NEWSLETTER
Patient Reported Outcomes
www.pro-newsletter.com
Number 46
Fall Issue 2011

Editorial

To celebrate the 20th anniversary of our PRO Newsletter, we gave carte blanche (our only request being that they mention this anniversary) to our long-standing partners, those who have been with us from the start, to share their thoughts on the history of our discipline. We've gathered their contemplations in the central supplement of this special-edition Newsletter as a way of sharing with you this important date in our history.

These narratives demonstrate the extraordinary development of what was originally designated as Quality of Life. Seen as utopian in the '70s and '80s, the research in this field has since established its validity and, under the name of "Patient-Reported Outcomes" (PRO), now plays an essential role in clinical research and the assessment of healthcare products.

20 YEARS SPECIAL EDITION

Today, PRO is an established element of the landscape of healthcare systems. Many people have participated in this journey, and the accounts shared here reflect the values and engagement of all.

In her contribution, Monika Bullinger describes objectively, clearly, and completely the evolution of the discipline. John Ware, Sergei Varshavsky, and Donald Patrick recount in their articles the groundswell of inventive undertakings to make Quality of Life an applied science. They discuss the remarkable development of the SF-36, the creation of companies specialized in the linguistic validation of questionnaires and the inexhaustible development of new questionnaires that address the continuously increasing specificity of issues in health. Finally, Dick Joyce, true to his independent nature, discusses how we may have wandered astray from the original ideals of our discipline. Together, their articles provide an interesting, stimulating and inspiring overview that we are happy to share with you, our loyal and fervent readers of the PRO Newsletter.

This special issue shows that since its origins as the "QoL Newsletter," our biannual publication has grown and matured. Might this also be a sign that the time has come to start writing new chapters in this continuing saga?

Your Editorial Board

INSTRUMENTS

The Revised FM Impact Questionnaire (FIQR): an Overview of its Development and Use in Recent Publications

Robert M. Bennett, MD, FRCP, MACR[1]; Ronald Friend, PhD[1,2]

[1] Fibromyalgia Research Unit, Oregon Health & Science University, 3455 SW Veterans Road, Portland, OR 97239, USA
[2] Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, USA

Abstract

The Revised FM Impact Questionnaire (FIQR) and its non-FM version, the Symptom Impact Questionnaire (SIQR), are recent major updates of the original FM Impact Questionnaire (FIQ) that was first published in 1991. Early experience with the FIQR/SIQR has been positive in terms of its properties, ease-of-use, translatability, and quickness of scoring. Herein we provide an overview of the FIQR/SIQR and its use in some recent publications.

KEYWORDS
FIQR, SIQR, FIBROMYALGIA SYMPTOM IMPACT QUESTIONNAIRE, QUESTIONNAIRE DEVELOPMENT, TRANSLATABILITY

Development of the FIQR

Fibromyalgia (FM) is the commonest cause of chronic musculoskeletal pain after osteoarthritis and low back pain. The original Fibromyalgia Impact Questionnaire (FIQ) was first published in 1991[1] and was based on the clinical experiences of the FM Treatment Team at Oregon Health & Science University (OHSU).

Its publication occurred just one year after the publication of the American College of Rheumatology Classification Criteria for FM in 1990,[2] an event that consolidated the previously disparate diagnostic schemes that were then in use.

During these some 20 years of use, several problems as regards its construct and scoring became apparent. For instance, the questionnaire was developed for use in patients seen at OHSU; this was predominantly a female Caucasian population living a typical American lifestyle. As the clinical spectrum of FM became more widely recognized, it became evident that the original seven symptom questions in the third domain needed to be expanded to include cognitive function, tenderness, balance, and overall sensitivity to environmental stimuli such as bright lights and loud noises.

Lastly, the scoring on the original FIQ was complicated by the need to use a ruler for the VAS measurement and unwieldy mathematical formulae to provide appropriate weighting to the domains and final scoring. With these issues in mind, we embarked on the development of an updated FIQ, subsequently called the Revised FM Impact Questionnaire (FIQR) in 2008; see the final version below.

The Symptom Impact Questionnaire (SIQR) has identical questions with the exception that the word "fibromyalgia" is replaced by "medical problems."

The Revised Fibromyalgia Impact Question (FIQR)

Domain 1: For each question, place an "X" in the box that best indicates how much difficulty you have experienced in doing the following activities during the past 7 days. If you did not perform a particular activity in the last 7 days, rate the difficulty for the last time you performed the activity. If you can't perform an activity, check the last box. If you can't perform an activity, check the last box.

Brush or comb your hair	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Walk continuously for 20 minutes	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Prepare a homemade meal	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Vacuum, scrub or sweep floors	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Lift and carry a bag full of groceries	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Climb one flight of stairs	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Change bed sheets	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Sit in a chair for 45 minutes	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult
Go shopping for groceries	No difficulty	□ □ □ □ □ □ □ □ □ □ □	Very difficult

Domain 2: For each of the following 2 questions, check the <u>one</u> box that best describes the overall impact of your fibromyalgia over the last 7 days:

Fibromyalgia prevented me from accomplishing goals for the week	Never	□ □ □ □ □ □ □ □ □ □ □	Always
I was completely overwhelmed by my fibromyalgia symptoms	Never	□ □ □ □ □ □ □ □ □ □ □	Always

Domain 3: For each of the following 10 questions, check the <u>one</u> box that best indicates the intensity of the following common symptoms over the last 7 days:

Please rate your level of pain	No pain	□ □ □ □ □ □ □ □ □ □ □	Unbearable pain
Please rate your level of energy	Lots of energy	□ □ □ □ □ □ □ □ □ □ □	No energy
Please rate your level of stiffness	No stiffness	□ □ □ □ □ □ □ □ □ □ □	Severe stiffness
Please rate the quality of your sleep	Awoke rested	□ □ □ □ □ □ □ □ □ □ □	Awoke very tired
Please rate your level of depression	No depression	□ □ □ □ □ □ □ □ □ □ □	Very depressed
Please rate your level of memory problems	Good memory	□ □ □ □ □ □ □ □ □ □ □	Very poor memory
Please rate your level of anxiety	Not anxious	□ □ □ □ □ □ □ □ □ □ □	Very anxious
Please rate your level of tenderness to touch	No tenderness	□ □ □ □ □ □ □ □ □ □ □	Very tender
Please rate your level of balance problems	No imbalance	□ □ □ □ □ □ □ □ □ □ □	Severe imbalance
Please rate your level of sensitivityto loud noises, bright lights, odors and cold	No sensitivity	□ □ □ □ □ □ □ □ □ □ □	Extreme sensitivity

Scoring: 1. Sum the scores for each of the 3 domains
2. Divide domain #1 score by three, divide domain # 2 score by one and divide domain score # 3 by two
3. Add the 3 resulting domains scores to obtain the total FIQR score (range is 0 to 100)

IN THIS ISSUE

Instruments

20 Years Special Edition

Testing and validation

The questionnaires were formatted for use on Survey Monkey (Portland, OR, USA), commercial online survey technology. In addition to the FIQR, the SIQR, the original questionnaire (FIQ) and the 36-Item Short Form Health Survey (SF-36) (Rand Corporation, Santa Monica, CA, USA) were also completed by all subjects. These surveys were completed by 202 FM patients, 51 patients with rheumatoid arthritis (RA) systemic lupus erythematosus (SLE), 11 patients with major depressive disorder (without concomitant FM), and 213 healthy controls. The focus group completed the FIQ and FIQR questionnaires in five different formats: the original paper version of the FIQ (FIQ-P); (2) an online version of the FIQ (FIQ-OL); (3) a paper version of the FIQR using 11 boxes scaled 0 to 10 (FIQR-P); (4) a paper version of the FIQR using a 100-mm VAS scoring (FIQR-P VAS); and (5) an online version of the FIQR (FIQR-OL). The online versions of the FIQR and FIQ were completed by the focus group at four weeks after completion of the paper versions.

Findings

The findings and analysis of these completed questionnaires were reported in detail in August 2009 in Arthritis Research and Therapy.[3] The major findings were:

1. For FM patients the mean total FIQR score was 56.6 ± 19.6 with a median score of 58. This compares with a total score of 60.6 ± 17.8 in the original FIQ ($P < 0.03$). Although this was a statistically significant difference, the change was only 4 points and the distribution curves were almost identical with a Shapiro-Wilk-skewness coefficient of 0.978 for the FIQR and 0.980 for the FIQ.
2. There was a good correlation between the total scores for the FIQR and the FIQ ($r = 0.88$, $P < 0.001$). Thus it should be possible to make reasonable comparisons between papers reporting the original FIQ and future papers reporting the FIQR; see figure 1.
3. The FIQR has adequate construct validity as shown by its strong correlation with the FIQ and good correlations with the SF-36. Indeed, multiple regression and partial correlation analyses showed that each of the three FIQR domains predicted corresponding construct-related SF-36 subscale domains.

OBITUARY

Donna Lamping (1953 – 2011)

Donna Lamping, who has died at age 58, was an international expert in the field of health psychology, health status, and quality of life assessment. Educated and trained in centers of excellence in Canada and the USA, she brought her cutting-edge knowledge to the UK in 1992. Despite her early death, her impact will be sustained by the cadre of young scientists she nurtured and developed over the past two decades.

Donna grew up in Toronto. After graduating in psychology from the University of Waterloo, she was awarded a prestigious doctoral fellowship from the Canadian Social Science & Humanities Research Council to study at Harvard University. Even at this early stage, her interest was as much on patient welfare as on methodological rigor, a combination that was to feature throughout her professional career. In her doctoral research she investigated how patients with chronic illness adapt and adjust to the inevitable stresses of their illness.

A research post at Harvard was followed by assistant professorships at McMaster University (Hamilton, Ontario, Canada), Fordham University (New York) and McGill University (Montreal). Throughout this period, her research focused on understanding the behavioral impact of chronic illness, initially considering patients with chronic kidney disease requiring hemodialysis but later shifting to the challenges facing those diagnosed with HIV and AIDS.

Her move to McGill was key to her subsequent career, as she left behind the disciplinary comfort of psychology and entered a multidisciplinary world, working with epidemiologists, statisticians, economists, and others. Here she had to demonstrate how essential psychology was to the understanding of health and disease. In turn, exposure to these other disciplines led her to what was to become her lasting concern, the measurement of patients' own perceptions of their condition.

During her time at McMaster, Donna met Itesh Sachdev, a social psychologist of language, who she married in Nepal in 1991. By then Itesh was at Birkbeck College (London) and the travails of sustaining an intercontinental relationship led Donna to join the staff of the London School of Hygiene & Tropical Medicine. She also had an appointment at the National Health Service that facilitated the practical application of newly developed measurement instruments. Apart from establishing psychometrics in the LSHTM, Donna made major managerial contributions both as Head of the Health Services Research Unit and in transforming the School's doctoral program.

Over the following two decades, Donna confirmed her position not only as a leader in the UK but also internationally. She was a long-standing member of the International Society for Quality Of Life Research (ISOQOL). She served as a member of its board from 2002-2005, on its executive committee from 2006 to 2008, and was its president in 2007. During her tenure as president Donna spearheaded a number of important initiatives designed to strengthen the society's infrastructure and to ensure its continuity and its growth.

Donna was delightful, intelligent, and indefatigable. Developing insulin-dependent diabetes in her forties was taken in her stride, something she was not going to allow to interfere with her zest for life, her optimism, and her extraordinary generosity to family, friends, colleagues and students. The same spirit was apparent when, in summer 2010, she discovered the bowel cancer to which she has succumbed. Perhaps her earlier studies of adjustment to illness had provided her with the insight to be able to remain so life-affirming until the end.

Donna is survived by her husband, Itesh, and her mother and sister, Helen and Gina.

Nick Black and Neil Aaronson

4. There was excellent internal consistency with a Cronbach's alpha of 0.95, indicating that the domain questions measure the same construct.
5. The correlations between all the various FIQR and FIQ formats completed by the focus group was >0.85 (P<0.001). These results were reassuring in that the use of numeric scoring was similar to a VAS format and that the online version behaved similarly to the paper version.
6. The goal of giving more weight to the first domain (function) compared to the other two domains appears to have achieved its desired effect. The imbalance between function and symptoms in the FIQ was reflected in weightings of 7% and 74% respectively. In the FIQR this ratio was 28% and 53% respectively; thus approximating the new weighting factors of 30% for function and 50% for symptoms that were adopted in the FIQR.
7. Test-retest reliability was not done in this original study apart from the 10 patients in the focus group who had test-retest reliability of 0.8. A subsequent study of a Turkish version of the FIQR reported test-retest of 0.83.[4]
8. The four new symptoms (memory, balance, tenderness, and environmental sensitivity) provided good discriminant validity between the FM group and the other three groups (healthy controls, RA/SLE, and MDD). The FM group had substantially higher scores for each individual symptom compared to subjects in the three control populations. All four new symptoms had strong correlations with the total FIQR score. There was a good discriminant validity between FM patients and the three other groups. The mean total score for the FIQR in the FM subjects (56.6) was significantly higher than the other three groups (healthy controls = 12.1, RA/SLE = 28.6, and MDD = 17.3). Thus a high total FIQR score has some diagnostic implications.
9. The aim of achieving an easily administered questionnaire with simplified scoring was realized. The average completion time for the FIQR in FM patients was 1.3 ± 0.02 minutes compared with 2.1 ± 0.03 minutes for the FIQ. The SF-36 took nearly four times as long to complete (4.1 ± 0.04 minutes). The time taken for scoring the FIQR was approximately one minute; this compares to a scoring time of approximately five minutes for the FIQ.

Results from recently published studies

As of writing this review in 2011, the FIQR has not yet been extensively used. There are five papers referencing the FIQR, namely a Turkish translation, its use as an outcome measure in a yoga study, a Spanish epidemiological study on suicidal ideation in FM, an objective evaluation study of balance problems, and an FM diagnostic usefulness study. These papers are briefly reviewed here:

Turkish translation of the FIQR

A Turkish translation of the FIQR along with its operating characteristics and validation was published in October 2010.[4] The Turkish FIQR was administered to 87 female (age 34.3 ± 10.2) patients along with the original FIQ and the short form SF-36 on two occasions one week apart. The total FIQR score on the first visit was 55.22 ± 21.96 and 57.16 ± 22.48 on the second visit; this provides a test-retest reliability of 0.835. The total scores of the FIQR and the FIQ were closely correlated (r = 0.87, P. <0.01); this correlation is similar to the correlation coefficient of 0.88 reported in our original paper.[3] Comparisons with the short form SF-36 showed significant inverse correlations with the physical component summary (PCS) and the mental component summary (MCS) with correlation coefficients of -0.63 and -0.51 respectively. Cronbach's alpha was 0.89 at the first visit and 0.91 and the second visit. The authors noted that the FIQR was much easier to translate into Turkish than the original FIQ, mainly due to the generalizability of the questions in the first domain. The mean time for completion of the Turkish FIQR was 2.4 minutes at the first visit and 2.2 minutes at the second visit. An average scoring time of about one minute was similar to that reported in the original paper.[3] It was concluded that "*the Turkish version of the FIQR is a reliable and valid instrument for the assessment of disease severity in FM. It may be used easily for both clinical practice and research use in the Turkish-speaking population in place of FIQ.*"

Yoga study

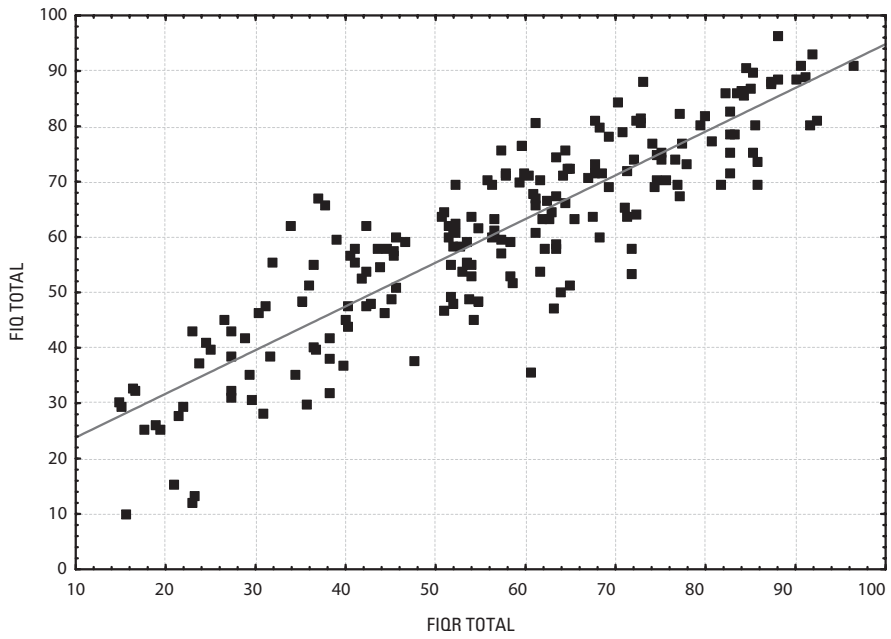
A study in which the FIQR was used as the major outcome variable in a FM treatment study has provided the first indication that the total FIQR and subscales are sensitive to change.[5] This study published in October 2010 reported the effect of Yoga of Awareness in 53 female FM patients who were randomized to the eight-week yoga program or to wait-listed standard of care control group. The total FIQR score improved by 12.83 in the yoga group compared to 0.57 in the control group (P. <0.0003) for a Cohen's effect size of 0.72. While the first domain of the FIQR (function) did not show statistical improvement, both the second domain (overall impact) and the third domain (symptoms) showed significant improvements compared to control (P<0.007). All 10 of the third domain symptoms, except for sleep, showed significant improvement. The original FIQ has shown good sensitivity to change in multiple studies[6] and it can be expected that the FIQR will have similar performance characteristics; however definitive verification of the FIQR's usefulness as an outcome measure in FM must await further studies.

Spanish suicidal ideation study

A July 2011 e-publication from Spain on suicide attempts and suicidal ideation in FM patients used a Spanish translation of the FIQR.[7] It was commented that *"Although the FIQR has not been yet validated in Spanish, the questionnaire showed an excellent internal consistency, with a Cronbach's alpha of 0.95, which is identical to that obtained in the original validation of the English version of the FIQR."* The total FIQR score in the 108 Spanish female FM patients who returned the questionnaire was 66.4 ± 20.7; this is some 10 points higher the values reported in our original paper and in the Turkish translation. This may well be related to a survey completion bias, in that those patients with suicidal ideation might have been more likely to return the completed survey. Indeed, there was a direct correlation between the total FIQR score and the Plutchik scale of suicide ris (r = 0.677, P. <0.0001). This study is another validation that the FIQR measures the severity of FM.

Figure 1.

