A review of quality of life instruments used in liver transplantation

Colleen L. Jay\textsuperscript{1}, Zeeshan Butt\textsuperscript{1,2,3}, Daniela P. Ladner\textsuperscript{1}, Anton I. Skaro\textsuperscript{1}, and Michael M. Abecassis\textsuperscript{1}

\textsuperscript{1} Division of Organ Transplantation, Department of Surgery, Feinberg School of Medicine, Northwestern University, 750 N. Lake Shore Drive, Chicago, IL 60611, USA

\textsuperscript{2} Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

\textsuperscript{3} Institute for Healthcare Studies, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

SUMMARY

With improvements in patient and graft survival after liver transplantation, recipient quality of life (QOL) has become an important focus of patient care and clinical outcomes research. To provide a better understanding of the instruments used to assess QOL in the adult liver transplant population, we conducted a systematic review of the MEDLINE database and Cochrane library. Our review identified 128 relevant articles utilizing more than 50 different QOL instruments. Generic health status instruments are the most commonly used, and among them the Medical Outcomes Study Short Form-36 (SF-36), the Hospital Anxiety and Depression Scale (HADS), and the Beck Depression Inventory (BDI) are the most prevalent. Few studies (16\%) included targeted, disease-specific instruments. The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Quality of Life questionnaire, the Liver Disease Quality of Life questionnaire, and the Chronic Liver Disease questionnaire are the most frequently employed targeted instruments; however, these instruments have been designed to assess QOL in patients with chronic liver disease rather than patients after liver transplantation. The present review focuses on the psychometric properties of the existing QOL instruments and discusses their individual strengths and limitations in evaluating liver transplantation recipients. The lack of a gold-standard QOL instrument for liver transplant recipients is an impediment to cross-study comparisons. We conclude that the development of a QOL instrument specifically for liver transplant recipients will improve QOL assessment in this population leading to a more nuanced understanding of the factors that influence transplant recipients’ well-being.

Keywords

Liver transplantation; Quality of life; Psychometrics; SF-36; LDQOL; Systematic review

BACKGROUND

Liver transplantation is a life-saving intervention for an increasing number of patients with end-stage liver disease [1]. Steady improvements in graft and patient survival have been
achieved over the past two decades. One year adjusted patient and graft survival rates were 87.9% and 82.3% respectively, for deceased donor liver transplants in 2005 [1]. This represents an increase of approximately 25% for both patient and graft survival since 1987. With significant improvements in survival and a more recent plateau of these gains [1], the focus on outcomes measures has shifted towards inclusion of patient-reported quality of life (QOL). The concept of QOL complements the World Health Organization (WHO) definition of health as “a state of complete physical, mental and social well-being and not merely the absence of disease and infirmity” [2]. Hence, patient-reported outcomes are being increasingly emphasized in recent years and have become an integral component of many ongoing clinical trials [3,4]. However, meaningful assessment of patient QOL relies on the ability to reliably and accurately assess well-being using psychometrically robust instruments.

Multiple studies have reported on QOL in patients with chronic liver disease, liver transplant candidates, and transplant recipients. However, the differences between the various QOL instruments in use and the degree to which these instruments capture the true impact of liver transplantation on QOL have not been rigorously assessed. In order to evaluate the currently available instruments, we have conducted a systematic review of the literature focusing on patient-reported QOL assessment in the context of liver transplantation. We critically evaluated the psychometric properties of these instruments and their ability to measure concerns specific to the liver transplant population. To date, more than 50 instruments have been used to assess QOL in liver transplant candidates or recipients. This is the first review to discuss the relevant merit of the most frequently-used instruments.

**Materials and methods**

A systematic search of the MEDLINE database (1969 to 2008) and the Cochrane library was performed to identify liver transplant articles that included a patient-reported QOL assessment. To capture any relevant articles, a search using the Medical Subject Headings (MeSH) terms “liver transplantation” AND “quality of life” was constructed. Citations associated with the keywords “QOL”, (health-related quality of life) “HRQL”, OR “HRQOL” were also included. The search was limited to original articles relevant to humans and available in English. Publications were included if they used a patient-reported QOL assessment in liver transplant candidates or recipients. Articles using ad hoc instruments or instruments that could not be identified were excluded. Studies were also excluded if they focused exclusively on pediatric patients or living donors. The full text was reviewed prior to exclusion. Furthermore, the bibliographies of review articles on this topic were reviewed for any relevant publications to verify the completeness of our search. This search was performed by a single reviewer (CJ). We reviewed the QOL instruments identified through this search according to psychometric criteria adapted from the Scientific Advisory Committee of the Medical Outcomes Trust criteria [5,6]. (Table 1)

