterion validity.

Responsiveness: The ability of an HR-PRO instrument to detect change over time in the construct to be measured	+ ? - 	(Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70) AND correlation with related constructs is higher than with unrelated constructs Solely correlations determined with unrelated constructs Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlation with related constructs is lower than with unrelated constructs
Floor and ceiling effect: The number of respondents who achieved the lowest or highest possible score	+ ? - 0	<15% of the respondents achieved the highest or lowest possible scores; Doubtful design or method; >15% of the respondents achieved the highest or lowest possible scores, despite adequate design and methods; No information found on interpretation.
Interpretability : The degree to which one can assign qualitative meaning to quantitative scores	+ ? 	+ Mean and SD scores presented of at least four relevant subgroups of patients and MIC or MID defined; ? Doubtful design or method OR less than four subgroups OR no MIC or MID defined;

	0	0 No information found on interpretation

[..] = reference number, MIC = minimal important change, ; MID = minimal important difference; SDC = smallest detectable change, LOA = limits of agreement, ICC = intraclass correlation coefficient,
DIF = differential item functioning, AUC = area under the curve
† + = positive rating, ? = indeterminate rating, - = negative rating, 0 = no information

7. References

1. Engel GL. The need for a new medical model: A challenge for biomedicine. *Science* 1977; 196: 129-136.

2. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

3. Home page | CDC health-related quality of life. Retrieved 2/12/2010, 2010, from http://www.cdc.gov/hrqol/ .

4. Matza LS, Swensen AR, Flood EM, Secnik K, and Leidy NK. Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value Health* 2004; 7: 79-92.

5. Carr AJ, Gibson B, and Robinson PG. Measuring quality of life is quality of life determined by expectations or experience? *Br Med J* 2001; 322: 1240-1243.

6. Ernst E, Hung SK. Great expectations: What do patients using complementary and alternative medicine hope for? *Patient* 2011; 4: 89-101.

7. Bowling A. *Measuring disease: a review of disease-specific quality of life measurement scales.* Buckingham, ENGLAND: Open University Press. 1995: 374 p.

8. Bowling A. What things are important in people's lives? A survey of the public's judgements to inform scales of health related quality of life. *Soc Sci Med* 1995; 41: 1447-1462.

9. Clinch J, Tugwell P, Wells G, and Shea B. Individualized functional priority approach to the assessment of health related quality of life in rheumatology. *J Rheumatol* 2001; 28: 445-451.

10. Hickey AM, Bury G, O'Boyle CA, Bradley F, O'Kelly FD, and Shannon W. A new short form individual quality of life measure (SEIQoL-DW): Application in a cohort of individuals with HIV/AIDS. *BMJ* 1996; 313: 29-33.

11. Moons P, Marquet K, Budts W, and De Geest S. Validity, reliability and responsiveness of the "schedule for the evaluation of individual quality of life-direct weighting" (SEIQoL-DW) in congenital heart disease. *Health Qual Life Outcomes* 2004; 2: 27.

12. Lhussier M, Watson B, Reed J, and Clarke CL. The SEIQoL and functional status: How do they relate? *Scand J Caring Sci* 2005; 19: 403-409.

13. Ruta DA, Garratt AM, Leng M, Russell IT, and MacDonald LM. A new approach to the measurement of quality of life. the patient-generated index. *Med Care* 1994; 32: 1109-1126.

14. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, and Townsend M. Measuring quality of life in children with asthma. *Qual Life Res* 1996; 5: 35-46.

15. Juniper EF, Guyatt GH, Ferrie PJ, and Griffith LE. Measuring quality of life in asthma. *Am Rev Respir Dis* 1993; 147: 832-838.

16. Paterson C. Measuring outcomes in primary care: A patient generated measure, MYMOP, compared with the SF-36 health survey. *BMJ* 1996; 312: 1016-1020.

17. Jenkinson C. MYMOP, a patient generated measure of outcomes. research into outcomes has moved away from symptom based assessments. *BMJ* 1996; 313: 626.

18. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, and de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.

19. Schellingerhout JM, Heymans MW, Verhagen AP, de Vet HC, Koes BW, and Terwee CB. Measurement properties of translated versions of neck-specific questionnaires: A systematic review. *BMC Med Res Methodol* 2011; 11: 87.

20. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, and de Vet HC. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.

21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HCW. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737-745.

22. Launois R. Quality of life and pharmaco-economics in clinical trials - spilker,B. *QUALITY OF LIFE RESEARCH* 6: 97-98.

23. Chen AY, Frankowski R, Bishop-Leone J, Hebert T, Leyk S, Lewin J, and Goepfert H. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: The M. D. anderson dysphagia inventory. *Arch Otolaryngol Head Neck Surg* 2001; 127: 870-876.

24. Guyatt GH, Kirshner B, and Jaeschke R. Measuring health status: What are the necessary measurement properties?. *J Clin Epidemiol* 1992; 45: 1341-1345.

25. Ganz PA, Lee JJ, and Siau J. Quality of life assessment. an independent prognostic variable for survival in lung cancer. *Cancer* 1991; 67: 3131-3135.

26. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. *Am J Public Health* 1990; 80: 446-452.

27. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HC. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international delphi study. *Qual Life Res* 2010; 19: 539-549.

28. Ashworth M, Shepherd M, Christey J, Matthews V, Wright K, Parmentier H, Robinson S, and Godfrey E. A client-centred psychometric instrument: The development of "PSYCHLOPS"("psychological outcome profile"). *Counselling and Psychotherapy Research* 2004; 4: 27.

29. Ashworth M, Evans C, and Clement S. Measuring psychological outcomes after cognitive behaviour therapy in primary care: A comparison between a new patient-generated measure "pSYCHLOPS" psychological outcome profiles and "hADS" hospital anxiety and depression scale. *J Ment Health* 2009; 18: 169-177.

30. Ashworth M, Robinson S, Evans C, Shepherd M, Conolly A, and Rowlands G,. What does an idiographic measure (PSYCHLOPS) tell us about the spectrum of psychological issues and scores on a nomothetic measure (CORE-OM)? *primary care and community psychiatry* 2007; 12: 7-12.

31. Ashworth M, Robinson S, Godfrey E, Parmentier H, Shepherd M, Christey J, Wright K, and Matthews V,. The experiences of therapists using a new client-centred psychometric instrument, "PSYCHLOPS" ("psychological outcome profile"). *Counselling and Psychotherapy Research* 2005; 5: 37-41.

32. Ashworth M, Robinson SI, Godfrey E, Shepherd M, Evans C, Seed P, Parmentier H, and Tylee A. Measuring mental health outcomes in primary care: The psychometric properties of a new patient-generated outcome measure, 'PSYCHLOPS' ('psychological outcome profiles'). *Prim Care Ment Health* 2005; 3: 261-270.

33. Chung VCH, Wong VCW, Lau CH, Hui H, Lam TH, Zhong LX, Wong SYS, and Griffiths SM. Using chinese version of MYMOP in chinese medicine evaluation: Validity, responsiveness and minimally important change. *Health Qual Life Outcomes* 2010; 8: .

34. Day A. The development of the MYMOP pictorial version. *Acupunct Med* 2004; 22: 68-71.

35. Day A, Kingsbury-Smith R. An audit of acupuncture in general practice. *Acupunct Med* 2004; 22: 87-92.

36. Evans C, Ashworth M, and Peters M. Are problems prevalent and stable in non-clinical populations? problems and test-retest stability of a patient-generated measure, PSYCHLOPS (psychological outcome profiles), in a non-clinical student sample. *Br J Guid Couns* 2010; 38: 431-439.

37. Helen Jane Cooke. Is the adapted "measure yourself medical outcome profile (MYMOP) questionnaire an appropriate tool to evaluate the bristol cancer help centre's supportive programme?
. 2000; .

38. Paterson C. Seeking the patient's perspective: A qualitative assessment of EuroQol, COOP-WONCA charts and MYMOP. *Qual Life Res* 2004; 13: 871-881.

39. Paterson C, Britten N. In pursuit of patient-centred outcomes: A qualitative evaluation of the 'measure yourself medical outcome profile'. *J Health Serv Res Policy* 2000; 5: 27-36.

40. Paterson C, Langan CE, McKaig GA, Anderson PM, Maclaine GDH, Rose LB, Walker SJ, and Campbell MJ. Assessing patient outcomes in acute exacerbations of chronic bronchitis: The measure your medical outcome profile (MYMOP), medical outcomes study 6-item general health survey (MOS-6A) and EuroQol (EQ-5D). *Qual Life Res* 2000; 9: 521-527.

41. Paterson C, Thomas K, Manasse A, Cooke H, and Peace G. Measure yourself concerns and wellbeing (MYCaW): An individualised questionnaire for evaluating outcome in cancer support care that includes complementary therapies. *Complement Ther Med* 2007; 15: 38-45.

42. Polley MJ, Seers HE, Cooke HJ, Hoffman C, and Paterson C. How to summarise and report written qualitative data from patients: A method for use in cancer support care. *Supportive Care Cancer* 2007; 15: 963-971.

43. MYMOP - welcome Retrieved 12/2/2011, 2011, from http://sites.pcmd.ac.uk/mymop/index.php?c=welcome .

44. Peace G, Manasse A. The cavendish centre for integrated cancer care: Assessment of patients' needs and responses. *Complement Ther Med* 2002; 10: 33-41.

45. MYMOP - sister Retrieved 12/2/2011, 2011, from http://sites.pcmd.ac.uk/mymop/index.php?c=sister&s=mycaw .

46. Seers HE, Gale N, Paterson C, Cooke HJ, Tuffrey V, and Polley MJ. Individualised and complex experiences of integrative cancer support care: Combining qualitative and quantitative data. *Supportive Care Cancer* 2009; 17: 1159-1167.

47. Psychlops | home Retrieved 12/2/2011, 2011, from http://www.psychlops.org.uk/ .

48. Terwee CB, Mokkink LB, Knol DL,Ostelo RW,Bouter LM, - de VH. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2011; 10: .

49. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. 2006; . Single data extraction generated more errors than double data extraction in systematic reviews.J Clin Epidemiol , 59: 7, 697-703.

50. Assessing health status and quality-of-life instruments: attributes and review criteria 2002; Quality of Life Research , 11: 3, 193-205.

51. Valderas JM. Ferrer M. Mendivil J. Garin O. Rajmil L. Herdman M. Alonso J. Scientific Committee on "Patient-Reported Outcomes" of the IRYSS Network. 2008; . Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures.Value in Health , 11: 4, 700-708.

52. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. 2011; . The Cochrane Collaboration's tool for assessing risk of bias in randomised trials.BMJ , 343: , 5928.

53. Juni P, Witschi A, Bloch R, Egger M. 1999; . The hazards of scoring the quality of clinical trials for meta-analysis.JAMA , 282: 11, 1054-1060.

54. Greenland S, O'Rourke K. 2001; . On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions.Biostatistics , 2: 4, 463-471.

55. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. 1996; . Assessing the quality of reports of randomized clinical trials: is blinding necessary?.Control Clin Trials , 17: 1, 1-12.

56. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, and Selby PJ. Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. *J Clin Oncol* 2004; 22: 714-724.

57. de Wit M, Delemarre-van de Waal HA, Bokma JA, Haasnoot K, Houdijk MC, Gemke RJ, and Snoek FJ. Monitoring and discussing health-related quality of life in adolescents with type 1 diabetes improve psychosocial well-being: A randomized controlled trial. *Diabetes Care* 2008; 31: 1521-1526.

58. Brewin T, Garrow J. Commissioning complementary medicine. evaluations of efficacy of treatments should be consistent. *BMJ* 1995; 311: 809.

59. Fisher P, Eden A. Commissioning complementary medicine. homoeopathic hospitals have unique skill. *BMJ* 1995; 311: 809.

60. PROMIS Retrieved 7/19/2012, 2012, from http://www.nihpromis.org/?AspxAutoDetectCookieSupport=1 .

Chapter 3
Individualized Health Related Quality of Life Measures for Children: A Systematic Review

Sana Ishaque[1, 2]
Jeffrey A. Johnson [1]
Charlotte Paterson [3]
Sunita Vohra [1, 2, 4]

[1] School of Public Health
University of Alberta, Edmonton, Alberta, Canada
[2] CARE Program for Integrative Health & Healing, University of Alberta, Edmonton, Alberta, Canada
[3]School of Social & Community Medicine, University of Bristol, The UK
[4]Faculty of Medicine and Dentistry and School of Public Health, University of Alberta

Corresponding author:
Sunita Vohra MD FRCPC MSc
Professor, [3] Faculty of Medicine and Dentistry and School of Public Health, University of Alberta
Email: svohra@ualberta.ca

Acknowledgements:
Sana Ishaque is an MSc Clinical Epidemiology student at the School of Public Health and is a Graduate Research Assistant at Complementary and Alternative Research and Education (CARE) Program, Department of Pediatrics, University of Alberta. Sunita Vohra is a Health Scholar with Alberta Innovates Health Solutions (AIHS). Jeffrey Johnson is a Senior Scholar with AIHS and is a Centennial Professor at the University of Alberta.

The authors gratefully acknowledge Dale Storie, and Soleil Sorette for their expert help with literature searching, Fatima Al Sayah for seconding on relevance and full text screening, and Tamara Dorfman for seconding data extraction.

Word count (excluding abstract, keywords, tables, and references): 2968

1. Abstract

Introduction: Health related quality of life (HRQL) of children is an essential outcome of patient-centered pediatric research. HRQL is a multidimensional concept and its measurement is challenging. Individualized HRQL (iHRQL) measures propose a measurement approach where patients can nominate and then score important aspects of their life from own perspective.

Objectives: To identify all available iHRQL measures for children, and to assess reporting of psychometric properties in studies where HRQL was primary outcome.

Methods: We searched six common databases from inception to March-April 2011. English language studies of any design including any children <= 18 years were included if they reported validation or use of an iHRQL measure.

Results: We identified 68 studies reporting on use or validation of eight iHRQL measures in children or mixed (adult and pediatric) population. Five disease-specific and three generic iHRQL measures were identified. Included studies reported 27 unique definitions of HRQL. The majority of the primary outcome studies did not report relevant references to support their choice of outcome measures.

Conclusion: Reported HRQL definitions were heterogeneous, and the majority of studies did not report relevant references of prior validation. Further assessment of their reported psychometric properties is needed before they can be recommended for pediatric use.

Keywords: patient-specific, patient-generated, psychometric properties, quality of life

2. Abbreviations

AQLQ: Asthma Quality of Life Questionnaire

GLQ-8: Global Quality of Life-8

HRQL: Health related quality of life

iHRQL: Individualized health related quality of life

MYMOP: Measure Yourself Medical Outcome Profile

PAQLQ: Pediatric Asthma Quality of Life Questionnaire

PASI-pg: Patient Specific Index

PCORI: Patient centered outcomes research institute

PGI: Patient Generated Index

RQLQ: Rhinoconjunctivities Quality of Life Questionnaire

SEIQoL-DW: Schedule for the Evaluation of Individual Quality of Life Direct Weighting

WHO: World Health Organization

3. Introduction

3.1 Health Related Quality of Life

Health related quality of life (HRQL) is as an essential outcome for patient-centered pediatric research [1]. As a result of advances in medical science and technology, more children are living with chronic diseases and disabilities, making their experience vital to assessment of therapeutic effectiveness [2]. Arguably, effective treatments not only ease the patient's complaints, but also make positive difference in the patient's HRQL. There is no universally agreed definition of HRQL; for adults it has been defined as *individual's subjective perception of the impact of their health status including disease and treatment, on physiological, psychological, and social functioning* [3]. This is based on World Health Organization's definition of health [4]. The same definition can be applied to children; it is important to note, however, that these aspects of life differ between children and adults, as well as within various pediatric developmental stages.

3.2 Measurement Approaches

HRQL is a multidimensional concept and its measurement is challenging. HRQL measures are generally classified as generic or disease-specific [5]. HRQL measures can also be divided into two categories: questionnaires with predetermined domains (or

traditional measures), and questionnaires where respondents/patient can nominate important aspects of their lives and score them (individualized measures). Each measurement approach has its strengths and weaknesses. With their standardized domains, traditional measures produce data that may be used for group comparisons and economic evaluations. However, traditional measures might miss issues that are important to a particular patient [6-8]. In comparison, individualized instruments can be beneficial in measuring issues that the patient feel/perceive valuable, N-of-1 trials, and personalized/alternative medicine.

A recent initiative toward patient-centered research is the establishment of Patient Centered Outcomes Research Institute (PCORI)[1]; providing information on how patients can improve outcomes *that are important to them* is one of the key drivers for this organization. The PCORI recognizes HRQL as an important patient-centered outcome. Individualized HRQL measures can help both patient and healthcare provider to tailor healthcare according to the patients' needs, values, and preferences. However, scores on individualized measures are specific to individual patient and they do not allow between group comparisons and hence cannot be used for economic evaluation.

3.3 Measurement of health status, quality of life, and/or health related quality of life of children

Our research team previously identified 15 systematic reviews on health status and quality of life (QOL) or HRQL measures in children for various clinical conditions[9-23]. These systematic reviews have been criticized for their methodology [24] including: narrow search strategy resulting in incomplete review; poor quality and reporting of search strategy; absence of risk of bias analysis; unclear differentiation between reviews of measurement properties and reviews conducted to identify all available measures for a particular condition; and poorly reported methods that often do not report if all steps were independently duplicated. To our knowledge there are no systematic reviews on current state of pediatric individualized health related quality of life (iHRQL) measures. As such, there is a need to conduct systematic reviews to identify pediatric iHRQL tools (or absence thereof). It is important to distinguish between the reviews of measurement properties and the reviews done to identify measure of a particular type or for a particular disease because, absence of risk of bias analysis may not be a quality issue for systematic reviews conducted to identify measurement tools.

Our objectives were to identify all available iHRQL (generic or disease specific) measures for children, and to assess reporting of psychometric properties in studies where iHRQL was primary study outcome.

4. Methods

4.1 Search Strategy for Identification of Studies

With guidance from health research librarians, a systematic search of the following databases was performed in March 2011:

1. MEDLINE (1950 – March 4th week 2011),
2. EMBASE (1980 –2011 week 13),
3. PsycINFO (1985 – Week 4 2011),
4. CINAHL (1982 – April 4, 2011), and
5. HAPI (March 30, 2011)
6. CENTRAL (1st quarter 2011), DARE (2nd quarter 2011), ACP journal club (1991 to March 2011), Cochrane Database of Systematic Reviews (2005 to March 2011), Health Technology Assessment (2nd quarter 2011)

The search was conducted using controlled vocabulary and keywords related to the terms: quality of life, individualized, and age group 0-18 years. We performed two searches in Medline: the first search with controlled vocabulary, keywords, and names of the individualized measures identified by previous exposure to literature; the second search was performed with the names of 79 HRQL measures of children. To identify these additional names of pediatric HRQL measures a search was performed in PROQOLID database with *HRQL* as type of instrument, and *adolescent/pediatrics* as population of interest. The names of 79 identified measures were included in the second Medline search strategy. The search was modified to adapt to the variation in indexing among several databases. Details of these searches can be found in Appendix 1.