Scatterplot of FIQ total against FIQR total



Legend: A scatter-plot of the total score for the Revised Fibromyalgia Impact Questionnaire (FIQR) and the Fibromyalgia Impact Questionnaire (FIQ) from 202 fibromyalgia patients (r = 0.88, P <0.001). *Reproduced from Bennett et al*[3]

Balance study

An August 2011 e-publication reported on objective evidence of balance problems in 26 FM subjects (88% female) and 27 healthy controls (88% female), using Computerized Dynamic Posturography (Neurocom, International, Inc.).[8] This apparatus evaluates sensory (Sensory Organization Test) and motor (Motor Control Test) systems and is currently the gold standard for balance assessment.[9] The main outcome of FM severity in this study was the FIQR. As in other studies employing non-FM patients, the SIQR was used as a surrogate for the FIQR. The total FIQR score was 54.06 ± 17.8 and the SIQR score was 5.00 ± 4.6. These total FIQR scores are very similar to the original study and the Turkish translation study. Cronbach's alpha was 0.976 with an inter-item correlation of 0.657. The total FIQR score predicted the primary posturography balance measure (the SOT composite score). Importantly, this study provided a more rigorous validation the FIQR question on balance ("*please rate your level of balance problems"),* as this question (#20 in FIQR) correlated with the dynamic posturography SOT score at r= -0.61, P < 0.0001.

Diagnostic usefulness

An April 2011 publication reported on the usefulness of selected questions in the FIQR and SIQR in distinguishing systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA) from FM.[10] The largest FIQR/SIQR differences between FM and RA/SLE were questions relating to "tenderness to touch," "difficulty cleaning floors," and "discomfort sitting for 45 minutes." When combined with pain locations, it was found that two FIQR/SIQR questions (*"tenderness to touch"* and *"difficulty sitting for 45 minutes")* plus pain in the low back, neck, hands, and arms provided the correct diagnosis in 97% of the subjects with only seven of 253 subjects being misclassified. It would appear that a high SIQR score (>45) along with high scores on "tenderness to touch" and "discomfort on prolonged sitting" should alert healthcare providers to a diagnosis of FM.

Conclusions

The FIQR and its non-FM version, the SIQR, are only just starting to be used in articles reported in peer-reviewed journals. So far there has only been one validated translation, the Turkish FIQR, but an unvalidated Spanish version of the FIQR was used in the suicidality study. Currently, several other translations are underway (Brazilian Portuguese, German, Japanese, and Chinese). All studies, so far reported, have confirmed that the FIQR correlates well with various other measures of FM severity. Finally there is some evidence that selected questions from the FIQR/SIQR may prove useful in designing epidemiological questionnaires looking at the prevalence of musculoskeletal pain.

For more information, please contact:
Robert Bennett, MD, FRCP, FACP, MACR
Oregon Health&Science University
Professor of Medicine and Nursing
Tel: 503-246-6944
Email: bennetrob1@comcast.net

REFERENCES

1. Burckhardt CS, Clark BD, Bennett R.M. The fibromyalgia impact questionnaire: development and validation. *J Rheumatol* 1991;18:728-733.
2. Wolfe F, Smythe HA, Yunus MB et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990;33(2):160-172.
3. Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. *Arthritis Res Ther* 2009;11(4):R120.
4. Ediz L, Hiz O, Toprak M, Tekeoglu I, Ercan S. The validity and reliability of the Turkish version of the Revised Fibromyalgia Impact Questionnaire. *Clin Rheumatol* 2010.
5. Carson JW, Carson KM, Jones KD, Bennett RM, Wright CL, Mist SD. A pilot randomized controlled trial of the Yoga of Awareness program in the management of fibromyalgia. *Pain* 2010;151(2):530-539.
6. Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. *Clin Exp Rheumatol* 2005;23(5 Suppl 39):S154-S162.
7. Calandre EP, Vilchez JS, Molina-Barea R et al. Suicide attempts and risk of suicide in patients with fibromyalgia: a survey in Spanish patients. *Rheumatology (Oxford)* 2011.
8. Jones KD, King LA, Mist SD, Bennett RM, Horak FB. Postural control deficits in people with fibromyalgia: a pilot study. *Arthritis Res Ther* 2011;13(4):R127.
9. Ben Achour LS, Missaoui B, Miri I, Ben Salah FZ, Dziri C. Role of the Neurocom Balance Master in assessment of gait problems and risk of falling in elderly people. *Ann Readapt Med Phys* 2006;49(5):210-217.
10. Friend R, Bennett RM. Distinguishing fibromyalgia from rheumatoid arthritis and systemic lupus in clinical questionnaires: an analysis of the revised Fibromyalgia Impact Questionnaire (FIQR) and its variant, the Symptom Impact Questionnaire (SIQR), along with pain locations. *Arthritis Res Ther* 2011;13(2):R58.

NEWS FROM...

Patient-Reported Outcomes Research in Chinese Medicine

Feng-bin Liu, Guangzhou University of Chinese Medicine, Guangzhou, China
Zheng-Kun Hou, Guangzhou University of Chinese Medicine, Guangzhou, China
Hong-Mei Wang, Zhejiang University, Hangzhou, China and University of Washington, Seattle, WA, USA

The fields of Chinese Medicine (CM) and Patient-Reported Outcomes (PRO) have similar theoretical and clinical underpinnings in that both focus on peoples' physical, psychological, social, and environmental characteristics and their mutual relationships. However, traditional CM holds its own unique view of PRO concept definition, assessment, and application. During its thousands of years of development, CM practitioners collected PRO data based on individual qualitative perceptions from patients and physicians without development of quantitative research methods.

The modern PRO was not introduced into CM until 1988.[1] Liu et al. developed the first CM-PRO instrument, *Quality of Life Questionnaire for Gastric-Intestinal Patients* in 1997.[2] Since 2000, research in CM PRO has developed rapidly, including the Chinese Quality of Life (QOL) measures (ChQOL),[3] the Health Status Scale of TCM (TCM-HSS),[4] the Spleen and Stomach Disease PRO scale (SSD-PRO) (The Spleen and Stomach diseases in CM are equivalent to the gastroenteric diseases mainly),[5] the TCM Stroke Scale for QOL Measurement,[6] the QOL scale for CAD,[7] and the Myasthenia Gravis patients of PRO scale (MG-PRO).[8] These instruments include generic and disease-specific types, and self and interviewer-administered versions. They can be used in the general population, and in patients with common or severe diseases, such as spleen and stomach disease, liver cirrhosis, hepatitis B, chronic obstructive pulmonary disease, pneumonia, stroke, rheumatoid arthritis, eczema, etc.

CM-PRO researchers in mainland China are paying close attention to emerging PRO technologies for application in clinical settings. For instance, computer-adaptive testing and Item Response Theory were used in the development and evaluation of the TCM-HSS[4] and MG-PRO.[8] The ChQOL[3] has been used in clinical trials and included in some instrument databases, and adapted into English, Italian, and Hong Kong Chinese versions. In addition, development of an Australian version and short version are in progress. Guidelines which aim to enhance the quality of CM-PRO research including instrument development, validation, adaptation, application, and interpretation are also being developed.[9]

It is expected that CM-PRO research in mainland China will continue to make significant progress though ongoing work in concept definition and instrument development in the context of Chinese Medicine. Future funding opportunities would facilitate relevant research and application.

1. Yang YC, He MT, Zhu CM, et al. Study on the relationship between happiness and mental health issues in elderly. *Zhongguo Xin Li Wei Sheng Za Zhi* 1988;2:9-12.
2. Liu FB. Study on the quality of life of the patients with gastroenteric diseases. *Guangzhou Zhong Yi Yao Da Xue Xue Bao* 1997;14:225-8.
3. Leung KF, Liu FB, Zhao L, et al. Development and validation of the Chinese Quality of Life Instrument. *Health Qual Life Outcomes* 2005;3:26.
4. Liu FB, Lang JY, Zhao L, et al. Development of Health Status Scale of Traditional Chinese Medicine(TCM-HSS). *Journal of SUN YAT-SEN University* (Medical Sciences) 2008;29:332-6.
5. Liu FB, Wang WQ. Establishment of the spleen-stomach patients reported outcomes scale in Chinese medicine and the corresponding item selection. *World Science and Technology* (Modernization of Traditional Chinese Medicine and Materia Medica) 2009;11:527-30.
6. Li H, Liang WX. Study and development of TCM Stroke Scale for quality of life measurement (1) - establishment of the scale. *Liaoning Zhong Yi Za Zhi* 2008;35:376-8.
7. Zhu T, Mao JY. The development and evaluation of Qol scale with TCM characteristics to patients with CAD. *Liaoning Journal of Traditional Chinese Medicine* 2008;35:854-5.
8. Chen XL, Liu FB, Guo L. Development of patient-reported outcome scale for myasthenia gravis: a psychometric test. *Journal of Chinese Integrative Medicine* 2010;8:121-5.
9. Hou ZK, Liu FB, Li XY, et al. Advertising on Preferred Reporting Items for Patient-Reported Outcome Instrument Development: The PRIPROID(C). *Proceedings of 2nd International Conference on Information Science and Engineering (ICISE2010).* Hangzhou: The Institute of Electrical and Electronics Engineers (IEEE), 2010:6434-

For more information, please contact:

Professor Feng-bin Liu at liufb163@163.com

Hongmei Wang, PhD, MPH
Visiting Scholar, SeaQoL
University of Washington, Health Services
Box 359455
Seattle, WA 98195-9455
Tel: 206-543-9932; Fax: 206-616-3135

SURVEY

PRO Newsletter Survey

The PRO Newsletter Editorial Board

Too many questionnaires or not the right PROs?

Results from the recent survey conducted among the PRO Newsletter readers.

As announced in the last issue of the PRO Newsletter, an electronic survey was conducted among our readers. This survey had two aims: firstly, to document the current use of PRO questionnaires by researchers and clinicians; secondly, to assess their level of satisfaction with existing questionnaires, and identify their priorities regarding the need for new measurement tools. The main results are described below.

Some interesting figures emerge from these very simple questions. The main surprise concerns the variety of PRO questionnaires used: we knew about the high variety of existing questionnaires, but we would not have hypothesized the absence of clear leadership by a shortlist of well-known generic and specific questionnaires. This finding reflects the eclecticism of our readers in their research and clinical work. In other words, the list of questionnaires used by our readers does not at all reflect the hypothesized existence of standards in our field. Less of a surprise is the confirmation by both researchers and clinicians that this huge variety of questionnaires does not fulfill their needs: about two-thirds of respondents feel the PRO measures they currently have for their work are not satisfactory. Interestingly, when asked about PRO questionnaires they would like to have available, readers' interest and priorities were not concentrated on a dominant concept, but included functioning, symptoms, quality of life, adherence, activities of daily living, and patient satisfaction, at comparable levels.

Researchers and Pharmaceutical Industry: results

Forty-four percent of pharmaceutical industry employees, 32% of clinical researchers, 36% other and a few CRO employees (7%) completed the survey. The respondents cited interest in all therapeutic areas (see Figure 1).

A vast majority of respondents (86%) use PROs in their research work.

More than 150 different questionnaires are being used by respondents, a large majority of which are disease-specific questionnaires.

More than 50% of responders cited at least three different questionnaires.

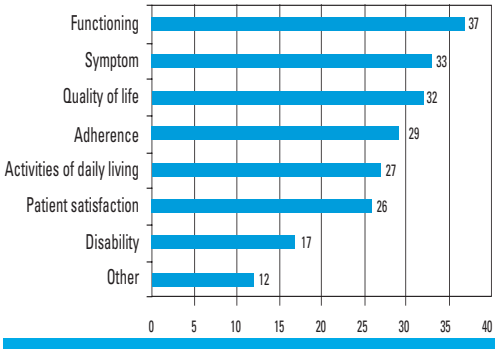
Sixty percent of respondents reported that the available questionnaires do not fully address their needs in a number of domains (see Figure 2). A large variety of diseases areas were cited, reflecting the variety in the research priorities of respondents.

Clinicians: results

Few clinicians completed the survey. However, their responses were similar to those of researchers.

Figure 2.

Type of assessment needed



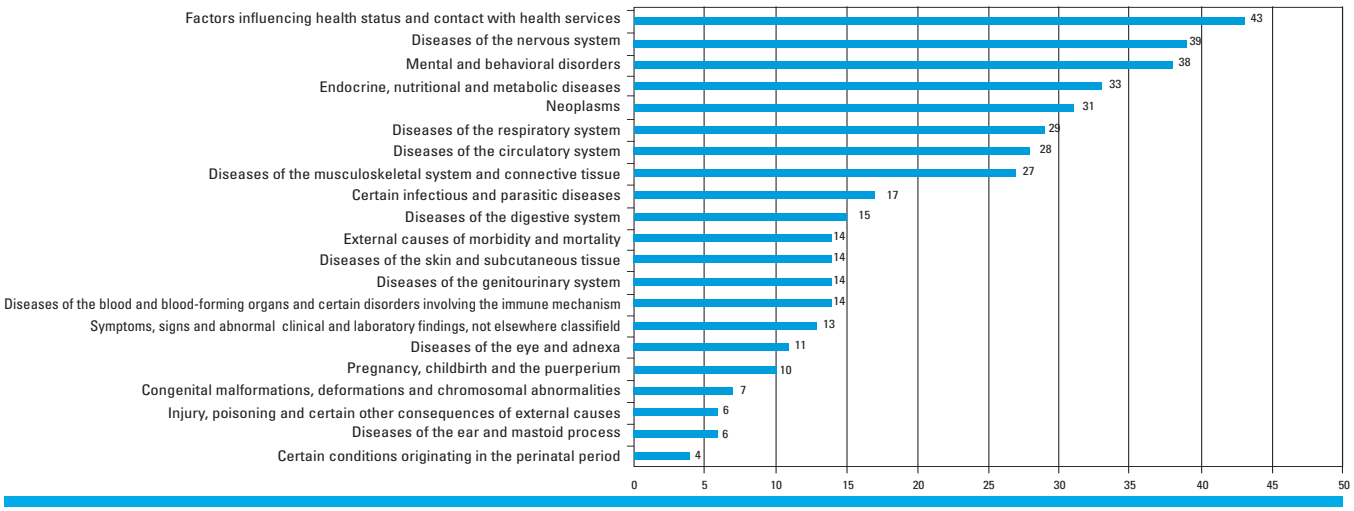
A majority (70%) of clinicians using questionnaires in clinical practice were not satisfied with the questionnaire(s) they used. They claim to need additional Patient-Reported Outcomes questionnaires for the following domains: symptoms, functioning, disability, activities of daily life, quality of life, patient satisfaction, and adherence.

Conclusion

The past decades have seen the development of two opposed strategies, both pretending to address the needs of researchers and clinicians: the multiplication of specific tools on one hand, and the consolidation of standards on the other. This modest survey suggests that neither of these two strategies have succeeded: researchers and clinicians express the need for a new generation of PRO standards.

Figure 1.

Therapeutic areas of interest



NEWS FROM...

Turkish National HRQOL Society (SAYKAD)

Prof. Dr. Erhan Eser

SAYKAD was founded in 2002 in Izmir, the third biggest city of Turkey, by a group of medical doctors working in the area of public health, psychiatry, and medical oncology. The founders of SAYKAD also belong to the core study group of WHOQOL Project Turkish team. The main target of SAYKAD has been to introduce and disseminate knowledge about HRQOL to diverse health disciplines and healthcare professionals in Turkey. To achieve this goal, the society has participated in and organized a number of scientific meetings and workshops and has provided methodological support to professionals in Turkey who work on HRQOL in clinical practice.

Organizational structure:

SAYKAD's executive committee is composed of five members elected for two years. It has 98 members since 2010. The organizational structure of SAYKAD has led to the establishment of a number of *working groups* such as Allergy, Dermatology, Pediatrics, Geriatrics, Urology, Pneumatology, Rheumatology, Cardiology, Orthopedics, Ear/Nose/Throat, Plastic Surgery, Rheumatology, Liaison Psychiatry, Nursing (caregivers), Public Health and Occupational Health, Endocrinology, Oncology, Nephrology, Pain, and Oral Health & Dentistry.

These working groups are in close collaboration with the relevant medical specialty societies in Turkey, ensuring that these societies include HRQOL in their regular scientific meetings in the shape of HRQOL plenary sessions and workshops (SAYKAD contributed about eight National Brach Congresses up to 2011, by giving conferences and workshops mainly on methodological topics).

Regular activities:

The main regular activity of SAYKAD has been to organize regular National HRQOL meetings every three years, beginning with 2004. Three national congresses were organized since 2004 (2004, 2007, and 2010). The next National HRQOL Meeting will be held in 2011 in Izmir, as usual.

A noticeable scientific development was apparent in the 2010 meeting compared to previous meetings. In addition to the plenary sessions there were 18 parallel symposia on various medical specialties in the scientific program. About 300 medical professionals and a few from different disciplines, such as health economics, participated in the congress. Some 187 papers were presented in the congress, 91 of which were presented in the thematic poster sessions. There were also eight workshops, two of which were given by invited world-renowned HRQOL experts who brought the international view of the subject to the audience. SAYKAD is thankful to ISOQOL who supported the SAYKAD congresses by giving financial travel support to such very impressive presenters as Madeleine King, Peter Fayers, Neil Aaronson, and the recently departed Donna Lamping.

Society website: www.saykad.org

The second regular activity of SAYKAD is its working group (WG) activites. Each working group regularly updates the scientific information on its area of expertise on SAYKAD's website. The recent goal of the WGs is updating the national HRQOL bibliography.

Irregular activities:

A number of irregular activities have been carried out such as participating in international and national research projects and contributing at national medical congresses with lectures, panel sessions, and workshops.

International Scientific Projects: SAYKAD founders are the developers of the core version of the WHOQOL. Currently SAYKAD is serving as the WHOQOL National Center. Following its core projects, SAYKAD members participated to various WHOQOL Module projects:

1- WHOQOL-Spirituality (Prof. Dr. Hayriye Elbi)
2- WHOQOL-OLD (EU FP 5 Project): (Prof. Dr Erhan Eser)
3- WHOQOL-DIS (EU FP 6 Project): (Prof. Dr Erhan Eser)

SAYKAD officially took part as a participating center of Turkey to the EU 6th FP Project titled DISQOL. This project was finalized in 2009 and generated three cross-cultural instruments for our disabled population (Quality of Care instrument, Attitudes to Disability questionnaire, and HRQOL instrument (WHOQOL-DIS)).

SAYKAD has also been involved in the national adaptation of KIDSCREEN Project.

National Scientific Projects: A number of national projects have been carried out by SAYKAD's membership including:

- National Lung Cancer and QOL Project (AKAYAK) officially carried out by SAYKAD and Turkish Chest Society (TORAKS).
- KINDL Module projects
- Various cross validation projects and instrument development such as VFQ, AQLQ, EORTC, ADAS-COG, etc.