**Results**

**Identified Studies**

We identified 468 citations in our original search. A total of 340 articles were eliminated after applying the exclusion criteria outlined above. One hundred and twenty eight articles met final inclusion criteria. (Figure 1, Appendix A) The first assessment of QOL outcomes in liver transplantation was published in 1988, more than 20 years after the first successful liver transplant performed by Starzl [7]. In the last decade, interest in this area has increased significantly. This is illustrated by the fact that MEDLINE identified more publications using the terms of “liver transplantation” and “quality of life” since 2000 than in the preceding two decades.
This review evaluates both generic and targeted QOL instruments that have been utilized in liver transplant candidates or recipients. Generic instruments are comprised of questions broad enough to be applicable to a wide variety of populations and disease states. Generic instruments were used in the majority of studies with the Medical Outcomes Study Short-Form 36 (SF-36) being the most frequently utilized. The SF-36 was administered in 54 of the 128 selected studies, and an additional nine publications incorporated a related short form instrument, such as the SF-12. Other commonly used generic instruments include the Hospital Anxiety and Depression Scale (n=13), the Beck Depression Inventory (n=10), the EuroQOL-5D (n=8), and the Sickness Impact Profile (n=8). Table 2 summarizes the generic instruments evaluated in this study.

Targeted QOL instruments include items focused on disease- or treatment-specific topics which are of particular interest to the population being assessed. Only 16% of the articles (n=21) included a targeted instrument. The most frequently applied targeted instruments included the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Quality of Life questionnaire (n=9), the Liver Disease Quality of Life questionnaire (n=7), and the Chronic Liver Disease questionnaire (n=4). Table 3 lists the characteristics of the targeted instruments reviewed.

Multiple cultural/language adaptations exist for both these generic and targeted instruments. A complete list of all QOL instruments used in adult liver transplant candidates or recipients is provided in Table 4.

Generic QOL Instruments

The Medical Outcomes Study Short Form-36 (SF-36)—The SF-36 is comprised of 36 items, which are used to derive eight health subscales. These scales assess physical functioning, role physical (limitations due to physical health), bodily pain, general health, vitality, social functioning, role emotional (limitations related to emotional problems), and mental health. The eight subscales can also be summarized into a physical and mental component summary score. The SF-36 was first validated in healthy populations in the United States, Great Britain, and Sweden,[8-12] and has since been used across many medical disciplines to assess health status [10,11]. An advantage of the SF-36 is the comparability of scores with norms published for many different cultural and disease populations. The SF-36 was first used to assess QOL in liver transplant recipients in 1993 and remains the most popular instrument, appearing in 11 of 13 studies in 2008.

Internal consistency estimates (Cronbach’s $\alpha$) of 0.77 – 0.94 were ascertained in a study of chronic liver disease patients referred for transplant evaluation [13]. In liver transplant recipients, SF-36 scores were significantly associated with rates of unemployment and disability post-transplant establishing construct validity [14]. Phillips et al. demonstrated predictive validity of the SF-36 in liver transplant recipients by showing a significant correlation between pre-transplant SF-36 scores and post-transplant morbidity, mortality, and resource utilization [15]. The number of longitudinal studies using the SF-36 is limited; however, several studies have demonstrated responsiveness of this instrument to changes in health over time, as demonstrated by improvements in scores following transplant [16–19]. The ability of the SF-36 to detect the sizable differences in health between end-stage liver disease patients and transplant recipients still leaves questions about the sensitivity of the SF-36 to detect smaller changes over time. Moreover, measurement with the SF-36 did not identify significant differences in QOL in transplant recipients with prolonged intensive care stays or need for hemodialysis compared to those with uncomplicated recoveries [9]. Similarly, Saab et al. and Kanwal et al. both demonstrated a lack of correlation between QOL measured in patients awaiting transplantation and disease severity according to Model for End-stage Liver Disease.
Disease (MELD) scores [20,21]. These findings highlight potential shortcomings in the SF-36 to capture aspects of health relevant and specific to liver transplantation.

**Hospital Anxiety and Depression Scale (HADS)**—The HADS [22] is a 14 item questionnaire designed to detect anxiety and depression [23,24]. Questions focus on mood, interest in activities, and anxiety and panic symptoms. Notably, the HADS does not include items addressing somatic symptoms to avoid overestimation of psychological distress related to physical illness. The HADS was originally validated in general medical (non-psychiatric) and surgical patients [22,25]. In support of the scale's criterion validity, authors found strong correlations between psychiatrists' ratings of disease severity and subscale scores (Spearman correlation, $r = 0.70$ for depression and 0.74 for anxiety) [22]. The HADS has been shown to discriminate mild degrees of anxiety and depression [25].