4.2 Study Selection

We included studies that (1) were published in English, (2) included any child (0-18 years of age), even if mixed (adult and pediatric) population, (3) were about any intervention-control pair, (4) were full text original research that developed, adapted,

validated, or used a HRQL measure that was partially or completely individualized. To identify HRQL measures that were individualized, we accepted statements such as “in the activity domain, three of the items are individualized [25].”

Two authors (SI, FA) independently performed screening of title, abstract, and keywords of searched articles. We anticipated that not all the authors of HRQL measures would describe the instrument fully in title, keyword, and abstract. Given that indexing does not identify if HRQL measure was individualized, we did not place strict restriction for the term *individualized* while performing relevance screen. The full text of the selected articles was retrieved, and two reviewers (SI, FA) independently applied pre-specified inclusion criteria to potentially relevant studies to identify studies for inclusion in this review. We identified HRQL measures with standard as well as individualized versions. In case of uncertainty, authors were contacted up to three times to for clarification. Any differences in extraction were resolved through consensus or through discussion with the third reviewer (SV).

4.3 Data Extraction

We anticipated two types of included studies in this review: studies that report use of HRQL measures to assess primary or secondary study outcome, and studies that report validation of HRQL measures. Two reviewers (SI, TD) independently extracted data from the included studies using a standardized extraction form. Data were extracted on

general characteristics of included studies, description of identified measures, and on reporting of primary outcomes. Data extraction form on the qualitative attributes of measures was based on the Qualitative Attributes of Physical Activity Questionnaire checklist [26].

4.4 Data analysis and synthesis, statistical issues

It was predetermined that combining data would be inappropriate for the purpose of this review therefore no meta-analysis was performed. Data were tabulated and graphed where appropriate.

5. Results

The search retrieved 4649 unique articles (see Figure 1, PRISMA flow diagram). The title, abstract, and keywords screen identified 736 potentially relevant articles. Full text screen of 729 articles resulted in inclusion of 68 studies[25,27-48][49-52] (see Figures 1 and 2). Seven articles could not be retrieved (Figure 2).

5.1 Characteristics of included studies

Sixty-eight studies reporting on eight iHRQL measures were included in this review. Table 1a summarizes studies on use of iHRQL measures, which consist of three

categories: 1) primary outcome[25,34-37,39-41,43,45,48,53-63]; 2) secondary outcome[32,64,65][49-52]; and 3) to validate other instruments[28-30,38,46,66-69] (Figure 2). The majority of studies (46/68 or 67.6%) are represented by these three categories; the remaining 22 studies (22/68 or 32.4%)[27,31,33,42,44,47,70-85] reported on validation of generic and disease specific iHRQL measures. Of these 22, nine studies concerning five iHRQL measures were of mixed (adult and pediatric) populations. Table 1b summarizes their baseline characteristics.

5.2 Identified Individualized HRQL measures

Our search identified eight iHRQL measures that were used for children: Pediatric Asthma Quality of Life Questionnaire (PAQLQ), Asthma Quality of Life Questionnaire (AQLQ), Global Quality of Life-8 (GLQ-8), Schedule for the Evaluation of Individual Quality of Life Direct Weighting (SEIQoL-DW), Patient Generated Index (PGI), Measure Yourself Medical Outcome Profile (MYMOP), Patient Specific Index (PASI-pg), and Rhiniconjunctivitis Quality of Life Questionnaire (RQLQ). Of eight, five were disease-specific, and three were generic iHRQL measures (Table 2). Of eight, two measures: PASI-pg, and RQLQ did not have any reported validation studies for children, yet both were used to assess primary outcome of pediatric studies (Table 2). Table 3 presents characteristics of included measures.

5.3 Identified definitions of health related quality of life, and/or quality of life

The definition of HRQL was reported in 27 (40%) studies[28,33,34,36,37,40,41,44-47,57-59,62,68,69,71,78,81,82,84,86]. There was no consensus on reported definitions across included studies. All 27 studies reported unique definitions of HRQL; of these, three studies reported two definitions[82,84,87], and one [36]reported three definitions. The identified definitions were classified into nine categories on the basis of the underlying concept (Table 4). Nine studies (9/27 or 33%) based the definition of the concept (quality of life or HRQL) on perception of individual patient. Another six studies (6/27 or 22%) based their definitions on World Health Organization (WHO) definition of health, although most of the studies did not explicitly refer to the WHO definition [4].

5.4 Reporting of psychometric properties in studies that used iHRQL instruments to assess primary outcome

Thirty studies used the six iHRQL measures (PAQLQ, AQLQ, SEIQoL-DW, RQLQ, PASI-pg) to assess their primary outcome; all of them reported references to previous validation work.

The references reported to support validation of these measures were based on exclusive pediatric studies (PAQLQ and SEIQOL-DW), mixed population studies (AQLQ, MYMOP) or adult-only studies (RQLQ, PASI-pg).

6. Discussion

The objective of this review was to identify all existing iHRQL measures for children and evaluate the reporting of their psychometric properties. Due to previous experience of our group with reporting of HRQL studies (unpublished), we conducted a sensitive (broad) search. We searched for studies reporting on iHRQL measure for children; both generic and disease specific measures were included regardless of their validation status. Our search identified 68 studies that used eight individualized HRQL measures in pediatric or mixed populations. The majority of the included studies employed the iHRQL measures for outcome assessment or to validate other measures; less than one third of included studies were about validation of iHRQL measures.

Studies that used HRQL measures for primary outcome assessment[25,34-37,39-41,43,45,48,53-63] rationalized their choice of measures by providing reference to previous validation work; the reported references were largely inappropriate, as except for two measures (PAQLQ, and SEIQoL-DW) they were about mixed population or adult studies. We screened the reported mixed population references for pediatric sample size and subgroup analysis, and it was not reported. This lack of evidence challenges their relevance/applicability to children. Outcome measures that are not valid and reliable can be a source of bias for clinical research, whether observational or experimental [94]. If outcome measures do not assess what they are supposed to assess validly and reliably, control of other sources of bias cannot assure integrity of the results.

Our review was not restricted to a particular study design: clinical trials and observational studies both were included. We found that item 6a of the most widely accepted checklist for the reporting of primary outcome measures in clinical trials [95](cite: CONSORT home page) recommends "completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed." The checklist does not explicitly discuss reporting validation of the primary outcome measure, although this is discussed in the explanation document [96] (CONSORT other page). Our systematic review corroborates recently identified deficiencies in the validation of primary outcome tools/instruments [94, 97]. Outcome assessment is wholly dependent on the sound psychometric properties of the assessment tools, without which even the results of most carefully conducted studies can be questioned. We recommend that the CONSROT checklist [95] be reworded to explicitly state if the outcome measures were validated in the population under study or if they were validated in similar population (s). Similarly, the STROBE statement for the reporting of observational studies does not discuss this important aspect of study design [98]. We recommend that future revisions for both reporting standards (CONSORT for clinical trials and STROBE for observational studies) stress the importance of reporting the validation of the primary outcome measure, as primary outcome measurement is a main focus of any study regardless of its design.

Absence of risk of bias analysis may be viewed as a limitation of this review. However, the purpose of this systematic review was not to make causal inferences about any intervention, or to review measurement propertied of identified measures. We did not evaluate the included studies with regard to effectiveness and hence were not concerned

on the potential risk of bias in estimates of treatment effect; our intention was to identify iHRQL measures used in current pediatric literature. Another potential weakness of this review is that in our protocol we intended to report only on HRQL measures, but we included measures of both HRQL and QOL. We recognize that HRQL and QOL are not equivalent [5,99]. Wherever possible, we extracted and reported on definitions of QOL and HRQL to help distinguish them. QOL is an ill-defined term; it includes everything about life. HRQL includes those components of life that directly related to health. Sixty percent of included studies, and half of the primary outcome studies, did not report definition of HRQL or QOL. Thirty-two unique definitions of HRQL were reported in 27 included studies; the identified heterogeneity of definitions warrants further research to reach consensus.

An additional limitation is the exclusion of non-English language studies, which may have used validated outcome measures. It is reasonable to assume that even if a well-validated measure existed in any other language, it could not be used in English speaking population without further validation. As our primary intent was to identify an English pediatric individualized HRQL tool, we excluded 61 non-English studies at the screening stage. These studies could be further assessed to see if a pediatric iHRQL measure has been reported in languages other than English.

We followed the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guideline to report this systematic review, following its checklist and recommendations for the flow diagram [100].We carefully considered

instruments/questionnaires such as Quality of Life Diary[101], Adolescent Quality of Life questionnaire (AQoL)[102,103], Visual Narrative Interviews (VNI)[104], and Subjective Quality of Life Profile for potential inclusion[105]. However, none of these measures were included. We found that open ended questions (diary), not being able to score subjective complaints (AQoL), qualitative interviews (VNI), and option for researchers to use certain domains of a questionnaire (SQLP) for a particular patient does not make a HRQL measure *individualized.* For inclusion in our review, we considered a measure individualized if respondents determined its items/domains and these self-selected items/domains were subsequently scored. In the instances where a questionnaire had both individualized and standard version, we looked for explicit statements to clarify the version used (e.g. PAQLQ: "in the activity limitation domain, 3 of the items are individualized according to the activities that affected the patient most because of their asthma[25]).

Currently there is no consensus on selection criteria for HRQL measures for a study. It is important to have measurement tools that are valid and reliable for the population under study. One would expect that child-specific tools be used in measurement of outcomes in research that includes any children. Careful consideration to select appropriate outcome measure(s) is required at the planning and methods stage of any study. This is challenging, particularly when it comes to patient reported outcomes (e.g. HRQL measures). In a recent paper McClimans et al have suggested that a suitable outcome measure selection be done by experts' judgment on available psychometric evidence for a particular purpose of measurement (discrimination or evaluation) [106]; the authors have

criticized choosing measures based on cumulative strength of their measurement properties, and hierarchal approach. The cumulative approach may be limited, as for a particular application all dimensions of measurement properties may not be important, and evidence of all measurement properties may not be available for comparatively new measures [106]. In the hierarchal approach one measurement property is preferred over another, for example a measure with evidence of content validity may be preferred over another measure with evidence of construct validity; this may result in selection of measures with no construct validity for a study[106].

There are several published checklists regarding selection of HRQL measures, such as those by Tian-hui et al[107], Greenhalgh et al[108], and Patient Reported Outcomes Measaurement Group[109]. However, the methods for developing their checklists were not reported; also these checklists focus on what needs to be done to select outcome measures, as opposed to how this should be done. Although, each of these checklist considers psychometric properties to be an important criteria for outcome measure selection, but all three of them provide different details of what needs to be checked making it challenging to use either one of them for practical purposes. Health science researchers from around the world are making efforts to find consensus to assess: i) which outcomes should be included for measurement in clinical trials [110], ii) which outcome measures already exist [111], and iii) how to assess psychometric properties of patient reported outcomes [112]. Together these initiatives will assist in outcome measure selection and decrease variability of measurement tools across studies.

It is recommended that researchers (i) come to consensus on HRQL definition or until this is developed, explicitly define HRQL as it is used in the context of their study, and (ii) select valid and reliable age appropriate outcome measures to assess it.

7. Conclusion

Five disease-specific and three generic pediatric iHRQL measures have been identified. Identified definitions of HRQL were heterogeneous, and the majority of studies did not report relevant references of prior validation. Further assessment of their reported psychometric properties is needed before they can be recommended for pediatric use.

Table 3-1a: Study characteristics (HRQoL measures used to assess study outcome or to validate other outcome measures)

Studies that use generic or disease-specific individualized measures to assess primary outcome					
Ref ID	Study (First author's last name & pub year), country	Instrument name	Condition	Sample size & age	Definition of HRQoL
390	Al-Akour 2008, Jordan[34]	PAQLQ	Asthma	n= 200 >12 yrs (51 %)and < 12 yrs (49 %), age range <12 to >12 yrs	QoL is 'a uniquely personal perception, denoting the way individual patients feel about their health status and/or non-medical aspects of their lives'
579	Marcos 2007, Spain[48]	PAQLQ-Spanish version	Asthma	n= 1103, age range 7-17 years mean=10.3(2.0)	Not reported
590	Gent 2007, The Netherlands[89]	PAQLQ	Community based sample of school children with or without physician's diagnosis of asthma	three groups: diagnosed asthma n= 81 age= 9.4 (0.7), undiagnosed n=130 age= 9.4(0.8), controls n= 202 age = 9.4 (0.7)	Not reported
626	Gent 2007, The Netherlands[88]	PAQLQ	Asthma	N= 404: 171- asthmatics normal weight, mean age 9.3 (0.8) 33- asthmatics excessive body weight, mean age 9.5 (0.8) 174- healthy control normal body weight, mean age 9.4 (0.7) 25- asthmatics excessive	Not reported

				body weight, mean age 9.6 (0.6) Range 7- 10 yrs.	
1046	G. Gonzalez Martin 2003, Chile[53]	PAQLQ	Asthma,	n= 21 (10 controls, 11 treatment group), age: 7 to 14 years, 9.9 (0.4 years)	Not reported explicitly. States that PAQLQ is a disease specific quality of life questionnaire that has been developed to measure the physical, emotional, and social impairments that are experienced by children with asthma
1106	Steven R. Ericson 2002, The USA [45]	PAQLQ	Asthma	n= 99, age= 12.6(2.3),	The effect of the disease and its treatment on the lives of patients and their caregivers
756	Rydstrom 2005, Sweden[61]	PAQLQ	Asthma,	N=226 children, 371 parents Age of children: 116 $\leq$12 yrs. And 105 $\geq$13 yrs. Parents- 147 $\leq$39 , 223 $\geq$40	Not reported
916	Warschburger 2004, Germany[91]	PAQLQ, German version	Asthma,	Asthma Intervention group=226 Control group=92 , total n=318 Age= Intervention group= 11.91 (±2.03) Control group =11.25 (±2.05) Range 8-16	Quality of life is defined as a multidimensional concept that encompasses broad domains of life and the individual's overall satisfaction with life and health.
2190	Carole J. Sapp 2003, The USA[62]	PAQLQ	Asthma	N: 99 Age: 12 to 17 years, mean age 14.26 (1.65)	Perception of functional impairments on daily life as a result of their illness
2791	Nogueira 2010, Brazil [58]	PAQLQ	Asthma	N=210 12-21 yrs.	Quality of life is defined 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

					of their goals, expectations, standards and concerns"
2964	Robert D. Annett 2003, Mexico[35]	PAQLQ	Asthma	n= 339 child-parent paits , age child= 9.3 (2.2) years	Not reported
2965	Robert D. Annet 2001, The USA[36]	PAQLQ	Asthma	n= 339, age= 9.3 (2.2)years at 12 month f/u	Reported three definitions and acknowledged uncertainty. (this tudy is a part of CAMP study) 1) Qol is an important measure of the patient's subjective experience with their disease. 2) Qol of life encompasses several domains of the patient's subjective experience with the illness: physical status, functional abilities, psychological well-being, social interactions, and economic factors. 3) Qol is a scientific outcome that represents a patient's functional status that results from a disease or its treatment.
2966	Annett 2010, The USA[37]	PAQLQ	Asthma	age= 11.4 (2.2) years at 36-month f/u, range 8–16 years, n= 280	Influence of disease on patient's life
3675	Zandieh 2006, Iran[93]	PAQLQ (Persian)	Asthma	n= 113, age= 7 to 17 y. o. a	Individual perception of their position in life in the context of culture and value systems in which they live and in relation to their goals, expectations, standards and concerns

4458	Walker 2008, The USA[87]	PAQLQ	Asthma	n= 221, age= 8.05(1.72) years	1) Quality of life (QOL) is defined by the individual and depends on many factors such as lifestyle, past experiences, hopes for the future, dreams and ambitions 2) Quality of life for a child with asthma has been defined as the measure of emotions, asthma severity/symptoms, emergency department visits, missed school days, activity limitations, and visits to the emergency department.
235	Green 2009, Poland [63]	PAQLQ	Asthma,	n= 137 (100 health, 37 asthmatics), age= 10.6, age range: 7-14 y.o.a	Not reported
243	Basaran, 2006, Turkey[25]	PAQLQ	Asthma	n= 62, age= 7-15 yrs, mean age = 10.4 (2.1)	Not reported
719	Riccioni 2006, Italy[60]	AQLQ	Asthma patient with/with out Nasal polyposis	24 NP cases, 65 controls, 89 total participants. Age: NP cases with mean age of 37 $\pm$ 14 yr. and 65 controls with a mean age of 25 $\pm$ 8 yr.	Not reported
1287	Busse, 1998, The USA[40]	AQLQ	Asthma	n= 538: Salmeterol group n= 263 Placebo group n= 275 Age	Quality of life assessment reflects patients' concerns and perception of how the disease and its management influence their lives.