SAYKAD has also begun work on a national HRQOL bibliography. In addition to this, for the sake of harmonizing the terminology used in the country, a new working team has started to work on creating a HRQOL Terms Dictionary for Turkish. Due to the lack of drug industry relationships, SAYKAD has no financial support other than membership dues and potential international project contributions. A new emerging target of SAYKAD is to establish a national ISOQOL chapter for Turkey. The negotiations between SAYKAD and ISOQOL on this issue have reached a very promising point.

Prof. Dr. Erhan Eser
Turkish Society of Quality of Health Research (SAYKAD) – Izmir, Turkey
Celal Bayar University Faculty of Medicine, Dept. of Public Health, Manisa, Turkey
E.eser@bayar.edu.tr – Erhaneser.md@gmail.com
www.saykad.org – www.bayar.edu.tr/halksagligi
Tel: +90 236 233 19 20

Sheffield Dignity Questionnaire (SDQ)

Simon Dixon

University of Sheffield and Devices for Dignity (D4D) Healthcare Technology Co-operative

KEYWORDS

SDQ, DIGNITY, HEALTH CARE, QUESTIONNAIRE DEVELOPMENT, WELL-BEING

Abstract

While dignity is a prominent issue in healthcare, no standardized questionnaire exists that captures the multi-faceted nature of it. A qualitative framework was used to identify dignity-related concepts and a set of questions was developed that linked to a conceptual model of dignity, healthcare, and well-being. The questions were given to 197 hospital inpatients. Factor analysis and reliability statistics suggested a structure that linked back to our conceptual model.

Background

The importance of dignity within healthcare has been increasingly recognized in recent years. Within the United Kingdom this become a policy imperative with the Dignity in Care Campaign, the Dignity Challenge,[1] guidelines from the Nursing and Midwifery Council[2] together with other initiatives from voluntary sector organizations (for example, Magee[3]). However, while dignity is a prominent issue in healthcare, no standardized questionnaire exists that captures the multi-faceted nature of it.[3] This study describes the work done to date on a program of research undertaken in collaboration between the University of Sheffield and the Devices for Dignity (D4D) Project, that aims to develop a patient-reported outcome measure relating to dignity that is capable of being used in health technology assessment.

Methods

Question development

Electronic searches were undertaken to identify how dignity has been used in relation to healthcare. This was supported by a pearl-growing strategy based on policy documents and gray literature known to the research team. Usage of the term "dignity" was described and different concepts that underlie its use were identified. Qualitative research techniques were used to identify broader themes that characterize the overarching concept of dignity and brought together in a conceptual map. Existing questions relating to the different dignity themes were highlighted, and other potential questions were constructed.

In order to draw all of these concepts together within a single framework we felt that Sen's capabilities approach was useful.[4] Sen's work has been used to re-examine well-being by economists and it distinguishes between the capabilities of individuals, their functionings, and the utility derived from both. So, functionings (e.g., the absence of pain) can only exist if associated capabilities are present (e.g., access to pain relief). Once this distinction is made, we can link "environment," "processes of care, " "capabilities," and "functionings" to "dignity" and "well-being" (Figure 1). A fuller description of the development of this framework and how it relates to the individual issues identified in the literature is given elsewhere.[5]

It was felt that questions did not need to be developed for "functionings" as these are captured by existing measures of health-related quality of life, such as the SF-36 or for cost-effectiveness analysis, the EQ-5D. We therefore focused on the identification of questions relating to environment, processes of care, and capabilities, which all feed through to dignity.

In order to do this, we re-examined our thematic analysis of the literature and identified 16 questions based either on existing questionnaires[6,7] or developed for this study. The questions covered safety, cleanliness, privacy, social inclusion, personal hygiene and appearance, views of others, independence, control, self-esteem, and processes of care. Each was constructed to have four levels of response.

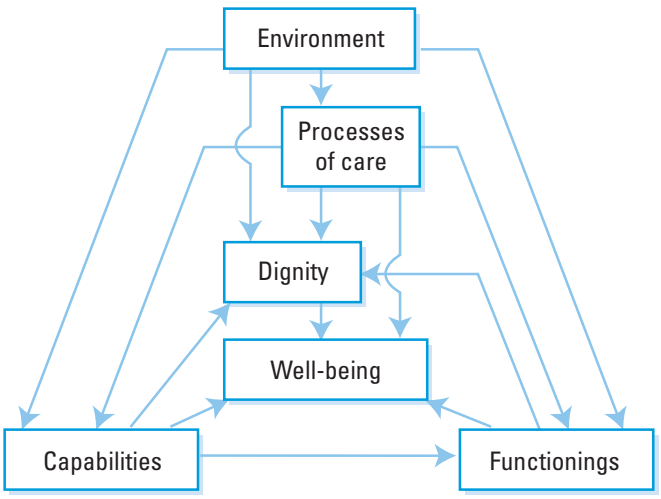
Survey

Patients were surveyed within a large teaching hospital. Participants were given a patient information leaflet describing the study, then approached by a trained interviewer the following day. Written consent way taken if the patient was interested in participating and the interview proceeded. The interview was undertaken using a personal digital assistant (PDA) device. The study was approved by an NHS Ethics Committee.

The interview was in three sections. In section 1, patients were asked to rate their life satisfaction using a single item taken from the Personal Well-being Index[8] and current health-related quality of life using the EQ-5D.[9]

Figure 1.

Conceptualizing dignity within Sen's capability approach



An additional domain was added to the EQ-5D that referred directly to dignity; the levels were, "I feel that I live with dignity," "I feel that I live with some dignity," and "I feel that I live with very little dignity." In section 2, patients were asked the 16 dignity-related questions based on our conceptual framework. In section 3, sociodemographic questions were given.

Analysis

The structure of the survey responses were examined using factor analysis with an orthogonal (varimax) rotation and the number of factors determined by the Kaiser criterion. Attribution of questions to latent factors (and, as such, possible questionnaire domains) was based on loadings with a magnitude of greater than 0.5 within the rotated component matrix. Reliability of responses within each construct/domain was based on Cronbach's alpha, with a value of 0.6 being considered sufficient for this exploratory study.[10] No imputation of missing data was performed.

Results

There were 197 patients in the survey. Based on the Kaiser criterion (and supported by an examination of the scree plot), a four-factor solution was suggested. The four factors appeared to relate to notions of "process," "control," "environment," and "external relationships." The rotated component matrix is shown in Table 1.

From the 16 questions, seven relate to "process," three relate to "control," three relate to "environment," and 2 relate to "external relationships." One question doesn't have a high loading on any of the identified domains. Alternative rotation methods, both orthogonal and oblique did not alter the interpretation of the results significantly.

The Cronbach's alpha coefficient for the four subscales were 0.871, 0.655, 0.733, and 0.510, respectively. There was no significant increase in coefficients with the deletion of individual questions within each of the factors/domains. The biggest improvement was from 0.655 to 0.686 for Factor 2 ("control"), with the removal of the question related to self-esteem ("How do you feel about yourself?").

Discussion

This paper describes the first attempt to produce a questionnaire focused on dignity in healthcare. While some patient experience and patient-satsifaction questions address some related issues (for example, the Picker Patient Experience Questionnaire[11] and Medical Interview Satisfaction Scale[12]), they do not capture all aspects in a cohesive framework. Though one dignity questionnaire does exists (the Patient Dignity Inventory[13]), this is very specific to end-of-life care, with many of its questions being irrelevant to other care situations.

The results from this first survey provide encouraging results with four clearly identifiable domains that relate to a conceptual framework of dignity developed through a review of the literature. The analysis suggests that only one of the questions should be removed from the resulting questionnaire, thereby, implying a 15-question instrument with four domains. However, it must be recognized that this is the first survey using these questions. Furthermore, reliability of the external relationships domain is poor for a definitive analysis.[10,14] This domain is also the most problematic in terms of its relationship to the underlying conceptual framework we developed; the latent construct doesn't directly match a separate concept within our framework.

Further work is clearly needed in two respects. Firstly, further analyses of the survey data are planned with respect to life satisfaction, the EQ-5D, and the additional dignity domain added to the EQ-5D. Secondly, other surveys are required using either the 15 or 16 questions, which should include settings other than inpatients.
In the longer term, it is our desire to generate an instrument that can be used in cost-effectiveness analyses of care interventions. At present, interventions that have a large, positive impact on a patient's dignity may not look cost-effective because the impacts are not captured by the outcome measures currently used within economic evaluation (such as the EQ-5D). To achieve this goal, a preference-based scoring algorithm needs to be developed which, in turn, will require a questionnaire much shorter than the current 15 questions.

Table 1.

Rotated component matrix *

	Component			
	1	2	3	4
How safe do you feel?			.782	
How clean and comfortable are your current surroundings?			.731	
How much privacy do you get in situations where you want it?	.649			
How much contact do you have with friends and family members?				.816
How clean and presentable do you feel?			.785	
How do others view you?				.715
How much independence do you have over your daily life?		.813		
How much control do you have over your daily life?		.793		
How do you feel about yourself?		-.659		
Information - process	.580			
Decisions - process				
Listened to - process	.705			
Warm and friendly - process	.764			
As a person - process	.693			
My problems - process	.794			
Respect - process	.826			

* Loadings with a magnitude of less than 0.5 have been removed to ease interpretation

Funding

The study was funded by Devices for Dignity (D4D), which is funded through the National Institute for Health Research Invention for Innovation Programme, the Technology Strategy Board, the Engineering and Physical Sciences Research Council, and the Medical Research Council.

For more information, please contact:
Simon Dixon
School of Health and Related Research
University of Sheffield
Regent Court
Sheffield. S1 4DA
Tel: +44 (0)114 2220 724
E-mail: s.dixon@sheffield.ac.uk

ACKNOWLEDGEMENTS

Other members of the research team are Simon Palfreyman, Phil Shackley, and John Brazier. The team has been supported by the D4D researchers at Sheffield Teaching Hospitals, in particular, Wendy Tindale and Nicola Heron. The survey was organized by Sue Butler and Jane Elliott at Sheffield Teaching Hospitals. The survey was processed by Picker Institute Europe.

REFERENCES

1. Cass E, Robbins D, Richardson A. Dignity in care. SCIE Guide 15. *SCIE*, 2009.
2. Nursing and Midwifery Council. Care and respect every time: new guidance for the care of older people. *NMC*, 2009.
3. Magee H, Parsons S, Askham J. Measuring Dignity in Care for Older People. *Help the Aged*, 2008.
4. Sen A. Commodities and Capabilities. Oxford: Oxford University Press, 1985.
5. Dixon S, Palfreyman S, Shackley P, Brazier J. What is dignity? A literature review and conceptual mapping. Health Economics and Decision Science Discussion Paper. Sheffield: *ScHARR*, 2011.
6. Malley J, Netten A. Putting people first. Development of the putting people first user experience survey. *PSSRU*, 2009.
7. Netten A, Burge P, Malley J, et al. Outcomes of Social Care for Adults (OSCA). Interim Findings. *PSSRU*, 2009.
8. International Wellbeing Group. Personal Wellbeing Index. Melbourne: Deakin University, 2006.
9. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33:337-43.
10. Hair J, Anderson R, Tatham R, Black W. Multivariate Data Analysis (5th ed). Engelwood Cliffs, NJ: Prentice Hall, 1998.
11. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *International Journal for Quality in Health Care* 2002;14:353-358.
12. Meakin R, Weinman J. The "Medical Interview Satisfaction Scale" (MISS-21) adapted for British general practice. *Family Practice* 2002;19:257-263.
13. Chockinov HM et al. The Patient Dignity Inventory: A Novel Way of Measuring Dignity-Related Distress in Palliative Care. *Journal of Pain and Symptom Management* 2008;36:559-571.
14. Nunnaly J. Psychometric theory. New York: McGraw-Hill, 1978.

NEWS FROM...

Mobile Patient Reported Outcomes – What it holds for India!

Thangaraj Nagasamy

With its booming economy and huge population, India possesses huge potential for the rapidly growing multinational pharmaceutical companies. Indian healthcare and pharmaceuticals are among the fastest growing sectors and the healthcare industry is expected to reach $200 billion worth by the year 2022. Since India holds such unassailable promise, it's no wonder that India is rapidly turning into a preferred destination for clinical trials.

Patient-reported outcomes (PROs) are increasingly used in clinical trials and research because of their utility in ascertaining the decisions of clinicians and patients about treatment alternatives. Patient-reported outcome measures (PROMs) play a big role in providing insight into the way patients perceive their health in addition to understanding the impact of treatments and adjustments to lifestyle on the quality of life.

While patient-reported outcomes are so important in improving healthcare, obtaining reliable and valid PROs hinges on having the right mix of patients. This task is made further difficult in a multi cultural and multi plural society like India. The rural-urban and the rich-poor divides throw formidable challenges in getting patient-reported outcomes that can really be called representative in character.

Given the vastness and topography of India, traditional hard copy method (using printed copy of PRO instruments) of obtaining patient-reported outcomes raises more questions than it answers. At the very first, it ignores the unlettered people of India who form about 34 percent of the total population and secondly it ignores majority of rural population who are faraway from the reach of clinicians and healthcare workers. The new-found ePROs could not either reach these hapless people who are in constant battle with poverty.

In this scenario, Mobile Patient Reported Outcomes (mPRO) holds out great promise in getting the so-far neglected lot into the realm of PROs. India with 851.70 million mobile phone subscribers at the end of June 2011, mPROs can be effectively used among the cross sections of population so far neglected. The interesting phemenon of greater rural mobile density than in the urban areas will make India undo the imbalance in the patient reported outcomes through the use of mobile technology. Mobile phones will also reduce the burden on the respondent and decrease data errors in patient-reported outcomes. Mobile Patient Reported Outcomes also ensure quick trial implementation with quicker results, real time analysis, remote monitoring and encrypted secure messaging. Mobile Patient Reported Outcomes technology is sure to bring about a sea change in the entire gamut of PROs in India.

For more information, please contact:
Thangaraj Nagasamy
New No.19/Old No. 53
Journalist Colony
Thiruvalluvar Nagar – Thiruvanmiyur– Chennai – 600041 – India
Mobile Ph No. +91 9952 0898 35

NEWS FROM...

The Translation and Cultural Adaptation Special Interest Group (TCA-SIG)

Katrin Conway, MAPI Research Trust, Lyon, France

The Translation and Cultural Adaptation Special Interest Group (TCA-SIG), established in 2004 during ISOQOL's annual meeting in Hong Kong, strives to identify and advance research in the fields of translation and cultural adaptation of Patient-Reported Outcomes (PRO) measures.

The TCA-SIG is chaired by Donald Patrick, University of Washington, Seattle, WA, USA, and Katrin Conway, MAPI Research Trust, Lyon, France, and its 25 members are divided into three subgroups, each pursuing individual objectives to meet the overall aim:

1. The cross-cultural issues subgroup is lead by Sonya Eremenco, United BioSource Corporation, Bethesda, MD, USA, and addresses issues related to the access of copyrighted instruments, the translation of PROs and their use in e-format.
2. The Translation Methodology subgroup is chaired by Mona Martin, Health Research Associates, Inc., Seattle, WA, USA, and pursues a research agenda for the development of methodologies in the field of translation and cross-cultural research.
3. The PRO Translation Certification subgroup coordinated by Mona and Katrin aims at establishing an international certification program for PRO translations.

After conducting exploratory work during the first five years of its existence, our group decided to focus its efforts on the publication of our findings. The following six topics were identified:

1. Copyright of translations of PRO measures, rules, and applications. Senior author: Caroline Anfray, MAPI Institute, Lyon, France. Submission target: 1st Quarter 2012
2. Translation of Patient Reported Measures: What type of certification should be offered? Senior author: Mona Martin, Health Research Associates, Inc., Seattle, WA. Submission target: 1st Quarter 2012
3. The process of reconciliation in the translation of quality of life questionnaires: Evaluation of existing procedures, criteria and outcomes. Senior author: Michael Koller, Center for Clinical Studies, University Hospital, Regensburg, Germany. Publication submitted.
4. Efficiency Of Translation Methodology in Error-Reduction for new Language Versions of PRO Instruments. Senior author: Mona Martin, Health Research Associates, Inc., Seattle, WA. Submission target: 2nd Quarter 2012
5. Translation Difficulties. Senior author: Catherine Acquadro, MAPI Research Trust, Lyon, France. Submission target: 2nd Quarter 2012
6. Pooling cross-cultural PRO data in global clinical trials: How much can a poor measure in a cultural subgroup affect the estimation in the overall population? Senior author: Antoine Regnault, MAPI Values, Lyon, France. Submission target: 1st Quarter 2012.

We are delighted to report the third topic (above) was submitted in June 2011 to Expert Review of Pharmacoeconomics & Outcomes Research and look forward to updating you on the progress of the remaining topics during our annual meeting during the Denver conference. This year's meeting of the TCA-SIG will be on Thursday, October 27th, between noon and 2 pm. The highlight of our annual meeting will be the following two presentations which, we are sure, will lead to a stimulating discussion:

1. "Linguistically validating PRO measures with populations that are difficult to interview" by Darren Clayson of PharmaQuest Ltd, Banbury, Oxfordshire, UK
2. "The study of different Spanish versions of the DTSQ" by Annarita Felici, Health Psychology Research Ltd, University of London, UK.

We hope you will join us in Denver. Please contact Tatiana Gauchon (tgauchon@mapigroup.com) for more information prior to the meeting or consult the ISOQOL website (www.isoqol.org) about our activities.

PUBLICATIONS...

Global Clinical Trials: Effective Implementation and Management

This book of Richard Chin and Menghis Bairu explores the great opportunities and challenges which exist in conducting clinical trials in developing countries. By exploring the various regulations specific to the major players and providing insight into the logistical challenges including language barriers, this book provides a working tool for clinical researchers and administrators to navigate the intricacies of clinical trials in developing countries. Important topics such as ethical issues are handled very carefully to highlight the significant differences of conducting this work in various jurisdictions. Overall, it presents a clear and comprehensive guide to the ins-and-outs of clinical trials in various countries to assist in design, development, and effectiveness of these trials. Contributors include high-profile, respected figures who have paved the way for clinical trials in developing countries. This title provides hands-on tools for regulatory and legal requirements and qualification, design, management, and reporting. Case studies outline successes, failures, lessons learned, and prospects for future collaboration. This title includes country-specific guidelines for the most utilized countries. It features a foreword by David Feigel, former Head of CDRH at FDA.

NEWSLETTER
Patient Reported Outcomes
www.pro-newsletter.com
Number 46
Fall Issue 2011
20 Years
Special Edition

Editorial

Happy 20th Birthday to Us!

2011 marks the 20th year of publication of the PRO Newsletter!

Keeping up the pace for 20 years is a feat unto itself!

Originally published as the Quality of Life Newsletter, our publication was the first of its kind, entirely focused on health-related quality of life (HRQL).

I must pay homage to Bernard Jambon for his visionary skills. He knew that asking the patients to self-assess their health or the impact of a treatment on their health would be crucial for the future, and he gave us the means to create a basis for communication between all those involved in HRQL assessment.