Forsberg et al. observed correlations between HADS scores and SF-36 scores in liver, heart, and kidney transplant recipients [26]. Nickel et al. also demonstrated a correlation between scores on these two scales in liver transplant recipients. Furthermore, they found an association between HADS scores and employment and involvement in volunteer activities [27]. However, Hellgren et al. reported contradictory findings with no association between HADS scores and employment in liver transplant recipients [9]. In a longitudinal study of patients on the waiting list, HADS scores showed little change over time, but adequate information is not available to determine if this is due to a lack of responsiveness of the instrument or the stability of these symptoms [28]. HADS scores were sensitive to the changes that occur following transplant [16]. Furthermore, HADS has been utilized as a predictor of psychosocial outcomes after liver transplant [29]. Since its first use in liver transplant recipients in 1998, HADS continues to appear in one to two studies annually in this population.

**Beck Depression Inventory (BDI)**—The BDI [30] was developed to measure depressive symptoms in adults and adolescents. It is a 21 item questionnaire including items assessing hopelessness, irritability, cognition, guilt, fatigue, weight loss, and sexual interest. Unlike the HADS, the BDI includes seven items that assess somatic symptoms, some of which are common in liver transplant recipients. The BDI was first validated in populations of hospitalized and outpatient psychiatric patients [30,31]. Scores on this patient-reported inventory showed strong correlation with clinician-assessed depression severity establishing criterion validity.

Singh et al. found a strong association between BDI and Karnofsky performance status scores in a prospective study of patients being evaluated for liver transplantation. BDI scores were also noted to be a significant predictor of waitlist mortality [32]. Santos et al. demonstrated that depression according to BDI scores was associated with worse scores on seven of eight SF-36 domains in a study of liver transplant recipients [33]. Depressive symptoms according to the BDI were also associated with reports of inability to perform daily or professional activities [34]. The BDI was responsive to changes occurring following transplant in several studies of patients undergoing liver transplant including those with alcoholic cirrhosis and hepatitis C [32,33,35]. However, the BDI has become less popular in studies of liver transplant recipients in recent years and has been used in only two studies since 2000. This may be related to the recognition that common somatic complaints in liver disease patients may artificially inflate estimates of depressive symptoms in this population.

**EuroQOL-5D (EQ-5D)**—The EQ-5D [36] assesses five single item concepts including mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. It also includes a visual analog scale (VAS) asking patients to rate their health on a scale of 0–100 from worst possible to best possible health. The EQ-5D was developed as a preference-based index measure to enable cross-national comparisons, form a single index value, and to be applied to
many different disease states [37,38]. In a survey of the British general population, Brazier et al. validated the EQ-5D by demonstrating expected health trends according to age, gender, and socioeconomic status and a strong correlation with SF-36 scores [38]. Some authors have noted ceiling effects, suggesting the need for cautious usage in patients with minor morbidity [39]. The short length of this instrument reduces responder burden but may limit precision compared to longer questionnaires [40]. Given its short length, this scale was intended to be used in conjunction with other measures [6].

Bryan et al. established construct validity of the EQ-5D in liver transplant recipients by showing significant differences in scores according to disease severity and duration [41]. Ratcliffe et al. found comparable improvements in scores according to disease severity and duration [41]. Ratcliffe et al. demonstrated responsiveness of EQ-5D scores to changes that occurred following transplant. No significant differences were identified in patients followed for three months on the waitlist, but unadjusted scores were responsive to changes occurring over the first two years post-transplant [42]. Over the past decade, the EQ-5D has remained a common method of estimating utilities for cost-effectiveness analyses of liver transplant recipients [43,44].

**Sickness Impact Profile (SIP)**—The SIP [45] assesses behaviors as a measure of the impact of a patient’s illness. It contains 136 items focusing on ambulation, mobility, body care, communication, alertness, emotional behavior, social interaction, sleep, eating, work, home management and recreation. Scores can be aggregated into a physical and a psychosocial dimension. Bergner et al. validated the SIP by comparing scores to a self and a clinician assessment of dysfunction and sickness. Authors also found an association between SIP scores and a disability index derived from the National Health Interview Survey [45]. SIP scores have also been shown to have moderate to high correlations with numerous other general health status measures [6,46].

Tarter et al. utilized the SIP to measure QOL in liver transplant candidates and recipients [47,48]. Self-reported SIP scores showed strong associations with Child’s class and caregivers’ perceptions of QOL obtained in interviews establishing construct validity. Reither et al. found correlations between scores for the SIP and the BDI and the State Trait Anxiety Inventory in liver transplant recipients [34]. Tarter et al. demonstrated responsiveness of the SIP to changes occurring in patients undergoing transplant [47]. Over time, the SIP has become less frequently used in studies of liver transplant populations, appearing in only two studies since 1998 [49]. The diminished use of the SIP may be a function of its length, which likely presents a significant burden to respondents.

