Quality of Life after Brain Injury (QOLIBRI) – Scale Development and Metric Properties

Nicole von Steinbuechel (1), Lindsay Wilson (2), Henning Gibbons (1), Graeme Hawthorne (3), Stefan Höfer (4), Silke Schmidt (5), Monika Bullinger (6), Andrew Maas (7), Edmund Neugebauer (8), Jane Powell (9), Klaus von Wild (10), George Zitnay (11), Wilbert Bakx (12), Anne-Lise Christensen (13), Sanna Koskinen (14), Jana Sarajuuri (14), Rita Formisano (15), Nadine Sasse (1), Jean-Luc Truelle (16) and the QOLIBRI Task Force

1.   Department of Medical Psychology and Medical Sociology, Georg-August-University, Goettingen, Germany

2.   Department of Psychology, University of Stirling, Stirling, UK

3.   Department of Psychiatry, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Australia

4.   Department of Medical Psychology, Medical University Innsbruck, Austria

5.   Department of Health and Prevention, Institute for Psychology, Ernst-Moritz-Arndt University of Greifswald, Germany

6.   Department of Medical Psychology, University Hospital Eppendorf, Hamburg, Germany

7.   Department of Neurosurgery, University Hospital Antwerp, Belgium

8.   IFOM, Private University of Witten/Herdecke, Köln, Germany

9.   Goldsmiths College, Department of Psychology, London, United Kingdom

10.   Department of Neurosurgery and Early Neurotraumatological Rehabilitation, Clemens Hospital, Westfalian Wilhelms University of Münster, Germany

11.   Association Martha Jefferson Hospital, Virginia, USA

12.   Hoensbroeck Rehabilitation Centre, Hoensbroeck and Maastricht University, Maastricht, Netherlands

13.   Centre for Rehabilitation of Brain Injury and Centre for Cognition and Memory, University of Copenhagen, Denmark

14.   Käpylä Rehabilitation Centre, Helsinki, Finland

15.   IRCCS, Rehabilitation Hospital, Santa Lucia Foundation, Rome, Italy

16.   Service de Medicine physique et réadaption, C.H.U. Raymond-Poincaré, Garches, France

The QOLIBRI Task Force consists of collaborating investigators in the following countries (*Steering Committee, †Methodological Group): Argentina (Armando Basso), Australia (Graeme Hawthorne†), Austria (Stefan Höfer†), Belgium (Christine Croisaux, Andrew Maas*), Brazil (Lucia Braga), China (Wai Poon, Zhang Tong), Denmark (Anne-Lise Christensen), Finland (Jaana Sarajuuri, Sanna Koskinen), France (Philippe Azouvi, Michelle Montreuil, Pierre North, Jean-Luc Truelle*), Germany (Monika Bullinger*†, Henning Gibbons†, Tanja Lischetzke†, Edmund Neugebauer*, Nadine Sasse, Silke Schmidt†, Nicole von Steinbüchel*†, Klaus von Wild*, Wolfgang Woerner†), Greece (Eva Tazopoulou), E. Wahjoepramono (Indonesia), Italy (Rita Formisano), Japan (Yoichi Katayama), Netherlands (Wilbert Bakx, Peter Vos), Poland (Maria Pachalska), Portugal (Sandra Guerreiro), Russia (Boleslav Lichterman), Taiwan (Wen-Ta Chiu), United Kingdom (Richard Greenwood, Jane Powell*, Lindsay Wilson†), United States of America (John DaVanzo, George Zitnay*).

***Corresponding author:** Professor Nicole von Steinbuechel, Department of Medical Psychology and Medical Sociology, Georg-August-University, Goettingen, Germany*

Abstract

The consequences of traumatic brain injury (TBI) for health-related quality of life (HRQoL) are poorly investigated, and a TBI-specific instrument has not previously been available. The cross-cultural development of a new measure to assess HRQoL after TBI is described.

An international TBI Task Force derived a conceptual model from previous work, constructed an initial item bank of 148 items, and then reduced the item set through two successive multi-centre validation studies. The first study with eight language versions of the QOLIBRI recruited 1528 participants with TBI and the second with six language versions 921 participants. The data from 795 participants from the second study who had complete GCS and GOS data were used to finalise the instrument.

The final version of the QOLIBRI consists of 37 items in six scales. Satisfaction is assessed in the areas of "Cognition", "Self", "Daily life and Autonomy", and "Social Relationships" and feeling bothered by "Emotions "and "Physical Problems". The QOLIBRI scales meet standard psychometric criteria (internal consistency, α = .75 to .89, test-retest reliability, r_{tt} = .78 to .85). Test-retest reliability (r_{tt} = 0.68 to 0.87) as well as internal consistency (α = .81 to .91) was also good in a subgroup of participants with lower cognitive performance. Although there is one strong HRQoL factor, a six scale structure explaining additional variance was validated by exploratory and confirmatory factor analyses and with Rasch modelling.

The QOLIBRI is a new cross-culturally developed instrument for assessing HRQoL after TBI that fulfils standard psychometric criteria. It is potentially useful for clinicians and researchers conducting clinical trials, assessing the impact of rehabilitation or other interventions, or carrying out epidemiological surveys.

Key words: Traumatic Brain Injury; Quality of Life; Outcome Instrument; Multi-National Study

Introduction

The ultimate goal of health care and rehabilitation following traumatic brain injury (TBI) is to return a person to full health or to enable him/her to maintain as high a health-related quality of life (HRQoL) as possible (Berger et al., 1999; Koskinen, 1998). TBI can result in lifelong physical, cognitive, emotional, and behavioural impairments, activity limitations and participation restrictions, affecting the person's self image and coping strategies (Dikmen et al., 2003; von Steinbuechel et al., 2005a). These consequences may have a major impact upon the person's quality of life (QoL) as defined by the World Health Organization QoL Group (1993, p. 153):"...an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns… ". The WHO definition views QoL as a broad concept, which is largely consistent with Corrigan and Bogner's (2004) analysis of critical dimensions in rehabilitation research emphasizing activities, social integration and participation as key outcomes.

The concept of health-related quality of life (HRQoL) refers to the specific effects of health on wellbeing and functioning (Guyatt et al., 1989; von Steinbuechel, 1995). HRQoL represents a person's perspective on his or her subjective health condition, functioning and wellbeing in the domains of physical, psychological (emotional and cognitive), social and daily life (von Steinbuechel et al., 2005a). The person is viewed as the best expert on his or her QoL, and the measurement of this multidimensional concept is usually accomplished via self rating; only in cases of severe cognitive impairment is an observer rating preferred (Bullinger and von Steinbuechel, 2001).