Twenty years ago, I was a trainee at MAPI and could not imagine how far this adventure would lead us: our participation in the IQOLA Project and many other linguistic validation projects, the ERIQA Group, the PRO Harmonization Group and the publication of regulatory guidances on the use of HRQL and PRO measures in clinical trials by the EMA and the FDA.

Twenty years of crucial changes in the field.

I am proud to see that our Newsletter has accompanied those changes so remarkably, and I hope that the 20 years to come will see many changes which will greatly be of benefit to patients worldwide!

Equally importantly, I join the editorial team to thank our readers and authors for keeping the PRO Newsletter going over the last 20 years. We are doing our best to carry on the work started 20 years ago. If you have any suggestions or comments on how we could improve our publication, please feel free to tell us.

Catherine Acquadro, MD

Looking Backward, Looking Forward

Reflections on the 20th Anniversary of the PRO Newsletter

John E. Ware, Jr., PhD
Barbara Gandek, MS
University of Massachusetts Medical School, Worcester, MA, USA

It does not seem possible that 20 years have passed since publication of the first issue of the PRO Newsletter, which is unique in its focus on patient-reported outcomes and its timely facilitation of commentary regarding progress in the field and editorializing about plans for the future. We welcome the opportunity of this noteworthy anniversary to look back and reflect on our PRO Newsletter contributions and the future of the PRO field.

The first issue of the PRO Newsletter in 1991 included the first publication about a new 36-item "Health Status Survey," a year in advance of its more thorough documentation in the APHA journal Medical Care.[1] That same PRO Newsletter issue included a front-page summary of Healthy People 2000, a U.S. initiative seeking to produce more healthy years of life.[2] The developers of the 36-item survey, also known as the SF-36® Health Survey, sought to standardize the metrics underlying frequently-measured generic health domains, so that results from very different applications such as the population health surveys that would be required for efforts like Healthy People 2000 could be compared with results from clinical trials and results for individual patients in everyday clinical practice. Healthy People 2000, like analogous international efforts, had to first come to grips with *measuring* health in preparation for the information systems that would enable those seeking to *manage* health.

Reflecting a sense of pride and hope for the future, that first PRO Newsletter article about the SF-36 noted that it was already in use by 52 organizations and in dozens of clinical trials, which seemed like a lot at the time. Like the Sickness Impact Profile (SIP), the SF-36 and other surveys from the Medical Outcomes Study (MOS) were constructed according to classical test theory;[3,4] as noted below, the landmark transition to modern psychometrics came later.

Coincidentally, the International Quality of Life Assessment (IQOLA) Project was formally launched in 1991, the same year that the first issue of the PRO Newsletter was published. The IQOLA name was coined by Bernard Jambon, the Newsletter's founder, during the inaugural meeting of the project. Subsequent PRO Newsletter articles in 1992 introduced the IQOLA Project, which was a public-private partnership sponsored by the pharmaceutical industry that included individuals and organizations from around the world, including the MAPI Research Institute in a major role coordinating IQOLA activities.[5,6] Other early issues of the PRO Newsletter announced forthcoming publications of manuals prepared for early adopters of the MOS short-form surveys as well as www.sf-36.org, a website designed to serve the needs of PRO researchers and facilitate royalty-free availability of MOS surveys.

From the perspective of 20 years later, it is striking how so much of what was first published in the PRO Newsletter grew far beyond what could have been reasonably anticipated at that time. For example, the original eleven SF-36 translations became more than 100, and the dozens of randomized controlled trials using the SF-36 turned into thousands. The International Society for Health-Related Quality of Life (ISOQOL), announced in a 1993 issue of the PRO Newsletter, now is celebrating its 18th annual meeting in Denver in October 2011. In addition, the PRO Newsletter publicized the new ISOQOL journal *Quality of Life Research*, dozens of generic and disease-specific PRO instruments that became legacy measures, and the activities of numerous groups and organizations that have influenced PRO science and related regulatory affairs.

In reading through archive issues of the Newsletter, one is struck by how much has progressed in the PRO field, the challenges that remain to be addressed, and the key role that the PRO Newsletter has played in calling attention both to what has been and what needs to be accomplished. Thinking about what might lie ahead for the next 20 years, with the perspective of the past 20 years, brings to mind the baseball team manager Casey Stengel, who famously said: "*Never make predictions, especially about the future.*" With this reminder in mind, our comments below are as much thoughts about what the PRO field needs to do in the next twenty years as they are forecasts as to what that future may hold.

Clearly one much-needed transition is away from quantifying each health domain using the different metrics of different health surveys towards cross-calibrating the metrics underlying the domains, so that results can be meaningfully compared across health surveys. This process is somewhat analogous to the successful cross-calibrations in thermometry which standardized the metrics underlying various thermometers hundreds of years ago. The role of survey item banks sampled to represent the content of widely-used health surveys and cross-calibrating them in relation to a common underlying "ruler" for measuring their common health domains was forecast in the PRO Newsletter, as was the more efficient administration and scoring of those survey items using computerized adaptive testing (CAT). The PRO Newsletter published articles about item banking and CAT as early as 1999[7] and devoted a special issue to this topic in 2004.[8] However, the PRO field still is marked by a plethora of instruments that measure the same domains with similar items but different metrics, which makes comparison of data across research studies and the interpretation of study results for the clinical community all the more problematic. While every domain may end up with a CAT, an equally important issue is that of deciding which domains are worth measuring and for what purposes. There are many generic domains and subdomains and dozens of disease-specific symptoms and other concepts that may help to understand disease burden and treatment benefits. More attention should be focused on the structure of all of these domains and how they should be sampled for different populations, as well as their clinical, economic and social consequences. What is the comprehensive conceptual framework that can guide health measurement in the future? With the CAT measurement wave, we seem to have forgotten about the conceptualization of health and how to decide what to measure.

Regardless of how often CAT proves to be better than the electronic administration of "static" (pre-selected) survey items (particularly those static forms that are based on the same IRT models that make CAT possible), IRT models are likely to play a strategic role in the evaluation of the equivalence of language translations, as demonstrated during the IQOLA Project.[9] Today, some health metrics have been translated well enough to enable their use in multinational clinical trials and population health surveys. At least as important going forward are comparisons of the effectiveness of different health policies for organizing and financing health care and the effectiveness of different treatment approaches within and between countries, in terms of the health outcomes that matter most to patients.

A third development will be changes in technology. As evidenced by Google's recent premium-priced offer to acquire Motorola Mobility, it seems likely that much of what we do with computers will be done with handhelds that serve many other purposes in addition to communications. Although we cannot guess what technology will be available in 20 years, any more than we could have envisioned today's wireless world in 1991, the world of 2031 no doubt will be faster and smarter. The challenge for the PRO field will be to incorporate new technology into routine use, which is only being done on a limited basis today.

Finally, it is our hope that the availability of much more efficient and user-friendly technology will enable more practical data collection and timely reporting of results and will lead to widespread adoption of PRO measures in everyday clinical practice. While there are small corners of the clinical world that value the information that can be obtained from PRO measures, these are not extensive. Among the major challenges to be overcome are ease of use, meaningfulness of PRO data to clinicians, and economic incentives. Although it may be difficult today to envision PRO measures being as much of a "vital sign" in clinical practice as temperature or blood pressure, it also was difficult twenty years ago to imagine today's widespread use of PRO measures in clinical trials.

The PRO Newsletter has chronicled the growth of the PRO field and has provided its contributors and readers with a forum and a source of considerable value. Its audience over the past 20 years has grown well beyond an initial small group of academic researchers to a diverse international community representing both public and private needs and interests. We wish it the best of health for another 20 years and look forward to it continuing to be a leading source–whether by the printed page or pages on the Internet–when it comes to information about the present and future of the PRO world.

For more information, please contact:
John E. Ware, Jr., PhD,
Barbara Gandek, MS
University of Massachusetts Medical School – Department of Quantitative Health Sciences
55 Lake Avenue North, Worcester, MA 01655
Email: john.ware@umassmed.edu, barbara.gandek@umassmed.edu

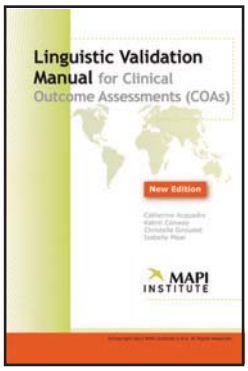
REFERENCES

1. Ware JE, McHorney CA. The SF-36 Health Status Survey: Brief overview. *Quality of Life Newsletter* June-September 1991; 1:4.
2. Erickson P. Improving quality of life: A major health policy goal for the U.S. in the coming decade. *Quality of Life Newsletter* June-September 1991; 1:1.
3. Bergner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS. The sickness impact profile: validation of a health status measure. *Med Care* 1976; 14:57-67.
4. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473-83.
5. Wagner A. International Quality of Life Assessment (IQOLA) Project. *Quality of Life Newsletter* February-May 1992; 3:2.
6. Gandek B. International Quality of Life Assessment (IQOLA) Project. *Quality of Life Newsletter* October 1992-January 1993; 5:10.
7. Ware JE, Bjorner J, Kosinski M. Dynamic health assessments: The search for more practical and more precise outcomes measures. *Quality of Life Newsletter* January-April 1999; 21:11-13.
8. Item Banking. *Quality of Life Newsletter* Fall 2004 Special Issue.
9. Gandek B, Ware JE (eds.) Translating functional health and well-being: IQOLA Project studies of the SF-36 Health Survey. *J Clin Epidemiol* 1998; 51:891-1203.

PUBLICATIONS...

MAPI Institute Linguistic Validation Manual New Expanded Edition

Publisher: MAPI Institute - Available January 2012



Translation of measures to be used in clinical research is essential in a global world where multinational trials are gaining momentum. Written by authors who are well respected within this field, The Linguistic Validation Manual for Clinical Outcome Assessments (COAs), New Expanded Edition, lays down procedures on translating Patient-Reported Outcome (PRO) Observer-Reported Outcome (ObsRO) and Clinican-Reported Outcome (ClinRO) measures.

The new edition is a revision and update of the book published in 2004. Three additional chapters have been added to reflect the methodological developments since the first publication:

- The MAPI Institute Checklist: to aid in the review of translations of Clinical Outcome Assessment (COA), MAPI Institute has developed a checklist to address FDA requirements for evidence that the COA measures have been adequately translated to ensure content comparability and validity for combined analysis of data from two or more language versions. The checklist helps to organize all information generated during translation and linguistic validation as well as recording reviewer comments.
- Translatability Assessment: a new way of assessing the translatability of instruments during the development phase.
- Copyright Issues: clarifications and recommendations about the copyright of measures used in clinical research.

20 Years of Quality of Life Research: A Russian Experience

Sergei Varshavsky, MD, PhD

Preference PRO LLC, President
San Francisco, California, USA

When I graduated medical school in Leningrad (now St. Petersburg) in 1983, medical practice in Soviet Union and pretty much worldwide was based on biological logic, opinions of established physicians, and popular beliefs. When I started my medical education in 1977 the total number of publications on randomized clinical trials in the whole world was 1248; and when I started my medical career six years later it was already up to 3050.

For comparison, in 2010, there were over 20,000 publications. Needless to say, the textbooks of the '70s hardly mentioned clinical trials at all.

During my graduation year I was lucky enough to have an internal medicine professor (Leonid Ermilov) who introduced me to the concept of the double-blind, randomized controlled experiment, and since then I have tried to use unbiased data on effectiveness of medical interventions. The same professor taught us some other important notions, which one could not find in textbooks, among them "outcome research" and "quality of life." Let me remind the readers that it was very difficult in the Soviet Union to get access to western medical literature, and almost impossible to travel abroad to attend professional conferences. International communication between Soviet physicians and their foreign colleagues simply did not exist.

The situation changed rapidly and drastically within one year; thanks to Mikhail Gorbachev, by the end of 1989 the Iron Curtain has collapsed. I received an invitation to participate in an international clinical trial, the European Myocardial Infarction Project (EMIP), and went to Lyon, France, to attend the investigator meeting. A few months later Soviet Union joined this study. It was the first randomized controlled clinical trial in this country. EMIP was led by two distinguished French trialists: Jean-Pierre Boissel and Alain Leizorovicz.

In 1990, when I was visiting Alain in Lyon, he introduced me to Bernard Jambon, the founder and chief executive of MAPI. We had lunch together and exchanged our business cards. Some months later Bernard sent me an e-mail asking if I would help with translation and adaptation of a quality of life questionnaire, the SF-36. I responded positively. It was the beginning of a beautiful friendship. Our company, Evidence Clinical and Pharmaceutical Research, established a quality of life (QOL) department, which was focused on translation and cultural adaptation of quality of life measures, educational projects and, to some extent, original research. Since its inception, the department is directed by Olga Sheinina.

In 2008, the QOL department was transformed into an independent company: Preference PRO. The original research was focused mostly on the evaluations of QOL in patients with chronic health conditions[1] such as congestive heart failure,[2,3] ischemic heart disease,[1] anemia,[4] rheumatoid arthritis,[5] psychiatric disorders,[6,7] etc. Together with Marianne Amir of Ben-Gurion University at Be'er-Sheva (Israel) we also conducted a fascinatiang sociological study.[8,9] Educational activity was through various publications in Russian medical journals, as well as by conducting conferences and seminars for Russian physicians, the clinical trial community, and regulators.[10,11]

MAPI, and personally Bernard Jambon, were very much involved in those activities; he and his colleagues visited St. Petersburg and Moscow numerous times. In education and original research, we also extensively collaborated with Prof. Yuri Krivolapov and Tatiana Ionova of the Russian Medical Military Academy, Prof. Natalia Petrova of St. Petersburg State University, as well as internationally with David Cella and Sonya Eremenco of CORE, John Ware and Barbara Gandek of the IQOLA project, and many others.

The most important part of our work during the past two decades has been, and still is, the cultural adaptation of QOL questionnaires and patient-reported outcome (PRO) measures. Hundreds of instruments have been translated and validated into tens of languages, primarily of Eastern Europe and Baltic states. We use well-established standard validation methodology, which includes forward and back translations, as well as patient testing and a physician review process. Besides that, we use the translatability analysis developed by Prof. Svetlana Kudria of St. Petersburg University.

Today, when we celebrate the 20th anniversary of the QOL Newsletter, I can proudly say, that we were at the very beginning of QOL research in Russia and Eastern Europe and have seen how this field of knowledge developed over the years.

For more information, please contact:
Sergei Varshavsky, MD, PhD
162 Duncan Street
San Francisco, CA 94110
Phone: 1 650 575 2044
e-mail: svar@preference-pro.com

REFERENCES

1. O.Loginovskaia, N.Tzai, S.Varshavsky. Evaluation of quality of life in stable angina patients with the Russian version of the SF-36 health survey. *Quality of Life Research* 1999, 8, 7: 624.
2. A.Nedoshivin, S.Varshavsky. Quality of life in congestive heart failure: Assessment with the Russian version of SF-36 health survey. *Quality of Life Research* 1999, 8, 7: 624.
3. A. Nedoshvin, A.Kutuzova, N.Petrova, S.Varshavsky, N.Perepech. Life Quality and Mental Status Assessment in Patients with Congestive Heart Failure. *Heart Failure (Serdechnaia Nedostatochnost)*, Vol. 1, No 4, 2000, 148-151.
4. E.Usacheva, S.Varshavsky, S.Eremenko, S.Plavinsky, C.H.Chang. Initial validation of the Russian version of the functional assessment of cancer therapy - anemia (FACT-An) questionnaire. *Quality of Life Research* 1999, 8, 7: 625
5. T..Ionova, A.Tzepkova, A.Maximov, S.Varshavsky, A.Novik. Quality of life in rheumatoid arthritis patients of different ages. *Quality of Life Research* 1999, 8, 7: 577.
6. O.Lapshin, R.Danko, V.Pashkovsky, S.Varshavsky, B.Gandek. Use of the Russian version of the SF-36 health survey in schizophrenic patients. *Quality of Life Research* 1999, 8, 7: 624.
7. T.Solokhina, E.Rytik, Y.Seiku, S.Varshavsky. Quality of life of relatives of patients with chronic schizophrenia. *Quality of Life Research* 1999, 8, 7: 601.
8. M.Amir, L.Ayalon, S.Varshvasky, N.Bulygina. Motherland or Home Country: A Comparative Study of Quality of Life Among Jews from the Former Soviet Union Who Emigrated to Israel, Jews in Russia and Israeli Nonimmigrants. *Journal of Cross-Cultural Psychology*, Vol. 30, No 6, November 1999, 712-721.
9. S.Varshavsky, M.Amir, N.Bulygina. Motherland or Home Country: Health-related Quality of Life Among Jews from the Former Soviet Union Who Emigrated to Israel and Jews in Russia as Compared to Israel Non-Immigrants. Abstracts: 4th Annual Conference of ISOQOL in *Quality of Life Research*, Vol 6, 1997, P. 736.
10. S.Varshavsky S. Quality of Life Assessment in Clinical Trials. Abstracts of the International Congress "GCP in Russia and the Baltic States," 22-24 September, 1997; P. 95-96.
11. S.Varshavsky. Assessment of Quality of Life with SF-36 Questionnaire in Evaluation of Effectiveness of Pharmacotherapy. *Achievements of Clinical Pharmacology.* Moscow 1999. P.30.

20 Years of Quality of Life Research

Monika Bullinger, PhD

Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Twenty years of quality of life research is certainly an occasion to look back to the beginning of addressing quality of life in medicine as well as looking ahead to the impact of this fascinating field of research and practice. And a 20-year anniversary is certainly a reason to congratulate the MAPI team on their consistent, continuous, and engaged work related to patient-reported outcomes.

As one of the presidents involved in the foundation of the International Society for Quality of Life Research (ISOQOL), I would like to give a personal account of the history and the future (which certainly will be a subjective view).

Personally I remember the foundation of the Society for Quality of Life Research in Brussels on a warm day with many enthusiastic researchers and clinicians in the early '90s. I also remember heated discussions about whether or not it was possible to measure quality of life in a Scottish castle in the mid '80s.

This debate was quite remarkable. The so-called theoreticians or philosophers insisted on the impossibility to measure quality of life without definition, while the pragmatists, who would not be searching for a nominal definition of quality of life, would be happy with a more modest operational definition. They found that this would make the quality of life concept more amenable to measurement.

Many months and years have passed since this early and quite impressive Scottish encounter. Certainly the pragmatists since then have moved ahead, resulting in a great variety of psychometrically sound instruments to access generic and conditions specific health-related quality of life in adults, adolescents, and children as well as that of their caregivers. However, it is also clear that insight into the more philosophical basis of health-related quality of life research was and is still needed to provide a sound theoretical foundation of the concept we are trying to address.