				Salmeterol group –mean= 37.2, range 12-80 Placebo group mean= 35.3 , range 12-75	
1302	Wenzel 1998, The USA[92]	AQLQ	Asthma	N=539 (Salmeterol : 264, Albuterol: 275) Age: Sal 12-73 mean 35.4, Alb 12-83 mean 33.8 (al)	Quality of life assessment is important to the health care provider because it provides additional, unique information from the patient's perspective regarding disease status and management. As such, it reflects patients' concerns and opinions regarding the impact of the disease and subsequent therapeutic interventions on their lives
3333	Lockey 1999, The USA[57]	AQLQ	Asthma	Salmeterol group n: 240, age: 12-76, mean age 40 yoa; Placebo group n: 234, age: 12-73, mean age 38 yoa	The impact of the disease on daily functioning and well-being
862	Wagner 2004, The U.S.A[86]	SEIQoL-DW	Diabetes Type 1	N=80 Age= 14 + 3.4	Aspects of life those are important to the individual.
4570	Chenhall 2010, Australia[41]	SEIQoL-DW	History of involvement in drugs, alcohol and criminal activity	Indigenous male youth 14-19 y. o. a N: 15	A person's quality of life can be defined as the extent to which "a person's valuable functioning are achieved and the extent to which they have the opportunity to choose from valuable options
442	Bernard 2008, The UK[39]	SEIQOL-DW	Type 1 Diabetes	age: 9-17 years, mean 12.07 (2.71)	Not reported
329	Bolkland 2011, Kenya[43]	SEIQOL-DW	Spina bifida	Health group n = 64 (24 parent interviews, 40 children interview); Spina bifida group n = 102	Not reported

				(63 parents, 39 children internviews) Age range (SB group) 6 mos-12+	
2256	Walker 2002, The UK[90]	SEIQoL	Type 1 diabetes	N=15 Age: mean 15.19, range=13.4 – 17.5	Not reported
177	Gooseens 2009, Belgium[54]	RQLQ	Seasonal allergic rhinitis	n=46, 14-68 years, mean= 36 (+-13.5)	No clear definition base it on RQLQ result
5540	John S. Rhee 2003, The USA[59]	RQLQ	Surgically treatable nasal airway obstructio n,	n= 40, age= 16 to 74 y. o. a	Disease specific qol are the aspects of patients' life affected by the disease.
724	Hull 2004, The USA[55]	MYMOP	Undergoin g acupunctu re for various reasons,	16-89 years group mean age: 54.3 (17.5), n= 110	Not reported
429	Jolles 2008, Switzerland[56]	PASI-pg	JRA	15-42 yrs, n= 14 (22 knees)	Not reported
Studies that use generic or disease-specific individualized measures to assess secondary outcome					
Ref ID	Study (First author's last name & pub year)	Instrument name	Country	Underlying condition/diagnosis, sample size & age	Definition of HRQoL

Mis 5	Vinson 2002, The USA	PAQLQ	Asthma	Age: 7 to 12.11 (mean age 3.5) y.o.a N: 235	Not reported
13 mis	Balon 1998, Canada	PAQLQ	Asthma	Active treatment group- n: 38, age: 11.4(2.5) Simulated treatment group- n: 42, age: 12.1 (2.7)	Not reported
2464	Wheeler 1997, The USA[32]	PAQLQ, and AQLQ	Atopy due to Asthma or rhinitis	Age: 13 to 45 (mean 33) years N= 15	Not reported
Mis 181	Hamre 2009, Germany	AQLQ ? German version	Asthma	N= 54 adults (18 to 68 years), 36 children (2-16 years)	Not reported
1201	H. Ramstrom 2000, Sweden[65]	SEIQOL-DW	Cystic Fibrosis	20 (18 adults, 2 children) in QOL Age- QOL group-Adult 29 (21 to 41) children 15	Not reported
1333	Paterson 1997, The UK[64]	MYMOP		New patients of acupuncture (10 to 77 yrs, mean 47; n= 29) and homeopathy (age: 3 to 76 yrs, mean: 42; n= 24)	Not reported

Mis 85	Cassale 2001, The USA	RQLQ	Seasonal allergic rhinitis	Age: 12- 75 years, mean: 34.5 (12.5) N: 529	Not reported
Studies that use generic or disease-specific individualized measures to validate other measures					
Ref ID	Study (First author's last name & pub year), country	Instrument name	Condition	Underlying condition/diagnosis, sample size & age	Definition of HRQoL
1338	Guyatt 1997, Canada[66]	PAQLQ	Asthma	N= 52 Age: 7 to 10 yr group, mean = 8.8 (1.3); 11 to 17 yr group, mean = 13.8 (2.2)	Not reported
2453	Shelley. C. Mishoe, 1998, The USA [29]	PAQLQ used to validate AMA	Asthma	N=35 Age= range= 6-12	Not reported
2458	Guyatt 1997, Canada[30]	PAQLQ (proxy issue) compared to global rating of change	Asthma	N=52 7-10 year olds group: 8.8 (1.3) y. o. a; 11-17 year olds group: 13.8 (2.2) y. o. a	Not reported
3145	Flapper 2006, The Netherlands[46]	PAQLQ	Asthma,	n= 298, age= 10(2) y. o. a	Quality of life can be measured, either as healh status, where health refers to actual problems and limitations in functioning; or as health related qol which is a different concept and includes the subject's own appraisal of health status.

3253	Juniper 1996, Canada[67]	PAQLQ	Asthma	N= 52 7 to 17 y.of.a	Not reported
3547	Somerville 2004, Germany[68]	PAQLQ German version	Asthma	Group A: n= 11, age= 14.7 (1.65) Group B: n= 56, age= 14.6 (1.8)	HRQL may be defined as "the functional effect of an illness and its consequent therapy upon a patient, as perceived by the patient."
3603	Ungar 2006, Canada[69]	PAQLQ	Asthma	N: 8-10 year group n= 11, 11-15 year group n= 5	HRQOL represents the patient's perception of four essential domains: disease state and physical symptoms, functional status, psychological functioning and social functioning.
2439	Barley 1998, The UK[28]	AQLQ used to validate other asthma measure	Asthma	N=90 Mean age =46 Range= 17-79	"generally speaking quality of life refers to a construct that is specific to an individual"
5243	E.A. Barley 1999, The UK[38]	AQLQ	Asthma	N: 74 Age: 17-76 years	Not reported

RQLQ: Rhino-conjunctivitis Quality of Life Questionnaire; PAQLQ: Pediatric Asthma Quality of Life Questionnaire; PASI-pg: Patient Specific index-patient generated, JRA: Juvenile rheumatoid arthritis

Table 3-1b: Study characteristics (studies reporting on psychometric properties)

Validation studies on disease specific measures					
Ref ID	Study (First author's last name & pub year), country	Instrument name	Condition	Sample size, and age (mean & range if provided)	Definition of HRQL
546	Ziora 2007, Poland[85]	PAQLQ	Asthma	N= 54 7-17 y. of. a (mean = 10.7)	Not reported
2376	E. .Tauler, 2001, Spain[27]	Spanish PAQLQ	Asthma	N=99, age= 7-17 years, mean age= 11.3 (2.9)	Not reported
2459	Juniper 1997, Canada[31]	PAQLQ (assessment of minimum age and reading skills required by children)	Asthma	N=52 Age Mean 12.0±3.1 Range (7–17)	Not reported
2469	Juniper 1996, Canada[33]	PAQLQ	Asthma	Age: 7-17 year, n= 52	Physical, social, emotional impairment that individuals may experience as a result of their illness. In this case asthma.
3123	Clarke 1999, Singapore[44]	PAQLQ-modified	Asthma	Group A- n: 19, age: 9.84 (2) Group B- n: 28, age: 9.96 (1.87)	Quality of life, is a multi-dimensional measure encompassing the physical, emotional, and social functioning of the child. It should measure the uniquely personal perspective of an individual on his/her health status and encompass the non-medical aspects of his/her life.
3252	Juniper 1996, Canada[70]	PAQLQ	Asthma	N=52, age= 7-17 [12.0 (3.1)] y. o. a	Not reported

3254	Juniper 1997, Canada[71]	PAQLQ	Asthma	N= 52, age= 7-17 [12.0 (3.1)] y. o. a	Functional impairments (physical, emotional and social) that are important to the patients in their everyday lives
3455	Raat 2005, The Netherlands[77]	PAQLQ (Dutch)	Asthma	N=238 , age= 6-18 y[10.8 (3.2)] y. o . a	NR
3467	Reichenberg 2000, Sweden[78]	PAQLQ-Swedish version	Asthma	61 participants 8.7 + 0.8 yrs.	Quality is defined as a grade of "goodness" and "quality of life" with regard to health and disease as "representing individual responses to the physical, mental and social effects of illness on daily living which influence the extent to which personal satisfaction with life circumstances can be achieved. The concept of health related quality of life (HRQL) when applied to children has mainly been limited to day-to-day functioning of the child, such as play, school-work, sports, ability to interact with family and peers, and sleep as well as worry and somatic consequences of the disease
3510	Sarria 2012, Brazil[80]	PAQLQ (Brazilian Portuguese version)	Asthma	Age: 8-17, mean 11 years N: 125	Not reported
3674	Yuksel 2009, Turkey[84]	PAQLQ (Turkish)	Asthma	N= 122, age= 7-16 [9.9 (2.2)] y. o. a	1) Health related quality of life is the effect of disease and its treatment as the patient perceives it. 2) It is a multidimensional issue including physical

					functioning, somatic sensation and social and emotional well-being.
5421	Juniper 1993, Canada[72]	AQLQ	Asthma	N: 39 Age: mean 42.0 (13.7)	Not reported
5492	Muntner 2000, Switzerland[73]	AQLQ	Asthma	N: 115 Age: 17 – 81, mean 44 (15)	NR
5500	Oga 2002, Japan[74]	AQLQ	Asthma	N=109, age= 16-87 y.o.a, mean age = 46.3 (17.8)	Not reported
5588	Spiric 2004, Serbia[81]	AQLQ (Serbian)	Atopy	Atopic n: 100, age: 16 to 63, mean age: 35; Nonatopic n: 60, age: 23 to 74 y. o. a, mean age: 49	The health related quality of life is a component of overall quality of life, which is primarily determined by the health of an individual and influenced by various clinical interventions.
5288	Coates 1990, Australia[42]	GLQ-8	Cancer (chemotherapy patients)	N: 166 Age: 17-81 years	Not reported

Validation studies on generic measures (ch:4)					
Ref ID	Study (First author's last name & pub year), country	Instrument name	Condition	Sample size and age (mean & range if provided)	Definition of HRQoL
943	Frick 2004, Germany[47]	SEQOL-DW	Multiple myloma, non-Hodgkin's lymphoma, and other	age: 10-70 years. N= 10-20: n= 1, total n= 79	QOL: can be defined as the difference or the gap between the current hopes and expectations of the individual and that individual's present experiences. Individual QOL: can be defined by

			tumors		asking the individual what is QOL or, more accurately: QOL is not merely what the patient says it is, but what he tells himself or herself what it is.
3629	Wagner 2004, Netherlands[83]	SEIQOL-DW	Type 1 Diabetes	N: 77 Age: 8-17, 12.4(2.5) y.o.a	Not repoerted
4457	Vinson 2010, The USA[82]	SEIQoL-DW	Cerebral palsy (CP)	CP group :n= 41 (8.8 +- 1.8 yoa), Other group: n = 60 (8.9 +- 1.7)	1) Qol as defined by the WHO stresses the importance of considering the individual's cultural context when examining the construct 2) Qol is a multidimensional, dynamic and person-centered construct which includes as assessment of subjective well-being.
1227	Paterson et al 2000, the UK[76]	MYMOP	CAM patients	New patients of 12 complimentary practitioners 176 patients completed MYMOP, 20 interviews; age: 16-86 years	Not Reported
3432 ch 4	Paterson 1996, The UK[75]	MYMOP	CAM patients	N= 265, 218 general practice patients, 47	Not reported

				CAM patients (2-84 years; mean & SD: 47 +- 17.6); 109 completed the f/u	
5553 ch 4	Ruta 1994, Scotland [79]	PGI	Low back pain patients	N= 777 (571 complete questionnaires), 16 to 85, mean= 43 years	Not reported

Table 3-2: Identified measures in order of frequency of published studies

		Type of studies identified			
			Use studies		
Instrument type	Instrument name	Validation studies	Pri	Sec	VD
Disease specific	Pediatric Asthma Quality of Life Questionnaire (PAQLQ)	11 studies	17	3	7
	Asthma Quality of Life Questionnaire (AQLQ)	4	4	1	2
	Patient Specific Index (PASI-pg)	None	1	None	None
	Rhinoconjunctivitis Quality of Life Questionnaire (RQLQ)	None	2	1	None
	Global Quality of Life questionnaire (GLQ-8)	1	None	None	None
Generic	Schedule for the Evaluation of Individual Quality of Life Direct Weight (SEIQoL-DW)	3	5	1	None
	Patient generated Index (PGI)	1	None	None	None
	Measure Yourself Medical Outcome Profile (MYMOP) 2+2	2	1	1	None

Pri: Primary outcome, Sec: secondary outcome, VD: to validate other measures

Table 3-3: Description of Included HRQL measures (based on included validation studies only)

Description or definition		**Included measure (s)***				
	Included measure (s)	**PAQLQ (original)**	**AQLQ**	**SEIQoL-DW**	**PGI**	**MYMOP**
Target population	Population: Mean age (SD and/or range), diagnosis/underlying condition	7-17 year asthma patients	17-70 years with asthma	-Type 1 Diabetes, 12.4(8-17, 2.5); -Cerebral palsy, CP group: 8.8 (1.8); TD: 8.9 (1.7)	Low back pain of varying severity, 43 (16-85)	N= 265 (2-84 years; mean & SD: 47 +- 17.6); 109 completed the f/u patients receiving CAM (P 1996); N= 81 (32.4-82.8 years) patients of chronic bronchitis with acute exacerbation (p 2000)
Description of what it is that the questionnaire intends of measure	Construct (e.g. Physical symptoms or psychological symptoms, quality of life etc)	Asthma specific quality of life	Health related quality of life in asthma	Quality of life (respondents are asked to identify five areas of life, affected by a particular condition that are important to overall quality of life)	Quality of life	Physical symptoms, activity, wellbeing
Names of those domains that allow responders/patients to identify/name their specific complaints. A domain or dimension refers to the area of behavior that we are trying to measure. For example: physical function domain can include questions about mobility and self care. And a domain on emotional function might ask about depression, anxiety, social issues etc	Name of Individualized Domain(s)	Activity domain: three of the five questions in the activity domain are individualized	Activity domain: five of the nine questions in the activity domain are individualized	Can be from any five domain as chosen by patient	Five most important (patient-specific/individualized) areas or activities of their life affected by current medical diagnosis	Symptom(s), and activity as chosen by patient;
	Recall period for individualized questions only	7-days	2-weeks	n/a	One month	7-days

Each HRQoL measure has different purpose, for detail refer to the section on psychometric properties of protocol	Purpose of the HRQoL measure used (evaluative, predictive, or discriminative)	Evaluation and discrimination	Evaluation and discrimination	Evaluative	Evaluation	Evaluation
Questionnaire can be filled/completed by patient or otherwise	Mode of administration (e.g. self, interview, proxy administered if proxy administered: name of proxy e.g. parent or etc)	Self and interview administered	Self and interview administered	Semi-structured interview	Self and interview administered	administered, follow-up self administered
	Abstract the individualized questions verbatim	Each child is asked to identify three activities in which they are limited due to their asthma	Questionnaire attached as Appendix	All questions are individualized	All questions are individualized	Questionnaire attached as Appendix
Number and type of response categories	Number and type of response options (Scale) for individualized questions only	7-point scale (7= no impairment, 1= sever impairment)	7- point scale (1= totally limited, couldn't do activity at all, 7= Not limited at all)	100 mm visual analogue scale (VAS); 5-point likert scale: high score equals worse outcome	11-point scale	0-6 point scale (Seven points)
scoring algorithms for individualized questions	Scoring of individualized questions (e.g. summed, averaged, weighted, overall or subscale?)	All questions are equally weighted; overall score is the mean all questions.	Scale scores are calculated as the average score among the items forming each scale. Overall score is calculated as the mean of all items	Total quality of life is calculated by multiplying each domain's importance weighting (proportion of 1.0) by its satisfaction score (0–100) then summing the products. The result is a total QOL score on a scale from 0–100, with higher scores indicating better quality of life. A Total SEIQoL Index score was calculated by multiplying each domain's importance ranking by its satisfaction score and then summing the products. The result	A index ranging from 0-100 is generated by multiplying scores of each of the six ratings by the proportion of points allocated to that area and summing	

				is a total score with a potential range from 15 to 75. The higher scores indicate a perceived better quality of life.		
Generic or disease-specific (generic measures will be evaluated using COSMIN and Terwee criteria ch 4)		Asthma specific	Asthma specific	Generic	Generic	Generic

* PASI-pg, and RQLQ were omitted from the table because there were no validation studies identified on these two measures.

* GLQ-8 was excluded because the age range of participants was 17-81 years, and it was unclear how many participants were under 18

Table 3-4 Definitions

Definition of HRQL or QOL	No of studies
Perception of individual	9
Effect of physical, social, and emotional status on health/QOL/everyday life (based on WHO's definition of health)	6
Effect of disease and/or treatment on life (or individual's health)	5
A construct that is specific to individuals	2
Subjective perception of difference between hopes and expectations from present experience	1
Satisfaction with life and health	1
Cultural context is important to consider; multidimensional person-centered construct that includes assessment of subjective wellbeing	1
Achievement of valuable functioning and opportunity to choose from valuable options	1
QOL= health status, HRQL is the subject's own appraisal of health status	1

APPENDIX 1

Medline March 28 2011

1. "Quality of Life"/
2. "Activities of Daily Living"/
3. Adaptation, Psychological/
4. depression/ or stress, psychological/
5. personal satisfaction/
6. Health Status/
7. happiness/
8. Anxiety/
9. (health status or functional status or personal satisfaction).ti,ab.
10. (quality of life or qol or hrqol).ti,ab.
11. (happy or unhappy or happiness or unhappiness or enjoy* or pleasur* or comfort* or optimis*).ti,ab.
12. ((activit* adj3 limit*) or (life* adj3 chang*) or function of satisfaction or well being).ti,ab.
13. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12
14. (patient centered or patient centred or patient generated).ti,ab.
15. (consumer generated or patient oriented or patient perspective*).ti,ab.
16. Self-Assessment/
17. (individuali?ed or personali?ed).ti,ab.
18. 14 or 15 or 16 or 17
19. 13 and 18
20. (mymop or seiqol or addqol or aqlq or womac or mycaw or psyclops or cmymop or w bq12).ti,ab.
21. patient generated index.mp.
22. (patient generated index or repertory grid or subjective quality of life profile or chronic respiratory disease questionnaire or disease repercussion profile or patient enablement instrument or well being questionnaire).ti,ab.
23. (measure yourself medical outcome profile or schedule for the evaluation of individual quality of life or audit of diabetes dependent qol).ti,ab.
24. (asthma quality of life questionnaire or "western ontario and mcmaster university osteoarthritis index" or "measure yourself concerns and wellbeing").ti,ab.
25. (pgi or pei or drp).mp. and 13
26. 20 or 21 or 22 or 23 or 24 or 25
27. 19 or 26
28. (child* or pe?diatric* or infant* or newborn* or neonat* or youth* or juvenile* or adolecen* or teenag*).mp.
29. 27 and 28

PROQOLID Search criteria

	Type of instrument = Health-Related Quality of Life - HRQL (incl. health status) (398 found)
AND	Population = Adolescent (75 found)
OR	Population = Pediatrics (56 found)

387	=	1936	AND	391
=	1903	OR	391	=
1910	AND	0	=	Choose a criterio
AND	0	=	Choose a criterio	AND
0	=	Choose a criterio	AND	0
=	Choose a criterio	AND	0	=
Choose a criterio	AND	0	=	Choose a criterio

Hapi

1. quality of life.de,ti,ab.
2. (qol or hrqol or hrql).de,ti,ab.
3. 1 or 2
4. (child* or adolescen* or infan*).mp. [mp=title, acronym, descriptors, measure descriptors, sample descriptors, abstract, source]
5. 3 and 4

EMBASE

1. "quality of life"/ or quality adjusted life year/ or "quality of life index"/
2. daily life activity/
3. happiness/
4. mental stress/
5. health status/ or functional status/ or physical mobility/
6. satisfaction/ or life satisfaction/ or patient satisfaction/
7. anxiety/
8. exp depression/
9. (quality of life or qol or hrqol).ti,ab.
10. (happy or unhappy or happiness or unhappiness or enjoy* or pleasur* or comfort* or optimis*).ti,ab.
11. ((activit* adj3 limit*) or (life* adj3 chang*) or function of satisfaction or well being).ti,ab.
12. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11
13. self evaluation/
14. (patient centered or patient centred or patient generated).ti,ab.
15. (consumer generated or patient oriented or patient perspective*).ti,ab.
16. (individuali?ed or personali?ed).ti,ab.
17. 13 or 14 or 15 or 16
18. 12 and 17
19. (mymop or seiqol or addqol or aqlq or womac or mycaw or psyclops or cmymop or w bq12).ti,ab.
20. (patient generated index or repertory grid or subjective quality of life profile or chronic respiratory disease questionnaire or disease repercussion profile or patient enablement instrument or well being questionnaire).ti,ab.
21. (measure yourself medical outcome profile or schedule for the evaluation of individual quality of life or audit of diabetes dependent qol).ti,ab.