Since English (1904), outcomes of TBI, however, have been traditionally assessed by functional indicators, such as disability recovery (Arango-Lasprilla et al., 2007; Jennett and Bond, 1975; Wilson et al., 1998), health status (e.g., Corrigan and Bogner, 2004; Findler et al., 2001; Klonoff et al.,1986; Lippert-Gruener et al., 2007; McCarthy et al., 2006), return to work or productivity (Klonoff et al., 2006; McCrimmon and Oddy, 2006; Shames et al., 2007), psycho-social and social functioning (Arango-Lasprilla et al., 2007; Schönberger et al., 2006; Tomberg et al., 2005) and community participation (Cattelani et al., 2008; Corrigan and Bogner, 2004; Mascialino et al., 2009; Powell et al., 1998). Whilst these outcomes are certainly related to HRQoL in TBI they do not incorporate the perspective of subjective wellbeing of HRQoL (von Steinbuechel et al., 2005a). The NIH consensus

development panel on rehabilitation of persons with TBI (1999) therefore recommended the use of generic QoL measures.

The most frequently administered generic health status measures in TBI have been the SF-36 Health Survey (Andelic et al., 2009; Callahan et al., 2005; Findler et al., 2001; MacKenzie et al., 2002; Ware et al., 1995; Ware et al., 1993;) and the Sickness Impact Profile (Corrigan et al., 1998; Klonoff et al., 1986; Pagulayan et al., 2008, 2006); both of these instruments show lower subjectively rated health in persons with TBI than in healthy persons. The measurement of life satisfaction in TBI – a component of HRQOL – has been most commonly performed with the Satisfaction with Life Scale (Diener et al., 1985), which shows improvements in satisfaction with life after treatment (Dahlberg et al., 2007) and an association of life satisfaction with healthy and productive life styles (Pierce and Hanks, 2006). Importantly, these generic QoL instruments allow for general comparative studies between different populations.

Disease or condition-specific HRQoL instruments in comparison are assumed to be more sensitive to specific health conditions and therefore allow more focused and more precise information. It is only recently that HRQoL has been recognized as a potentially important outcome variable in TBI (Neugebauer et al., 2002). In TBI, dimensions such as cognition, self-perception and self-esteem are likely to be particularly important (von Steinbuechel et al., 2005a). Although the negative impact of TBI on QoL has been reported as described above, positive factors, such as adaptation to a new life after TBI, have not been evaluated (Dijkers, 2004). This emphasises the need for complementary assessment of positive aspects of HRQoL in persons after TBI (von Steinbuechel et al., 2005b). Riemsma et al. (2001) and von Steinbuechel et al. (2005b) point out potentially important validity problems with generic health status and HRQOL instruments when administered to patients with cognitive impairment. That is, in subgroups of persons after TBI this assessment may be invalid, or at least has unknown validity and reliability, as the patients may differ in the domains (such as cognition, self perception and daily life activities) important to them. Critically, in severely impaired persons, awareness of cognitive and other deficits may be reduced. Measuring a construct such as HRQOL in persons with cognitive deficits via the cognitive domain (self-rated questionnaires) represents a major methodological challenge. One solution is to develop an appropriate patient-based disease-specific

questionnaire, and to check whether it has satisfactory psychometric properties also in more severely impaired individuals.

As a result of the recommendations of the Trauma Consensus Group (TCG; Neugebauer et al., 2002) a new TBI-specific HRQoL measure, the Quality of Life after Brain Injury instrument in TBI (QOLIBRI) was developed (see Figure 1). This paper reports the construction and validation of the questionnaire. Additional background to the development of the QOLIBRI can be found in von Steinbuechel et al. (2005a, 2005b) and Truelle et al. (2008). Aspects of the validity of the QOLIBRI, including relationships with the GCS, GOSE, SF-36 and other clinical, demographic, and outcome variables are described in von Steinbuechel et al (submitted). The QOLIBRI shows systematic relationships e.g. with the GOSE, the SF-36, the HADS and other clinical variables indicating good validity'.

< Insert Figure 1 about here >

Methods and Materials

Initial development of QOLIBRI

The conceptual model for the QOLIBRI was developed on the basis of a TBI literature review and consensus meetings of an international consortium (later referred to as "the QOLIBRI Task Force"; hereafter Group). The Group reviewed measures that could be used to assess subjective experience after TBI for appropriateness, relevance, applicability, and psychometric quality. The following instruments were selected: the QOLBI (Quality of Life of the Traumatic Brain Injured; Tazopoulou et al., 2005), the Profile de la Qualité de la Vie Subjective (Gerin et al., 1989), the BICRO-39 (Brain Injury Community Rehabilitation Outcome Scale; Powell et al., 1998), and the EBIQ (European Brain Injury Questionnaire; Teasdale et al., 1997). The items were pooled into an item bank of 148 items and translated into English if necessary. The initial review revealed substantial communality in the areas and items covered by different instruments. Fifteen QOLIBRI Task Force members – neurosurgeons, neurologists, neuropsychologists, psychologists and other health care professionals working in neuro-rehabilitation were involved in the selection process of 56 of these items (for details see v. Steinbuechel et al. 2005b) and allocated them to seven HRQoL domains: physical condition, thinking activities, feelings and emotions, functioning in daily life, relationships and social/leisure activities,

current situation and future prospects. The items were either formulated as satisfaction items ("How satisfied are you with...") or as bothered items ("How bothered are you with..."). Items were self-rated on a five-point scale ("Not at all/ Slightly/ Moderately/ Quite/ Very"). Additional items and open ended questions were included to assess the relevance of the items to participants.

The item bank was translated into Danish, Dutch, Finnish, French, German, Italian, and Spanish using linguistic validation guidelines (Aquadro et al., 1996). Guidelines for cognitive debriefing and language harmonisation were followed (von Steinbuechel et al., 2002). The draft instrument was subsequently administered to 1528 persons after TBI across the centres collaborating in the QOLIBRI-TBI Task Force. The participants were recruited predominantly as convenience samples in a cross-sectional study.

The aim was to enrol 250 patients per language. Inclusion and exclusion criteria were the same as in the final validation study (see next paragraph). Four languages reached this goal (Finnish, French, German, and Italian). Psychometric testing of the items was performed on these cohorts with classical and modern test theory methods. The other languages underwent confirmatory analyses. Following the analyses of 1050 cases with the GOSE (Glasgow Outcome Scale – Extended; Wilson et al., 1998) and GCS (Glasgow Coma Scale; Teasdale and Jennett 1974) and subsequent item reduction and refinement of scores, an interim version of the QOLIBRI instrument with 49 items was constructed. We considered it essential to retest, revalidate and finalise this preliminary instrument in clinical settings.

Development and validation of the final QOLIBRI

Psychometric properties of the preliminary 49-item QOLIBRI version were investigated in a cross-sectional sample, designed to cover a heterogeneous range of head injury outcomes. Nine countries (= six languages) participated in the second validation study: Australia (N=64 cases); Belgium (N=33), Finland (N=171), France (N=135), Germany (N=172), Italy (N=150), the Netherlands (N=118), UK (N=41), and the USA (N=25). Most countries/centres recruited convenience samples from rehabilitation facilities. Australia and Germany randomly sampled outpatients of hospital registries as possible participants. The inclusion criteria were: a minimum age of 15 years at time of injury, time since trauma between 3 months to 15 years, diagnosis of TBI according to ICD-10, and informed

consent. Exclusion criteria were: a recorded GOSE < 3, a spinal cord injury, the presence of a significant current or pre-injury psychiatric condition or ongoing severe addiction, a diagnosed terminal illness, and inability to understand, cooperate and answer questions in the respective language. Ethics clearance was obtained by each of the participating centres.