**Targeted QOL instruments**

**National Institute of Diabetes and Digestive and Kidney Diseases Liver Transplant Database Quality of Life (NIDDK QOL) questionnaire**—The NIDDK QOL questionnaire [50,51] was developed for the seven year NIDDK multi-center study of liver transplant recipients. There are 63 items which are organized into the domains of general health, personal function, psychological status, social and role function, and measures of liver disease. Items were drawn from multiple established general health questionnaires and a few instruments previously used in other transplant populations including kidney and heart transplant recipients. Some items were modified from their original form, limiting comparison to previous studies. The questionnaire includes 21 disease-specific items assessing symptoms related to chronic liver disease. Several items also address side-effects related to immunosuppressant agents. Specifically, two items deal with distress related to overeating and changes in facial appearance addressing concerns associated with chronic corticosteroid use.
Belle et al. first reported results for the NIDDK QOL questionnaire in a longitudinal study of transplant recipients from three centers [50]. Internal consistency of subscales ranged from $\alpha = 0.71$ to 0.86 [51,52]. These authors found that the greatest differences between recipients were related to disease-specific symptoms demonstrating a possible sensitivity to concerns particular to liver patients [51]. In a study of patients with cholestatic liver disease, strong test-retest reliability ($r = 0.82–0.99$) and internal consistency estimates ($\alpha = 0.87–0.94$) were demonstrated [53]. Cowling et al. demonstrated significant improvements in scores at one year post-transplant establishing some degree of responsiveness [54]. There were no significant differences in second year and first year scores, but information is too limited to determine if this is due to a lack of responsiveness or stability of patients’ health state. This questionnaire was utilized consistently following the completion of the NIDDK study appearing in four studies published in 2003–2004.

**Liver Disease QOL questionnaire (LDQOL)**—The LDQOL [13] is a targeted instrument which incorporates the SF-36 as well as 76 additional items comprising 12 multi-item scales. The 12 domains are symptoms of liver disease, effects of liver disease, concentration, memory, quality of social interaction, health distress, sleep, hopelessness, loneliness, stigma of liver disease, sexual functioning and sexual problems.

Gralnek et al. established the validity of this instrument for measuring QOL in patients with chronic liver disease in a multi-center study of patients referred for liver transplant evaluation [13]. They demonstrated strong internal consistency of the disease-specific scales ($\alpha = 0.62–0.95$). Significant associations were noted between self-reported disease severity and 11 of the 12 disease-specific scales. The number of disability days was also associated with scores on 10 of the 12 scales. In addition, Kanwal et al. reported associations between Child's class and the effects of liver disease, sexual problems and sexual functioning domains [20]. More recently, a prospective validation of a short form version of the LDQOL including 36 targeted items representing nine domains in addition to the SF-36 has been published [55]. To date, no studies have included a longitudinal assessment of QOL with this instrument, so the instrument's responsiveness to change remains unknown. Although it was introduced in 2000, the LDQOL has been more frequently utilized in recent years, appearing in three studies since 2007.

**Chronic Liver Disease Questionnaire (CLDQ)**—The CLDQ [56] is a 29 item questionnaire developed for measuring aspects of QOL in patients with chronic liver disease. It includes items representing six domains measuring fatigue, activity, emotional function, abdominal symptoms, systemic symptoms and worry. These items were selected based on responses from 60 chronic liver disease patients, 20 liver experts, and a review of the literature.

Younossi et al. established construct validity according to significant differences in CLDQ scores according to Child's class [56]. They also demonstrated responsiveness of the instrument with worsening scores corresponding to increasing disease severity over time according to Child's class. A correlation between Child's class and CLDQ scores was again demonstrated in another study by Saab et al [21]. Younossi also established discriminant validity of the scale by comparing waitlisted patients to the general population as well as to other patient/disease populations [57]. They reported a strong association between CLDQ scores and resource utilization. This instrument was first introduced in 2000, and most recently appeared in a study of liver transplant recipients in 2005.

**Discussion**

Health care outcomes can be divided into three fundamental categories: survival (how long people live), cost (how much the intervention costs), and quality of life (how well people live).
While QOL is an important and established health care outcome [58,59] its measurement in liver transplant recipients has not been standardized or rigorously studied. However, QOL measurement has the capacity to obtain “a full appreciation of the impact of illness and treatment” [58] given its reliance on the patient's perspective. The premise of organ transplantation in general and liver transplantation in particular is to return people to a state of health wherein they can return to a productive, fulfilling existence. This notion is at the heart of QOL measurement.

