Quality of Life Assessed with the Medical Outcomes Study Short Form 36-Item Health Survey of Patients on Renal Replacement Therapy: A Systematic Review and Meta-Analysis

Ylian S. Liem, MD, MSc,1 Johanna L. Bosch, PhD,1,2 Lidia R. Arends, MSc, MA,1 Majanka H. Heijenbrok-Kal, PhD,1 M. G. Myriam Hunink, MD, PhD1,3

1Erasmus University Medical Center Rotterdam, the Netherlands; 2Massachusetts General Hospital, Boston, MA, USA; 3Harvard School of Public Health, Boston, MA, USA

ABSTRACT

Objectives: The Medical Outcomes Study Short Form 36-Item Health Survey (SF-36) is the most widely used generic instrument to estimate quality of life of patients on renal replacement therapy. Purpose of this study was to summarize and compare the published literature on quality of life of hemodialysis (HD), peritoneal dialysis (PD), and renal transplant (RTx) patients.

Methods: We used random-effects regression analyses to compare the SF-36 scores across treatment groups and adjusted this comparison for age and prevalence of diabetes using random-effects meta-regression analyses.

Results: We found 52 articles that met the inclusion criteria, reporting quality of life of 36,582 patients. The unadjusted scores of all SF-36 health dimensions were not significantly different between HD and PD patients, but the scores of RTx patients were higher than those of dialysis patients, except for the dimensions Mental Health and Bodily Pain. Point differences between dialysis and RTx patients varied from 2 to 32. With adjustment for age and diabetes, the differences became smaller (point difference 2–22). The significance of the differences of both dialysis groups compared with RTx recipients disappeared for the dimensions Vitality and Social Functioning. The significance of the differences between HD and RTx patients disappeared on the dimensions Physical Functioning, Role Physical, and Bodily Pain.

Conclusion: We conclude that dialysis patients have a lower quality of life than RTx patients, but this difference can partly be explained by differences in age and prevalence of diabetes.

Keywords: hemodialysis, meta-analysis, peritoneal dialysis, quality of life, renal transplantation.

Introduction

Because survival among patients with end-stage renal disease (ESRD) is improving, health-related quality of life is becoming more important as an outcome measure in the evaluation of the various renal replacement therapies (RRTs) and other therapeutic interventions for these patients. Moreover, it has been argued that quality of life of patients on RRT can predict their future morbidity and mortality [1–4]. In general, measurement of health-related quality of life is becoming more important; not only as an outcome measure in chronic disease but also as an adjustment factor in economic evaluations.

Reflecting the increasing interest, the body of literature on quality of life among patients on RRT has expanded rapidly in recent years. Of the RRTs, renal transplantation (RTx) is generally accepted as the preferred treatment for ESRD. As early as the 1980s, Evans et al. reported that quality of life is higher among RTx recipients compared with dialysis patients [5]. Other authors, however, reported that this might be explained by pre-existing differences between patients selected for the different forms of RRT [6], including differences in age, sex, ethnicity, primary renal disease, and comorbidity. This study however, included only a small number of PD and RTx patients. Studies on the difference in health-related quality of life of hemodialysis (HD) compared with that of peritoneal dialysis (PD) patients remain controversial. Some studies show a higher quality of life for PD patients as compared with hospital HD patients [5,7], whereas others found similar physical quality of life for PD and HD patients, but higher mental quality of life for PD patients [8]. Thus, considerable uncertainty remains as to the differences in quality of life of HD and PD patients and as to the differences in quality of life of dialysis patients and RTx patients when adjusted for covariates.
When confronted with such a large body of literature with disparate results it can be helpful to perform a systematic review and meta-analysis, adjusting for covariates where possible. Previously, Cameron and colleagues performed a meta-analysis of the literature on quality of life associated with RRTs. Nevertheless, this study reviewed the literature on emotional distress and/or psychological well-being measures and did not include measures of health-related quality of life [6]. The Medical Outcomes Study Short Form 36-Item Health Survey (SF-36) is the most widely used generic quality-of-life assessment instrument to estimate quality of life of patients with ESRD [9]. It has shown to be valid [10], sensitive to treatment changes [11], and to be accepted by ESRD patients [12].