In this second study we aimed to recruit 120 to 150 participants per language. Patients were stratified by clinical severity of TBI as assessed by the GCS. Test-retest reliability was investigated in a subset of at least 30 participants per language by second administration of the questionnaire after a two week interval.

The questionnaires were administered in one of four modes: by self-report (mail), self-report (participant present at the clinic), face-to-face interview, or administration over the telephone.

Measures

The preliminary *QOLIBRI version* for this second validation study consisted of 49 items, 43 of which were arranged in seven scales, with the remaining six items composed as an overall scale for use as an independent screening instrument (von Steinbuechel et al., in prep.). Four scales assessed satisfaction with "Cognition" (7 items), "Emotion and self-perception" (8), "Activities of daily living (ADLs) and Autonomy" (8), and "Social relationships" (6). Three scales assessed feeling bothered by "Negative feelings" (5), "Restrictions and problems" (4), and "Physical condition" (4). Responses were recorded on a 5-point scale: "Not at all/ Slightly/ Moderately/ Quite/ Very". The "bothered" items had an additional (filter) category "Does not apply". During scoring, the endorsements of this option were coded as "Not at all". Scale scores were presented on a percent scale (0-100)

Demographic characteristics of participants were collected, including country, language, gender, age, relationship status, educational attainment, labour force participation, and social participation.

Subjective Health status was assessed through administration of the SF-36 Health Survey (Version 1, Ware et al., 1993; one country used version 2 with subsequently transformed data; Ware and Kosinski, 1996; Ware, 2000). Two summary scales are reported, the Physical Component Summary (PCS) and the Mental Component Summary (MCS).

Current health conditions were assessed through a list of 28 common health conditions (adapted from the WHOQoL project; von Steinbuechel et al., 2006). *Clinical data* extracted from participants' medical records included the date of injury, cause of injury, site of major head injury, and the worst GCS score in the first 24 hours, classified into minor (13-15), moderate (9-12) and severe (3-8; Rimel et al., 1982). The presence of comorbid health conditions was recorded (epilepsy, hemiparesis, vision and hearing problems, extra cerebral injuries, communication, attention and memory dysfunction, executive function, and affective and behavioural disorders), as were details of participation in rehabilitation and the use of anticonvulsive, psychotropic and recreational drugs.

Depression and anxiety were assessed with the Hospital Anxiety and Depression (HADS) scale (Zigmond and Snaith, 1983), which was available in all languages. A score of greater than 8 was taken to indicate probable morbidity (Olsson et al., 2005).

Cognitive status was measured in a subsample of participants using either the Mini Mental State Examination (MMSE; Folstein et al., 1975) or the Telephone Interview for Cognitive Status (TICS; Brandt et al., 1988) which was translated into the particular languages. Cut offs on the TICS and MMSE of 32/33 and 27/28, respectively, were used to define groups with low cognitive performance and normal performance. These cut offs have been found to be equivalent in an elderly sample, and scores in the lower range are taken as indicating 'borderline' or 'impaired' performance, while the upper range is regarded as normal (Brandt and Folstein, 2003). In a younger group, 'borderline' scores are more likely to be indicative of impairment (Crum et al., 1993).

Disability recovery was assessed by the GOSE (Wilson et al., 1998). Users were trained to administer the GOSE following the manual published by Wilson et al. (1998). The GOSE classification was Severe Disability (3-4), Moderate Disability (5-6), and Good Recovery (7-8).

Statistical approaches

To finalise the QOLIBRI scales, metric properties were examined (a) on the item level and (b) on the scale level with respect to internal consistency, test-retest reliability, and factor structure.

(a) Item level

- We first checked whether responses to items were distributed over the whole range. For item frequency analysis we used the endorsement index devised by the WHOQOL group (1998). Distributions were checked for frequency problems to determine whether any two adjacent response categories had a sum of less than 10% of total number of responses, and did so for at least half of the language versions. Floor and/or ceiling effects (>60% of cases at the maximum or the minimum of the scale) were also checked.

- `Skewness is common in the responses to clinical scales. For example, many people will tend to report satisfactory to good QoL on items in a scale. Extreme skewness can, however, create problems for analysis using correlation, reducing both the probability that a scale will show strong relationships with other measures, and reliability (or, the precision of measurement) of a scale. Skewness (conventionally, items with skewness >1 are considered for removal) was checked using SPSS 16.0. In the QOLIBRI scales, however, some moderately skewed items (1.0 - 1.3) were included, to capture a range of impairment (Schmidt et al., 2006).

(b) Scale level

- The internal consistency of the scales, which reflects the extent to which all items of a scale "belong together" (i.e., measure the same psychological construct), was assessed using Cronbach's α, and the fit of individual items to each scale was examined by correlating the item with the total for the other items in the scale. ´ Cronbach's α and corrected item-total correlations (CITCs) were calculated using the SPSS 16.0 reliability procedure. It is conventionally accepted that CITCs should be over 0.4 (WHOQOL group, 1998). An α of 0.70 is often regarded as the lower boundary of acceptability for measures used in group comparisons (Moosbrugger and Kaleva, 2007), and over 0.90 for clinical application to individuals (Bland and Altman, 1997).

- Test-retest reliability is one of the most important measures of reliability for questionnaires. The test-retest reliability of the QOLIBRI scales was assessed using the intra-class correlation coefficient (ICC) calculated between the scale means on two occasions (retest on average 14 days after initial testing). The conventional interpretation of the ICC is that values of 0.4 to 0.75 are fair to good, and values over 0.75 are excellent (Fleiss, 1986).

- In addition to correlating adequately with its own scale it is also important that an item does not correlate similarly with other scales. SPSS 16 was used to calculate statistics comparable to the Multitrait Analysis Program (MAP; Hays et al., 1988). The MAP criterion for a definite scaling problem is a corrected item-home scale correlation that is > 2 SEs below the correlation of the item with another scale. A probable scaling problem is indicated by a corrected item-home scale correlation within 2 SEs of the correlation of the item with another scale (Hays et al., 1988).

- Item response theory approaches are now widely advocated as an alternative to classic correlational techniques for assessing the fit of items to scales. Rasch analysis was carried out using Winsteps 3.66. The data were examined to ensure that items were suitable for Rasch analysis. All categories had 10 or more responses, and no items in the analysis were very skewed (Bond and Fox, 2007). Furthermore the average category responses for all items were in the expected order. Winsteps produces two fit statistics: "Infit" weights responses of people whose performance is close to the item value, while "outfit" is sensitive to outlying scores (Bond and Fox, 2007). Deviation of "infit" from expectation is generally regarded as more important for measurement than "outfit" deviation. For large sample sizes the mean square is preferred to the Z statistic as a measure of fit, and satisfactory fit is indicated by values between 0.7 and 1.3 (Smith et al., 2008). Values greater than 1.3 indicate lack of item fit with the unidimensional model, while values less than 0.7 suggest items are over-fitting or redundant.