This paper is the first to critically evaluate the existing QOL instruments that have been used in liver transplantation. We believe that a thorough appreciation of the strengths and weaknesses of existing QOL questionnaires is necessary to advance further research in this area. Of note, Tome and colleagues recently conducted a systematic review of quality of life outcomes after liver transplantation [60]. They identified 44 longitudinal studies that used a validated QOL instrument, 19 of which used the SF-36. Using a sign test on common domains for the longitudinal studies, they concluded that there was significant post-transplant improvement in general QOL, social functioning, physical health, and psychological health. While the Tome et al review is illustrative in its summary of this literature, it also highlights the diversity of QOL instruments in current use. Importantly, the lion’s share of these instruments has not been designed to assess the key symptoms and issues facing liver transplant recipients. With so many QOL instruments in use and without a clear sense of their psychometric strengths and weaknesses, it will continue to be difficult to aggregate liver transplant QOL findings in a meaningful way.

To address this need, we reviewed the wide variety of instruments that have been used to assess QOL in the liver transplant population. Generic health assessment questionnaires were most common and included the SF-36, the HADS, the BDI, the EQ-5D, and the SIP. These generic instruments have enabled researchers to make comparisons between both patients with chronic liver disease and liver transplant recipients and the general public. Initial validation studies of the generic health status instruments reported strong psychometric properties. However, these instruments were not developed specifically for liver transplant recipients, and consequently there is limited data available to assess their consistency, reliability and validity in this patient population. One notable exception is a study by Gralnek et al. demonstrating good internal consistency and construct validity of the SF-36 in a group of patients referred for transplant evaluation.

In addition, a particular advantage of the SF-36 and the EQ-5D is the ability of these instruments to obtain a utility index score. Utility measures are important for the determination of quality-adjusted life years (QALYs) used in cost-effectiveness studies. Of note, the SF-6D has been developed as a utility measure from the SF-36 [61]. This score can be determined from individual patient responses [61] or reported population scores on the eight domains [62]. The EQ-5D has also been frequently used for the computation of QALYs [63,64].

In contrast to the generic instruments, targeted instruments are used less frequently (16% of included studies). However, the targeted instruments were evaluated specifically in patients with chronic liver disease and in liver transplant recipients. As such, they include elements that focus on disease-specific items. Therefore, estimates of consistency and reliability are more accurate for the population of interest: liver transplant recipients. However, the targeted instruments developed thus far are associated with their own shortfalls. For instance, the CLDQ and LDQOL were both designed to measure symptoms specific to chronic liver disease patients, not patients undergoing liver transplantation. Although the NIDDK QOL questionnaire attempted to overcome such deficiencies by including items addressing symptoms related to corticosteroid use, immunosuppressive regimens have changed significantly over time making these questions considerably less relevant.
None of the current targeted instruments address several key aspects specific to transplant recipients including issues related to post-surgical complaints such as incisional pain, hernia formation, scarring and disfigurement. Moreover, surgical patients may experience stresses related to concerns about associated complications of general anesthesia, the procedure itself, as well as concerns about disease-transmission from the donor. Transplant recipients may also experience ongoing anxiety related to both acute and chronic graft failure. Cancer risks, opportunistic infections, other side effects of immunosuppressive agents (e.g. new onset diabetes mellitus), and the risk of recurrence of primary liver disease are other highly relevant concerns. As stated above, the NIDDK QOL questionnaire attempted to address some specific concerns by including items assessing changes in facial appearance and appetite related to chronic corticosteroid use, but these concerns have become less relevant in the current immunosuppression era as most liver transplant recipients are not maintained on high dose corticosteroids in the long term with few notable exceptions. In contrast, a more relevant concern is that of nephrotoxicity leading to chronic renal insufficiency and ultimately the need for dialysis as a result of the pervasive utilization of calcineurin inhibitors [65].

Additionally, the landscape of potential complications has changed over time due to the increasing use of expanded criteria livers. For instance, livers from living donors as well as those from donation after cardiac death donors are associated with increased rates of biliary complications, which may result in significant negative effects on QOL [66–68]. In a sense, liver transplantation has become a victim of its own success, resulting in higher demand in the context of a limited supply of suitable donor organs. As a consequence of this growing discrepancy, sicker recipients are receiving expanded criteria grafts defined by inferior intrinsic organ quality [69,70]. Although this use of expanded criteria grafts among sicker patients may still achieve acceptable survival outcomes, many of these patients require re-transplantation or survive with chronic complications and therefore diminished QOL. In addition, risk predictors in transplant recipients including the etiology of liver disease, recipient age and associated co-morbidities have changed over the past two decades. As the field of transplantation continues to evolve, patient-reported QOL outcomes need to be responsive to the consequences of different graft types, technical and surgical factors, immunosuppressive regimens, and recipient characteristics that may impact survivorship.