22. (asthma quality of life questionnaire or "western ontario and mcmaster university osteoarthritis index" or "measure yourself concerns and wellbeing").ti,ab.
23. (pgi or drb or pei).mp. and 12 [mp=title, abstract, heading word, table of contents, key concepts]
24. 19 or 20 or 21 or 22 or 23
25. 18 or 24
26. limit 25 to (infant or child or preschool child <1 to 6 years> or school child <7 to 12 years> or adolescent <13 to 17 years>)
27. (child* or pe?diatric* or infant* or newborn* or neonat* or youth* or juvenile* or adolecen* or teenag*).mp.
28. 25 and 27
29. 26 or 28

Medline Search with proqolid identified terms April 5 2011

1. Adolescent Asthma Quality of Life Questionnaire.mp.
2. Asthma Control Questionnaire.mp.
3. asthma Control Test.mp.
4. Acne Disability Index.mp.
5. About My Asthma.mp.
6. Assistive Technology Device Predisposition Assessment.mp.
7. Adolescent Rhinoconjunctivitis Quality of Life Questionnaire.mp.
8. Brace Questionnaire.mp.
9. Brief Illness Perception Questionnaire.mp.
10. Childhood Asthma Questionnaires.mp.
11. Children's Dermatology Life Quality Index.mp.
12. Adolescent Asthma Quality of Life Questionnaire.mp.
13. Asthma Control Questionnaire.mp.
14. asthma Control Test.mp.
15. Acne Disability Index.mp.
16. About My Asthma.mp.
17. Assistive Technology Device Predisposition Assessment.mp.
18. Adolescent Rhinoconjunctivitis Quality of Life Questionnaire.mp.
19. Brace Questionnaire.mp.
20. Brief Illness Perception Questionnaire.mp.
21. Childhood Asthma Questionnaires.mp.
22. Children's Dermatology Life Quality Index.mp.
23. Cystic Fibrosis Questionnaire.mp.
24. Childhood Health Assessment Questionnaire.mp.
25. "Child Health and Illness Profile".mp.
26. child-oidp.mp.
27. Child Health Questionnaire.mp.
28. Child Oral Health Quality of Life Questionnaire.mp.
29. College of Optometrists in Vision Development Quality of Life Outcomes Assessment.mp.
30. Chronic Respiratory Disease Questionnaire.mp.

31. Client Satisfaction Questionnaire.mp.
32. Dutch Eating Behavior Questionnaire.mp.
33. Diabetes Family Behavior Scale.mp.
34. Dermatitis Family Impact questionnaire.mp.
35. Diabetes Quality of Life for Youth scale.mp.
36. Dermatology Quality of Life Scales.mp.
37. Diabetes Self-Management Profile.mp.
38. Family System Test.mp.
39. "Health and Daily Living Form".mp.
40. Hypoglycemia Fear Survey.mp.
41. Health Status Questionnaire.mp.
42. Infants' Dermatitis QOL Index.mp.
43. IMPACT III.mp.
44. Impact of Weight on Quality of Life.mp.
45. Juvenile Arthritis Quality of Life Questionnaire.mp.
46. KIDSCREEN.mp.
47. kindl.mp.
48. lsia.mp.
49. Low Vision Quality-of-Life Questionnaire.mp.
50. Mental Health Inventory.mp.
51. McMaster Health Index Questionnaire.mp.
52. Minneapolis-Manchester Quality of Life instrument.mp.
53. The Miami Pediatric Quality of Life Questionnaire: Parent Scale.mp.
54. Memorial Symptom Assessment Scale.mp.
55. oidp.mp.
56. Pediatric Restless Legs Syndrome Severity Scale.mp.
57. Paediatric Asthma Quality of Life Questionnaire.mp.
58. pedi mcat.mp.
59. pedi.mp.
60. Patient-Oriented Eczema Measure.mp.
61. howru.mp.
62. Youth Quality of Life Instrument.mp.
63. Wong-Baker FACES Pain Rating Scale.mp.
64. WeeFIM instrument.mp.
65. "Warwick Child Health and Morbidity Profile".mp.
66. Well-Being Questionnaire.mp.
67. Vineland Adaptive Behaviour Scales.mp.
68. Quality of Life measure for children.mp.
69. TNO-AZL Preschool children Quality of Life questionnaire.mp.
70. TNO AZL Children's Quality of Life.mp.
71. Stoma-QOL.mp.
72. sf 36.mp.
73. sf 12.mp.
74. Rivermead Behavioral Memory Test.mp.
75. qolie.mp.
76. Quality of Life Inventory.mp.

77. Quality of Life in Childhood Epilepsy Questionnaire.mp.
78. "Quality of Life Enjoyment and Satisfaction Questionnaire".mp.
79. Piers-Harris Children's Self-Concept Scale.mp.
80. pedsql.mp.
81. Pictorial Thai Quality of Life.mp.
82. Paediatric Rhinoconjunctivitis Quality of Life Questionnaire.mp.
83. Patient-Oriented Eczema Measure.mp.
84. pi ed.mp.
85. 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84
86. exp reliability/
87. exp validity/
88. psychometry/
89. (reproducib* or psychometr* or clinimetr* or clinometr* or reliab* or valid* or coefficient or internal consistency or cronbach alpha* or item correlation* or item selection* or item reduction* or agreement or precision or imprecision or precise values or test retest or predict*).mp.
90. ((test and retest) or stability or interrater or inter rater or intrarater or intra rater or interobserver or inter observer or intraobserver or intra observer or intertechnician or inter technician or intratechnician or intra technician or interexaminer or inter examiner or intra examiner or intra examiner or inter assay or interassay or intra assay or intra assay or interindividual or inter individual or intra individual or intraindividual or interparticipant or inter participant or intraparticipant or intra participant).mp.
91. 86 or 87 or 88 or 89 or 90
92. 85 and 91
93. limit 92 to "all child (0 to 18 years)"
94. (child* or pe?diatric* or infant* or newborn* or neonat* or youth* or juvenile* or adolecen* or teenag*).mp.
95. 92 and 94
96. 93 or 95
97. (patient centered or patient centred or patient generated).ti,ab.
98. (consumer generated or patient oriented or patient perspective*).ti,ab.
99. Self-Assessment/
100. (individuali?ed or personali?ed).ti,ab.
101. 97 or 98 or 99 or 100
102. 96 and 101
103. or/1-84
104. limit 103 to "all child (0 to 18 years)"
105. 103 and 94
106. 104 or 105
107. 106 and 101

108. 107 not 102

EBM Child HRQL

1. (child or children or childhood or adolescen* or youth or pediatric* or paediatric* or infant* or neonat* or newborn*).mp.
2. exp Quality of Life/
3. (quality of life or qol or hrqol).ti,ab,kw.
4. health status.mp. or Health Status/
5. "Activities of Daily Living"/ or functional status.ti,ab,kw.
6. personal satisfaction/
7. Adaptation, Psychological/
8. Stress, Psychological/
9. (well being or wellbeing).ti,ab,kw.
10. ((activit* adj3 limit*) or (life* adj3 change*)).ti,ab,kw.
11. ((function* or satisfaction) adj3 (measur* or assess* or evaluat* or instrument* or tool* or battery or scale* or outcome* or survey* or questionnaire*)).ti,ab,kw.
12. (distress* or cope or coping or disab* or handicap*).ti,ab.
13. quality of life research.jn.
14. (patient centered or patient centred or patient generated).ti,ab.
15. (consumer generated or patient oriented or patient perspective*).ti,ab.
16. Self-Assessment/
17. (individuali?ed or personali?ed).ti,ab.
18. or/14-17
19. or/2-13
20. 1 and 18 and 19

PsycINFO

1. "quality of life"/
2. "activities of daily living"/ or ability level/ or activity level/ or daily activities/ or habilitation/ or physical mobility/ or self care skills/
3. happiness/
4. psychological stress/
5. satisfaction/ or client satisfaction/ or life satisfaction/ or dissatisfaction/ or physical comfort/
6. well being/
7. exp adjustment/ or adaptive behavior/
8. anxiety/
9. "depression (emotion)"/ or major depression/ or sadness/
10. (quality of life or qol or hrqol).ti,ab.
11. (happy or unhappy or happiness or unhappiness or enjoy* or pleasur* or comfort* or optimis*).ti,ab.
12. ((activit* adj3 limit*) or (life* adj3 chang*) or function of satisfaction or well being).ti,ab.
13. or/34-45
14. self evaluation/ or self report/

15. (patient centered or patient centred or patient generated).ti,ab.
16. (consumer generated or patient oriented or patient perspective*).ti,ab.
17. (individuali?ed or personali?ed).ti,ab.
28. 47 or 48 or 49 or 50
19. 46 and 51
20. (mymop or seiqol or addqol or aqlq or womac or mycaw or psyclops or cmymop or w bq12).ti,ab.
21. patient generated index.mp.
22. (patient generated index or repertory grid or subjective quality of life profile or chronic respiratory disease questionnaire or disease repercussion profile or patient enablement instrument or well being questionnaire).ti,ab.
23. (measure yourself medical outcome profile or schedule for the evaluation of individual quality of life or audit of diabetes dependent qol).ti,ab.
24. (asthma quality of life questionnaire or "western ontario and mcmaster university osteoarthritis index" or "measure yourself concerns and wellbeing").ti,ab.
25. (pgi or pei or drp).mp. and 46
26. or/53-58
27. 52 or 59
28. limit 60 to (childhood or adolescence <13 to 17 years>)
29. (child* or pe?diatric* or infant* or newborn* or neonat* or youth* or juvenile* or adolecen* or teenag*).mp.
30. 60 and 62
31. 61 or 63

CINAHL

S25: 23 or 24

S24 S17 or S18 or S19 or S20 or S21 — False — S24

S23 S14 and S22 — False — S23

S22 S15 or S16 — False — S22

S21 (pgi or pei or drp) and S14 — False — S21

S20 asthma quality of life questionnaire or "western ontario and mcmaster university osteoarthritis index" or "measure yourself concerns and wellbeing" — False — S20

S19 measure yourself medical outcome profile or schedule for the evaluation of individual quality of life or audit of diabetes dependent qol — False — S19

S18 patient generated index or repertory grid or subjective quality — False

	of life profile or chronic respiratory disease questionnaire or disease repercussion profile or patient enablement instrument or well being questionnaire	S18
S17	mymop or seiqol or addqol or aqlq or womac or mycaw or psyclops or cmymop or w bq12	False S17
S16	(MH "Self Assessment")	False S16
S15	(patient centered or patient centred or patient generated) or (consumer generated or patient oriented or patient perspective*) or (individuali?ed or personali?ed)	False S15
S14	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13	False S14
S13	(activit* N3 limit*) or (life* N3 chang*) or function of satisfaction or well being	False S13
S12	happy or unhappy or happiness or unhappiness or enjoy* or pleasur* or comfort* or optimis*	False S12
S11	quality of life or qol or hrqol	False S11
S10	health status or functional status or personal satisfaction	False S10
S9	(MH "Quality of Life+") OR (MH "Quality-Adjusted Life Years")	False S9
S8	(MH "Anxiety+")	False S8
S7	(MH "Happiness")	False S7
S6	(MH "Health Status Indicators") OR (MH "Health Status+")	False S6
S5	(MH "Personal Satisfaction+")	False S5
S4	(MH "Depression+")	False S4
S3	(MH "Stress, Psychological+")	False S3

False

S2

S2 (MH "Adaptation, Psychological+")

S1 (MH "Activities of Daily Living+")

Figure 3-1: PRISMA 2009 Flow Diagram

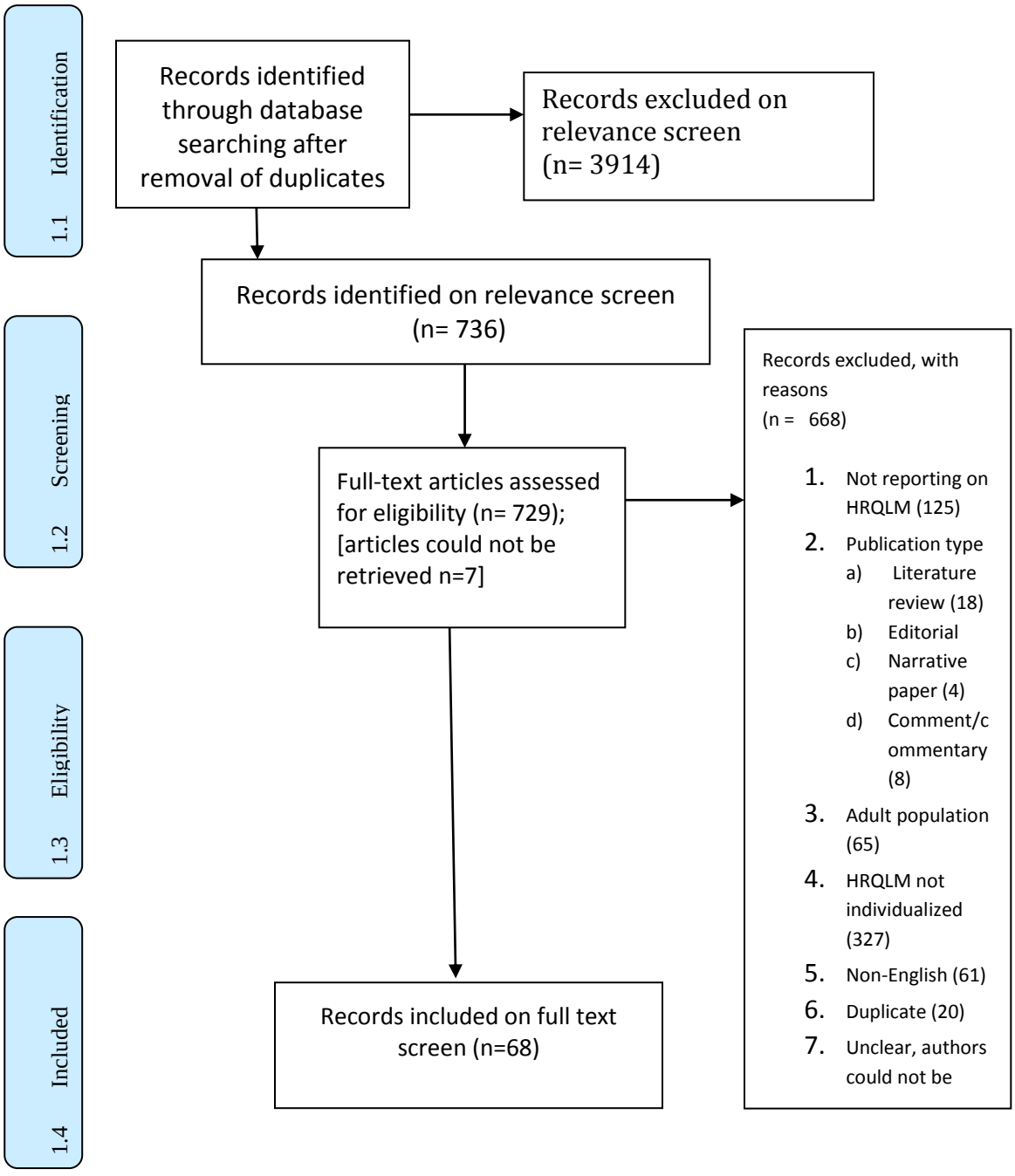
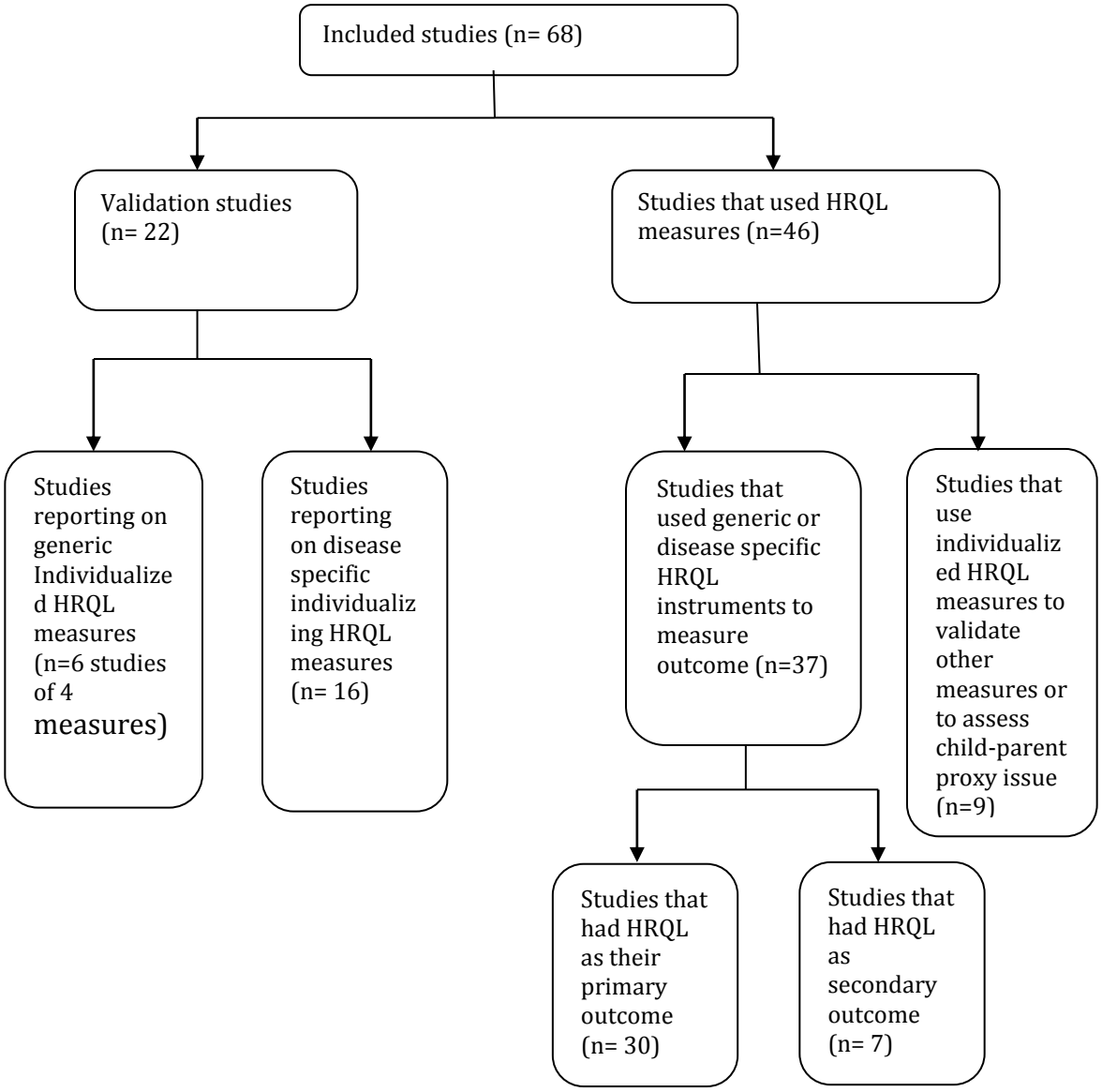


Figure 3-2: Description of included studies



8. References

1. Admins. Patient-centered outcomes research institute amends draft research agenda in response to public comment. *Patient-Centered Outcomes Research Institute* 2012; .