- To study whether the structure of the questionnaire (specifically the division into separate scales) was justified we used factor analysis. The dimensionality of the final version was examined using principal component analysis (PCA). PCA using both a forced one-factor solution and a six-factor solution that resulted when Kaiser's criterion (Eigenvalues > 1) was applied, and oblique rotation (promax method with the assumption of correlated scales) was performed, again using SPSS 16.

- Finally we studied the structure of the questionnaire using confirmatory factor analysis, an approach which allows various statistics for overall fit to be calculated. Good fit indicates that the assumed grouping of items into scales adequately reflects the empirical patterns of relationships between items. Within confirmatory factor analysis, observed variables are represented by single

Results

Descriptives

A total of 921 participants were enrolled. There were 126 cases with missing GCS in the data, which were excluded from subsequent analysis. In the remaining 795 cases missing data were still present to varying degrees. For demographics concerning gender and age 100% of the data were present, for living arrangements, employment status and relationship status 93% were present. In the clinical data 100% were present for GCS and GOSE, for major lesion location and years since injury 99%, 98% for numbers of comorbid health conditions, and 92% for self reported health situation.

There were less than 5% missing responses for single QOLIBRI items (e.g., "participation in work", 4.4%). HADS anxiety and depression scores were present for 99% and SF-36 scores for 96% of the sample. For the QOLIBRI means were calculated for each scale, and prorated if less than 33% of responses were missing. Missing responses were imputed per participant by substituting the missing value by the scale mean rounded to an integer. Means for the SF-36, the HADS and the comorbid health conditions list were calculated, using prorating if less than 33% of responses were missing.

Demographic and clinical characteristics of the final validation study (N=795) are presented in Table 1. Percentages are given with respect to the number of participants with complete data. As is typical for a TBI sample there are a greater number of men than women. Within the range covered (from 17 to 68 years), three age groups were formed (17-30 years, 31-40y, 45-68y) of almost equal size. More than half of the sample was severely injured by GCS criteria, and for half the injury occurred four or more years previously. Less than a quarter of the sample was in full-time employment, and only half was currently in a relationship. Over half of the sample was living independently, that is, did not "need help for daily life tasks". Over half report four or more comorbid health conditions; in contrast only 28% described themselves as being "unhealthy" at the moment. According to the GOSE, the majority (72%) were disabled by the consequences of their TBI.

(Table 1 about here)

Finalising the QOLIBRI scales and item set

The final resulting QOLIBRI instrument consists of 37 items in four satisfaction scales, "Cognition" (7 Items), "Self" (7), "Daily Life and Autonomy" (7), and "Social Relationships" (6), and two bothered scales, "Emotions" (5) and "Physical Problems" (5).

Item characteristics

The properties of the 43 preliminary QOLIBRI items (resulting from the initial validation and excluding the overall scales items) are shown in Table 2. In general there were very few missing responses; the maximum value was 4.4% for "participation in work". All items except three met the endorsement criterion (cf. Power et al., 2005), and did so for at least 50% of the language versions. The exceptions were the two bothered items "epileptic seizures" and "problems with smelling/tasting", and an item concerning satisfaction with the "ability to look after basic personal needs". These three items were removed.

At this point, seven items were left to form a second bothered scale "Physical problems". One item concerning "ongoing legal actions" had a low corrected item-total correlation (CITC, < 0.40) with the interim scale and was excluded. This resulted in CITC dropping below 0.40 for a further item

pertaining to "restrictions with driving", which was also removed. The resulting five-item "Physical problems" scale had a satisfactory internal consistency (Cronbach's α = 0.75). Finally, a satisfaction item from the "Self" scale, satisfied with "ability to control emotions", was excluded; of the eight initial items in this scale, it had the lowest CITC (0.58) and MAP analysis revealed several substantial correlations > 0.40 with other scales. For the remaining 37 items reliability analysis indicated that all items had CITCs of greater than 0.40, and the majority had CITCs greater than 0.60. On the MAP analysis all items correlated more strongly with their home scale than any other scale, and scaling success for each of these scales was 100%. Two of the retained 37 items had skewness indices slightly above 1 (bothered by "feelings of loneliness", "problems with seeing/hearing"), but were kept because of their clinical importance, reasonable scale fits, and ability to differentiate between (strongly impaired) persons in the low QOL range.

(Table 2 about here)

Internal consistency

Internal consistency was assessed for each scale, and for each language version of the QOLIBRI (see Table 3). Cronbach's α ranges from 0.75 ("Physical problems") to 0.89 ("Cognition"; "Self"). The individual scales thus fulfil criteria for use in research studies, and the total QOLIBRI score provides a reliable assessment at the level of the individual with Cronbach's α of 0.95, ranging from 0.92 (French; N = 147) to 0.97 (English; N = 96).

Individual scale scores exceed α = 0.70 for all language versions except for the "Emotions" and "Physical problems" scales of the Dutch version (α = 0.64 and α = 0.69, respectively) and the "Physical problems" scale of the French version (α = 0.64). The results indicate that the QOLIBRI scales generally have good internal consistency. Also in a subgroup of persons with low cognitive performance (MMSE < 28 or TICS < 33; N = 84) internal consistency was comparable to persons with normal cognitive status (MMSE >27 or TICS > 32; N = 121); in the former group, lowest α was 0.81 for the "Physical problems" scale where it was 0.76 in the latter group.

(Table 3 about here)

Test-retest reliability

Table 4 shows intra-class correlation coefficients (ICCs) in the sample of the 381 participants retested after two weeks (27 Dutch, 119 German, 56 English, 49 Finnish, and 126 French cases). Mean age in the group retested was 36 years (SD=12.5) compared to 40 years (SD=14.0) in those not retested; 52% of the retested group had severe injuries, 11% had moderate injuries, and 37% had minor injuries. For the retested sample as a whole ICCs ranged from 0.78 ("Emotions") to 0.85 ("Physical Problems"), indicating that all scales show good test-retest reliability. The reliability of the total score was r_{tt} = 0.91. Scores on the MMSE or TICS were available for 181 participants who were retested. Participants were divided into those with low performance (borderline or impaired; N=84) versus those with scores in the normal range (N=121). ICCs for these two subgroups are shown in Table 4. In general, most scales showed good reliability in both subgroups. The lowest ICCs were for the "Social relationships" and "Emotions" scales, and the values here are still consistent with good reliability.

(Table 4 about here)

Test-retest reliability was good for the four language versions where sufficient numbers of retested participants were available (N ≥ 48). For the QOLIBRI total score, ICCs range from 0.87 (Finland) to 0.91 (France; see Table 5). Single-scale test-retest reliability is mostly in the range of 0.75 to 0.80.

(Table 5 about here)

QOLIBRI scale means on first and second test are shown in Table 6; differences between first and second assessment reached significance at the .05 level for the "Cognition" scale and the "Physical Problems" scale. When the changes were calculated as effect sizes using Cohen's *d* (Cohen, 1988), none exceeded 0.10, indicating very small differences.