In addition to patient and graft survival, metrics based on QOL measurement, capturing the more subtle short- and long-term implications of liver transplantation, will be essential for providing improved patient care and for informing improved organ allocation policies. The recent Institute of Medicine report on survivorship highlights the ongoing health care needs of patients who survive cancer [71]. Similarly, patients who survive otherwise fatal, end-stage liver disease through organ transplantation may face issues best addressed through a chronic illness model of care. Transplant survivorship elicits a need for studies to adequately address the long-term health-related QOL needs of transplant recipients. Careful collection of QOL outcomes will provide for a more accurate assessment of both donor and recipient variables, including the use of expanded criteria grafts in specific recipient populations. Moreover, several studies have demonstrated that physicians are typically inaccurate in their estimates of patient QOL [40,72–74]. Thus, a better understanding of the QOL implications of these complex circumstances should result in enhanced physician-patient communication, which has been shown to translate into improved treatment adherence and greater patient satisfaction [40,73,75].

Some limitations of this review must be noted. Several important indices necessary to evaluate the psychometric properties of the various QOL measures are not reported in the literature. For example, there is scarce data on the responsiveness over time for many of the instruments reviewed. Additionally, reliability estimates, while available in initial validation studies in non-transplant populations, were often lacking when the scale was used for liver transplant patients.
Reliability is a property of a scale in a population and should be re-estimated when used in a new population. These shortcomings are primarily a function of the current state of the transplant QOL literature. The major strength of this review is its expansive look at the most prominent instruments in use, with special consideration for their strengths and deficiencies for measuring QOL in liver transplant candidates and recipients.

Conclusions

Based on our review, we conclude that there are no available instruments that allow for the precise and reliable assessment of the full QOL impact of liver transplantation. We have pointed out both the strengths and weaknesses of the generic and targeted instruments that have been commonly used in liver transplant candidates and recipients. We believe that these instruments do not address important issues such as surgical concerns, the effects of immunosuppressant medications, and donor and recipient variables that may impact QOL in liver transplant recipients. The most direct approach to remedy this shortcoming is the development of a liver transplant specific QOL instrument that draws on aspects of existing instruments for end-stage liver disease with modifications and/or additions based on input from liver transplant recipients, surgeons, hepatologists, and the existing literature. The consistent application of validated, treatment-specific QOL instruments will allow for a more accurate assessment of the impact of different donor, recipient and treatment factors. A full understanding of not only survival benefit but of QOL benefit will guide us towards improved patient care and more judicious organ allocation decisions.

Acknowledgement

Research for this paper was done in part while the lead author was a National Research Service Award postdoctoral fellow with the Division of Organ Transplantation at Northwestern University, Feinberg School of Medicine under an institutional award from the Agency for Healthcare Research and Quality, 5 T32 DK077662–02 (PI: Michael Abecassis, MD MBA).

Appendix A: Included articles


Abbreviations

QOL, Quality of Life
SF-36, Short Form-36
REFERENCES


62. Ara R, Brazier J. Predicting the Short Form-6D Preference-Based Index Using the Eight Mean Short Form-36 Health Dimension Scores: Estimating Preference-Based Health-Related Utilities When Patient Level Data are not Available. Value Health. Jul 18;2008 [Epub ahead of print]

Figure 1.
Search strategy and results.
Table 1

Psychometric properties of QOL instruments.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual and measurement model</td>
<td>This involves the rationale for the concepts to be measured and an operationalization of these concepts.</td>
</tr>
<tr>
<td>Reliability</td>
<td>This is the degree to which the instruments are free from random error. Reliability is further divided into the concepts of:</td>
</tr>
<tr>
<td></td>
<td>a. <strong>Internal consistency</strong> is the degree to which items on a scale measure the same concept, reported in terms of Cronbach alpha coefficients (minimum suggested $\alpha = 0.5$ to 0.7 for group comparisons and $\geq 0.85$ for individual).6,7,6.</td>
</tr>
<tr>
<td></td>
<td>b. <strong>Test-retest reliability</strong> is the extent to which similar scores are obtained from the same individuals over short time periods, expressed as correlation coefficients (coefficients of 0.6 or greater have been suggested as a measure of “good” or “substantial” reliability).7,7,8.</td>
</tr>
<tr>
<td>Validity</td>
<td>The extent to which an instrument measures what it purports to measure is expressed as validity.</td>
</tr>
<tr>
<td></td>
<td>a. <strong>Criterion-related validity</strong> is how well the instrument compares with accepted “gold-standard” measures.</td>
</tr>
<tr>
<td></td>
<td>b. <strong>Construct-related validity</strong> is the degree to which an instrument behaves as expected when compared with other measures of health. Construct validity also involves the ability to distinguish between known groups who should report different measures of health.6.</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>The criteria assesses the ability to detect changes over time corresponding to changes in patients’ health status</td>
</tr>
<tr>
<td>Burden</td>
<td>The time and effort demanded to complete the survey is a measure of the instruments’ burden.</td>
</tr>
<tr>
<td>Cultural/Language Adaptations</td>
<td>Information about translational process and evidence regarding conceptual and linguistic equivalence enable assessment of the quality of these adaptations.</td>
</tr>
</tbody>
</table>