Quality of life research in my view has evolved over four phases, the first being discussions about the measurability and the definition of quality of life in the 1980s. The second phase involved the many approaches to measure health-related quality of life beginning around the 1990s. The third phase, around the year 2000, related to including these measures in different types of studies, ranging from epidemiological surveys to randomized clinical trials. More recently, from the year 2010, discussions about the impact of health-related quality of life assessment in medicine began. These do not only consist of operationalization of a patient-reported outcome but also addresses the question whether quality of life assessment can be used to define patient needs for care and also to evaluate quality even in an individual case.

Over the years, health-related quality of life research has inspired many researchers and clinicians, healthcare professionals, and health politicians. Nevertheless, the concept of health-related quality of life is in danger of loosing impact on the background of the role it plays in health economy and in comparison to classical medical outcomes. Although there is consensus to include patient views in assessing patients' needs and outcomes, there is a conceptual challenge of the use this information with regard to benefits in health economy and morbidity indicators in medicine. I would not call it a struggle; it is more a sense of having to contribute to the conviction, clarify and make colleagues understand why health-related quality of life is important, how it can be measured, and how this measurement produces reliable and valid information to be used not only for research but also in practice.

Certainly health-related quality of life research has made health-outcomes research more sound because it became clear that not only patient-reported but also other outcome measures should comply with rigorous methodological standards in terms of measurement properties. It has also opened the doors for a more patient-centered approach to medical care explicitly, including psychometric and mental health issues as well as disabilities and its joined disciplines that did not talk to each other so easily in earlier times, as we all know from the stereotypes of the typical physician and typical psychologist.

Personally I think of quality of life research as a rewarding and fulfilling field characterized by a humanistic impetus and a methodological expertise contributing to a better understanding of well-being and functioning in medicine. I am happy to have been part of this process.

The MAPI team's highly competent, energetic and impressive staff is not just another company in the field. It has played and continues to play a crucial role in making the field move, prosper, and succeed. For this I would like to express my gratitude and best wishes for the future. I am sure that the next 20 years will be as rewarding and challenging to MAPI as the previous 20 years have been.

Happy anniversary!

For more information, please contact:
Prof. Dr. Monika Bullinger
Universitätsklinikum Hamburg-Eppendorf
Institut und Poliklinik für Medizinische Psychologie
Haus W 26
Martinistraße 52
20246 Hamburg
Tel: +49 (040) 7410 52863
+49 (040) 7410 56430
Fax: +49 (040) 7410 54940
E-mail: bullinger@uke.uni-hamburg.de

IS QoL a COA?

Dick Joyce, PhD
Allschwill, Switzerland

Some Background

The first article of the MAPI Research Trust's Quality-of-Life Newsletter (MAPI QoL-NL), twenty years ago, argued for close study of the cognitive processes involved in evaluating QoL.[1] QoL was a major subject of interest for MAPI and other institutions. having by then become firmly established within the rapidly increasing field of Health Sciences (HS, of which Medicine itself was clearly the most important).

QoL as a process has always received even less attention than it has as a state. It is not an easy area for study, and the difficulties of both aspects have probably encouraged the movement away from enquiries into "true" QoL. This was reflected in autumn 2004 by a change in name of the newsletter, which became the Patient-Reported Outcomes Newsletter (MAPI PRO-NL).[2] Interest in the definition of QoL and development of methods for its study have continued to stagnate since the change of title. The overwhelming majority still rely on the questionnaire format, often paying obsessive attention to semantic detail in order to facilitate their use in ever larger clinical trials.

The most outstanding example of a study of healthy subjects on a nationwide epidemiological scale was proposed in 1972 by the King and subsequently executed in the Kingdom of Bhutan. In this happy country, the measurement of Gross National Happiness substitutes for the calculation of Gross Domestic Product made by most national authorities. The current British Prime Minister rather belatedly proposed something of the kind in 2006 and has been assisted in his ambition by the London School of Economics at a cost of £2,000,000 per year.[3] (Recent events may have dampened British enthusiasm for the project; Canada and France appear currently to be making greater progress in this direction.) Large-scale applications such as these surely require more basic attention to the improvement of their methodology than do even the largest of clinical trials.

Signs and Symptoms—and COAs

In the Anglo-Saxon tradition, PROs (including QoL) used to be called Symptoms; aspects of the patient that could be directly or instrumentally observed by a witness were labelled Signs. The distinction was rarely made by non-Anglophones, and was seldom adopted outside non-Anglo-Saxon medical terminologies.
The whole area has been further extended (or, as some may think, complicated) in a recent definition by the FDA of Clinical Outcome Assessments (COAs).[4] These include patient-reported outcome (PRO) measures, to which they add clinician-reported outcome (ClinRO) measures and observer-reported outcome (ObsRO) measures.

The changes in labelling signify not merely semantic quibbles, but yet another cycle in an epic, centuries-long struggle in behavioral research: the oscillation between so-called "subjective" and "objective" methodologies, a polarity also referred to as "idealist/materialist," "psychological/physiological," and "mind/brain," among many.

The theorist Goethe and others were followed by the experimentalists Wundt and Wertheimer. Later, the explosion of interest in psychoanalysis begun by Freud and Jung near the turn of the last millennium was succeeded by its polar opposite in the behaviorism of fewer than fifty years later, a movement led by Watson and Skinner. All of these to varying extents looked at both ends of the spectrum, usually turning away from one or the other. None took as pure a position as the British Royal Society, founded in the later half of the seventeenth century, whose motto "Nullius in Verbia" could be loosely translated as "Words mean nothing."

Since then, the emphasis has swung repeatedly from studies of internal process through words to deduction and measurement and back again. Erwin Schrödinger among others struck a balanced view that "The observer's senses have to step in eventually. The most careful record, when not inspected, tells us nothing."[5]

Conceptual Shortcomings

The relative polar utilities are indeed difficult to define, but at the moment insufficient attention is being given to this task, which must precede that of measurement. The interest of the patient in the medical process is based upon attitudes to her own health, an entirely subjective matter; members of the medical profession, on the other hand, are preoccupied with the accuracy of diagnosis (and, to a lesser extent, of prognosis) and the success of treatment, all of which are at the very least uneasy combinations of the subjective and objective. The patient's interest is in QoL as a continuing experience; although she can report upon this self-consciously it is not even commensurate, let alone identical with the abdominal pain or insomnia of which it may in part be a consequence.

Much of the early work on QoL, which originated in the architectural interests of Le Corbusier and others in the 1930s, was an attempt to quantify what was not merely unquantifiable but (as it is still) largely inaccessible. Some progress, useful in practice, could be made if the investigator clearly identified his own interest as that of the patient (but not by completing questionnaires intended for the patient, as sometimes used to occur).
Other important methodological errors are more frequent.[6] For example, though clinical workers may pay attention to the need to maintain constant environmental conditions between occasions, as between individuals, when conducting trials, there is seldom adequate control of other aspects,

such as the intrinsic behavioral variability of even a consistent examiner, or the degree and stability of the rapport established between the members of each clinician-patient pair.

Whether consciously or not, investigators and their professional colleagues seem at present to be taking refuge from such difficulties by a flight into the objective, or even pseudo-objective. The parking of QoL among PROs is a matter of convenience, but an unfortunate example: QoL is not what the patient is able to externalize as a PRO, but what she tells herself it is. This may be reported to the doctor with a limited degree of success; but it may in any event be substituted thereafter as a brief response to an enquiry about whether the patient is feeling better, worse, or the same. This has been the most common medical method of assessing QoL since the predecessors of Hippocrates, and is still regarded as adequate for practical purposes by many of his successors.

For the patient, it is not the increase of systolic blood-pressure from 135 to over 200 or a quantifiable fall in blood potassium below, say, 3.5 mEq/l that change her QoL. This is due to her anxiety about the awareness of cardiac irregularities to which these changes have given rise. The inclusion of attempts to categorize these feelings among hundreds of PROs and to label all of them, not as influences upon QoL but as instruments for its measurement, is misguided. The steady, sigmoid increase in the number of studies reputedly associated with (few of them of) QoL from the mid-1930s until recently seems now to be in decline, though it is accompanied by a more than compensatory increase of studies that employ one or more reputed PROs.

Escape from the Subjective and the Growth of "Administrationism"

These developments amount to the current cyclical escape from the pandemic and intransigent issues of subjectivity; the deconstruction of the increasing complexity of all areas of life, in medical science, in research as a whole and even more widely. Those of a harder scientific bent may regard it as yet another instance of the growing realization that there is no biological factor, from single gene or intensity of ambient light to social organization, that operates in total isolation from any other: each is inseparable from the influence of adjacent genes, the intensity of ward illumination or the phlegmatic cough of a neighboring patient.

The broadening of diagnostic categories and multiplication of subcategories in PROs seen in the increasing number of diagnoses listed in successive editions of the International Classification of Diseases (ICD)[7] arises as much from confusion as from the desire for clarification. Either way, it further muddies the far from transparent waters of QoL.

Many other important and unsettling events, outside medical science as well as inside, illustrate the contemporary flight from subjectivity. For instance, among vastly influential non-medical problems that members of the medical profession have recently faced or are likely to face in the near future is a substantial reduction in their income. This is largely due to incessant political interference with the old structures of welfare assurance; for example, the introduction, under insurer pressure, of such absurdities as the Disease Related Group (DRG) system of fixing medical fees by diagnosis rather than by treatment. These changes are likely to lead to increased reliance upon robot-driven or other automatic methods of testing, and consequently decreased interest in personal responsibility for the patient. Such limitations upon diagnosis and choice of treatment may, however, diminish as the proportion of women medical practitioners increases.

Each application of administratively redesigned medical policy increases the number of administrators required to oversee the additional production of quantifiable PROs. The time available for actual practice is thus reduced and a decline in medical service follows. This further reduces already limited information about QoL.

Similar anxieties can be legitimately voiced about developments in nursing and other specialities. Here too, the benefits of better training and increased responsibility and status are accompanied by increasing administration time and redirection of competent practitioners into blind-alley administration. Other sub-specialities, such as physiotherapy or clinical psychology are experiencing growing criticism of the competency of their practitioners (not necessarily a bad thing in itself). Here again, this has resulted in an increase in administrative work that practitioners must undergo and undertake in order to allow patients to claim reimbursement from overweening providers of medical insurance.

Other Developments

These examples are inimical to the critical examination of the subjective. The effects of other contemporary developments are even more difficult to investigate. Among them are such important changes in teaching as the development of interdisciplinary modules and distance learning, clearly desirable for underprivileged communities, although medically qualified parents are becoming less likely to advise their offspring to follow them.

Influences from the information revolution are especially important but hard to predict. They already affect patients, doctors, the pharmaceutical industry and regulatory authorities, and will continue to do so more and more profoundly. Among patients, Expert Patient, Self-Help and other Support Groups. They are made possible by the exploding use of social networking or the individual development of so-called multitasking that these also make possible.[8] On the other hand, the overall increase of Total Data Collection (e.g., by Google and Klout), may dehumanize us all yet further; or it may turn out to be beneficial.

Whether any of these developments will improve or further impede the study of QoL, however, will depend upon whether the next cycle will favor the development of more sophisticated methods of studying subjectivity. There seem to be few signs of such movement,[9] although the work of Damasio is a notable exception.[10]

Drug regulatory authorities, such as the American FDA and European EMA continually increase their demands for information about new entrants to the market, as well as existing products. These have included, until recently, insistence on studies of "QoL," contributing to a volume of paper already too heavy to be adequately studied. The FDA, it must be admitted, is coming to pay less and less attention to claims about the effects of interventions upon QoL.

Meanwhile, the pharmaceutical industry, encouraged by increasing problems with regulatory control, will probably be able to increase promotion of potentially dangerous off-label prescribing;[11] sometimes, it must be admitted, serendipitously beneficial.

The influence of changes in social policies about energy, if any, and their influence upon atmospheric CO_2 (and respiration!) are totally imponderable. Reflections upon corruption and other forms of pollution—the decline of responsibility, collegiality and basic decency—are not relevant in the present context, except to suggest that their implications may be even greater than those discussed above.

A final note on the problem discussed above: different timelines of medical discovery between 1990 and 2010 report the same three inventions: cloning, and successful methods of treating papilloma virus and hepatitis A. All three focus upon the molecule, and none upon the patient as an individual.

Conclusions

The main intention of measuring QOL, as with other so-called outcomes, remains that of making comparisons between the state of an individual patient at different times or between groups of patients participating in clinical trials. Structured questionnaires have been considered essential for these purposes because few medical practitioners can take enough time to get to know the individual with the desirable degree of intimacy. A potential major advantage of comparative or so-called alternative practitioners and their methods is that in general they devote much more time to the client, about whom they learn much more.

A PRO may be able to compare an individual patient's own views of symptom progress as well as these with the doctor's assessment. But, in general, grouped comparisons of any COA are unlikely to be reliably informative because of individual differences in the interpretation and relevance of the questions posed.

Methods of measurement for both QOL and COAs are defined by the needs and assumptions of investigators, not patients. The deconstruction of QOL, whether or not it remains at its present rather primitive level of development, can never be captured by the simplicity of a questionnaire. It inhabits its own universe of discourse.

A Proposal

Is QoL merely one of many COAs? It seems no longer to be considered a real PRO at all. But this does not mean that it should be completely neglected. On the contrary. It is the topic of greatest significance to the patient, and the importance of reviving its study can hardly be over-emphasized.

Might it be worthwhile for many reasons to celebrate the 20th anniversary of the MAPI Newsletter by reviving a Quality of Life Newsletter alongside its congeners?

For more information, please contact:
Dick Joyce
charles.joyce@cantab.net

REFERENCES

1. Quality of Life Newsletter (1991) 1, 1
2. Patient Reported Outcomes Newsletter (2004) 32, 1
3. Allegra Stratton. David Cameron aims to make Happiness the new GPD. *The Guardian*, November 14, 2010
4. Federal Register 78, 45271 28 July 2011
5. E Schroedinger: The Mystery of the Sensual Qualities. In: What is Life? (1967). Cambridge: CUP p 176.
6. Joyce CRB. Non-Specific Aspects of Treatment from the Point of View of a Clinical Pharmacologist. In: Non-Specific Aspects of Treatment (ed Shepherd M and Sartorius N) (1989) 57-94. Toronto, Huber
7. World Health Organization. Geneva. 1990 – . The International Classification of Diseases.
8. See, for instance, Sanjay Gupta (accessed 23.07.11 at 13:02): http://archives.cnn.com/2001/US/08/07/gupta.debrief.otsc/index.html Gupta
9. Wallace AB (2007) Contemplative Science. New York: Columbia University Press
10. Damasio A. (2010) Self Comes to Mind: The Evolution of Consciousness. New York: Pantheon
11. Angell M The Illusions of Psychiatry. *New York Review of Books*, July 14 2011

Message to a Good Friend

From QOL NL to PRO NL: Dick Joyce has strongly questioned this in his article, although it is in fact the current approach that he is criticizing. Extract from our correspondence about the article:

"...Dick, in theory one can only agree with your reasoning, which reminds us of what our goal is supposed to be. However, striving for perfection should not stand in the way of "mere" improvement. And over the past 20 years, we have all succeeded in implementing improvements by generalizing the PRO evaluations, despite a healthcare system that is continuously increasing its structural rigor.

With all the respect and affection I have for you, Dick.

Bernard"

Reviewing my comments, Dick confirmed that he agreed with this statement which briefly and "beautifully" stated the purpose of his article.

Bernard Jambon

10 Years of the Neuropsychological Test Battery (NTB)

John E Harrison BSc (Hons), PhD [1,2]
Angela Caveney, PhD[3,4]

[1] Metis Cognition Ltd., Kilmington, United Kingdom
[2] Imperial College, Division of Neurosciences & Mental Health, London, UK
[3] CogState Ltd., New Haven, CT, USA
[4] University of Michigan, Ann Arbor, MI, USA

KEYWORDS
NTB, VERBAL FLUENCY, DEMENTIA, COGNITION, DRUG TRIALS, ALZHEIMER'S DISEASE

Abstract

The Neuropsychological Test Battery (NTB) has now been used as a cognitive outcome measure in dementia clinical drug trials for more than 10 years. This period has witnessed changes in both the composition and administration of the NTB, as well as a good deal of further information with respect to its psychometric properties. Data have also been reported that describe the NTB's sensitivity to cognitive decline in patients with Alzheimer's disease. In this article we will describe and discuss use of the NTB, summarize the various psychometric developments and discuss prospects for current and future use of the battery.

A brief history of 'the' NTB

The Neuropsychological Test Battery, or "NTB" (Table 1) as it is more commonly known, is a composite cognitive measure comprised of standardized tests that have been in use in the field of clinical psychology for, in some cases, more than 60 years. All of the selected measures have been extensively individually validated and should not by any means be considered new tests. What is novel is the combination of these measures into a *portmanteau* assessment designed to yield a single efficacy measure of cognitive change.[1] This is a common approach in Alzheimer's disease (AD) drug trials, where issues of Type I error and potentially small effect sizes bias outcome selection toward single composite cognitive measures. In this respect the NTB is similar to other traditionally employed measures, such as the MMSE and ADAS-cog. What distinguishes the NTB is its focus on (1) associative learning as a paradigm for assessing episodic memory and (2) the assessment of executive functions (EF), such as planning, strategy and working memory. The determination to focus on associative learning was derived from academic research[2] and phase 2a proof of concept studies[3] that suggested this paradigm may be of particular utility for assessing cognition in patients with dementia. Interest in assessing executive function was prompted by recognition that these functions can be compromised early in AD. Measures of EF are also recognized to be robustly correlated with instrumental activities of daily living,[4] and thus it was considered to be worth including EF measures in the hope that efficacy on these tests could be linked to functional preservation or improvement. In the following section we will summarize the use and evolution of the NTB in clinical drug trials over the past decade.

It is sometimes wrongly reported that the NTB was first designed for use in studies of the immunotherapy Bapineuzumab. In fact, the NTB predates the development of Bapineuzumab by some years and was first designed for use in the study of compound AN1792.[5] Given the evolution of the NTB, it is worth commenting that it was originally intended as a hybrid of computerized and "paper and pencil" assessments, however, computerized assessments were not added at that time due to time and logistical constraints. The specific combination of tests that comprise the NTB (see Table 1) employed in AN1792-201 has since been used in a number of further studies, including trials of Bapineuzumab and Lecozotan. However, variants of the NTB, often sharing a close 'family resemblance' to the original version, have been employed to meet the specific needs of various trial sponsors. For example, in the case of Prana Biotechnology, preclinical data suggested that procognitive effects might be expected on tests of EF. Consequently the NTB content for the Prana study was changed to remove one test of memory which was replaced by a further test of executive function, the Trail Making Test.[6,7] Over the years, several other versions of the NTB have been employed. Of the six tests that comprise the original NTB, the EF elements are rarely altered. They are more often augmented, such as occurred in the aforementioned Prana trial.