(Table 6 about here)

Another approach considering the test-retest reliability of *differences* between QOLIBRI scale scores (being rather unrelated to test-retest reliability per se) already providing support for the suggested six-factor structure of the QOLIBRI (see below), is reported now: For this purpose, for each pair of the six subscales, differences between subscale scores were computed for test and retest data. Test-retest correlations between pairwise compared scale differences ranged from 0.55 ("Self" x "Emotions") to 0.75 ("Emotions" x "Physical"). The averaged test-retest correlations of QOLIBRI scale differences - computed based on Fisher-z transformed coefficients - amounted to 0.62, which is an extremely high value given that differences between measures cannot be as reliable as the measures themselves. Thus, in a multi-scale approach the different scales of the QOLIBRI data reliably assess information related to the different QoL domains; clearly, there is not only error variance beyond the g-factor. However, further research is needed to determine whether the QOLIBRI can be used to reliably assess QoL profiles at an individual level.

Rasch analyses

Each scale was analysed in turn using Rasch analysis. Items within each scale are shown in Table 7 ordered by item 'difficulty'. Thus, for example, in the "Cognition" scale participants were most likely to express satisfaction with "finding way about", and least likely to be satisfied with "remember". The range covered by items in each scale was fairly narrow, from 0.59 logits for the "Emotion" scale to 1.51 logits for "Social relationships", indicating that in general items were quite similar in their likelihood of being endorsed positively. It should be borne in mind that item difficulty is calculated at the middle of the rating categories. Rasch analysis of individual QOLIBRI scales (Table 7) showed that infit was in the required range for all items in each of the scales. Rasch analysis thus confirms that items have a satisfactory fit with their home scales. Weaker items are "self-perception" with an infit value of 0.7 suggesting a certain amount of redundancy, and "run personal finances" with an outfit value of 1.33, which indicates misfitting outliers in the data.

Rasch analysis was also performed with all items combined to examine whether QOLIBRI items fit a single unidimensional scale. Item difficulty measures ranged from -0.47 to 0.61 logits. PCA

of the residuals showed that the Rasch model explained 38.2% of the variance, indicating that a unidimensional model explains only a moderate amount of the variance. The infit values indicated that the majority of QOLIBRI items fitted an overall Rasch dimension. However, there were five items with infit values of 1.3 or more: "partner" (infit=1.41), "sex life" (1.30), "other injuries" (1.30), "pain" (1.31), and "seeing/hearing" (1.36). The results of this analysis give moderate support to an unidimensional model, but also indicate that some of the items in the "Social relationships" and "Physical problems" scales have a poor fit with an unidimensional model.

(Table 7 about here)

Factor analyses

The results of two principal components analyses (PCA) are shown in Table 8. Loadings on the first component of a single-factor solution indicate that items in the first three scales generally have a good fit (loadings >0.6) with a unidimensional QoL model descriptive system. Items in the last three scales have a weaker fit with this single-factor descriptive system, and two items ("partner" and "see/hear") have a poor fit (loading <.45). The single-factor PCA is thus consistent with the Rasch analysis conducted on all items combined, and indicates that there is a unidimensional component to the QOLIBRI, primarily based on the items in the first three scales, which are concerned with cognitive function, self-perception and independent living. The items from the last three scales, with the two exceptions described above, have moderate fit with this descriptive system model.

The results from the second PCA where six factors with eigenvalues > 1 were extracted, on the other hand, nicely confirm the overall structure of the QOLIBRI: all items have the highest loadings on their home scales, and there is relatively little cross-loading >0.25.

The PCA single-factor solution accounted for only 37% of the variance, while the six factors accounted for 59%. These scales are moderately correlated, as shown in Table 9.

(Tables 8 and 9 about here)

Structural equation modelling

Structural equation modelling (SEM) was used to confirm the structure of the QOLIBRI. The SEM model with six factors showed substantial inter-correlation of latent factors (range of r=.469-.796). Thus, a second-order factor underlying the six original factors on first level was specified to account for these intercorrelations of latent variables and in keeping with conceptual considerations (general QoL as a higher-order construct). The final model (Figure 2) consisted of the six latent variables (factors) on one level and a "second-order" latent variable (a general QoL factor). Fit statistics indicated a reasonable fit with this model (CFI = 0.896, RMSEA = 0.055, χ^2 = 2115.96, df = 623, p[χ^2]<.001). In fact, the model meets the RMSEA criterion, rather than the CFI criterion for satisfactory fit (Hu & Bentler, 1999). Combination of cut-off rules are more appropriate for lower sample sizes (<250) and with larger sample sizes tend to over-reject models (see Hu & Bentler, 1999). A similar problem arises if chi-square statistics are applied to larger sample sizes. We believe RMSEA is more appropriate than CFI because it compensates for model complexity.

(Figure 2 about here)

Discussion

The parallel, consensual and cross-cultural development of this new-specific instrument for the assessment of HRQoL of persons after TBI represents a distinctive and unusual approach.

The different language samples were generally collected as convenience samples (with two exceptions: Germany and Australia) which is a common strategy in this type of validation study (Hawthorne et al., 2006). The wide variation represents an advantage for a validation study because it ensures future applicability to a wide range of patients. The relatively small number per language sample however presents a limitation for analyses on the language level.

For use in international multi-centre studies, the content as well as metric equivalence of an instrument needs to be demonstrated (Bullinger et al., 1993; Stucki et al., 1997). Thus the translation of HRQOL instruments requires a standardized process including translation, back translation, review, cognitive debriefing, and harmonisations of the different language translations. We ensured comparability of the different translations by following accepted guidelines concerning cognitive debriefing in each language and harmonisation procedures (von Steinbuechel et al., 2002; Acquadro et al., 1996). Cross-cultural development also includes the demonstration of comparable metric properties for the whole sample as well as for the different languages, in particular with respect to internal consistency, reliability, factor structure, and validity (see von Steinbuechel et al., this issue).

The results of the present study indicate favourable psychometric properties of the QOLIBRI. In spite of the variation in demographic and clinical characteristics, internal consistency and test-retest reliability are acceptable to good both in the total sample and in different language groups. The test-retest reliability of the QOLIBRI is similar to that of the EBIQ (Sopena et al., 2007), a patient-reported outcome measure developed for TBI.

Reliabilities for the individual QOLIBRI scales were also good in the subgroup with poorer cognitive performance. The lowest reliability was recorded for the "Emotions" scale, an area in which some fluctuation might be expected.

A concern with self-report instruments in TBI is the potential lack of insight which may be experienced by those with cognitive impairment. Experience in diseases such as dementia suggests that subjective QoL judgments can be obtained reliably even in people with substantial cognitive impairment (Brod et al., 1999; Novella et al., 2001; Wlodarczyk et al., 2004). The relationship between reliability of reports and cognitive status, however, is an issue that has not been properly addressed in head injury (Riemsma et al., 2001). This study showed that internal consistency (i.e., Cronbach's α) for participants with poor cognitive performance was satisfactory and test-retest reliability was good for the QOLIBRI. This study however did not include substantial numbers of severely cognitively impaired patients, which presents a topic for future research. On the other hand the majority of participants were severely injured. We conclude that the reliability of the QOLIBRI is satisfactory to good even when some cognitive impairment is present, and may be appropriate for the great majority of persons after TBI.