*adapted from review criteria developed by the Scientific Advisory Committee of the Medical Outcomes Trust.5,6
Table 2
Characteristics of Generic Health Status Measurements.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Medical Outcomes Study Short Form-36 (SF-36)</th>
<th>Hospital Anxiety and Depression Scale (HADS)</th>
<th>Beck Depression Inventory (BDI)</th>
<th>EuroQOL 5D (EQ-5D)</th>
<th>Sickness Impact Profile (SIP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burden: # items</td>
<td>36</td>
<td>14</td>
<td>21</td>
<td>6</td>
<td>136</td>
</tr>
<tr>
<td>Subscales</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Domains (subscales)</td>
<td>Physical functioning</td>
<td>Anxiety</td>
<td>Depression</td>
<td>Mobility</td>
<td>Ambulation</td>
</tr>
<tr>
<td></td>
<td>Role-physical</td>
<td>Depression</td>
<td></td>
<td>Self-care</td>
<td>Body care/movement</td>
</tr>
<tr>
<td></td>
<td>Role-emotional</td>
<td></td>
<td></td>
<td>Usual activities</td>
<td>Mobility</td>
</tr>
<tr>
<td></td>
<td>Mental health</td>
<td></td>
<td></td>
<td>Pain/discomfort</td>
<td>Social interaction</td>
</tr>
<tr>
<td></td>
<td>Vitality</td>
<td></td>
<td></td>
<td>Anxiety/depression</td>
<td>Alertness behavior</td>
</tr>
<tr>
<td></td>
<td>Bodily pain</td>
<td></td>
<td></td>
<td></td>
<td>Emotional behavior</td>
</tr>
<tr>
<td></td>
<td>General health</td>
<td></td>
<td></td>
<td></td>
<td>Communication</td>
</tr>
<tr>
<td></td>
<td>Social Functioning</td>
<td></td>
<td></td>
<td></td>
<td>Sleep/rest</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eating</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Work</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Home management</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Recreation</td>
</tr>
<tr>
<td>Summary Scores</td>
<td>yes (2) (physical/mental)</td>
<td>---</td>
<td>---</td>
<td>yes</td>
<td>Yes (2) (physical/psychosocial)</td>
</tr>
<tr>
<td>Internal Consistency\textsuperscript{a}</td>
<td>&gt; 0.738 (0.77–0.94)\textsuperscript{b}</td>
<td>&gt; 0.8080</td>
<td>0.8630</td>
<td>Not applicable\textsuperscript{c}</td>
<td>0.84–0.9445,46</td>
</tr>
<tr>
<td>Test-retest reliability\textsuperscript{a}</td>
<td>0.60–0.818</td>
<td>&gt; 0.8400</td>
<td>---</td>
<td>0.70–0.8581</td>
<td>0.85–0.9245,46</td>
</tr>
<tr>
<td>Utility measure (for QALY)</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Cultural/Language adaptions</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

\textsuperscript{a} values are reported from original validation studies (not in liver disease or transplant population)

\textsuperscript{b} in chronic liver disease population being evaluated for transplant13

\textsuperscript{c} non-applicable due to one item per domain
### Table 3
Characteristics of Liver Disease-Targeted Health Status Measurements.