A trend across the past 20 years has been to add further measures to dementia drug trial protocols in a bid to capture pharmacodynamic effects. A legacy of this trend is that cognition is now typically assessed using a variety of measures, including the Clinical Dementia Rating (CDR) scale, NTB and ADAS-cog. These assessments can require in excess of three hours to administer, giving rise to concerns regarding the validity of data from sometimes fatigued and frustrated study participants. The use of multiple measures also results in single cognitive domains being assessed repeatedly. For example, the NTB and ADS-cog between them contain five measures of episodic memory. One response to these issues of redundancy and extensive assessment time has been to alter the administration parameters of these tests. To illustrate, the Rey Auditory Verbal Learning Test (RAVLT) is most often used in its traditional format of:

Table 1.

Original composition of the NTB

Test	Domain	1st outcome measure	2nd outcome measure
Visual Paired Associates	Memory	Immediate recall	Delayed recall
Verbal Paired Associates	Memory	Immediate recall	Delayed recall
Rey Auditory Verbal Learning Test	Memory	Immediate recall	Delayed recall
Controlled Oral Word Association Test	Executive function	Total number of acceptable words	-
Category Fluency Test	Executive function	Total number of acceptable words	-
Digit Span	Executive function	Sequences correctly recalled	-

- Five trials of immediate recall for List A
- One trial of immediate recall for List B
- One short-term recall trial for List A
- Delayed recall for List A
- Recognition memory for List A

The original NTB employed the standard 30-minute delay between the immediate recall phase and the commencement of delayed recall. However, a problem with this methodology is that a significant number of patients with AD recall no words after 30 minutes. Floor effects such as this are problematic when analyzing and interpreting test data. Therefore subsequent studies have either employed only the first five learning trials of list A or have significantly reduced the delay period, sometimes to as little as one minute.

The remaining two NTB memory tasks, the Visual and Verbal Paired Associates (PA) tests, also feature a 30-minute interval between immediate and delayed recall. Floor effects on both of these tasks have also been reported. Shortening the delay period has been a common change to the Verbal PA, though it is not at all uncommon for these PA tasks to be entirely omitted. As justification the argument has been offered that verbal memory is adequately assessed by the RAVLT, and so inclusion of the Verbal PA test represents redundancy. Recently computerized measures of assessing episodic visual memory have become a common replacement for the Visual PA test, though sponsors who are disinclined to use computerized measures have been willing to abandon use of the Visual PA entirely. Part of the justification for doing so has been a perceived lack of sensitivity of this test. This is especially true for delayed recall where the possible range of performance is very limited (0-6) and where patients typically score at chance levels (1 point). Further disincentives are the difficulties of training test administrators and the need to acquire specialist materials with which to administer this test. However, perhaps the most challenging aspect of employing both the Visual and Verbal PA tests is the time required to administer these measures. We have anecdotal accounts of patients requiring as much as 20 minutes per test to complete the immediate recall phases. This markedly extends the testing, leading to frustration, and in about 10% of cases, refusal to continue with testing.

Omitting measures such as the PA tests and delayed recall components of the RAVLT reduces the time needed to administer the NTB from an average of approximately 70 minutes to circa 15 minutes. As a general rule it is our inclination to keep the time needed to assess cognition to less than 30 minutes, and this marked reduction in administration time permits the inclusion of other measures. For example, in the trials of Tarenflurbil the Visual PA test was replaced with the Digit Cancellation Test, a common ADAS-cog+ test capable of indexing attention and yielding gross estimates of psychomotor speed[8]. Other versions of the NTB, such as the one employed by EnVivo Pharmaceuticals, have exclusively featured EF measures and supplemented them with computerized measures of attention, psychomotor speed, visual learning, and working memory.[9]

Further validation of the NTB

The paper published by Harrison et al.[1] described key characteristics of the NTB, including test-retest reliability in patients with AD at 6 (r=0.92) and 12 (r=0.88) months and internal consistency (Cronbach's alpha = 0.84). Factor analysis of baseline performance yielded two Eigen factors, one showing robust correlations between the three EF measures and a second factor showing clear correlations between the memory scores. These findings have since been replicated.[10,11] The reporting of these psychometric characteristics, together with the evidence of sensitivity to decline over time and efficacy are presumed to have facilitated acceptance of the NTB as a primary cognitive outcome measure by the EMA and FDA.[12]

In addition to replicating levels of test-retest reliability, data have been presented describing further psychometric characteristics of the NTB.[11] NTB tests were, in part, selected as objective psychometric measures and it was expected that levels of inter-rater reliability would be high. This was investigated by Rentz et al. and inter-class correlations (ICC) ranging from 0.983 (95% CI 0.958-1.0) for the Category Fluency test to 0.999 (95% CI 0.999-1.0) for RAVLT Immediate Recall were observed. The overall ICC for the NTB was reported to be 0.998 (95% CI 0.996-1.0). Rentz et al. also helpfully reported the within-subject standard deviations (WSD) for the NTB composite and each of the contributing tests. This metric is a valuable estimate of performance stability and helpfully augments other measures of reliability. The WSDs observed by Rentz et al. are reported in Table 2.

Further validation of the NTB has occurred with respect to use in different cultures. Translations of the NTB have most often been conducted by MAPI and a key aspect of this process is the harmonious integration of NTB tasks to facilitate use in the target language. A good example of this principle is the selection of appropriate COWAT stimuli. The traditional English version of the COWAT is based on a systematic evaluation of all 26 letters and selection being made according to the frequency with which healthy volunteers are able to successfully generate responses.[13] This process has yielded the well-known "F," "A," and "S" versions of the COWAT for use with English-speaking individuals. Rarely a similar process will have been conducted in the target language and stimuli can be selected appropriately[14]. More typically, local versions will have been derived on the basis of intuition and experience. A preferred methodology has been to select these versions based on their appearance in peer-reviewed articles. Occasionally the target language is such that material changes need to be applied in order to administer homologous versions of the task.[15]

The evidence so far suggests that the NTB can capture efficacy in circumstances in which the ADAS-cog exhibits only a trend in favor of efficacy.[5] This may be particularly true in studies of patients with mild dementia (MMSE range 20-26) at baseline. The NTB tends to exhibit efficacy in patients with moderate impairment (baseline MMSE of c.14-20), though the ADAS-cog seems to exhibit greater sensitivity to cognitive change in this population.[10] This may be due to the tendency of patients in the moderately impaired cohort to show impairment on more than just the episodic verbal memory components of the ADAS-cog. The ADAS-cog and NTB are often selected for use in the same studies, which runs the risk of imposing an unnecessary testing burden on study participants as there is a good deal of

Table 2.

Stability of NTB metrics and composites

NTB test measure	Alzheimer's disease patients	Mild Cognitive Impairment	Healthy volunteers
NTB Overall	0.22	0.26	0.31
NTB All Memory	0.28	0.35	0.44
NTB Immediate Memory	0.35	0.42	0.52
NTB Delayed Memory	0.34	0.39	0.48
NTB Executive Function	0.27	0.29	0.27
Category Fluency	2.36	2.41	3.22
COWAT	4.21	5.03	4.84
Rey Auditory Verbal-Immediate	4.35	5.91	7.06
Rey Auditory Verbal-Delayed	2.5	3.16	3.18
Digit Span	4.35	5.91	7.06
Wechsler Verbal PA-Immediate	2.08	2.68	2.6
Wechsler Verbal PA-Delayed	1.00	0.93	0.91
Wechsler Visual PA-Immediate	2.37	2.93	2.88
Wechsler Visual PA-Delayed	1.16	0.99	0.85

redundancy in this approach.[16] Both the NTB and the ADAS-cog feature two measures of episodic verbal memory and, as discussed above, the removal of two or more of these episodic verbal memory tests can significantly reduce the testing burden to patients. An analysis of baseline and change from baseline data has the potential to determine whether these measures are truly associated measures of the construct episodic verbal memory and to determine which of the four measures is most sensitive to impairment, exhibits the most robust psychometric characteristics and the greatest sensitivity to drug effects.

Summary and conclusions

Contemporary versions of the NTB now commonly include measures of working memory, attention, and psychomotor speed in addition to the more traditionally measured functions such as memory, praxis, and language. Recently, tests have been drawn from computerized cognitive testing systems.[17] Consequently, the newer NTB variants are typically hybrid systems, featuring traditional "paper-and-pencil" as well as computerized measures, selected according to merit. Several ongoing trials feature NTBs of this kind and an example of their typical composition is shown in Table 3.

The past 10 years have yielded valuable information with respect to the measurement of cognition in AD clinical drug trials. This experience has helped us to evaluate the value of the tasks that comprise the NTB and to refine its content. Contemporary versions of the NTB have tended to cover a greater number of cognitive domains than either the ADAS-cog or the original NTB and, through judicious test selection, reduce the administration time to more acceptable levels. Experience has also permitted us to advise regarding the removal of redundant items and to improve traditional components to reduce the chances of floor performance. A clear trend in contemporary AD trial design is to evaluate patients in the very earliest stages of the disease and even in the prodromal phase. In this context is worth mentioning that the NTB with the composition shown in Table 3 will also likely be resistant to ceiling effects, a further important consideration. Variants of the NTB have found use in pharmaceutical trials, nutraceutical trials,[18] and in clinical research programs. Ongoing programs of research are seeking to capture information with respect to NTB task characteristics. For example, unreliable measures can often yield substantial variability in performance and so researchers such as Suzanne Hendrix[19] are using historic datasets to characterize the impact of this variability. Others are analyzing historic data in a bid to determine which NTB measures best predict cognitive change as AD progresses, as well as seeking to establish which NTB tests tend to demonstrate efficacy for different drug classes.

A good deal of information is now available for both the original NTB and the more recent NTB variants. It seems likely that considerable further information will be published and that these data will further inform us with respect to how best to test for efficacy in the cognitive domains known to be impaired in AD.

For more information, please contact:
J Harrison, Park House, Kilmington Common, Wiltshire, BA12 6QY, UK
e-mail: john@metiscog.com

REFERENCES

1. Harrison JE, Minassian, SL, Jenkins L. The NTB: A Neuropsychological Test Battery for use in Alzheimer's disease Clinical Trials. *Archives of Neurology* 2007; 64(9):1323-9.
2. Fowler K, Saling MM, Conway EL et al. Computerized neuropsychological tests in the early detection of dementia: prospective findings. *Journal of the International Neuropsychological Society* 1997;3(2):139-46.
3. Harrison JE, Kirby L, Baumel B et al. A double-blind, placebo-controlled study of Phenserine in Alzheimer's disease. Presented at 8th Annual Montreal/Springfield Symposium on Alzheimer's disease, Montreal, 2004.
4. Szlyk JP, Myers L, Zhang YX et al. Development and assessment of a neuropsychological battery to aid in predicting driving performance. *Journal of Rehabilitation Research and Development* 2002;39(4):483-496.
5. Gilman S, Koller M, Black RS et al. Clinical effects of A-beta immunization (AN1792) in patients with AD in an interrupted trial. *Neurology* 2005;64:1553-1562.

Table 3.

A contemporary example of an NTB

Test	Domain	Outcome metric
Rey Auditory VLT	Episodic memory	Immediate recall
Controlled Oral Word Association Test	Executive functions (planning, strategy, working memory)	Total number of acceptable words
Category Fluency Test	Executive function	Total number of acceptable words
Reaction time tasks	Psychomotor speed and attention	Latency in milliseconds
One back memory tests	Working memory and episodic visual memory	Number of correct responses

6. Lannfelt L, Blennow K, Zetterberg H et al. Targeting A as a modifying therapy of Alzheimer's disease: safety, efficacy and biomarker findings of a Phase IIa Randomised, Double-Blind Placebo-Controlled trial of PBT2. *Lancet Neurology* 2008;7(9):779-786.
7. Faux NG, Ritchie CW, Gunn, A. et al. PBT2 rapidly improves cognition in Alzheimer's Disease: additional phase II analyses. *Journal of Alzheimer's Disease* 2010;20:509-516.
8. Mohs R and the ADCS (1997) Development of cognitive instruments for use in clinical trials of anti-dementia drugs: additions to the Alzheimer's Disease Assessment Scale (ADAS) that broaden its scope. The Alzheimer's Disease Cooperative Study. *Alzheimer's Disease and Associated Disorders* 1997;11: S13 – S21.
9. Hilt D, Safirstein B, Hassman D, et al. EVP-6124 - Safety, Tolerability and Cognitive Effects of a Novel alpha7 Nicotinic Receptor Agonist in Alzheimer's Disease Patients on Stable Donepezil or Rivastigmine Therapy. ICAD, Vienna, Austria, 2009.
10. Black R, Harrison J, Li D et al. A comparison of Neuropsychological Test Battery (NTB) and ADAS-cog based on Alzheimer's disease clinical trials. Presented at ICAD, Vienna Austria. 2009.
11. Rentz D, Eng E, Harrison J et al. The Neuropsychological Test Battery (NTB) demonstrates high test-retest and Inter-rater reliability. Presented at CTAD, Las Vegas, 2009.
12. Harrison JE. Commentary on the proposed new guidelines for medicinal products for the treatment of Alzheimer's disease and other dementias. *CRFocus* 2007;18(11):24-7.
13. Borkowski JG, Benton AL, Spreen O. Word fluency and brain damage. *Neuropsychologia* 1967;5:135-140.
14. Mimica N, Žaki Milas D, Joka S et al. A Validation Study of Appropriate Phonological Verbal Fluency Stimulus Letters for Use with Croatian Speaking Individuals. *Collegium Anthropologicum* 2011;35 (Suppl. 1):235-8.
15. Shen JHQ, Yu H, Shen Q et al. Alzheimer's disease instrument validation study in Asia. P4-170. Presented at ICAD, Paris, 2011.
16. Harrison JE. Measuring Cognitive Change in Alzheimer's disease Clinical Drug Trials. *Journal of Nutrition, Health & Aging* 2007;11(4):327-9.
17. Harrison JE, Maruff, P. Measuring the mind: Assessing cognitive change in clinical drug trials. *Expert Reviews in Clinical Pharmacology* 2008;1(4):471-3.
18. Vellas B, Kamphuis P, Harrison JE et al. The Souvenaid II study: Rationale and study design. P4-245. Presented at ICAD, Paris, 2011.
19. Hendrix SB. Current clinical development strategy: What have we learned in the past five years? Presented at CTAD, Toulouse, 2010.

NEWS FROM...

Linguistic Validation of Quality of Life Instruments Intended for Children in Russia

Svetlana Koudria, Olga Sheinina, Sergei Varshavsky

While HRQOL instruments intended for children often are similar in content and structure to those used with adults, our experience of testing children in Russia shows that the construct of "health-related quality of life" is strikingly different in children and adults. Careful observation of linguistic behavior of over 50 children during the face-to-face interviews using several questionnaires undergoing linguistic validation from English into Russian led to the following findings.

First, children regularly answer items about difficulties in terms of their personal likes and dislikes. A child suffering from asthma would say that s/he has difficulty walking but has no difficulty running—just because s/he does not like to walk, but likes to run.

Second, the parents' role in a child's health-related quality of life seems sometimes underestimated in pediatric questionnaires, at least as it pertains to Russian realities. In Russia, a parent often stands between the child and the community that treats the child's disease; therefore, parents' influence should be considered one of the major health-related quality of life factors. For example, for a child, the notion of being "limited in activities due to disease" regularly boils down to limitations imposed by parents, and not imitations caused by the disease. When interpreting the "limited in activities due to disease" items, the majority of children spoke about what they are not allowed to do, and not about what they appear unable to do because of their condition. This difference is very important for measuring HRQOL.

There is yet another observation related to the role of parents: namely, the parents can greatly influence children's vocabulary. In Russia, some parents intentionally avoid the name of the diagnosis when talking to their child so that the child could develop and socialize normally without thinking of himself or herself as of "ill," "asthmatic," or disadvantaged in any other way. This is a reason why some children may not know what their condition is called (e.g., "asthma"), and consequently are unable to explain this word, whereas others are.

Also, we found good evidence supporting the existing linguistic notion that children of pre-school and early school age cannot generalize their experience verbally—they operate with specific examples and facts, and they are only learning to generalize at school. Consequently, concepts like "asthma," "asthma symptoms," "limited," "activities," "on average," and "shortness of breath"— words denoting generic concepts—might be obscure for many children. That is why instead of "asthma" a 6-10 years old child typically would say "I coughed," and s/he is not able to explain what "asthma" means because "asthma" is, in essence, a verbal and academic generalization.

Furthermore, we observed that children often notice only the most salient clinical signs. For example, they would notice "coughing"—for asthma, "running nose"—for rhino-conjunctivitis. However, children would not—and this indeed seems to be a tendency—pay attention to such a sign as wheezing. Wheezing is a salient clinical sign for a doctor because it helps to diagnose the condition but it seems that for a child wheezing is just a part of difficult breathing, and, therefore, a child would not distinguish wheezing as a separate distinctive sign of his or her disease. Consequently, a child may not know what exactly "wheezing" refers to and how to assess the effect of wheezing on his/her life.

Last but not least, we observed that many children have no words/expressions to describe certain disease-specific concepts because they did not experience the corresponding symptoms, or experienced them at their "pre-speaking" age. For example, we interviewed a 6-year-old boy who had the most severe asthmatic experiences at the age of 3, when he could not yet describe them. So he had not had time to shape his experience into words. Therefore, he could not make any comments on severe asthmatic experiences—although he'd been through it.

We hope that the above findings will allow us to better understand the specific demands of the cultural adaptation and, maybe, development of pediatric questionnaires. We believe that cultural adaptation of the existing questionnaires should begin with studying children's way of speaking about their disease and with examining the role parents have in the communications between the doctors and children within the given culture. We think that translatability assessment is particularly appropriate while working with the instruments intended for children. This may include but is not limited to reviewing by pediatricians and parents as the first step of the validation process. We are planning to continue our research on pediatric questionnaires to examine what other factors should be taken into consideration and to find out which steps can be taken to improve our standard validation methodology for QoL instruments intended for children.

Olga Sheinina
Director, Patient Reported Outcomes – Cultural Adaptation and Linguistic Validation
Tel: +1 202 441 6482 cell – Fax: +1 641 354 1852
http://www.preference-pro.com

"Glass Half Full or Glass Half Empty": the Youth Quality of Life (YQOL) Generic and Condition-Specific Modules

Donald Patrick, PhD, MSPH[1,2], Todd Edwards, PhD[1,3], Anne Skalicky, MPH[1,4]

[1] Seattle Quality of Life Group
[2] Director, Seattle Quality of Life Group, and Professor, Departments of Health Services and Epidemiology, University of Washington, Seattle, WA, USA
[3] Research Assistant Professor, Department of Health Services, University of Washington, Seattle, WA, USA
[4] Research Scientist, University of Washington, Seattle, WA, USA

KEYWORDS
QUALITY OF LIFE, CHILDREN, ADOLESCENTS, YOUTH-REPORT, PARENT-REPORT

Abstract

Measures of quality of life for youth often contain questions that ask "how empty is your glass?" and thus assess what is wrong with the health and life of youth. Asking "how full is your glass?" or what is going right for youth provides a balanced perspective toward improving their lives or "filling up the glass." The authors propose a combination of a generic measure with positive items and condition-specific modules with deficit items for assessing youth quality of life (QoL). We describe youth-report and observer-reported instruments based on the conceptual models that guided their development.