The conceptual model of HRQoL, on which the QOLIBRI descriptive system was based, initially suggested a seven-, then a six-dimensional model with four "satisfaction" scales and three "bothered" scales which were then collated into two scales following psychometric examination. The PCA and SEM analyses support this six-factor structure of the QOLIBRI in a second-order latent variable model. Also RASCH analysis based on the presented results does not support unidimensionality for all the items, and instead suggests a multi-factorial structure. However, there is potential to develop, in future research, a QOLIBRI short form with reduced number of items and dimensions.

The final descriptive system of the QOLIBRI provides both for a HRQoL profile across six domains of life, and also a total index of HRQoL. The data from the various analyses suggest that where a rich profile description of HRQoL is required, the individual domain scores will provide this. When used in evaluation studies shifts in individual domains will reflect areas of life where gains consequent upon treatment are made. In contrast, where an index of HRQoL is required, this index can be used to assess the impact of treatments on overall HRQoL as measured by QOLIBRI total score.

A limitation of the study was that the number of participants was too small for most language samples to confirm the QOLIBRI factor structure within each language. Examination of the factor structure in some of the larger samples of the first validation study however confirms the structure. Rasch analysis suggests a certain amount of redundancy at the item level, and that it would be worth in the future exploring the possibility of an abbreviated version of the QOLIBRI. A short six-item screening measure is already available with the "Overall" items (psychometric properties will be reported in von Steinbuechel et al, in prep.); the use of the different versions will depend on the research questions to be answered. Other limitations of the study include the use of convenience samples, and the fact that it was beyond the scope of the study to carry out detailed cognitive testing.

The development of the QOLIBRI descriptive system in two large multi-national samples of persons after TBI has resulted in an instrument with good psychometric properties. Further work will explore the sensitivity and responsiveness of the measure and further cross-cultural comparisons of scale structures and scores. The psychometric properties of the QOLIBRI suggest that it is a practical and

reliable instrument which can be considered for use in studies examining HRQoL after TBI in clinical as well as in research settings (Zitnay et al., 2008).

References

Acquadro, C., Jambon, B., Ellis, D., and Marquis, P. (1996). Language and translation issues, in: *Quality of Life and Pharmacoeconomics in Clinical Trials*. B. Spilker (ed). Raven: New York, pps. 575-585.

Andelic, N., Hammergren, N., Bautz-Holter, E., Sveen, U., Brunborg, C., and Roe, C. (2009). Functional Outcome and Health-related Quality of Life 10 Years after Moderate-to-Severe Traumatic Brain Injury. Acta Neurol. Scand. (in press).

Arango-Lasprilla, J.C., Rosenthal, M., Deluca, J., Komaroff, E., Sherer, M., Cifu, D., and Hanks, R. (2007). Traumatic Brain Injury and Functional Outcomes: Does Minority Status Matter? Brain Inj. 21, 701-708.

Berger, E., Leven, F., Pirente, N., Bouillon, B., and Neugebauer, E. (1999).Quality of Life after Traumatic Brain Injury: A Systematic Review of the Literature. Restor. Neurol. Neurosci. 14, 93-102.

Bland, J.M. and Altman, D.G. (1997). Statistics Notes: Cronbach´s alpha. BMJ 314, 572.

Bond, T.G. and Fox, C.M. (2007). *Applying the Rasch Model*, 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Brandt, J., Spencer, M., and Folstein, M. (1988). The Telephone Interview for Cognitive Status. Neuropsychiatry Neuropsychol. Behav. Neurol. 1, 111-117.

Brandt, J. and Folstein, M.F. (2003). Telephone Interview for Cognitive Status: Professional Manual. Psychological Assessment Resources, Inc., Lutz, FL.

Brod, M., Steward, A.L., Sands, L., and Waltson, P. (1999).Conzeptualization and measurement of quality of life in dementia: The Dementia Quality of Life Instrument (DQoL). Gerontologist 39, 25-36.

Bullinger, M. and von Steinbuechel, N. (2001). Quality of Life – Measurement and Outcome, in: *Comprehensive Care for People with Epilepsy*. M. Pfäfflin, T. Fraser, R. Thorbecke, U. Specht, and R. Wolff (eds). John Libbey and Company Ltd: London, pps. 277-292.

Bullinger, M., Anderson, R., Cella, D., and Aaronson, N. (1993). Developing and evaluatin cross-cultural instruments from minimum requirements to optimal models. Qual. Life Res. 2, 451-459.

Callahan, C.D., Young, P.I., and Barisa, M.T. (2005). Using the SF-36 for Longitudinal Outcomes Measurement in Rehabilitation. Rehabil. Psychol. 50, 65-70.

Cattelani, R., Roberti, R., and Lombardi, F. (2008). Adverse Effects of Apathy and Neurobehavioral Deficits on the Community Integration of Traumatic Brain Injury Subjects. Eur. J. Phys. Rehabil. Med. 44, 245-251.

Cohen J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale: Lawrence Erlbaum.

Corrigan, J.D. and Bogner, J. (2004). Latent Factors in Measures of Rehabilitation Outcomes after Traumatic Brain Injury. J. Head Trauma Rehabil. 19, 445-458.

Corrigan, J.D., Smith-Knapp, K., and Granger, C.V. (1998). Outcomes in the first 5 years after traumatic brain injury. Arch. Phys. Med. Rehabil. 79, 298-305.

Crum, R.M., Anthony, J.C., Bassett, S.S., and Folstein, M.F. (1993). Population-based norms for the Mini Mental State Examination by age and educational level. JAMA 12, 2386-2391.

Dahlberg, C.A., Cusick, C.P., Hawley, L.A., Newman, J.K., Morey, C.E., Harrison-Felix, C.L., and Whiteneck G.G. (2007). Treatment Efficacy of Social Communication Skills Training after Traumatic Brain Injury: A Randomized Treatment and Deferred Treatment Controlled Trail. Arch. Phys. Med. Rehabil. 88, 1561-1573.

Diener, E., Emmons, R.A., Larsen, R.J., and Griffin, S. (1985). The Satisfaction with Life Scale. J. Pers. Assess. 49, 71-75.

Dijkers, M.P. (2004). Quality of Life after Traumatic Brain Injury:A Review of Research Approaches and Findings. Arch. Phys. Med. Rehab. 85, 21-35.

Dikmen, S.S., Machamer, J.E., Powell, J.M., and Temkin, N.R. (2003). Outcome 3 to 5 Years after Moderate to Severe Traumatic Brain Injury. Arch. Phys. Med. Rehab. 84, 1449-1457.

English, T.C. (1904). The after-effects of head injuries. Lancet 20, 485-489.