2. Trends in disability prevalence and their causes: Proceedings of the national disability statistics and policy forum (4th, washington, DC, may 16, 1997). disability forum report 4. Retrieved 7/20/2012, 2012, from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED425606&ERICExtSearch_SearchType_0=no&accno=ED425606 .

3. Matza LS, Swensen AR, Flood EM, Secnik K, and Leidy NK. Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value Health* 2004; 7: 79-92.

4. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

5. Guyatt GH, Feeny DH, and Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622-629.

6. Bowling A. *Measuring disease: a review of disease-specific quality of life measurement scales.* Buckingham, ENGLAND: Open University Press. 1995: 374 p,374 p.

7. Bowling A. What things are important in people's lives? A survey of the public's judgements to inform scales of health related quality of life. *Soc Sci Med* 1995; 41: 1447-1462.

8. Clinch J, Tugwell P, Wells G, and Shea B. Individualized functional priority approach to the assessment of health related quality of life in rheumatology. *J Rheumatol* 2001; 28: 445-451.

9. Baca CB, Vickrey BG, Hays RD, Vassar SD, and Berg AT. Differences in child versus parent reports of the child's health-related quality of life in children with epilepsy and healthy siblings. *Value in Health* 2010; 13: 778-786.

10. Brouwer CN, Maille AR, Rovers MM, Grobbee DE, Sanders EA, and Schilder AG. Health-related quality of life in children with otitis media. *Int J Pediatr Otorhinolaryngol* 2005; 69: 1031-1041.

11. Carlon S, Shields N, Yong K, Gilmore R, Sakzewski L, and Boyd R. A systematic review of the psychometric properties of quality of life measures for school aged children with cerebral palsy. *BMC Pediatrics* 2010; 10: 81.

12. Cremeens J, Eiser C, and Blades M. Characteristics of health-related self-report measures for children aged three to eight years: A review of the literature. *Quality of Life Research* 2006; 15: 739-754.

13. Davis E, Waters E, Mackinnon A, Reddihough D, Graham HK, Mehmet-Radji O, and Boyd R. Paediatric quality of life instruments: A review of the impact of the conceptual framework on outcomes. *Developmental Medicine & Child Neurology* 2006; 48: 311-318.

14. de Wit M, Delemarre-van de Waal HA, Pouwer F, Gemke RJ, and Snoek FJ. Monitoring health related quality of life in adolescents with diabetes: A review of measures. *Arch Dis Child* 2007; 92: 434-439.

15. Grange A, Bekker H, Noyes J, and Langley P. Adequacy of health-related quality of life measures in children under 5 years old: Systematic review. *J Adv Nurs* 2007; 59: 197-220.

16. Janssens L, Gorter JW, Ketelaar M, Kramer WL, and Holtslag HR. Health-related quality-of-life measures for long-term follow-up in children after major trauma. *Quality of Life Research* 2008; 17: 701-713.

17. Klassen AF, Stotland MA, Skarsgard ED, and Pusic AL. Clinical research in pediatric plastic surgery and systematic review of quality-of-life questionnaires. *Clin Plast Surg* 2008; 35: 251-267.

18. Maity S, Thomas AG. Quality of life in paediatric gastrointestinal and liver disease: A systematic review. *Journal of Pediatric Gastroenterology & Nutrition* 2007; 44: 540-554.

19. Pal DK. Quality of life assessment in children: A review of conceptual and methodological issues in multidimensional health status measures. *Journal of Epidemiology & Community Health* 1996; 50: 391-396.

20. Pickard AS, Topfer LA, and Feeny DH. A structured review of studies on health-related quality of life and economic evaluation in pediatric acute lymphoblastic leukemia. *J Natl Cancer Inst* 2004; Monographs: 102-125.

21. Quittner AL, Modi A, and Cruz I. Systematic review of health-related quality of life measures for children with respiratory conditions. *Paediatric Respiratory Reviews* 2008; 9: 220-232.

22. Ravens-Sieberer U, Erhart M, Wille N, Wetzel R, Nickel J, and Bullinger M. Generic health-related quality-of-life assessment in children and adolescents: Methodological considerations. *Pharmacoeconomics* 2006; 24: 1199-1220.

23. Solans M, Pane S, Estrada MD, Serra-Sutton V, Berra S, Herdman M, Alonso J, and Rajmil L. Health-related quality of life measurement in children and adolescents: A systematic review of generic and disease-specific instruments. *Value in Health* 2008; 11: 742-764.

24. Terwee CB, Mokkink LB, Knol DL,Ostelo RW,Bouter LM, - de VH. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2011; 10: .

25. Basaran S, Guler-Uysal F, Ergen N, Seydaoglu G, Bingol-Karakoc G, and Ufuk Altintas D. Effects of physical exercise on quality of life, exercise capacity and pulmonary function in children with asthma. *J Rehabil Med* 2006; 38: 130-135.

26. Terwee CB, Mokkink LB, van Poppel MN, Chinapaw MJ, van Mechelen W, and de Vet HC. Qualitative attributes and measurement properties of physical activity questionnaires: A checklist. *Sports Medicine* 2010; 40: 525-537.

27. Spanish version of the pediatric asthma quality of life questionnaire. *Quality of Life Research* 2001; 10: 81-91.

28. Asthma quality of life questionnaire. *Respir Med* 1998; 92: 1207-1214.

29. About my asthma. *Journal of Asthma* 1998; 35: 553-563.

30. Paediatric asthma quality of life questionnaire [pediatric asthma quality of life questionnaire]. *Pediatrics* 1997; 99(2: 165-168.

31. Paediatric asthma caregiver's quality of life questionnaire [pediatric asthma caregiver's quality of life questionnaire]. *European Respiratory Journal* 1997; 10: 2285-2294.

32. Asthma quality of life questionnaire. *Annals of Allergy, Asthma, and Immunology* 1997; 79: 229-233.

33. Pediatric asthma quality of life questionnaire [paediatric asthma quality of life questionnaire]. *Quality of Life Research* 1996; 5: 35-46.

34. Al-Akour N, Khader YS. Quality of life in jordanian children with asthma. *Int J Nurs Pract* 2008; 14: 418-426.

35. Annett RD, Bender BG, DuHamel TR, and Lapidus J. Factors influencing parent reports on quality of life for children with asthma. *J Asthma* 2003; 40: 577-587.

36. Annett RD, Bender BG, Lapidus J, Duhamel TR, and Lincoln A. Predicting children's quality of life in an asthma clinical trial: What do children's reports tell us?. *J Pediatr* 2001; 139: 854-861.

37. Annett RD, Bender BG, Skipper B, and Allen C. Predicting moderate improvement and decline in pediatric asthma quality of life over 24 months. *Qual Life Res* 2010; 19: 1517-1527.

38. Barley EA, Jones PW. A comparison of global questions versus health status questionnaires as measures of the severity and impact of asthma. *European Respiratory Journal* 1999; 14: 591-596.

39. Barnard KD, Speight J, and Skinner TC. Quality of life and impact of continuous subcutaneous insulin infusion for children and their parents. *Practical Diabetes International* 2008; 25: 278-284.

40. Busse WW, Casale TB, Murray JJ, Petrocella V, Cox F, and Rickard K. Efficacy, safety, and impact on quality of life of salmeterol in patients with moderate persistent asthma. *Am J Manag Care* 1998; 4: 1579-1587.

41. Chenhall RD, Senior K, Cole D, Cunningham T, and O'Boyle C. Individual quality of life among at risk indigenous youth in australia. *Applied Research in Quality of Life* 2010; 5: 171-183.

42. Coates A, Glasziou P, and McNeil D. On the receiving end--III. measurement of quality of life during cancer chemotherapy. *Ann Oncol* 1990; 1: 213-217.

43. Cornege-Blokland E, Jansen HE, De Jong-De Vos Van Steenwijk,C.C.E., and Poenaru D. Quality of life of children with spina bifida in kenya is not related to the degree of the spinal defects. *Tropical Medicine and International Health* 2011; 16: 30-36.

44. Elizabeth C, Suzanna S, Tim CF, Chi SL, Mital R, and Bee-Wah L. Pediatric asthma quality of life questionnaire: Validation in children from singapore. *Asian Pac J Allergy Immunol* 1999; 17: 155-161.

45. Erickson SR, Munzenberger PJ, Plante MJ, Kirking DM, Hurwitz ME, and Vanuya RZ. Influence of sociodemographics on the health-related quality of life of pediatric patients with asthma and their caregivers. *Journal of Asthma* 2002; 39: 107-117.

46. Flapper BC, Koopman HM, ten NC, and van der Schans CP. Psychometric properties of the TACQOL-asthma, a disease-specific measure of health related quality-of-life for children with asthma and their parents. *Chron respir dis* 2006; 3: 65-72.

47. Frick E, Borasio GD, Zehentner H, Fischer N, and Bumeder I. Individual quality of life of patients undergoing autologous peripheral blood stem cell transplantation. *Psychooncology* 2004; 13: 116-124.

48. Garcia-Marcos L, Carvajal Uruena I, Escribano Montaner A, Fernandez Benitez M, Garcia De La Rubia S, Tauler Toro E, Perez Fernandez V, and Barcina Sanchez C. Seasons and other factors affecting the quality of life of asthmatic children. *Journal of Investigational Allergology and Clinical Immunology* 2007; 17: 249-256.

49. Pediatric asthma quality of life questionnaire. *Pediatr Nurs* 2002; 28: 149-158.

50. Pediatric asthma quality of life questionnaire. *N Engl J Med* 1998; 339: 101.

51. Casale TB, Condemi J, LaForce C, Nayak A, Rowe M, Watrous M, McAlary M, Fowler-Taylor A, Racine A, Gupta N, Fick R, and Cioppa GD. Effect of omalizumab on symptoms of seasonal allergic rhinitis: A randomized controlled trial. *J Am Med Assoc* 2001; 286: 2956-2967.

52. Hamre HJ, Witt CM, Kienle GS, Schnurer C, Glockmann A, Ziegler R, Willich SN, and Kiene H. Anthroposophic therapy for asthma: A two-year prospective cohort study in routine outpatient settings. *Journal of asthma and allergy* 2009; 2: 111-128.

53. Gonzalez-Martin G, Joo I, and Sanchez I. Evaluation of the impact of a pharmaceutical care program in children with asthma. *Patient Educ Couns* 2003; 49: 13-18.

54. Goossens M, Laekeman G, Aertgeerts B, and Buntinx F. Evaluation of the quality of life after individualized homeopathic treatment for seasonal allergic rhinitis. A prospective, open, non-comparative study. *Homeopathy* 2009; 98: 11-16.

55. Hull SK, Page CP, Skinner BD, Linville JC, and Coeytaux RR. Exploring outcomes associated with acupuncture. *Journal of Alternative and Complementary Medicine* 2006; 12: 247-254.

56. Jolles BM, Bogoch ER. Quality of life after TKA for patients with juvenile rheumatoid arthritis. *Clin Orthop* 2008; 466: 167-178.

57. Lockey RF, DuBuske LM, Friedman B, Petrocella V, Cox F, and Rickard K. Nocturnal asthma: Effect of salmeterol on quality of life and clinical outcomes. *Chest* 1999; 115: 666-673.

58. Nogueira KT, Lopes CS. [The association between common mental disorders and quality of life in adolescents with asthma]. *Rev bras epidemiol* 2010; 13: 476-486.

59. Rhee JS, Book DT, Burzynski M, and Smith TL. Quality of life assessment in nasal airway obstruction. *Laryngoscope* 2003; 113: 1118-1122.

60. Riccioni G, Bucciarelli T, Di Ilio C, Corradi F, Della Vecchia R, Staniscia T, Mattei PA, and D'Orazio N. Brief communication: Is nasal polyposis a determinant of bronchial

hyperresponsiveness and altered quality of life in asthmatic subjects? A case-control study. *Ann Clin Lab Sci* 2006; 36: 170-173.

61. Rydstrom I, Dalheim-Englund A-, Holritz-Rasmussen B, Moller C, and Sandman P-. Asthma - quality of life for swedish children. *J Clin Nurs* 2005; 14: 739-749.

62. Sapp CJ. Adolescents with asthma: Effects of personal characteristics and health-promoting lifestyle behaviors on health-related quality of life. 2003; 120 p-120 p.

63. Trzcieniecka-Green A, Bargiel-Matusiewicz K, and Wilczynska-Kwiatek A. Quality of life and activity of children suffering from bronchial asthma. *Eur J Med Res* 2009; 14: 7.

64. Paterson C. Complementary practitioners as part of the primary health care team: Consulting patterns, patient characteristics and patient outcomes. *Fam Pract* 1997; 14: 347-354.

65. Ramstrom H, Erwander I, Mared L, Kornfalt R, and Seiving B. Pharmaceutical intervention in the care of cystic fibrosis patients. *J Clin Pharm Ther* 2000; 25: 427-434.

66. Guyatt GH, Juniper EF, Griffith LE, Feeny DH, and Ferrie PJ. Children and adult perceptions of childhood asthma. *Pediatrics* 1997; 99: 165-168.

67. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, and Townsend M. Measuring quality of life in the parents of children with asthma. *Qual Life Res* 1996; 5: 27-34.

68. Somerville A, Knopfli B, and Rutishauser C. Health-related quality of life in swiss adolescents with asthma. validation of the AAQOL-D and comparison with australian adolescents. *Swiss Med Wkly* 2004; 134: 91-96.

69. Ungar WJ, Mirabelli C, Cousins M, and Boydell KM. A qualitative analysis of a dyad approach to health-related quality of life measurement in children with asthma. *Soc Sci Med* 2006; 63: 2354-2366.

70. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, and Townsend M. Measuring quality of life in children with asthma. *Qual Life Res* 1996; 5: 35-46.

71. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, and Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments for asthma: Comparison of measurement properties. *Eur Respir J* 1997; 10: 2285-2294.

72. Juniper EF, Guyatt GH, Ferrie PJ, and Griffith LE. Measuring quality of life in asthma. *Am Rev Respir Dis* 1993; 147: 832-838.

73. Muntner P, Sudre P, and Perneger TV. Comparison of the psychometric properties of the asthma quality of life questionnaire (AQLQ) among 115 asthmatic adults assessed during acute hospitalization and as outpatients. *Quality of Life Research* 2000; 9: 987-995.

74. Oga T, Nishimura K, Tsukino M, Sato S, Hajiro T, and Mishima M. Comparison of the responsiveness of different disease-specific health status measures in patients with asthma. *Chest* 2002; 122: 1228-1233.

75. Paterson C. Measuring outcomes in primary care: A patient generated measure, MYMOP, compared with the SF-36 health survey. *BMJ* 1996; 312: 1016-1020.

76. Paterson C, Britten N. In pursuit of patient-centred outcomes: A qualitative evaluation of the 'measure yourself medical outcome profile'. *Journal of Health Services Research and Policy* 2000; 5: 27-36.

77. Raat H, Bueving HJ, de Jongste JC, Grol MH, Juniper EF, and van der Wouden JC. Responsiveness, longitudinal- and cross-sectional construct validity of the pediatric asthma quality of life questionnaire (PAQLQ) in dutch children with asthma. *Qual Life Res* 2005; 14: 265-272.

78. Reichenberg K, Broberg AG. Quality of life in childhood asthma: Use of the paediatric asthma quality of life questionnaire in a swedish sample of children 7 to 9 years old. *Acta Paediatr* 2000; 89: 989-995.

79. Ruta DA, Garratt AM, Leng M, Russell IT, and MacDonald LM. A new approach to the measurement of quality of life. the patient-generated index. *Med Care* 1994; 32: 1109-1126.

80. Sarria EE, Rosa RC, Fischer GB, Hirakata VN, da Rocha NS, and Mattiello R. Field-test validation of the brazilian version of the paediatric asthma quality of life questionnaire. *J Bras Pneumol* 2010; 36: 417-424.

81. Spiric VT, Bogic M, Jankovic S, Maksimovic N, Miljanovic SM, Popadic AP, Raskovic S, and Milic N. Assessment of the asthma quality of life questionnaire (AQLQ): Serbian translation. *Croat Med J* 2004; 45: 188-194.

82. Vinson J, Shank L, Thomas PD, and Warschausky S. Self-generated domains of quality of life in children with and without cerebral palsy. *Journal of Developmental and Physical Disabilities* 2010; 22: 497-508.

83. Wagner J. Acceptability of the schedule for the evaluation of individual quality of life-direct weight (SEIQoL-DW) in youth with type 1 diabetes. *Qual Life Res* 2004; 13: 1279-1285.

84. Yuksel H, Yilmaz O, Kirmaz C, and Eser E. Validity and reliability of the turkish translation of the pediatric asthma quality of life questionnaire. *Turk J Pediatr* 2009; 51: 154-160.

85. Ziora D, Madaj A, Wieckowska E, Ziora K, and Kozielski J. Correlation of spirometric parameters taken at a single examination with the quality of life in children with stable asthma. *Journal of Physiology and Pharmacology* 2007; 58: 801-809.

86. Wagner JA, Abbott G, and Lett S. Age related differences in individual quality of life domains in youth with type 1 diabetes. *Health and Quality of Life Outcomes* 2004; 2: .