Introduction

Measures of quality of life for youth often contain questions that ask "how empty is your glass?" and thus assess what is wrong with the health and life of youth. Asking "how full is your glass?" or what is going right for youth provides a balanced perspective toward improving their lives to "fill up the glass". We propose a combination of a generic measure with positive items and condition-specific modules with deficit items for full assessment of youth quality of life (QoL).

Adolescents present formidable challenges to QoL assessment in finding the most important and relevant domains, in conceptualizing their thinking and feelings to indicate the "positive" and "negative" aspects of their lives, and in measuring these perceptions with a reliable and valid instrument. Measures of mortality, morbidity, and youth behavioral risks are important in tracking health trends and in identifying social, cultural, and economic differences[1]. Self-reported QoL measures, both generic and condition specific, add the critical voices of the youth themselves and provide the means for comparing the perceived well-being of special populations such as children and adolescents with disabilities or special healthcare needs, indigent youth, and high-risk teens including adolescents with depression, behavioral challenges, and in living in disadvantaged environments.

Who is SeaQoL? How did the Youth Quality of Life Development Arise?

The Seattle Quality of Life Group at the University of Washington was organized in the early 1990s to develop the Youth Quality of Life instrument, a generic instrument the Youth Quality of Life Instrument - Research Version (YQOL-R). When the group began, the Centers for Disease Control requested development of a Youth QoL measure that was relevant to youth with and without disabilities and did not equate QoL with functional status. Equating QoL with function can work against youth with disabilities, many of whom may enjoy a high quality of life when using a wheelchair and/or many other wide-ranging accommodations for functional challenges. Subsequent support came from many public and private sources to the University of Washington to further work on the generic instrument and to create condition-specific modules, all of which are in the public domain (see acknowledgements section). To obtain all instruments as well as available translations and cultural adaptations, go to www.seaqolgroup.org.

Methods

Conceptual Approach to Quality of Life

Contemporary guidelines for measurement development[2,3] guided the design and conduct of the study.

Early decisions were made regarding the development of the YQOL instruments after extensive reviews of the adolescent health-related quality of life literature, which revealed a shortage of instruments that met criteria considered essential for the understanding and assessment of adolescent QoL,[4] generally, and in a number of condition-specific populations (see below). We defined QoL as a subjective evaluation of one's own life in line with the *needs-based* model that identifies QoL as the degree to which most or all human needs are met. The needs-based approach has multiple origins rising from Maslow's needs hierarchy[5,6] and pioneering papers by Hornquist,[7] and Hunt and McKenna.[8] Of particular influence was the World Health Organization's WHOQOL Group, who defined QoL as an evaluation of one's *position in life.* Position in life means the youth's perceived status within his or her social, economic, and cultural environment and based on his or her values, concerns, and standards.[9] Another important influence was the work of Raphael, whose conceptualization of youth QoL as *being, belonging, and becoming* rang true as our qualitative development ensued.[10] Thus, the items comprising the YQOL instruments have been selected to represent the areas of greatest importance to youth in relation to their own values and goals in moving through the important life transition from childhood to adulthood.

As instrument development ensued, and work began in earnest across the life course from birth through early adulthood, different perspectives were adopted depending on the measurement context (purpose) and source of the report. Where possible, the SeaQoL group emphasizes self-reporting. With younger age respondents, where high variability exists in the development of abstract reasoning and ability to self-report reliably, the perspectives of parents, caregivers, clinicians, teachers, and peers are valuable and necessary. Thus the source of report became an added measurement consideration.

Typology of Instrument Content

As shown below, YQOL instruments have the following four types of items listed by source of report:

I. **Youth self-report, perceptual items (sensations and feelings)**. Perceptual items are known only to and reported only by the youth themselves. By definition, the SeaQoL group does not use any parent report of youth perceptual items. Perceptual items are asked only of youth ages 11-18 years. Items may be feelings, symptoms, or general perceptions.[4,11] For example, how youth see themselves, feel about their medical or other condition, view their life, or feel about their family, friendships, school or neighborhood. These perceptual items are contained in both generic and condition-specific modules.

II. **Youth self-report, contextual items (behaviors and events)**. These items are self-reports of behaviors or events. Potentially verifiable by others, these reports may or may not be related to the youth's perceptions of her or his quality of life. For example, reports of missing out on activities may or may not be correlated with feelings of social inclusion.

III. **Observer-reported, perceptual items (perceptions).** These items are reports of perceptions or attitudes made by a parent or other observer for children birth through adolescence. These items are reports of how parents or observers view infant, child, or youth. For example, peer or parent judgment of the facial differences of youth are perceptual and known only to those making the judgments.

IV. **Observer-reported, contextual items (signs, behaviors, and events).** These items are reports of signs, behaviors and events made by a parent or other observer for children birth through adolescence. These items are reports of what an infant, child, or youth did or did not do, and events that occurred over a specified recall period. These can also be actual observations made during pre-specified periods, usually by diary, and reported according to frequency and sometimes intensity. These reports may be based on observations by the parent, but they may also be based on the child or youth telling the parent about events, or information gleaned from other adults such as teachers. For example, parent observation of communication challenges or antisocial behavior or the frequency of coughing during the past 24 hours along with the intensity of the cough.

Description of instruments

(see Table 1)

1. Youth Self-Report, Perceptual items (sensations and feelings)

A. **Generic Quality of Life:** The YQOL generic instrument was developed using three types of data: (a) in-depth interviews with youth ages 11-18 with and without disabilities, residing in from many different settings, including homeless youth, and asking them what was important to their life; (b) focus groups with youth ages 11-18 with and without disabilities, with primary caregivers of youth with and without disabilities, and with youth health and welfare professionals; and (c) consultation with experts and a survey of existing assessment instruments, such as the National Longitudinal Study of Adolescent Health.[12] To the maximum possible extent, the content of the items was defined by adolescents themselves with items written in their own language and colloquial expressions.

• *Youth Quality of Life Instrument – Research version (YQOL-R).* The YQOL-R is a self-administered questionnaire with two types of items, contextual (i.e., can be reported by others), and perceptual (i.e., known only to the youth themselves). The YQOL-R Contextual items are discussed in section III below. Table 1 shows the modular applications which can be used in addition to the generic YQOL instruments. In addition to the 19 individualized facets and a total score, four broader domains measured by 41 perceptual items have been identified from the YQOL-R: Sense of Self, Social Relationships, Culture and Community, and General Quality of Life. "*I can handle most difficulties that come my way"* is an item example from the Self domain; "I feel I am *important* to others" and "I feel I am *getting* a good education" are, respectively, items from the Relationship and Culture and Community domains. The YQOL scores are transformed to a 0-100 scale for ease of interpretation, with higher scores indicating better QoL.

• *Youth Quality of Life Instrument – Short Form (YQOL-SF).* A 10-item short form that is representative of the longer YQOL-R is currently being developed using Rasch analysis. When completed by 2012, this form will be convenient for co-administration with the condition-specific modules as a means for comparing generic QoL across populations of adolescents with varied chronic conditions and disabilities.

• *Youth Quality of Life Instrument – Surveillance version (YQOL-S).* The YQOL-S is comprised of eight items selected from the questions on the YQOL-R to reflect issues of most likely importance to policy makers, rather than to be representative of the YQOL-R as a whole. The YQOL-S is designed for monitoring leading indicators of QoL in adolescent populations, and is not scored by domain, as each question is regarded as a social indicator in itself. The YQOL-S has been used to examine the relationship between QoL and health risk behavior,[13] and mobility disability,[14] and is an appropriate tool for assessing and monitoring QoL indicators in diverse adolescent populations. It requires only one minute to complete and can be easily added to ongoing school-based or other surveys, such as the Youth Risk Behavior Survey.[15]

B. **Generic Quality of Life – Individualized Measures.** Using all items in the YQOL-R these two measures assess the generic areas of life that are most important to youth and areas that they most want to change: (1) "You have just answered some questions about how your life is now. Which areas listed below are most important to you? Youth are asked to select five areas in their life which are most important to them, ranked in order of importance, and five areas in their life which they would like to change for the better, ranked by order of importance. A list of 19 areas or facets of quality of life are presented to youth ranging from community and culture facets of "being safe" or "going to a good school and learning" to relationships facets like "getting along with my family" and "having good friends" to sense of self facets of "looking good" and "believing in myself".[17]

C. **Condition-specific modules:** The YQOL-condition-specific instruments have been developed as modules to be used in conjunction with the generic YQOL instrument, depending on the objective of assessment.

• *Youth Quality of Life Instrument – Facial Differences module (YQOL-FD).* The YQOL-FD is a craniofacial-specific quality of life instrument designed for youth with

Table 1.

Overview of SeaQoL Group Instrument Development*

Youth Self-Report, Perceptual	Domains Measured	# Items	Scores	Purpose
Youth Quality of Life Instrument – Research Version (YQOL-R)	Sense of Self, Social Relationships, Culture & Community, General QoL	41 perceptual, 2 individual	Domain, total, 2 individualized	Research & Intervention Studies
Youth Quality of Life Instrument – Short Form (YQOL-SF)	Generic QoL	10	Total	Research & Intervention Studies
Youth Quality of Life Instrument – Surveillance (YQOL-S)	Policy-relevant generic QoL	8	Individual items	Population Surveillance
Youth Quality of Life Instrument – Facial Differences Module (YQOL-FD)	Coping, Positive Consequences, Stigma, Negative Consequences, Negative Self-Image	30	Domain	Research & Intervention Studies
Youth Craniofacial Surgery Attitudes Measures (CSAM)	Attitudes toward craniofacial surgery: past & future surgery	20	Domain	Research & Intervention Studies
Youth Quality of Life Instrument – Weight Module (YQOL-W)	Self, Social, Environment	21	Domain, total	Research & Intervention Studies
Youth Quality of Life Instrument – Deafness and Hard-of-Hearing Module (YQOL-DHH)	Participation, Self-Acceptance/Advocacy, Stigma	32	Domain	Research & Intervention Studies
Cystic Fibrosis Symptoms and Impacts (CFRSD)	Cystic Fibrosis respiratory symptoms/signs, emotional impacts, activity impacts	16	Under development	Research & Intervention Studies
Youth Self-Report, Contextual				
Youth Behaviors and Events (Y-ROBE)	Generic youth behaviors and events	15	Individual indicators	Research & Intervention Studies
Youth Facial Differences Behaviors and Events (FD-YROBE)	FD-specific behaviors and events	18	Individual indicators	Research & Intervention Studies
Youth Deafness and Hard-of-Hearing Behaviors and Events (DHH-YROBE)	Deafness and hard-of-hearing-specific behaviors and events	28	Individual indicators	Research & Intervention Studies
Youth Disability Screener (YDS)	Disability status	4	Total	Screening
Lymphatic Malformation Function (LMFA-Y)	LM-specific function	25	Under development	Research & Intervention Studies
Observer-Report, Perceptual				
First Impressions Ratings	Social judgments of others' personal attributes, social attributes, appearance and intelligence	26	Domain	Research & Intervention Studies
Observer-Report, Contextual				
Generic Youth Behaviors and Events (OROBE)	Generic youth behaviors and events	15	Individual indicators	Research & Intervention Studies
Youth Facial Differences Behaviors and Events (FD-OROBE)	Facial differences-specific behaviors and events	20	Individual indicators	Research & Intervention Studies
Youth Deafness and Hard-of-Hearing Behaviors and Events (DHH-OROBE)	DHH-specific behaviors and events	19 (ages 5-7)-23 (ages 8-10)	Individual indicators	Research & Intervention Studies
Cystic Fibrosis Signs and Impacts (CFRSignD)	CF respiratory symptoms/signs, emotional impacts, activity impacts	12	Under development	Research & Intervention Studies
Lymphatic Malformation Function (LMFA-O)	LM-specific function	25 items	Under development	Research & Intervention Studies
ADHD Behaviors and Events (ADHD-OROBE)	ADHD-specific behaviors and events	19 items	Individual indicators	Research & Intervention Studies

*To access SeaQoL instruments visit www.seaqol.org

congenital *and* acquired facial differences ages 11-18. Thirty perceptual items from the YQOL-FD assesses five domains: negative consequences of having a facial difference, negative self image, experienced stigma, positive consequences of having a facial difference, and coping.[16] An example of an item is "*I feel uncomfortable meeting people for the first time because of how my face looks.*"

• *Youth Quality of Life Instrument – Weight module (YQOL-W).* The YQOL-W is a 21-item weight-specific quality of life instrument designed for youth 11-18 to evaluate how they view their weight in relation to their quality of life. The instrument development was conducted in three phases: Phase I, item generation and selection; Phase II, psychometric validation; and Phase III, pilot study to determine ability to detect change (responsiveness). An example item is "*Because of my weight I worry about what people say about me.*" The instrument was simultaneously developed in both US English and Mexican Spanish.[17] The YQOL-W has been used to evaluate weight change and QoL in youth enrolled in immersion treatment programs in weight-loss camps.[18] A version is being used in treatment evaluation in a clinical setting with youth and family members (information available from the authors).

• *Youth Quality of Life Instrument – Deafness and Hard-of-Hearing module (YQOL-DHH).* The YQOL-DHH is a 32-item self-administered, condition-specific quality of life measurement designed for all youth ages 11-18 who are deaf or hard of hearing. The items on the YQOL-DHH were developed with the participation of DHH youth and experts. An example item is *"Because I am deaf or hard-of-hearing, I have to work harder than other youth to do the things I want to do."* At this point there is also an American Sign Language translation available for administration through DVD with an accompanying answer booklet in US English.[19]

• *Craniofacial Surgery Attitudes Measure (CSAM).* The 20-item CSAM assesses two domains: attitudes regarding past surgical experience and attitudes toward planned or potential future surgeries, including perceived need for more surgery. The measure is being further developed and is available from the authors.

• *Cystic Fibrosis Respiratory Symptoms Diary (CFRSD).* We have developed a CF-specific respiratory symptom diary for use in clinical trials and clinical care using standard qualitative methods. The CFRSD is designed for self-report, ages 12 years through adulthood.[20] In total, the new CF symptom diary contains 16 questions, with eight specifically addressing respiratory symptom severity and frequency. Additional questions address activity and emotional impacts of these symptoms. The CFRSD is being used in numerous clinical trials of pharmacologic treatments for CF. The CFRSD is currently in development for smart phone administration and testing for changes in frequency of administration in relation to changes in disease state.

• *Lymphatic Malformation Functional Assessment (LMFA).* **The LMFA** is a 25-item measure which assesses functional status related to lymphatic malformations of the head and neck in patients ages 0-19 years, self-report ages 11-19. This measure has just begun validation.

II. Youth Self-Report, Contextual items (behaviors and events)

• *Youth Report of Generic Behaviors and Events (YROBE).* Self-report of generic behaviors and events considered important by youth ages 11-18. Youth-ROBE includes 15 individual items that can be used with the YQOL-R perceptual measures.

• *Youth Report of Facial Differences Behaviors and Events (FD-YROBE).* The FD-ROBE includes 18 items for self-report of craniofacial-specific behaviors and events considered important by youth with craniofacial malformations ages 11-18. These items can be used with the YQOL-FD facial difference-specific perceptual instrument.

• *Youth Report of Deafness and Hard-of-Hearing Behaviors and Events (DHH-YROBE).* The DHH-ROBE includes 28 items for self-report of behaviors and events considered important by youth with hearing impairments ages 11-18. This measure is currently submitted for publication.

• *Youth Disability Screener (YDS).* The YDS is a 4-item screener for obtaining self-reported disability status from youth ages 11-18, partly based partly on the 1994 National Health Interview Survey on Disability (NHIS-D),[21] and on the Questionnaire for Identifying Children with Chronic Conditions,[22] both of which are parent reported. This measure is available on our website.

• *Youth Report of Lymphatic Malformation Function (LMFA-Y).* The 25-item LM-Function measures youth report of the functional impacts of head and neck lymphatic malformations most important to adolescents with LM. The instrument is currently under development.

III. Observer-Report, Perceptual items (perceptions)

• *First Impressions Rating Scale (FIRS).* The FIRS is a 26-item semantic differential measure that quantifies the social judgments made by people of youth with facial differences. Severity of facial appearance significantly impacts social attribute perceptions, possibly resulting in stigma and altered social experiences.[23]

IV. Observer-Report, Contextual items (signs, behaviors, and events)

• *Observer Report of Generic Youth Behaviors and Events (OROBE).* The OROBE is a 15-item observer-report of generic behaviors and events considered important by youth ages 11-18. Observers are usually either teachers or parents.

• *Observer Report of Facial Differences Behaviors and Events (FD-OROBE).* The FD-OROBE is a 20-item observer-report of craniofacial-specific behaviors and events considered important by youth with craniofacial malformations ages 11-18. Observers are clinicians, teachers, or parents.

• *Observer Report of Deafness and Hard-of-Hearing Behaviors and Events (DHH-OROBE).* The DHH-OROBE is a 19-item (5-7 year olds) or 23-item (8-10 year olds) observer-report of hearing-specific behaviors and events considered important by youth with hearing impairments ages 5-10[24]. Observers in this instance are parents.

• *Cystic Fibrosis Respiratory Sign Diary (CFRSignD).* We are currently developing a 12-item CF-specific respiratory sign diary for use in clinical trials and clinical care using standard qualitative methods. The CFRSignD is designed for observer-report, ages 0 through 11 years. Observers are parents.

• *Observer Report of Lymphatic Malformation Function (LMFA-O).* The

25-item LM-Function measures observer report of the functional impacts of head and neck lymphatic malformations most important to adolescents and parents of young children, with LM. The instrument is currently under development.

• *Observer Report of Attention-Deficit Hyperactivity Disorder Events and Behaviors (ADHD-OROBE).* The 19-item ADHD-OROBE measure was designed to assess the everyday impact on the family of living with a child who has ADHD including the frequency, time-of-day, and bothersomeness of observable ADHD behaviors and events, and the amount of time free from these experiences.[25] Observers are parents.

Discussion

Context or Purpose of assessment (Research, Surveillance, Clinical Practice).

Most applications of the YQOL instruments have been in research contexts to understand group-level attributes and relationships of quality of life to other variables.[13,26] In several cases, YQOL instruments have been used in interventions to examine and identify changes in quality of life.[18,27,28]

The YQOL-S has been used to examine the relationship between QoL and health-risk behavior,[13] and mobility disability,[14] and is an appropriate tool for assessing and monitoring QoL indicators in diverse adolescent populations. It requires only one minute to complete and can be easily added to ongoing school-based or other surveys, such as the Youth Risk Behavior Survey.