<table>
<thead>
<tr>
<th></th>
<th>NIDDK Liver Transplant Database QOL questionnaire</th>
<th>Liver Disease Quality of Life questionnaire (LDQOL)</th>
<th>Chronic Liver Disease questionnaire (CLDQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors</strong></td>
<td>Belle, 199550</td>
<td>Gralnek, 200013</td>
<td>Younossi, 199956</td>
</tr>
<tr>
<td>Burden: # items</td>
<td>63</td>
<td>112</td>
<td>29</td>
</tr>
<tr>
<td>Subscales</td>
<td>6</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>Summary Scores</td>
<td>yes (for 21 disease-specific items)</td>
<td>yes (3) (SF-36 PCS/MCS &amp; disease-specific)</td>
<td>yes</td>
</tr>
<tr>
<td>Internal Consistency$^a$</td>
<td>0.62−0.9952.53</td>
<td>0.62−0.9513</td>
<td>&gt; 0.7956</td>
</tr>
<tr>
<td>Test-Retest Reliability$^a$</td>
<td>0.87−0.9453</td>
<td>0.5913</td>
<td>---</td>
</tr>
<tr>
<td>Domains (subscales)</td>
<td>General health</td>
<td>Measures of liver disease</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Role function</td>
<td></td>
<td>Fatigue</td>
</tr>
<tr>
<td></td>
<td>Social function</td>
<td></td>
<td>Activity</td>
</tr>
<tr>
<td></td>
<td>Psychological status</td>
<td></td>
<td>Emotional function</td>
</tr>
<tr>
<td></td>
<td>Personal function</td>
<td></td>
<td>Abdominal symptoms</td>
</tr>
<tr>
<td></td>
<td>Measures of liver disease</td>
<td></td>
<td>Systemic symptoms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Worry</td>
</tr>
<tr>
<td>Cultural/Language Adaptations</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

$^a$ from studies done in liver disease / liver transplant population
**Table 4**

List of all quality of life instruments included in this review.

| Anamnestic Comparative Self Assessment Scale (ASCA) |  |
| Austin Quality-of-Life Scale |  |
| Beck Anxiety Inventory (BAI) |  |
| Beck Depression Inventory (BDI) |  |
| Berlin Mood Questionnaire |  |
| Berlin questionnaire of current moods (BSF) |  |
| Billing and Moos Inventory of coping with illnesses |  |
| Brief Symptom Inventory (BSI) |  |
| Center for Epidemiologic Studies Depression Scale (CES-D) |  |
| Chronic Liver Disease questionnaire (CLDQ) |  |
| European Organization for Research and Treatment of Cancer QOL questionnaire (EORTC QLQ) |  |
| EuroQOL-5D (EQ-SD) |  |
| Everyday life abilities questionnaire (ALL) |  |
| Fatigue Severity Scale (FSS) |  |
| Ferrans and Powers questionnaire |  |
| Functional Assessment of Cancer Therapy (FACT-G) |  |
| Functional Living Index Cancer |  |
| General condition questionnaire (GCQ) |  |
| General quality of life inventory (GQOLI-74) |  |
| Giessen Complaint Questionnaire |  |
| Hospital Anxiety and Depression Scale (HADS) |  |
| Index of General Affect |  |
| Index of Well Being (IWB) |  |
| Interdisciplinary Group for Cancer Care Evaluation in Italy (GIVIO) questionnaire |  |
| Jalowiec Coping Scale (JCS-40) |  |
| Leeds Anxiety and Depression Scale |  |
| LEIPAD quality of life questionnaire |  |
| Life Orientation Test (LOT) |  |
| Linear Analogue Self-Assessment scales (LASA) |  |
| Liver Disease QOL instrument (LDQOL) |  |
| Liver Disease Symptom index |  |
| Medical Outcomes Study Short Form-36 (SF-36) |  |
| MOS-20 (and other MOS short form versions) |  |
| Memphis survey |  |
| Mental Health Inventory |  |
| Multi-dimensional Fatigue Inventory (MFI) |  |
| NIDDK QOL questionnaire |  |
| Nottingham Health Profile (NHP) |  |
| PICAVIRES (profile of QOL related to health) |  |
| Profile of Mood Survey (POMS) |  |
| Psychological general well-being Index (PGWI) |  |
| Psychosocial Adjustment to Illness Scale (PAIS) |  |
| Quality of Life Index-Liver Transplant Version (QLI-LT) |  |
| Quality of Life Inventory                |
| Questionnaire on Life Satisfaction     |
| Questions on Life Satisfaction survey |
| Sense of Coherence Scale (SOC)         |
| Sickness Impact Profile (SIP)          |
| Social embedding questionnaire (SEB)   |
| Social Reintegration Questionnaire     |
| STALX anxiety questionnaire            |
| State Trait Anxiety Survey (STAI)      |
| Symptoms Check List – 90 (SCL-90-R)    |
| Transplant Effects Questionnaire TxEQ-D|
| UCLA Loneliness Scale-Revised         |
| Visual Analog Scale measuring QOL     |
| Ways of Coping Scale (WOC)            |
| World Health Organization QOL questionnaire (WHO-QOL) |