87. Walker J, Winkelstein M, Land C, LewisBoyer L, Quartey R, Pham L, and Butz A. Factors that influence quality of life in rural children with asthma and their parents. *Journal of Pediatric Health Care* 2008; 22: 343-350.

88. van Gent R, van der Ent CK, Rovers MM, Kimpen JLL, van Essen-Zandvliet LEM, and de Meer G. Excessive body weight is associated with additional loss of quality of life in children with asthma. *J Allergy Clin Immunol* 2007; 119: 591-596.

89. Van Gent R, Van Essen LEM, Rovers MM, Kimpen JLL, Van Der Ent CK, and De Meer G. Quality of life in children with undiagnosed and diagnosed asthma. *Eur J Pediatr* 2007; 166: 843-848.

90. Walker J, Bradley C. Assessing the quality of life of adolescents with diabetes: Using the SEIQoL, DQoL, patient and diabetes specialist nurse ratings. *PRACT DIABETES INT* 2002; 19: 141-144.

91. Warschburger P, Busch S, Bauer CP, Kiosz D, Stachow R, and Petermann F. Health-related quality of life in children and adolescents with asthma: Results from the ESTAR study. *Journal of Asthma* 2004; 41: 463-470.

92. Wenzel SE, Lumry W, Manning M, Kalberg C, Cox F, Emmett A, and Rickard K. Efficacy, safety, and effects on quality of life of salmeterol versus albuterol in patients with mild to moderate persistent asthma. *Annals of Allergy, Asthma and Immunology* 1998; 80: 463-470.

93. Zandieh F, Moin M, and Movahedi M. Assessment of quality of life in iranian asthmatic children, young adults and their caregivers. *Iran j allergy asthma immunol* 2006; 5: 79-83.

94. Johnston BC, Shamseer L, da Costa BR, Tsuyuki RT, and Vohra S. Measurement issues in trials of pediatric acute diarrheal diseases: A systematic review. *Pediatrics* 2010; 126: e222-31.

95. The CONSORT group > home. Retrieved 9/24/2012, 2012, from http://www.consort-statement.org/.

96. The CONSORT group > CONSORT statement > 3-12 - methods > 6a - outcomes. Retrieved 9/24/2012, 2012, from http://www.consort-statement.org/consort-statement/3-12---methods/item6a_outcomes/.

97. Sinha I, Jones L, Smyth RL, and Williamson PR. A systematic review of studies that aim to determine which outcomes to measure in clinical trials in children. *PLoS Medicine / Public Library of Science* 2008; 5: e96.

98. STROBE statement: Available checklists. Retrieved 9/24/2012, 2012, from http://www.strobe-statement.org/index.php?id=available-checklists.

99. Guyatt GH, Kirshner B, and Jaeschke R. Measuring health status: What are the necessary measurement properties?. *J Clin Epidemiol* 1992; 45: 1341-1345.

100. PRISMA. Retrieved 7/20/2012, 2012, from http://www.prisma-statement.org/.

101. Ream E, Gibson F, Edwards J, Sepion B, Mulhall A, and Richardson A. Fatigue and quality of life diary. *Cancer Nurs* 2006; 29(4: 317-325.

102. WardSmith P, Hamlin J, Bartholomew J, and Stegenga K. Quality of life among adolescents with cancer. *Journal of Pediatric Oncology Nursing* 2007; 24: 166-171.

103. WardSmith P, McCaskie B, and Rhoton S. Adolescent-evaluated quality of life: A longitudinal study. *Journal of Pediatric Oncology Nursing* 2007; 24: 329-333.

104. McDonough P, Walters V, and Strohschein L. Chronic stress and the social patterning of women's health in canada. *Soc Sci Med* 2002; 54: 767-782.

105. Freyer G, Ligneau B, Schlumberger M, Blandy C, Contedevolx B, Trillet-Lenoir V, Lenoir GM, Chau N, and Dazord A. Quality of life in patients at risk of medullary thyroid carcinoma and followed by a comprehensive medical network: Trends for future evaluations. *Annals of Oncology* 2001; 12: 1461-1465.

106. McClimans LM, Browne J. Choosing a patient-reported outcome measure. *Theoretical Medicine & Bioethics* 2011; 32: 47-60.

107. Chen TH, Li L, and Kochen MM. A systematic review: How to choose appropriate health-related quality of life (HRQOL) measures in routine general practice? *Journal of Zhejiang University SCIENCE B* 2005; 6: 936-940.

108. Greenhalgh J, Long AF, Brettle AJ, and Grant MJ. Reviewing and selecting outcome measures for use in routine practice. *J Eval Clin Pract* 1998; 4: 339-350.

109. Patient-reported outcomes measurement - instrument selection - selection criteria. Retrieved 6/25/2012, 2012, from http://phi.uhce.ox.ac.uk/inst_selcrit.php .

110. Home :: Core outcome measures in effectiveness trials initiative (COMET) Retrieved 7/21/2012, 2012, from http://www.comet-initiative.org/ .

111. PROMIS. Retrieved 7/19/2012, 2012, from http://www.nihpromis.org/?AspxAutoDetectCookieSupport=1 .

112. Cosmin | cosmin. Retrieved 7/19/2012, 2012, from http://cosmin.nl/ .

Ishaque Thesis
Chapter 4

A Review of the Psychometric Properties of the Generic Individualized Health Related Quality of Life Measures (iHRQL) for Children

Sana Ishaque[1, 2]
Xiuyun Wu[1]
Jeffrey A. Johnson[1]
Charlotte Paterson [3]
Sunita Vohra[1,2,4]

[1]School of Public Health
University of Alberta, Edmonton, Alberta, Canada
[2] CARE Program for Integrative Health & Healing, University of Alberta, Edmonton, Alberta, Canada
[3]School of Social & Community Medicine, University of Bristol, The UK
[4]Faculty of Medicine and Dentistry and School of Public Health, University of Alberta

Corresponding author:
Sunita Vohra MD FRCPC MSc
Professor, Faculty of Medicine and Dentistry and School of Public Health, University of Alberta
Email: svohra@ualberta.ca

Acknowledgements: Authors are grateful to Dr. Caroline Terwee for her scholarly explanations of COSMIN checklist and measurement properties throughout this study.

Sana Ishaque is a MSc Clinical Epidemiology student at the School of Public Health and is a Graduate Research Assistant at Complementary and Alternative Research and Education (CARE) Program, Department of Pediatrics, University of Alberta. Sunita Vohra is a Health Scholar with Alberta Innovates Health Solutions (AIHS). Jeffrey Johnson is a Senior Scholar with AIHS and is a Centennial Professor at the University of Alberta.

Word count (excluding abstract, keywords, tables, and references): 2908

1 Abstract

***Introduction*:** Evaluation of health related quality of life (HRQL) in children requires application of child-specific HRQL measures, as adult HRQL measures are typically not appropriate in children. Individualized HRQL (iHRQL) measures are widely applied in children to assess primary and secondary outcomes. Our objective was to review the evidence-base and quality of the measurement properties of pediatric iHRQL questionnaires.

Methods: Six different databases were searched to identify generic iHRQL measures for children. Methodological quality (or risk of bias) was assessed using COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist and scoring system. The quality of measurement properties was assessed using modified Terwee criteria.

Results: Six studies on three iHRQL measures were included. Measurement properties reported in order of frequency were: content validity, construct validity, responsiveness, and reliability. Only two studies on content validity were assessed as excellent. Methodological quality ranged from poor to good for the rest of the properties.

Conclusion:

Pediatric generic iHRQL measures have been developed, but further validation is required before their use can be recommended.

Keywords: individualized, health related quality of life, pediatric, psychometric properties

2 Abbreviations

COSMIN: COnsensus-based Standards for the selection of health Measurement INstruments

iHRQL: Individualized health related quality of life

HRQL: Health related quality of life

MYMOP: Measure Yourself Medical Outcome Profile

PGI: Patient Generated Index

SEIQoL-DW: Schedule for the Evaluation of Individual Quality of Life Direct Weighting

3 Introduction

Greater attention is now given to quality of life of children with chronic diseases. Advances in medical research have remarkably affected the epidemiology of childhood illnesses[1,2]. Improvements in survival rates for many childhood conditions have resulted in increased long-term demands on health care[3]. It is important to deliver interventions that effectively address patients' needs and expectations. Arguably, the best therapy not only alleviates the patient's signs and symptoms, but also makes a significant difference in the patient's health related quality of life (HRQL).

There is no consensus on defining HRQL; for adults it may be defined as individual's subjective perception of the impact of their health status including disease and treatment, on physiological, psychological, and social functioning[4]. The same can be applied to children; however, it is important to note that underlying aspects of pediatric HRQL are different from adults. The physiological, psychological, and social functioning of children is different from adults and it also varies within different pediatric age groups. This difference is even more pronounced when adults' HRQL is compared to children[5-7]. Because of this, HRQL measures used in adults may not be appropriate to assess children.

Evaluation of HRQL in children requires consideration of their reading comprehension, their understanding of disease state, proxy versus self-report, time perception, and their rapidly changing developmental state[5]. Some of these issues can be satisfactorily

addressed if children (or appropriate proxies) are included in the development and evaluation of psychometric properties of self-report HRQL instruments [2].

Psychometric properties are not the [8] inherent properties of any measurement instrument, but rather of the data collected with application of the instrument. The measurement properties should be considered meaningful within the context of population and setting in which they were applied. There are three domains of psychometric properties: reliability, validity, and responsiveness [9]. Within the reliability domain, there are three subdivisions: internal consistency, reliability, and measurement error. The validity domain includes: content, construct, and criterion validity. The important psychometric properties depend on the purpose of the outcome measure[9-11]. Measurement tools are generally applied for three purposes: evaluation, discrimination, and prediction.

Traditional HRQL measures contain multiple dimensions, but contain standardized items/content for all respondents. Individualized HRQL (iHRQL) measures offer a different measurement approach, where patients can nominate and then score important aspects of their life from own perspective. These measures are commonly used for evaluation, thus their most important properties are responsiveness, and test-retest reliability[12]. The objective of this paper was to obtain an overview of the methodological quality of studies on the measurement properties of pediatric generic iHRQL questionnaires and to assess the quality of measurement properties themselves.

4 Methods

4.1 Criteria for studies for this review

All original studies that evaluated the measurement properties of generic iHRQL measures in pediatric or mixed adult and pediatric population, and published in English were included in this review.

4.2 Strategy for identification and selection of studies

All relevant studies were identified by searches in the electronic databases: MEDLINE (1950 to March 4th week 2011), EMBASE (1980 to 2011 week 13), PsycINFO (1985 to Week 4 2011), CINAHL (1982 to April 4, 2011), HAPI (March 30, 2011) CENTRAL (1st quarter 2011), DARE (2nd quarter 2011), ACP journal club (1991 to March 2011), Cochrane Database of Systematic Reviews (2005 to March 2011), Health Technology Assessment (2nd quarter 2011). The search strategy for MEDLINE is attached in Appendix 1. For other databases this strategy was modified to adapt to fit the databases' characteristics.

A study was included if it (1) was published in English, (2) included any child (0-18 years of age), even if mixed (i.e. adult and pediatric) population, (3) was about any intervention-control pair, (4) was full text original research that developed, adapted, or validated a generic iHRQL measure.

4.3 Data extraction and analysis

Two reviewers (SI, XW) independently extracted data on measurement properties, interpretability (percentage of missing items, percentages of respondents with highest and lowest possible scores, minimal important change/minimal important difference), and generalizability (median/mean age, gender distribution, setting in which study was conducted, country, language or instrument, sampling methods, and percentage of missing responses). Authors of included studies were contacted up to three times for clarification and to obtain available additional data on measurement properties.

4.4 Assessing risk of bias

Two reviewers (SI, XW) independently assessed risk of bias using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist[9-11,13]. Disagreements between the reviewers were resolved by consensus. The COSMIN checklist was developed by an international Delphi study in which consensus was reached on definitions of measurement properties; and standards were set for satisfactory design and statistical analysis of a study on measurement properties of health related patient reported outcomes[9,11]. The checklist consists of 10 boxes, and a three-step process to assess the methodological quality of studies evaluating the measurement properties of health status measurement instruments[9-11,13]. In the first step, COSMIN taxonomy and definitions are used to determine which measurement properties were evaluated in each included article. The corresponding COSMIN boxes are marked for each article. In step2 corresponding COSMIN boxes are completed for each article. In

step3 each item is scored on a 4-point rating scale i.e. “poor”, “fair”, “good” or “excellent”. An overall score for the methodological quality of a study is determined by taking the lowest rating of any of the items in a box.

A study on a measurement property is rated as having excellent quality if all relevant items on COSMIN are marked adequate. A study is rated as having ‘good’ methodological quality if some aspects of a measurement property are not reported, but one can deduce that these issues would have been adequate. A study is rated as having ‘fair’ quality if there are minor flaws in methods or statistical analysis. A study is rated as having ‘poor’ quality if there are major flaws in methods or statistical analysis. Sample size also affects the rating of a study for any particular measurement property. Details on 4-point rating of COSMIN checklist can be found on www.cosmin.nl [13].

Quality of measurement properties was assessed by the modified Terwee criteria[14,15]. The original Terwee quality criteria were published in 2007[14], the authors later modified it to be comparable to COSMIN; Appendix 2. The modified version has been used to evaluate the measurement properties of health measurement instrument in systematic reviews[16,17]. The possible overall rating for a measurement property is “positive” (+), “indeterminate” (?), “negative” (-), or “no information available” (0).

5 Results

Six studies concerning validation of three iHRQL measures on pediatric or mixed adult/pediatric populations [18-23] were included for risk of bias analysis (see Figure 1).

The included measures were: Schedule for the Evaluation of Individual Quality of Life Direct Weighting (SEIQoL-DW)[18,22,23], Measure Yourself Medical Outcome Profile (MYMOP)[19,20], and Patient Generated Index (PGI)[21].

Table 1 presents the general characteristics of these studies. Of the included measures, SEIQoL-DW was the only measure with reported validation studies on an exclusively pediatric population[22,23]. MYMOP[19,20] and PGI[21] had mixed adult/pediatric population studies only. For each measure, Table 2 and Table 3 present summary data on psychometric properties and risk of bias analyses respectively.

5.1 Content validity

Content validity was evaluated in five of the six included studies[19-23]. Of five, two studies evaluated SEIQoL-DW[22,23], another two evaluated MYMOP[19,20], and one evaluated PGI[21]. Methodological quality of one of the studies on MYMOP [19] and one on PGI[21] scored as excellent in terms of content validity. The other study on MYMOP scored good[20], because the purpose of the instrument was not clearly reported, but was assumed by the reviewers to be evaluative. Of the two studies that evaluated SEIQoL-DW, one scored fair [23] and the other scored poor [22]because it was not reported if all items were assessed for relevance or if all items comprehensively reflected the construct to be measured. Content validity was scored negative and indeterminate in the fair and poor quality studies respectively (Table 2, Table 3); it was scored positive for the rest of the studies.

5.2 Construct validity

Construct validity (hypothesis testing) was evaluated in three studies[18,19,21], one on each of the included measures. The study assessing construct validity of SEIQoL-DW compared it to the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30[18]. The *a priori* hypotheses were reported, and correlation statistics were calculated. However there was only one participant under 18 years of age included in the sample of 79. We therefore excluded this study from COSMIN analysis because of the limited relevant data to a pediatric population. Construct validity evaluation of MYMOP [19] ranked as fair because of absence of reported prior hypotheses tested. The construct validity of PGI was also ranked fair[21], because it was not clear how the missing items were handled.

5.3 Responsiveness

MYMOP was the only measure for which responsiveness was evaluated in one of the included studies[19]. The study on responsiveness of MYMOP was ranked fair, because prior hypotheses were not explicitly reported.

5.4 Reliability

Reliability was evaluated only for PGI[21]. The study was ranked fair, because it did not report how missing items were handled and Pearson correlation was reported.

6 Discussion

We identified exceedingly few generic iHRQL measures with reported validation studies in children. The three measures identified with applications in pediatric populations were: SEIQoL-DW, MYMOP, and PGI[18-23]. None of these have any data from pediatric studies on measurement error. Reliability was only reported for PGI[21]. Of six included studies, four were of mixed adult and pediatric populations[18-21]. Since they did not report subgroup results for children, their results cannot be applied to pediatric populations. The two studies that evaluated children exclusively both involved SEIQoL-DW[22,23], and reported only on content validity in terms of any measurement properties. In one of the included studies on MYMOP, parents completed the questionnaire for child participants, as a proxy assessment [19].

Evaluation of methodological quality on psychometric properties is essential to know the applicability of an instrument and to conduct further research. To our knowledge this is the first study that evaluated the methodological quality of generic iHRQL measures for children. While developing or conducting validation studies for a pediatric measure, additional important considerations beside psychometric properties include: developmental stage, respondent's level of comprehension, and respondent and administrator's burden[5,24]. One limitation of our study might be that we did not evaluate these qualitative features of included measures. Generally qualitative features (e.g. length, questionnaire format, and response options) are described to facilitate selection of HRQL measures for application. However, because of the limited evidence we identified regarding measurement properties, we feel such assessment is premature -

we do not recommend routine pediatric use of these measures before further validation is done. A checklist on qualitative attributes of physical activity questionnaires has recently been published[25]; further research to test measurement properties and applicability of the checklist to other patient-reported outcome measures is needed.

Absence of evidence on internal consistency and criterion validity should not be viewed as weaknesses of iHRQL measures as both of these measurement properties are not relevant to them. Internal consistency is the interrelatedness among the items[9-11]; since individualized measures do not have predetermined items, this may be considered relevant. While criterion validity has usually been evaluated against some external gold standard, the COSMIN group acknowledges the lack of "gold" standard for patient-reported outcomes and therefore suggests criterion validity only be assessed if a long version of a questionnaire is tested against its short version[9,11]. In all other cases, construct validity can be tested, where a priori hypotheses are set before comparison of "external gold" is done with a questionnaire.

Measurement error is the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured[9]. Minimum requirement to assess measurement error is to have at least two measurements taken under similar conditions with appropriate time interval[9]. The suggested statistic is Standard Error of Mean, Smallest Detectable Change, or Limits of Agreement[10,11]. We recommend that more studies be performed to assess measurement error of included measures. Reliability is the proportion of the total variance in the measurements, which is due to "true"

differences between patients[9]. It requires same study design and data as measurement error. COSMIN group's recommended reliability statistic for continuous scores is intra-class correlation coefficient. We recommend that more studies be performed to assess the reliability of included measures.

Content validity is the degree to which the contents of a measurement tool are an adequate reflection of the construct to be measured[9,10]. It is evaluated qualitatively by consulting experts and a convenience sample of the target population to provide their judgment about the relevance and comprehensiveness of the items. Given our results, further examination of the content validity of the included measures is required for them to be recommended for pediatric use.