The YQOL-SF 10-item version (current versions of the SF are available for use) of the generic is highly recommended for use with the condition-specific modules. Comparisons across populations are thus possible, as well as assessment addressing the specific concerns of youth with particular conditions. Instruments such as the YQOL-W (Weight) apply to all youth regardless of their actual weight, since the QoL issues are highly related to body shape and perceptions. The YQOL-DHH (Deaf and Hard of Hearing), however, apply only to those with hearing problems, even though communication in general is a challenge for all adolescents, hearing or otherwise.

YQOL instruments should be chosen based on the purpose of assessment, target population, the measurement properties (reliability, validity, interpretation, and practical considerations) of self- or observer-report.

Challenges of moving self-report down the age range.

Some developers have encouraged self report below age 11, using items that are more concrete and different aids to response, such as smiley faces or circles of different sizes to indicate quantity of feeling.[29,30] The measurement properties of such applications may meet recommended measurement properties in *some* populations. The SeaQoL Group has found wide variation, however, in development among youth, primarily using cognitive interviews and readability testing. Some younger children (age 8 or 9) may understand items as intended when asked to "think aloud," but other older youth (ages 10 and above) may not have similar understanding. The wide range of development suggests the importance of cognitive interviewing, readability, and usability testing.[31]

Future directions

Application of Modern Test Theory.

Similar to adult measures, applications of modern test theory, i.e., item-response theory (IRT) and Rasch modeling are being applied to child and youth outcome measures. The PROMIS (Patient-Reported Outcomes Measurement Information System) initiative funded in the US32 is systematically developing instruments applying item-response theory with pediatric populations. The SeaQoL group is using Rasch modeling to assure that the unidimensional YQOL-SF represents the full range of items across the generic QoL continuum. Future work will be conducted using both IRT and Rasch to explore dimensionality of youth measures and to help calibrate different measures.

Cross-cultural applications.

The YQOL-W (Weight) was developed in the US with white, African-American, and Latino youth and in Mexico. Cross-cultural comparison among youth in these different cultural contexts is currently being conducted and reported. A major effort is underway to adapt YQOL measures in Mandarin Chinese, with both qualitative and quantitative work in China.

The future of youth quality of life assessment parallels the challenges and promise of adolescence itself. As the field becomes more varied and evidence accumulates using different instruments in different cultural and economic contexts, many possibilities exist. Perhaps nothing is more challenging or important in identifying how some youth remain resilient in the face of physical, economic, and social hardship, while other youth do not. To "fill up the glass" requires systematic attention to the social determinants of youth QoL, to the burgeoning field of genomics, and to personality studies. Most of all, researchers and users must listen to the youth themselves, involving them in every stage of research and constantly checking with them about the relevance, the language, the intent, and the use.

For more information, please contact:
Donald L. Patrick, PhD, MSPH
Seattle Quality of Life Group
Department of Health Services
University of Washington
4555 Brooklyn Ave NE
Seattle, WA 98195
donald@uw.edu

ACKNOWLEDGEMENTS:
Work funded by grants to Donald L. Patrick, PhD, MSPH (PI) and University of Washington from the National Institute of Deafness and Communication Disorders (R01 DC008144); National Institute of Diabetes and Digestive and Kidney Diseases (Grant #R01 DK071101); R01 DE13546 from the National Institute of Dental and Craniofacial Research, contract from Eli Lilly and Company; Cystic Fibrosis Foundation research grant (GOSS05A0); the Leroy Matthew's Physician Scientist Award; NIH (RR-00037-39) and NIH/NHLBI (K23 HL72017); Cooperative agreement with the Centers for Disease Control and Prevention (#U48/CCU009654 and CCU006992-06); the Sie-Hatsukami Research Endowment; and a grant awarded to the Klamath Tribes by the Substance Abuse and Mental Health Services Administration (#1 HD4 SP0770D).

REFERENCES

1. Truman BI. Rationale for Regular Reporting on Health Disparities and Inequalities — United States, 2011. In: Prevention CfDCa, ed., Morbidity and Mortality Weekly Report. Wash, DC: U.S. Department of Health and Human Services, 2011.
2. Atkinson MJ, Lennox RD. Extending basic principles of measurement models to the design and validation of Patient Reported Outcomes. *Health Qual Life Outcomes.* 2006; 4: 65.
3. Solans M, Pane S, Estrada MD, et al. Health-related quality of life measurement in children and adolescents: a systematic review of generic and disease-specific instruments. *Value Health.* 2008; 11: 742-64.

4. Edwards TC, Huebner CE, Connell FA, et al. Adolescent quality of life, part I: conceptual and measurement model. *J Adolesc.* 2002; 25: 275-86.
5. Maslow AH. A Theory of Human Motivation. *Psychological Review.* 1943; 50: 370-96.
6. Doyal L. Thinking About Human-Need. *New Left Rev.* 1993: 113-28.
7. Hornquist JO. The concept of quality of life. *Scand J Soc Med.* 1982; 10: 57-61.
8. McKenna SP, Hunt SM. A new measure of quality of life in depression: testing the reliability and construct validity of the QLDS. *Health Policy.* 1992; 22: 321-30.
9. World Health Organization (WHO). The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization.PG - 1403-9. *Soc Sci Med.* 1995; 41.
10. Raphael D, Rukholm E, Brown I, et al. The Quality of Life Profile--Adolescent Version: background, description, and initial validation. *J Adolesc Health.* 1996; 19: 366-75.
11. Patrick DL, Edwards TC, Topolski TD. Adolescent quality of life, part II: initial validation of a new instrument. *J Adolesc.* 2002; 25: 287-300.
12. Udry JR. The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994-1996. (Data Sets 48-50, 98, A1-A3, Kelley, M.S. & Peterson J.L.) In: Carolina Population Center UoNC, ed. Los Altos, CA: Sociometrics Corporation, American Family Data Archive, 1998.
13. Topolski TD, Patrick DL, Edwards TC, et al. Quality of life and health-risk behaviors among adolescents. *J Adolesc Health.* 2001; 29: 426-35.
14. Edwards TC, Patrick DL, Topolski TD. Quality of life of adolescents with perceived disabilities. *J Pediatr Psychol.* 2003; 28: 233-41.
15. Washington State Dept of Health. Healthy Youth Survey, Form B Youth Risk Behavior Survey Centers for Disease Control and Prevention. Olympia, WA: Available at: www.cdc.gov/yrbs. , 2008.
16. Edwards TC, Patrick DL, Topolski TD, et al. Approaches to craniofacial-specific quality of life assessment in adolescents. Cleft Palate Craniofac J. 2005; 42: 19-24.
17. Morales LS, Edwards TC, Flores Y, et al. Measurement properties of a multicultural weight-specific quality-of-life instrument for children and adolescents. *Qual Life Res.* 2011; 20: 215-24.
18. Patrick DL, Skalicky AM, Edwards TC, et al. Weight loss and changes in generic and weight-specific quality of life in obese adolescents. *Qual Life Res.* 2011; 20: 961-8.
19. Patrick DL, Edwards TC, Skalicky AM, et al. Validation of a quality-of-life measure for deaf or hard of hearing youth. *Otolaryngol Head Neck Surg.* 2011; 145: 137-45.
20. Goss CH, Edwards TC, Ramsey BW, et al. Patient-reported respiratory symptoms in cystic fibrosis. *J Cyst Fibros.* 2009; 8: 245-52.
21. National Center for Health Statistics. National Health Interview Survey on Disability. http://www.cdc.gov/nchs/data/nhis/dfs_100_1994pdf. Hyattsville, MD: Centers for Disease Control and Prevention, 1994.
22. Stein MB, Forde DR, Anderson G, et al. Obsessive-compulsive disorder in the community: an epidemiologic survey with clinical reappraisal. *Am J Psychiatry.* 1997; 154: 1120-6.
23. Patrick DL, Edwards TC, Topolski TD, et al. How People Perceive Youth with Facial Differences: Development and Assessment of the First Impression Rating Scale (FIRS). Seattle Quality of Life Report. Seattle, WA: http://depts.washington.edu/yqol/reports, 2011.
24. Edwards TC, Skalicky AM, Patrick DL, et al. Development of a Parent-Report Observational Measure for Assessing Communication and Social Functioning of Children who are Deaf or Hard-of-Hearing. Report from Seattle Quality of Life Group. Seattle, WA: http://depts.washington.edu/yqol/reports, 2011.
25. Seattle Quality of Life Group. Family Experience Pilot Study, Final Report. http://deptswashingtonedu/cdpr/reportshtml. Seattle, WA: Departments of Health Services and Epidemiology, University of Washington, 2002.
26. Strauss RP, Ramsey BL, Edwards TC, et al. Stigma experiences in youth with facial differences: a multi-site study of adolescents and their mothers. *Orthod Craniofac Res.* 2007; 10: 96-103.
27. Findling RL, Childress AC, Cutler AJ, et al. Efficacy and safety of lisdexamfetamine dimesylate in adolescents with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry.* 2011; 50: 395-405.
28. Janssens L, Gorter JW, Ketelaar M, et al. Health-related quality-of-life measures for long-term follow-up in children after major trauma. *Qual Life Res.* 2008; 17: 701-13.
29. Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life?: an analysis of 8,591 children across age subgroups with the PedsQL 4.0 Generic Core Scales. *Health Qual Life Outcomes.* 2007; 5: 1.
30. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care.* 1999; 37: 126-39.
31. Bevans KB, Riley AW, Moon J, et al. Conceptual and methodological advances in child-reported outcomes measurement. *Expert Review of Pharmacoeconomics & Outcomes Research.* 2010; 10: 385-96.
32. Bevans K. Progress in the Development of 10 Pediatric Instruments for Children and Youth. An update from the Children's Hospital of Philadelphia PROMIS Research Site. Bethesda, MD: University of Pennsylvania, 2011.

NEWS FROM...

The Ibero-American Chapter of ISOQOL, Brief Historical Sketch

Juan Dapueto

The Ibero-American group of ISOQOL informally met for the first time in 1999 during the 6th ISOQOL Annual Conference, in Barcelona. In August 2001, the 1st Ibero-American Meeting was held in Montevideo, Uruguay. This meeting, sponsored by ISOQOL, gathered together around 100 researchers from different countries including Argentina, Brazil, Cuba, Chile, Spain, Mexico, UK, USA, Uruguay, and Venezuela.

In the following years, two encounters took place with the participation of researchers from Argentina, Brazil, Chile, Spain, and Uruguay. The initial encounter, the second Brazilian meeting of HRQL researchers, was in São Paulo in 2002; it was the first regional educational activity. The other meeting was the first South American Workshop, held in Montevideo, Uruguay, in 2003.

In 2004, Porto Alegre hosted the 2nd Ibero-American Meeting. During the encounter, participants agreed to hold regional meetings every two years and to pursue becoming an ISOQOL official chapter. At that time, the group was recognized as a Significant Interest Group (SIG) by the ISOQOL Board of Directors. The SIG gathered officially for the first time in San Francisco's ISOQOL Meeting in 2005, and two co-chairs were elected and assigned the responsibility to lead SIG. The 3rd Ibero-American Meeting took place in Buenos Aires, Argentina in August 2006, with significant economic and scientific support of ISOQOL.
A second SIG meeting was held during ISOQOL 13th Annual Conference in Lisbon. In that meeting all participants agreed putting forward the foundation of the Ibero-American Chapter of Quality of Life Researchers, following ISOQOL directives. The Chapter bylaws were approved in 2008, during the 15th Annual Meeting of ISOQOL and 4th Ibero-American Meeting that took place in Montevideo, in 2008.
During these years, the Chapter has already held two elections of co-chairs and members of the Regional Committee. The group has kept close contact among its membership through its website and mailing list. The 5th Ibero-American meeting was held in Santiago, Chile, in September 2011. The number of papers increased as did the scientific quality of the research presented at the meetings. Hopefully, this will lead to a larger number of papers published in international journals in the next years.

The Chapter is now devoted to the organization of the 6th Ibero-American Meeting to be held in Mexico in August 2012.
This brief historical sketch provides evidence of the Chapter's ability to sustain itself and to support future activities and developments.

Dr. Juan J. Dapueto
Profesor. Director del Departamento de Psicología Médica
Facultad de Medicina - Universidad de la República – Hospital de Clínicas, piso 15 – Montevideo, Uruguay
Tel: 2487 3423

Conferences Congresses Workshops Meetings

ANNOUNCEMENTS

The Patient Reported Outcomes Methods Group of the Cochrane Collaboration

The Cochrane Collaboration Meeting will be held in Madrid, Spain on October 19-22. On Thursday, the 20th, Donald Patrick and Gordon Guyatt, co-convenors of the PROMG, will facilitate a workshop entitled "Making results of patient-reported outcomes interpretable". The Patient Reported Outcomes Methods Group Annual Meeting will take place on Friday, October 21, from 5:45 to 7:15 pm.

At present, the PROMG is focusing on methods for pooling continuous data from outcome measures in meta-analyses. A recently published article describes summarizing data as minimal important difference units. A paper in the revision stage describes 12 available approaches from a statistical point of view. Another paper, soon to be submitted, is part of series of articles describing the GRADE approach to rating and summarizing quality of evidence that has been adopted by Cochrane. The article places the 12 approaches into five categories that will be informative from the point of view of Cochrane review groups, and illustrates how each can be presented in Summary of Findings tables. Finally, the PROMG is preparing an article that expands on the information available in the Cochrane Handbook PRO chapter and describes a number of aspects related to collecting, analyzing, and summarizing data from PROs in systematic reviews and meta-analyses.

Another activity is to provide evaluations of PROs in the actual Cochrane treatment review areas and you can locate them at www.cochrane-pro-mg.org/index.html.

In the coming year, the modifications of the PRO chapter in the Cochrane Handbook for Systematic Reviews of Interventions shall begin.

If you wish to join the Methods Group, please complete the membership form on this website at www.cochrane-pro-mg.org/index.html.

The ISOQoL Translation and Cultural Adaptation special Interest Group (TCA-SIG)

The ISOQoL Translation and Cultural Adaptation Special Interest Group (TCA-SIG) is delighted to announce that its annual meeting will take place on Thursday, 27 October between 12:30 and 2:00 pm during the ISOQOL 18th Annual Conference, Denver, October 26-29, 2011, The Sheraton Denver Downtown Hotel , Tower Court C. The highlight of our annual meeting will be the following two presentations which, we are sure, will lead to a stimulating discussion:

1. "Linguistically Validating PRO measure with populations who are difficult to interview" Darren Clayson of PharmaQuest Ltd, Banbury, Oxfordshire, UK
2. "The study of different Spanish versions of the DTSQ" Annarita Felici, Health Psychology Research Ltd, University of London, UK.

We hope you will join us in Denver. Please contact Tatiana Gauchon tgauchon@mapigroup.com for more information prior to the meeting or consult the ISOQoL website www.isoqol.org

The primary goal of MAPI Research Trust's Patient Reported Outcomes Newsletter is to encourage and facilitate the rapid dissemination and exchange of information on health outcomes within the scientific community.
The views expressed in this Newsletter are those of the authors and do not necessarily represent those of MAPI Research Trust.

Call for Papers and Articles
PRO Newsletter 47

Any news and information on Patient-Reported Outcomes are welcome (e.g., short articles on on-going Quality of Life research, announcements of publications, meetings, websites, etc.)
Please refer to PRO Newsletter Online website at www.pro-newsletter.com for submission information.

Deadline for Submission:
March 1st, 2012

Please send your paper by post, fax or e-mail to Barbara Wolf, MAPI RESEARCH TRUST, 27, rue de la Villette, 69003 Lyon, France.
Fax: +33 (0)4 72 13 66 82
E-mail: bwolf@mapigroup.com

Director of Publication:
Bernard Jambon

Editors-in-Chief:
Benoit Arnould
Martine Le Gal
Marie-Pierre Emery

Guest Editor:
Catherine Acquadro

Production Editor:
Barbara Wolf

Assistant:
Mathilde Charnay

MAPI RESEARCH TRUST
27, rue de la Villette
69003 Lyon, France
Tel +33 (0)4 72 13 65 75
Fax +33 (0)4 72 13 55 73
E-mail: contact@mapi-trust.org

Patient Reported Outcomes

Please send the PRO Newsletter to:

Name: _____

Company: _____

Position: _____

Address: _____

_____ Country: _____

Tel: _____ Fax: _____ E-mail: _____

Return to: Barbara Wolf, MAPI RESEARCH TRUST, 27 rue de la Villette, 69003 Lyon, France. Fax +33 (0)4 72 13 55 73.

Conferences Congresses Workshops Meetings

CALENDAR

October 19, 2011

FDA PUBLIC WORKSHOP

The Food and Drug Administration (FDA) is announcing a public workshop to discuss measurement principles for clinical outcome assessments (COAs) for use in clinical trials for new drugs.

Date and time: October 19, 2011, from 8:30 a.m. to 5 p.m. Eastern Time (US).

Location: FDA's White Oak Campus, 10903 New Hampshire Ave., Building 31 Conference Center, the Great Room (Room 1503), Silver Spring, MD 20993-0002.

There will also be a live **Webcast** at https://collaboration.fda.gov/coaworkshop/ Registration is free but limited to 150 attendees.

For more information, visit: http://www.federalregister.gov/articles/2011/07/28/2011-19140/review-and-qualification-of-clinical-outcome-assessments-public-workshop

NOTE: **Proceedings of the workshop will be published by the FDA and summarized on the PRO Newsletter Forum (www.forum.pro-newsletter.com). You can participate in our live discussion starting October 20th. To learn more on how to participate click on www.pro-newsletter.com.**

October 26-29, 2011

ISOQOL 18th Annual Conference
Denver, Colorado, USA
Sheraton Denver Downtown
www.isoqol.org

October 28-30, 2011

6th Asian Conference on Pharmacoepidemiology
Beijing, China
International Convention Center and Continental Grand Hotel
www.acpe-beijing.org

November 5-8, 2011

ISPOR 14th Annual European Congress
Madrid, Spain
Hotel Auditório Madrid
www.ispor.org

December 2-6, 2011

65th Annual Meeting of the American Epilepsy Society
Baltimore, Maryland, USA
The Baltimore Convention Center
www.aesnet.org

February 23-26, 2012

American Academy of Pain Medicine 28th Annual Meeting
Palm Springs, California, USA
www.painmed.org/annualmeeting/main.aspx

March 26-28, 2012

DIA's 24th Annual EuroMeeting
Copenhagen, Denmark
www.diahome.org/diahome/FlagshipMeetings/home.aspx?meetingid=25205

April 21-23, 2012

ISPE Mid-Year Meeting
Miami Beach, Florida, USA
Eden Roc Hotel
www.pharmacoepi.org/meetings/midyear12/ISPE2012MidYrFlyer.pdf

June 2-6, 2012

ISPOR 17th Annual International Meeting
Washington Hilton, Washington, DC, USA
www.ispor.org/meetings/washingtondc0512/symposiumopportunities.asp

KEYWORDS

MEDCOM, Lyon, France. ISSN 1774-8828