Findler, M., Cantor, J., Haddad, L., Gordon, W., and Ashman, T. (2001). The Reliability and Validity of the SF-36 Health Survey Questionnaire for Use with Individuals with Traumatic Brain Injury. Brain Inj. 15, 715-723.

Fleiss, J.L. (1986). Analysis of Data from Multiclinic trials. Control. Clin. Trials 7, 267-275.

Folstein, M.F., Folstein, S.E., and McHugh, P.R. (1975). Mini mental state. A Practical Method for Grading the Cognitive State of Patients for the Clinician. J. Psychiatr. Res. 12, 189-198.

Gerin, P., Dazord, A., Boissel, J.P., Hanauer, M.T., Moleur, P., and Chauvin, F. (1989). Quality of Life Assessment in Therapeutic Trials. Conceptual Aspects and Presentation of a Questionnaire. Therapie 44, 355-364.

Guyatt, G.H., Veldhuyzen van Zanten, S.J., Feeny, D.H., and Patrick, D.L. (1989). Measuring quality of life in clinical trials: a taxonomy and review. Can. Med. Assoc. J. 140, 1441-1448.

Hawthorne, G., Davidson, N., Quinn, K., McCrate, F., Winkler, I., Lucas, R., Kilian, R., and Molzahn, A. (2006). Issues in conducting cross-cultural research: implementation of an agreed international procotol designed by the WHOQOL Group for the conduct of focus groups eliciting the quality of life of older adults. Qual. Life Res. 15, 1257-1270.

Hays, R.D., Hayashi, T., Carson, S., and Ware Jr., J.E. (1988). *User's guide for the multitrait analysis program.* Santa Monica, CA: RAND Corporation.

Hu, L. and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling 6, 1-55.

Jennet, B. and Bond, M. (1975). Assessment of Outcome after Severe Brain Damage. Lancet, 480-484.

Klonoff, P.S., Costa, L.D., and Snow, W.G. (1986). Predictors and Indicators of Quality of Life in Patients with Closed-head Injury. J. Clin. Exp. Neuropsych. 8, 469-485.

Klonoff, P.S., Watt, L.M., Dawson, L.K., Henderson, S.W., Gehrels, J.-A., and Voreis Wethe, J. (2006). Psychosocial Outcomes 1-7 Years after Comprehensive Milieu-oriented Neurorehabilitation: The Role of Pre-Injury Status. Brain Inj. 20, 601-612.

Koskinen, S. (1998). Quality of Life 10 Years after a Very Severe Traumatic Brain Injury (TBI): The Perspective of the Injured and the Closest Relative. Brain Inj. 12, 631-648.

Lippert-Grüner, M., Maegele, M., Haverkamp, H., Klug, N., and Wedekind, C. (2007). Health-related Quality of Life during the First Year after Severe Brain Trauma with and without Polytrauma. Brain Inj. 21, 451-455.

MacKenzie, E.J., McCarthy, M.L., Ditunno, J.F., Forrester-Staz, C., Gruen, G.S., Marion, D.W., and Schwab, W.C. (2002). Using the SF-36 for Characterizing Outcome after Multiple Trauma Involving Head Injury. J. Trauma 52, 527-534.

Mascialino, G., Hirshson, C., Egan, M., Cantor, J., Ashman, T., Tsaousides, T., and Spielman, L. (2009). Objective and Subjective Assessment of Long-term Community Integration in Minority Groups Following Traumatic Brain Injury. NeuroRehabilitation 24, 29-36.

McCrimmon, S. and Oddy, M. (2006). Return to Work Following Moderate-to-Severe Traumatic Brain Injury. Brain Inj. 20, 1037-1046.

McCarthy, M.L., Dikmen, S., Langlois, J.A., Selassie, A..W., Gu J.K., and Horner, M.D. (2006). Self-reported Psychosocial Health among Adults with Traumatic Brain Injury. Arch. Phys. Med. Rehabil. 87, 953-961.

Moosbrugger, H. and Kelava, A. (2007). *Testtheorie und Fragebogenkonstruktion* (Test theory and questionnaire construction). Berlin: Springer.

Neugebauer, E., Bouillon, B., Bullinger, M., and Wood-Dauphinee, S. (2002). Quality of Life after Multiple Trauma - Summary and Recommendations of the Consensus Conference. Restor. Neurol. Neurosci. 20, 161-167.

NIH Consensus Development Panel on Rehabilitation of Persons with Traumatic Brain Injury (1999). Rehabilitation of Persons with Traumatic Brain Injury. JAMA 282, 974-983.

Novella, J.L., Jochum, C., Ankri, J., Morrone, I., Jolly, D., and Blanchard, F. (2001). Measuring general health status in dementia: practical and methodological issues in using the SF-36. Aging (Milano) 13, 362-369.

Olsson, I., Mykletun, A., and Dahl, A.A. (2005). The Hospital Anxiety and Depression Rating Scale: A Cross-sectional Study of Psychometrics and Case Finding Abilities in General Practice. BMC Psychiatry 14, 46.

Pagulayan, K.F., Hoffmann, J.M., Temkin, N.R., Machamer, J.E., and Dikmen, S.S. (2008). Functional Limitations and Depression after Traumatic Brain Injury: Examination of the Temporal Relationship. Arch. Phys. Med. Rehabil. 89, 1887-1892.

Papulayan, K.F., Temkin, N.R., Machamer, J., and Dikmen, S.S. (2006). A longitudinal study of health-related quality of life after traumatic brain injury. Arch. Phys. Med. Rehabil. 87, 611-618.

Pierce, C.A. and Hanks, R.A. (2006). Life Satisfaction after Traumatic Brain Injury and the World Health Organisation Model of Disability. Am. J. Phys. Med. Rehabil. 85, 889-898.

Powell, J.H., Beckers, K., and Greenwood, R.J. (1998).Measuring Progress and Outcome in Community Rehabilitation after Brain Injury with a New Assessment Instrument - the BICRO-39 Scales. Arch. Phys. Med. Rehab. 79, 1213-1225.

Power, M., Quinn, K., Schmidt, S., and the WHOQOL-OLD Group (2005). Development of the WHOQOL-old module. Qual. Life. Res. 14, 2197-2214.

Riemsma, R.P., Forbes, C.A., Glanville, J.M., Eastwood, A.J., and Kleijnen, J. (2001). General Health Status Measures for People with Cognitive Impairment: Learning Disability and Acquired Brain Injury. Health Technol. Assess. 5, 1-100.

Rimel, R.W., Giordani, B., Barth, J.T., and Jane, J.A. (1982). Moderate head injury: completing the clinical spectrum of brain trauma. Neurosurgery 11, 344-351.

Schönberger, M., Humle, F., Zeemann, P., and Teasdale, T.W. (2006). Working alliance and Patient Compliance in Brain Injury Rehabilitation and their Relation to Psychosocial Outcome. Neuropsychol. Rehab. 16, 298-314.

Shames, J., Treger, I., Ring, H., and Giaquinto, S. (2007). Return to Work Following Traumatic Brain Injury: Trends and challenges. Disabil. Rehabil. 29, 1387-1395.