In 2006, the U.S Food and Drug Administration published a guideline on patient-reported outcomes that required inclusion of patients (or target population) to achieve comprehensive content coverage, i.e. content validity[26]. The guideline was focused on measurement instruments intended for adults, not children. In pediatrics, the issue of *consulting the target population* needs consideration of children's linguistic and cognitive skills, and developmental stage. Preverbal (very young) or non-verbal (e.g. developmental delays or autism) children are unlikely to be able to take part in instrument development. In such situations, proxy involvement and/or use of observational measures are appropriate options. While evaluating studies of pediatric HRQL measures, reviewers should consider potential legitimate reasons for children not being consulted in the

instrument development; the current version of the COSMIN checklist does not consider this [9-11].

The issue of self versus proxy reporting is important in application of pediatric HRQL measures. Proxies sought to evaluate HRQL are generally parents (or primary caregivers), healthcare providers, and teachers[27,28]. However, there is evidence that the parent-child agreement about children's quality of life and functioning may not be optimal[29]. Self-report HRQL measures are preferred whenever possible[27], and proxy assessment should only be reserved where self-assessment is not possible because of the child's age or developmental issues; should this occur, proxy measures should be coupled with observational scales, which generally rely on external (visible) attributes and can have inter-rater reliability assessed to protect against observer bias.

In addition to age and developmental stage of children, proxy assessment may be required if the measures are long or written at a higher comprehension level such that children may not be able to complete them. A possible solution to this is to develop and validate face-based tools to measure individualized HRQL of children. Self-report face-based tools have been used to assess pain and nausea in children as young as three and seven years of age respectively[30,31]. To our knowledge, there is no generic individualized self-report measure of HRQL for children as of yet. Of the three identified measures, MYMOP has a reported face-based version [32,33]; this version could be further validated for pediatric use.

Our previous work (ch 3 SR) shows that SEIQoL-DW [34-38] and MYMOP [39] have been used as a primary outcome in children and mixed population studies. Assessment of the primary outcome is the objective of any research study; primary outcome measures that are not valid and reliable affect both the study's internal validity, resulting in erroneous conclusions, as well as its external validity (i.e.generalizability)[40,41]. The findings of a study using invalid outcome measures are likely to be incorrect and thus cannot be applied to general population. Studies that used MYMOP and SEIQoL-DW in children justified their choice of measure by citing adult studies or relied on studies evaluated in this paper, which we feel provide limited evidence of their pediatric validation[19,23,34]. It is important to have valid/appropriate measures for all study participants; if there are any children in a study, it is essential that valid child-specific measures be used to assess them.

We recognize assessing validity in pediatric populations is further challenged by the lack of consensus on the upper bound of what is "pediatric". For example, the American Academy of Pediatrics decided the upper age limit to be anywhere between 16 to 18 years in 1938 [42][43][44][43]. This limit was extended to 21 years in 1988[44]. Setting an age limit for *pediatric* is arbitrary from research point of view. We recommend i) authors be explicit about the upper and lower bounds of the target population of a measures, and ii) that the instrument be limited in use to the age group in which it has been tested.

For evaluation of HRQL (or any other patient-reported outcome), investigators need to consider sound psychometric properties of the instrument for the entire population under study. To rely on reported validation work, it is important to know if the validation study was conducted on children exclusively or mixed populations with adequate pediatric sample size and clearly reported subgroup analysis. The lack of validated disease-specific individualized HRQL measures available for pediatric research may be because of the large number of diseases which occur in infants, children, and youth, compounded by rapidly changing developmental stages. Development or adaptation and validation of face-based generic iHRQL measures, such as the MYMOP-Pictorial, may help overcome this gap.

7 Conclusion

No generic pediatric iHRQL tools are sufficiently validated to recommend their use in pediatric research. There is room for improvement in the methodological quality of identified tools. Given the complexity of measurement in children and issues such as proxy assessment, age, and developmental stage, we recommend that face-based individualized HRQL measures be developed and validated for pediatric research and clinical use.

Table 4-1: General Characteristics of included studies

Instrument	Study/ country	Underlying condition	Mean age (SD or range)	Setting	Language of instrument	Response rate	Measurement property evaluated
SEIQOL-DW	Frick 2004, Germany[18]	Participants undergoing autologous Peripheral Blood Stem Cell Transplantation (PBSCT)	(10-20),<18 n=1, 3% of study population; total n= 79	Hospital/ rehabilitation care	German	100%	Hypothesis testing
	Wagner 2004, The Netherlands[23]	Diabetes type 1	12.4 (2.5, 8-17)	Campers with diabetes	English	60%	Content validity
Modified SEIQoL-DW	Vinson 2010, The USA[22]	Cerebral palsy	CP group: 8.8 (1.8); TD: 8.9 (1.7)	School children of grade 1-6	English	100%	Content validity, interpretability
MYMOP	Paterson 1996, The UK [19]	N= 265, 218 general practice patients, 47 CAM patients; only six participants were <18 years of age	47 (17.6, 2-84)	Primary care (general practice and CAM providers)	English	At one month n=215 completed their 3rd MYMOP (81%), of 135 followed up for 4-months 109 (76%) completed f/u.	Content validity, construct validity, Responsiveness
	Paterson et al 2000, the UK [20]	New patients of 12 complimentary practitioners 176 patients completed MYMOP,	(16-86)	Primary care (general practice and CAM providers)	English	20 interviews were completed with six refusals	Content validity

		20 interviews					
PGI	Ruta 1994, Scotland [21]	Low back pain of varying severity	43 (16-85)	Primary care (general practitioners offices)	English	(571/777) 74%	Content validity, Hypothesis testing, Reliability, Interpretability

CAM: Complementary and alternative medicine, CP: cerebral palsy, TD: typically developing group

Table 4-2: Psychometric evidence for generic iHRQL measures for children

Instrument	Reliability (test-retest)	Content validity	Construct validity (SV, HT, CCV)	Interpretability
SEIQoL-DW[18]**			HT*	
SEIQoL-DW[23]		SEIQoL-DW was administered to 77 children. The SEIQoL-DW was administered among 77 campers with diabetes. 67 participants with valid data nominated total 335 domains, only 19 domains (6%) were nominated with the assistance of the standard list. Every respondent could nominate at least two domains without use of the list, and only a handful needed help with any of the additional three domains. All participants with valid data understood how to use the pie chart and were easily engaged in the task. 1/3 under 12 year participant (n=10) could not complete the questionnaire		
Modified SEIQoL-DW[22]		All children (n=101) generated at least two domains that were important to their quality of life. Ninety-one children (90%) generated five domains, as instructed		100% completion rate. Total mean scores TD: 65.98 (7.72), CP: 67.7 (7.27)
MYMOP [19]		Total 265 patients completed MYMOP. At one month 215 patients (18%) returned their third MYMOP questionnaire. Of 135 patients followed for	MYMOP's validity was supported by its ability to detect different degrees of change in relation to change scores in acute and	Change in MYMOP score at two weeks and four weeks showed a consistent gradient across the spectrum of clinical

		four months, 103 (76%) completed f/u. Practitioners reported that MYMOP was quick and easy to complete and was popular with patients.	chronic conditions, and by its correlation with SF-36	change.
MYMOP [20]		A qualitative study used a constant comparative method to analyze semi-structured interviews, which were then compared to the results of MYMOP questionnaires. Variable oriented analysis, and case oriented analysis was used. 20 interviews were conducted.		
PGI[21]	To test the reliability of the instrument, 167 patients were sent a second questionnaire; 111 patients (67%) completed and returned it. Of 111, 69 reported no change in health since completing the first questionnaire. Reliability coefficients were estimated for stage 1, stage 2, and final score of PGI; the values were 0.65, 0.75, and 0.70 respectively.	-PGI was tested initially in interviews with 20 patients. The measure was found comprehensible; all three stages could be satisfactorily completed with minimal prompting by the interviewer. -For self-completion, PGI was posted to 20 patients who were subsequently interviewed. As a result to these interviews, minor alterations were made to the wording, and the layout of the questionnaire was simplified. -Finally, the questionnaire was piloted to 74 patients suffering from back pain. Of the 49 questionnaires returned, 23 (47%) were correctly completed, 11 (27%) were partially completed, and 15 (31%) were returned blank. Some minor changes were make to the PGI afterwards.	HT: Seven a priori hypotheses were reported in the method section of the paper. Results of the study confirm majority of the a priori hypotheses	63% participants correctly completed the questionnaire, Mean PGI score was 33, distribution of score for 11 is reported in a bar chart. Percentage of respondents with the lowest and highest possible score were 11.96% (30/359), and slightly over 0% respectively. MIC or MID was not reported.

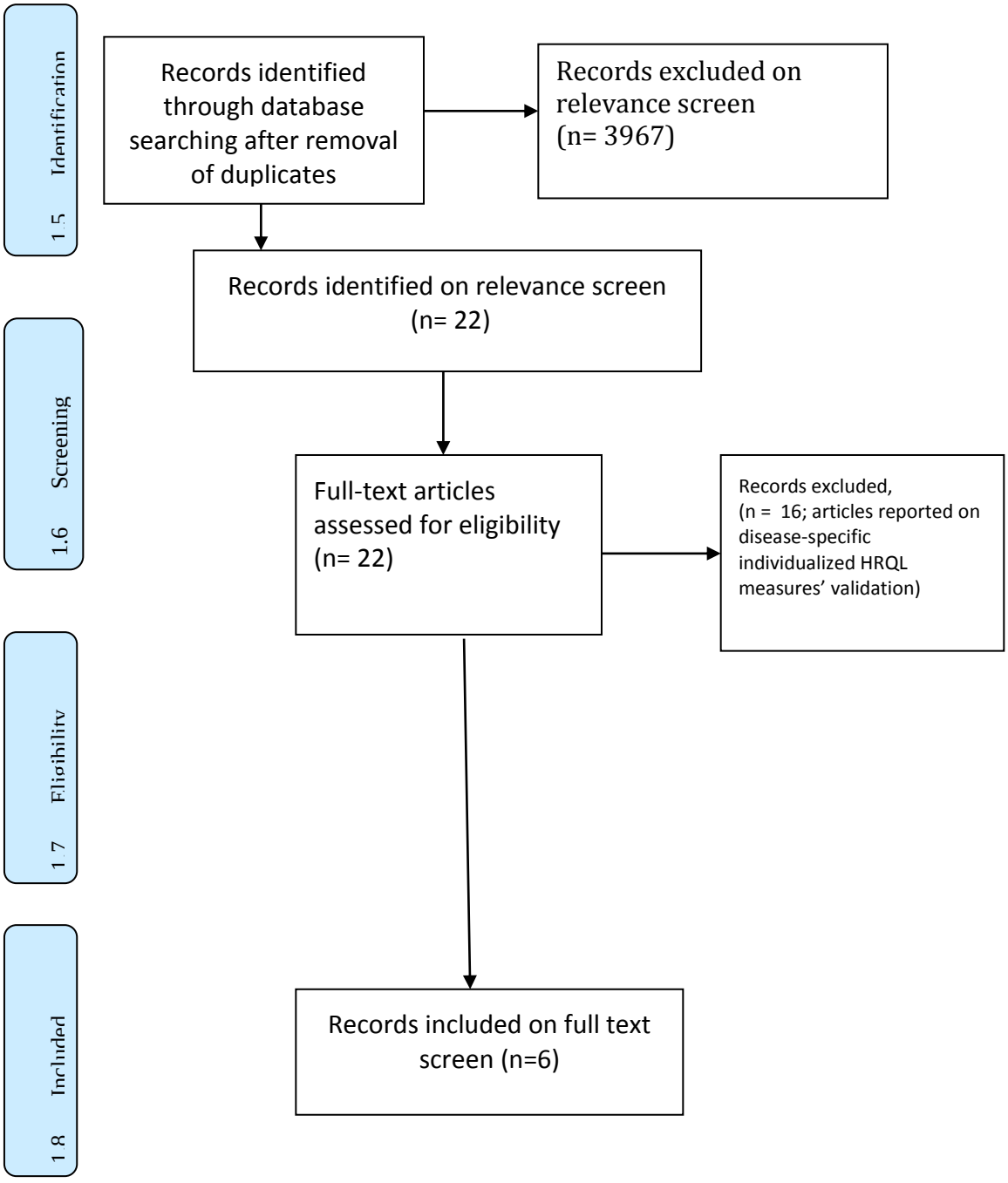
**excluded from analysis because there was only one participant <18 years
*Measurement properties excluded from the table as none of the included studies reported on them: Internal consistency, measurement error, criterion validity, and responsiveness

Table 4-3: Methodological Quality of Studies per measurement property

		Measurement properties							
Instrument	Study	Internal consistency	Measurement error	Reliability	Content validity	Construct validity (Hypothesis testing)	Criterion validity	Responsiveness	Interpretability
SEQOL-DW	Frick 2004					*			
	Wagner 2004				Fair/-				
	Vinson 2012				Poor/?				?
MYMOP	Paterson 1996				Excellent /+	Fair/+		Fair/+	
	Paterson 2000				Good/+				
PGI	Ruta 1994			Fair/-	Excellent /+	Fair/+			?

*excluded from COSMIN analysis, design does not comply with COSMIN descriptions

Figure 4-1: PRISMA 2009 Flow Diagram



APPENDIX 1

Medline March 28 2011

1. "Quality of Life"/
2. "Activities of Daily Living"/
3. Adaptation, Psychological/
4. depression/ or stress, psychological/
5. personal satisfaction/
6. Health Status/
7. happiness/
8. Anxiety/
9. (health status or functional status or personal satisfaction).ti,ab.
10. (quality of life or qol or hrqol).ti,ab.
11. (happy or unhappy or happiness or unhappiness or enjoy* or pleasur* or comfort* or optimis*).ti,ab.
12. ((activit* adj3 limit*) or (life* adj3 chang*) or function of satisfaction or well being).ti,ab.
13. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12
14. (patient centered or patient centred or patient generated).ti,ab.
15. (consumer generated or patient oriented or patient perspective*).ti,ab.
16. Self-Assessment/
17. (individuali?ed or personali?ed).ti,ab.
18. 14 or 15 or 16 or 17
19. 13 and 18
20. (mymop or seiqol or addqol or aqlq or womac or mycaw or psyclops or cmymop or w bq12).ti,ab.
21. patient generated index.mp.
22. (patient generated index or repertory grid or subjective quality of life profile or chronic respiratory disease questionnaire or disease repercussion profile or patient enablement instrument or well being questionnaire).ti,ab.
23. (measure yourself medical outcome profile or schedule for the evaluation of individual quality of life or audit of diabetes dependent qol).ti,ab.
24. (asthma quality of life questionnaire or "western ontario and mcmaster university osteoarthritis index" or "measure yourself concerns and wellbeing").ti,ab.
25. (pgi or pei or drp).mp. and 13
26. 20 or 21 or 22 or 23 or 24 or 25
27. 19 or 26
28. (child* or pe?diatric* or infant* or newborn* or neonat* or youth* or juvenile* or adolecen* or teenag*).mp.
29. 27 and 28

Appendix 2: Quality criteria for measurement properties (Based on Terwee et al 2007)

Property (definitions are based on COSMIN taxonomy)		Rating	Quality Criteria
Reliability: The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions	Internal consistency: The degree of the interrelatedness among the items	+	Sub)scale unidimensional AND Cronbach's alpha(s) ≥ 0.70
		?	Dimensionality not known OR Cronbach's alpha not determined
		-	(Sub)scale not unidimensional OR Cronbach's alpha(s) < 0.70
	Measurement error: The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured	+	MIC > SDC OR MIC outside the LOA
		?	MIC not defined
		-	MIC ≤ SDC OR MIC equals or inside LOA
	Reliability: The proportion of the total variance in the measurements which is due to 'true' differences between patients	+	ICC/weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80
		?	Neither ICC/weighted Kappa, nor Pearson's r determined
		-	ICC/weighted Kappa < 0.70 OR Pearson's r < 0.80
Validity: The degree to which an HR-PRO instrument measures the construct(s) it purports to measure	Content validity: The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured	+	The target population considers all items in the questionnaire to be relevant AND considers the questionnaire to be complete
		?	No target population involvement

			-	The target population considers items in the questionnaire to be irrelevant OR considers the questionnaire to be incomplete
	Construct validity: The degree to which the scores of an HR-PRO instrument are consistent with hypotheses	Cross-cultural: The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument	+ ? -	Original factor structure confirmed OR no important DIF Confirmation original factor structure AND DIF not mentioned Original factor structure not confirmed OR important DIF
		Structural: The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the	+ ? -	Factors should explain at least 50% of the variance Explained variance not mentioned Factors explain < 50% of the variance

		construct to be measured		
		Hypothesis testing: Idem construct validity	+ ? -	(Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses) AND correlation with related constructs is higher than with unrelated constructs; Solely correlations determined with unrelated constructs; Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR correlation with related constructs is lower than with unrelated constructs
	Criterion validity: The degree to which the scores of an HR-PRO instrument are an adequate reflection of a ‘gold standard’		+ ? - 0	Convincing arguments that gold standard is ‘‘gold’’ AND correlation with gold standard >0.70; No convincing arguments that gold standard is ‘‘gold’’ OR doubtful design or method; Correlation with gold standard <0.70, despite adequate design and method;

			0 No information found on criterion validity.
Responsiveness: The ability of an HR-PRO instrument to detect change over time in the construct to be measured		+ ? -	(Correlation with an instrument measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70) AND correlation with related constructs is higher than with unrelated constructs Solely correlations determined with unrelated constructs Correlation with an instrument measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlation with related constructs is lower than with unrelated constructs
Floor and ceiling effect: The number of respondents who achieved the lowest or highest possible score		+ ? -	<15% of the respondents achieved the highest or lowest possible scores; Doubtful design or method; >15% of the respondents achieved the highest or lowest possible scores, despite adequate

	0	design and methods; No information found on interpretation.
Interpretability : The degree to which one can assign qualitative meaning to quantitative scores	+	+ Mean and SD scores presented of at least four relevant subgroups of patients and MIC or MID defined;
	?	? Doubtful design or method OR less than four subgroups OR no MIC or MID defined;
	0	0 No information found on interpretation

[..] = reference number, MIC = minimal important change, ; MID = minimal important difference; SDC = smallest detectable change, LOA = limits of agreement, ICC = intraclass correlation coefficient,
DIF = differential item functioning, AUC = area under the curve
† + = positive rating, ? = indeterminate rating, - = negative rating, 0 = no information

8 References

1. Glendinning C, Kirk S, Guiffrida A, and Lawton D. Technology-dependent children in the community: Definitions, numbers and costs. *Child: Care, Health & Development* 2001; 27: 321-334.

2. Perrin JM. Health services research for children with disabilities. *Milbank Q* 2002; 80: 303-324.

3. Trends in disability prevalence and their causes: Proceedings of the national disability statistics and policy forum (4th, washington, DC, may 16, 1997). disability forum report 4. Retrieved 7/20/2012, 2012, from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED425606&ERICExtSearch_SearchType_0=no&accno=ED425606 .

4. Matza LS, Swensen AR, Flood EM, Secnik K, and Leidy NK. Assessment of health-related quality of life in children: A review of conceptual, methodological, and regulatory issues. *Value Health* 2004; 7: 79-92.

5. Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics* 1999; 16: 605-625.

6. Drotar D. Measuring child health: Scientific questions, challenges, and recommendations. *Ambul Pediatr* 2004; 4: 353-357.

7. Drotar D. Validating measures of pediatric health status, functional status, and health-related quality of life: Key methodological challenges and strategies. *Ambul Pediatr* 2004; 4: 358-364.

8. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research* 2002; 11: 193-205.

9. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HCW. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737-745.

10. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, and de Vet HC. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.

11. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HC. The COSMIN checklist for assessing the methodological quality of

studies on measurement properties of health status measurement instruments: An international delphi study. *Qual Life Res* 2010; 19: 539-549.

12. Launois R. Quality of life and pharmaco-economics in clinical trials - spilker,B. *QUALITY OF LIFE RESEARCH* 6: 97-98.

13. Cosmin | cosmin Retrieved 7/19/2012, 2012, from http://cosmin.nl/ .

14. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, and de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.

15. Schellingerhout JM, Heymans MW, Verhagen AP, de Vet HC, Koes BW, and Terwee CB. Measurement properties of translated versions of neck-specific questionnaires: A systematic review. *BMC Med Res Methodol* 2011; 11: 87.

16. Smith SC, Lamping DL, and Maclaine GD. Measuring health-related quality of life in diabetic peripheral neuropathy: A systematic review. *Diabetes Research & Clinical Practice* 2012; 96: 261-270.

17. Mutsaers JH, Peters R, Pool-Goudzwaard AL, Koes BW, and Verhagen AP. Psychometric properties of the pain attitudes and beliefs scale for physiotherapists: A systematic review. *Man Ther* 2012; 17: 213-218.

18. Frick E, Borasio GD, Zehentner H, Fischer N, and Bumeder I. Individual quality of life of patients undergoing autologous peripheral blood stem cell transplantation. *Psychooncology* 2004; 13: 116-124.

19. Paterson C. Measuring outcomes in primary care: A patient generated measure, MYMOP, compared with the SF-36 health survey. *BMJ* 1996; 312: 1016-1020.

20. Paterson C, Britten N. In pursuit of patient-centred outcomes: A qualitative evaluation of the 'measure yourself medical outcome profile'. *J Health Serv Res Policy* 2000; 5: 27-36.

21. Ruta DA, Garratt AM, Leng M, Russell IT, and MacDonald LM. A new approach to the measurement of quality of life. the patient-generated index. *Med Care* 1994; 32: 1109-1126.

22. Vinson J, Shank L, Thomas PD, and Warschausky S. Self-generated domains of quality of life in children with and without cerebral palsy. *Journal of Developmental and Physical Disabilities* 2010; 22: 497-508.

23. Wagner J. Acceptability of the schedule for the evaluation of individual quality of life-direct weight (SEIQoL-DW) in youth with type 1 diabetes. *Qual Life Res* 2004; 13: 1279-1285.

24. Ravens-Sieberer U, Erhart M, Wille N, Wetzel R, Nickel J, and Bullinger M. Generic health-related quality-of-life assessment in children and adolescents: Methodological considerations. *Pharmacoeconomics* 2006; 24: 1199-1220.

25. Terwee CB, Mokkink LB, van Poppel MN, Chinapaw MJ, van Mechelen W, and de Vet HC. Qualitative attributes and measurement properties of physical activity questionnaires: A checklist. *Sports Medicine* 2010; 40: 525-537.

26. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research. U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research. U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims: Draft guidance. *Health & Quality of Life Outcomes* 2006; 4: 79.

27. Eiser C, Morse R. Can parents rate their child's health-related quality of life? results of a systematic review. *Qual Life Res* 2001; 10: 347-357.

28. Cremeens J, Eiser C, and Blades M. Characteristics of health-related self-report measures for children aged three to eight years: A review of the literature. *Quality of Life Research* 2006; 15: 739-754.

29. Sattoe JN, van Staa A, Moll HA, and On Your Own Feet Research,Group. The proxy problem anatomized: Child-parent disagreement in health related quality of life reports of chronically ill adolescents. *Health & Quality of Life Outcomes* 2012; 10: 10.

30. Baxter AL, Watcha MF, Baxter WV, Leong T, and Wyatt MM. Development and validation of a pictorial nausea rating scale for children. *Pediatrics* 2011; 127: e1542-9.

31. Stinson JN, Kavanagh T, Yamada J, Gill N, and Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain* 2006; 125: 143-157.

32. Day A. The development of the MYMOP pictorial version. *Acupunct Med* 2004; 22: 68-71.

33. Day A, Kingsbury-Smith R. An audit of acupuncture in general practice. *Acupunct Med* 2004; 22: 87-92.

34. Wagner JA, Abbott G, and Lett S. Age related differences in individual quality of life domains in youth with type 1 diabetes. *Health and Quality of Life Outcomes* 2004; 2: .

35. Cornege-Blokland E, Jansen HE, De Jong-De Vos Van Steenwijk,C.C.E., and Poenaru D. Quality of life of children with spina bifida in kenya is not related to the degree of the spinal defects. *Tropical Medicine and International Health* 2011; 16: 30-36.

36. Barnard KD, Speight J, and Skinner TC. Quality of life and impact of continuous subcutaneous insulin infusion for children and their parents. *Practical Diabetes International* 2008; 25: 278-284.

37. Walker J, Bradley C. Assessing the quality of life of adolescents with diabetes: Using the SEIQoL, DQoL, patient and diabetes specialist nurse ratings. *PRACT DIABETES INT* 2002; 19: 141-144.

38. Chenhall RD, Senior K, Cole D, Cunningham T, and O'Boyle C. Individual quality of life among at risk indigenous youth in australia. *Applied Research in Quality of Life* 2010; 5: 171-183.

39. Hull SK, Page CP, Skinner BD, Linville JC, and Coeytaux RR. Exploring outcomes associated with acupuncture. *J Altern Complement Med* 2006; 12: 247-254.

40. Johnston BC, Shamseer L, da Costa BR, Tsuyuki RT, and Vohra S. Measurement issues in trials of pediatric acute diarrheal diseases: A systematic review. *Pediatrics* 2010; 126: e222-31.

41. Squires JE, Hutchinson AM, Bostrom AM, O'Rourke HM, Cobban SJ, and Estabrooks CA. To what extent do nurses use research in clinical practice? A systematic review. *Implementation Science* 2011; 6: 21.

42. Age limits of pediatrics. *Pediatrics* 1972; 49: 463.

43. Litt IF. Age limits of pediatrics, american academy of pediatrics, council on child health, pediatrics, 1972;49:463. *Pediatrics* 1998; 102: 249-250.

44. American academy of pediatrics council on child and adolescent health: Age limits of pediatrics. *Pediatrics* 1988; 81: 736.

Chapter 5: Summary, conclusions, and implications

This chapter summarizes the major findings of this thesis, and provides recommendations for future research and practice. To our knowledge, we are the first to critique iHRQL measures for pediatric application.

1. Objective of the thesis

The main objective of this thesis was to identify a valid and reliable generic individualized health related quality of life (iHRQL) measure for pediatric use. We started by evaluating an adult generic iHRQL measure (the MYMOP) and its adaptations, to see if a valid pediatric version existed. While we identified four adaptations, including a pictorial MYMOP that may be suitable for future pediatric use, there was insufficient evidence of pediatric validation. As such, we conducted a systematic review to identify all available iHRQL measures for children (0-18 years) and to assess quality of reporting in studies where HRQL was primary outcome. Our systematic review included studies on generic and disease-specific iHRQL measures; no restriction was placed based on reported psychometric properties (or validation status). Since our objective was to identify valid and reliable generic iHRQL measure for pediatric use, the third study included in this thesis was a risk of bias analysis to evaluate the validation (or development) of generic iHRQL measures in studies that included any child.

2. Summary of Methods

The first study involved the review of psychometric properties of MYMOP and its adaptations using modified Terwee quality criteria[1-3]. To identify studies for inclusion, adaptations of MYMOP were identified by personal communication with the instrument's developers, supplemented by a SCOPUS search.

The Terwee quality criteria were first published in 2007[1] to assist in evaluation of measurement properties of health status questionnaires; later the criteria were modified to match the international consensus on definitions, and divisions of measurement properties. The modified Terwee criteria were not published separately, but were applied in systematic reviews of measurement properties [2-4]. To assess the translation procedure of identified Chinese-MYMOP, Box G of the "COnsensus based Standards for the evaluation of health Measurement Instruments" (COSMIN) checklist was used[5-7].

The second study was a systematic review of iHRQL measures used in children. The review included both partial and/or fully individualized generic and disease specific measures regardless of their reported validation. Six common databases were searched from inception to 2011 for English language articles.

The third study was the risk of bias analysis of studies that reported development or evaluation of generic iHRQL measures for children or mixed adult and pediatric populations. To evaluate the methodological quality of included studies, the COSMIN

checklist was used[5-7]; and to evaluate their measurement properties, the modified Terwee criteria [1]were applied.

3. Main Results and Discussion

The MYMOP and its adaptations are individualized measures that allow respondents to measure their most important health issues from their own perspective. These tools may be used for evaluation of patients in clinical encounters. Four adaptations of MYMOP were identified in 16 publications; they were developed to evaluate individualized therapies in cancer, psychiatry, acupuncture, and the Chinese population. Content validity based on qualitative methods was the most frequently reported psychometric property of the included measures. Reported evidence was ranked/scored "positive," "indeterminate," or "negative" based on the level and response of the target population to the relevance and comprehensiveness of the items of included questionnaires [1]. Evidence on content validity was limited in two of the five studies due to unclear reporting on the target population's assessment of the contents' comprehensiveness and relevance. Construct validity was the second commonly reported measurement property, but was often not accompanied by reporting *a priori* hypotheses. The current recommendations for testing construct validity suggest that prior hypotheses about strength and direction of correlations should be reported; in the absence of advance hypotheses, there is risk of alternate explanations being sought[5,8]. At times, authors described "criterion validity", but we reclassified it as construct validity, based on modern day recommendations to limit criterion validity to the comparison of a long vs. short version of the same

questionnaire. Our reclassification of evidence on criterion validity did not affect the rating of the reported evidence. We did not identify satisfactory evidence regarding the measures' reliability, responsiveness, measurement error, or interpretability. Our findings on the imprecise use of definitions of measurement properties are consistent with a previous review by Mokkink et al[9].

In the second study, a systematic review identified 68 articles reporting on use or validation of eight iHRQL measures in children or mixed (children and adults) population. Of eight, five were disease-specific and three were generic iHRQL measures. The majority (67.6%) of the included studies were on the use of iHRQL measures; the identified instruments were used to assess primary outcome, secondary outcome, and to validate other measures. Only 32.4% reported on validation of iHRQL measures; validation studies were about six of the identified measures. No validation studies were identified for two of the five included measures (RQLQ, and PASI-pg). Thirty-two unique definitions of HRQL were identified from 27 included studies.

Our second objective was to assess the reporting of psychometric properties of identified measures in studies that used them for primary outcome assessment. Six iHRQL measures were used in thirty studies to assess primary endpoint; all of them cited literature to backup validation of their chosen outcome measure. We screened the references for their appropriateness (or applicability) to pediatric population; appropriate references were reported for only two measures (PAQLQ, and SEI-QoL), references for

the rest of the measures were based on adult or mixed (child and adult) populations. The mixed population references did not report any pediatric subgroup analysis.

In third study, six articles on three generic iHRQL measures were included. Of six, four studies were of mixed adult and pediatric populations. Since they did not report subgroup results for children, their results cannot be applied to pediatric populations. The two studies that evaluated children exclusively reported on SEIQoL-DW's content validity. Evidence on the content validity of SEIQoL-DW was limited because it was not reported if the target population considered all items to be relevant and comprehensive.

4. Limitations

This thesis has some identified weaknesses. First, our review of MYMOP and its adaptations may be limited because it was not a systematic review. Although each step was not independently duplicated, which can be a source of potential error; however, no important differences were found in the effect estimates between single and double data extraction [10]. Moreover, lack of independent seconding of the critical appraisal of the identified measures did not threaten this review's conclusions, which identified the absence of a pediatric MYMOP. Our search strategy was carefully developed and was comprehensive. First we identified the names of the tools by direct communication with the instrument developers, then we searched the identified instruments' websites. Finally a SCOPUS search was performed. Although we only included studies published in English, a Chinese tool was identified in the review, demonstrating the adequacy of our

search. To further strengthen our review, we used objective checklist criteria to evaluate the quality of measurement properties that enhances the reproducibility of our results.

Second, no risk of bias analysis was performed in the systematic review conducted to identify iHRQL measures for children. While the absence of a risk of bias analysis may be considered a weakness of systematic reviews of measurement properties [9], primary objective of our systematic review was to identify iHRQL measures used in children, hence we did not assess psychometric properties of identified measures in this review. Also, we did not evaluate the included studies with regard to effectiveness of any intervention and hence were not concerned about the potential risk of in estimates of treatment effect.

A third potential weakness of this review is inclusion of both iHRQL and individualized quality of life (QOL) measures when the actual protocol discussed (and intended) identification of iHRQL measures only. We recognize that HRQL and QOL are different; QOL includes every aspect of life, and HRQL includes those aspects of life that are directly related to health. We extracted the definition of HRQL or QOL whenever reported to help differentiate between the two.

Fourth, we did not include non-English language studies in the systematic review, suggesting it is possible that we may have missed a valid generic iHRQL measure applicable to children published in a language other than English. However, our intention

was to identify valid iHRQL measure for an English speaking population. If a valid iHRQL measure exists in non-English language, it would not be applicable to English speaking population without further validation work.

Fifth, with regard to COSMIN and Terwee analysis of identified generic iHRQL measures (chapter 4), we did not report on qualitative attributes of the identified measures. Qualitative attributes are generally reported to assist in selection outcome measures for research or clinical application. Given our results that the identified generic iHRQL measures have limited evidence of their pediatric application, reporting of qualitative attributes may be premature.

5. Implications for practice

With some limitations, MYMOP and its adaptations offer patient-centered assessment for adults using individualized therapies, such as those typically offered in complementary, alternative, or integrative medicine.

The only iHRQL measure with some evidence of content validity for child research is SEIQoL. The clinical application of SEIQoL is limited because of the lack of empirical evidence of its measurement properties in this setting, and its requirement for specialized equipment for direct weighting. Outcome measures used in clinical research may also be appropriate for use in a clinical setting, but require specific evaluation for such use Depending on their format, some tools may be easier to adapt for clinical use than others.

MYMOP has had some evaluation in clinical settings; further evaluation is recommended for this purpose, as its ease of use makes it amenable for clinicians.

6. Implications for research

6.1 MYMOP and its adaptations

Future research needs to:

1. Develop evidence of measurement error, and floor and ceiling effect of MYMOP and all adaptations.
2. Develop and report *a priori* hypotheses before checking for construct validity and responsiveness.

6.2 Systematic review of iHRQL measures for children

1. Find consensus on definition of health related quality of life for children.
2. Develop checklist through a consensus process, such as Delphi, to help researcher and clinicians choose valid and reliable age appropriate outcome measures for their study.
3. When selecting pediatric outcome measures, researchers should rely on validation studies based on child data; validation studies of mixed populations (adults and children) may only be relied upon if they have an adequate pediatric sample size, and if a pediatric subgroup analysis is reported.

6.3 COSMIN and Terwee evaluation of generic iHRQL measures for pediatric use

1. Researchers should explicitly report the upper and lower age limits of the target population for a measure
2. For other populations in general, and for pediatrics in particular, researchers should state developmental age limits, any developmental delays, and if children are preverbal.
3. In situations where self-report HRQL measures cannot be applied (e.g. preverbal, non-verbal children, comatose patients, developmental delays), observational scales should be used. Observational scales are instruments that measure externally observable (objective) features. These measures are generally completed by proxies such as parents or health care providers. In case of individualized proxy measurement of HRQL, the individuals completing the questionnaire may not agree on what constitutes important components of life that are related to health (HRQL) for the child in question. We recognize the challenge of proxy reporting and if no other alternative is appropriate, we suggest a standardized observational scale be used. If individualization is sought to reflect the heterogeneity in observable behaviors between individuals, this should likely supplement, rather than replace, a standardized observational tool. These scales must be evaluated for their inter-rater and intra-rater reliability.
4. The use of a given instrument or measure should be the age group in which it has been validated

5. Development of a valid and reliable face-based generic pediatric iHRQL measure is recommended to promote self-assessment, and decrease the reliance on proxy assessment wherever possible.
6. For development of pediatric HRQL measures, children should be consulted whenever their age and developmental state allows.

7. Conclusion

There is evidence that MYMOP and its adaptations offer individualized assessment for individualized therapies, such as those typically offered in complementary, alternative, or integrative medicine; further validation work is recommended. The second study identified five disease-specific and three generic individualized HRQL measures. Reported HRQL definitions were heterogeneous, and the majority of studies did not report relevant references of prior validation. Pediatric generic individualized HRQL measures have been developed, but further validation is required before they can be recommended for use. A face-based generic pediatric iHRQL is recommended for development and validation so as to maximize its utility across children of different ages and developmental abilities, and decrease reliance on proxy reporting.

8. References

1. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, and de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.

2. Mutsaers JH, Peters R, Pool-Goudzwaard AL, Koes BW, and Verhagen AP. Psychometric properties of the pain attitudes and beliefs scale for physiotherapists: A systematic review. *Man Ther* 2012; 17: 213-218.

3. Smith SC, Lamping DL, and Maclaine GD. Measuring health-related quality of life in diabetic peripheral neuropathy: A systematic review. *Diabetes Research & Clinical Practice* 2012; 96: 261-270.

4. Schellingerhout JM, Heymans MW, Verhagen AP, de Vet HC, Koes BW, and Terwee CB. Measurement properties of translated versions of neck-specific questionnaires: A systematic review. *BMC Med Res Methodol* 2011; 11: 87.

5. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HC. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international delphi study. *Qual Life Res* 2010; 19: 539-549.

6. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, and de Vet HC. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.

7. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HCW. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737-745.

8. Cosmin | cosmin Retrieved 7/19/2012, 2012, from http://cosmin.nl/ .

9. Terwee CB, Mokkink LB, Knol DL,Ostelo RW,Bouter LM, - de VH. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2011; 10: .

10. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. 2006; . Single data extraction generated more errors than double data extraction in systematic reviews.J Clin Epidemiol , 59: 7, 697-703.