Schmidt, S., Debensason, D., Mühlan, H., Petersen, C., Power, M., Simeoni, M., and Bullinger, M. (2006). The DISABKIDS generic quality of life instrument showed cross-cultural validity. J. Clin. Epidemiol. 29, 587-598

Smith, A.B., Rush, R., Fallowfield, L.J., Velikova, G., and Sharpe, M. (2008). Rasch Fit Statistics and Sample Size Considerations for Polytomous Data. BMC Medical Research Methodology, 8.

Sopena, S., Dewar, B.K., Nannery, R., Teasdale, T.W., and Wilson, B.A. (2007). The European Brain Injury Questionnaire (EBIQ) as a reliable outcome measure for use with people with brain injury. Brain Inj. 21, 1063-1068.

Stucki, G., Stucki, S., and Sangha, O. (1997). Patientenzentrierte Evaluation der Krankheitsauswirkungen bei musculoskeletalen Erkrankungen. Adaptation und Neuentwicklung von Outcome Instrumenten. Z. Rheumatologie 56, 266-275.

Teasdale, G. and Jennett, B. (1974). Assessment of Coma and Impaired Consciousness. A practical scale. Lancet 2, 81-84.

Teasdale, T.W., Christensen, A., Willmes, K., Deloche, G., Braga, L., Stachowiak, F., Vendrell, J.M., Castro-Caldas, A., Laaksonen, R., and Leclercq, M. (1997). Subjective Experience in Brain-injured Patients and their Close Relatives: a European Brain Injury Questionnaire study. Brain Inj. 11, 543-563.

Tomberg, T., Toomela, A., Pulver, A., and Tikki, A. (2005). Coping Strategies, Social Support, Life Orientation and Health-related Quality of Life Following Traumatic Brain Injury. Brain Inj. 19, 1181-1190.

Tazopoulou, E., Truelle, J.-L., North, P., and Montreuil, M. (2005). Generic and Specific Measures of Quality of Life after Traumatic Brain Injury: Initial Validation of a New Specific Measure, The QOLBI. Acta Neuropsychol. 3, 13-24.

Truelle, J.-L., von Wild, K., Höfer, S., Neugebauer, E., Lischetzke, T., von Steinbuechel, N., and the QOLIBRI Group (2008). The QOLIBRI - Towards a Quality of Life Tool after Taumatic Brain Injury: Current Developments in Asia. Acta Neurochir. Suppl. 101, 125-129.

von Steinbuechel, N. (1995). Gesundheitsbezogene Lebensqualität als Beurteilungskriterium für Behandlungseffekte bei Patienten mit Epilepsie. Prävention Rehabilitation 7, 39-146.

von Steinbuechel, N., Richter, S., Morawetz, C., and Riemsma, R. (2005a). Assessment of Subjective Health and Health-related Quality of Life in Persons with Acquired or Degenerative Brain Injury. Curr. Opin. Neurol. 18, 681-691.

von Steinbuechel, N., Petersen, C., Bullinger, M., and the QOLIBRI Group (2005b). Assessment of Health-related Quality of Life in Persons after Traumatic Brain Injury – Development of the QOLIBRI, a Specific Measure. Acta Neurochir. 93, 43-49

von Steinbüchel N., Lischetzke T., Gurny M., and Eid, M. (2006) Assessing quality of life in older people: psychometric properties of WHOQOL-BREF. Eur. J. Ageing 3, 116-122

von Steinbüchel, N., Petersen, C., and Bullinger, M. (2002). Study Protocol, Cognitive Debriefing Protocol and Harmonization Guidelines for the QOLIBRI Development in TBI Study. Unpublished.

von Steinbuechel N, Wilson L, Gibbons H, Hawthorne G, Höfer S, Schmidt S, Bullinger M., Maas A., Neugebauer E., Powell J., von Wild K, Zitnay G., Bakx W, Christensen A.-L., Koskinen S., Formisano R., Jana Saarajuri J., Sasse, N., Truelle J.-L., and the QOLIBRI Task Force (submitted) Quality of Life after Brain Injury (QOLIBRI) - scale validity and correlates of quality of life. Journal of Neurotrauma

Ware, J.E., Snow, K.K., Kosinski, M., and Gandek, B. (1993). *SF-36 Health Survey Manual and Interpretation Guide*. New England Medical Center, The Health Institute: Boston.

Ware, J.E., Keller, S.D., Gandek, B., Brazier, J., and Sullivan, M. (1995). Evaluating Translations of the Health Status Questionnaires. Methods from the IQOLA project. Int. J. Technol. Assess. Health Care 11, 525-551.

Ware, J.E. and Kosinski, M. (1996). The SF-36 Health Survey (Version 2.0) Technical Note. Boston, MA: Health Assessment Lab, September 20.

Ware, J.E. (2000). SF-36 Health Survey Update. Spine 25, 3130-3139.

Wlodarczyk, J.H., Brodathy, H., and Hawthorne, G. (2004). The relationship between quality of life, Mini-Mental State Examination, and the Instrumental Activities of Daily Living in patients with Alzheimer's Disease. Arch. Gerontol. Geriat. 39, 25-33.

WHOQOL group (1998). The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. Social Science & Medicine 46, 1569-1585.

Wilson, L., Pettigrew, L., and Teasdale, G. (1998). Structured Interviews for the Glasgow Outcome Scale and the Extended Glasgow Outcome Scale: Guidelines for their Use. J. Neurotrauma 15, 573-585.

Zigmond, A.S. and Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. Acta Psychiatr. Scand. 67, 361-370.

Zitnay, G.A., Zitnay, K.M., Povlishock, J.T., Hall, E.D., Marion, D.W., Trudel, T., Zafonte, R.D., Zasler, N., Nidiffer, F.D., DaVanzo, J., and Barth, J.T. (2008). Traumatic brain injury research priorities: The Conemaugh International Brain Injury Symposium. J. Neurotrauma 25, 1135-1152.

Table Legends

Table 1: Demographic and clinical characteristics of the study population with TBI (N=795).

Table 2: Item descriptives and scale reliability analyses.

Table 3: QOLIBRI scale means and SDs, and Cronbach's α for all data, and separately for each language version and for subgroups with low and high MMSE/TICS scores.

Table 4: Test-retest intra-class correlations with 95% confidence intervals for all retested participants, and those with low scores on MMSE or TICS (borderline or impaired cognition) versus those with scores in the normal range.

Table 5: QOLIBRI scale percentage scores and standard deviations (SD) for retested participants at first and second test.

Table 6: Test-retest intra-class correlations for the four languages with sufficient number of retested participants (N ≥ 48).

Table 7. Measures of item difficulty and fit from Rasch analysis of each of the QOLIBRI scales. Items within each scale are ordered by item 'difficulty' (i.e. from most likely to least likely to be endorsed positively).

Table 8: Principal components analysis of the items in the QOLIBRI. Factor loadings >.25 are shown.

Table 9: Intercorrelations (Spearman's rho) between QOLIBRI scale scores.