The purpose of our study was therefore to summarize the published literature on the SF-36 as a measure of health-related quality of life of patients receiving HD, PD, and RTx. Furthermore, our aim was to compare the SF-36 scores across treatment groups adjusted for age and prevalence of diabetes as comorbidity.

Methods

Study Retrieval

An English literature search was performed using MEDLINE (United States National Library of Medicine, Bethesda, MD, USA) and PsycINFO (Web SPIRS, Silver Platter, New York, NY, USA). All articles from peer-reviewed journals, published before June 2005 were considered for inclusion. Additional studies were identified through the bibliographies of the articles. Studies were included if they met the following criteria: 1) they reported all SF-36 dimension scores; 2) they included at least one of the forms of RRT specified as HD, PD or RTx; 3) data were collected prospectively; and 4) the sample size was at least 10 patients per treatment group. Articles were excluded if the data were provided by proxies. We also excluded articles on quality of life of combined pancreas-kidney transplant recipients. Of articles with similar or overlapping researchers or articles from the same center, we evaluated their independence by determining when, where and how many subjects were included. If more than one published article reported data from the same subjects, the most recent article was selected, unless its sample was smaller or less information on covariates was reported.

SF-36

The SF-36 consists of eight dimensions, generating a profile of health-related quality of life [13]. These dimensions are: 1) Physical Functioning; 2) Role Limitations due to Physical Functioning; 3) Bodily Pain; 4) General Health Perceptions; 5) Vitality; 6) Social Functioning; 7) Role Limitations due to Emotional Functioning; and 8) Mental Health. Raw scores are transformed into a score between zero and hundred for each dimension. Higher scores indicate better health.

Data Extraction

A standardized data sheet was used to collect the data from the studies. Data were extracted by one reader (Y.S.L.) and independently verified by two others (J.L.B., M.H.H.). Discrepancies were resolved by discussion. Readers were not blinded to information about the authors, author affiliation, and journal name, because this has been shown to be unnecessary [14]. The extracted study characteristics included publication year, country and center of authors and patients, number of patients included, demographic and clinical patient characteristics and the eight SF-36 dimension scores. If discrepancies in numbers existed between text and tables, we extracted the number reported in the table. If SF-36 scores had to be read from a graph, we rounded off to the nearest 0.5 points.

If the study reported quality of life at multiple time-points, we chose one time-point closest to the mean time on therapy for the treatment group. For studies evaluating interventions, such as immunosuppressive regimens or exercise programs, we selected the baseline time-point to minimize the effect of interventions on the mean quality-of-life estimates, unless the intervention started at the initiation of the RRT. If the time of interview in relation to time on treatment was not reported we chose the time-point for which sample size and age were reported and if this information was available for all time-points we chose the time-point for which most demographic or clinical information was available.

If treatment groups were split up according to covariates, we preferred to use data of the total group, if reported. If, however, more demographic or clinical information was available for the split groups we included these groups as separate entries into the meta-analysis.

Data Synthesis and Analysis

We explored the data, testing for homogeneity of the variables age, sex, diabetes, time on RRT, and SF-36 dimension scores within the three treatment groups separately. After, we calculated pooled weighted means and 95% confidence intervals for these variables using random-effects models, also for the three treatment groups separately. Random-effects models weigh the outcomes of the study according to the within-trial as well as the between-trial variance [15]. We tested for statistically significant differences between the groups, using Students’ t-tests and chi-square tests. In a random-effects meta-regression analysis [16,17] we corrected the differences in SF-36 scores between the treatment groups for the covariates age and comorbidity (diabetes mellitus), for the subgroup of studies for
which this information was available. These covariates have been reported to be independent predictors of quality of life among RRT patients. We also performed a subanalysis of all studies reporting time on RRT to evaluate its effect on differences in SF-36 scores between the treatment groups. To account for multiple testing, we considered a P-value < 0.01 to reflect statistical significance for all statistical tests and models. For the analyses we used SAS 8.02 (SAS Institute Inc., Cary, NC, USA).

Results

Our literature search resulted in 192 articles. Of these articles we excluded 35 on the basis of the abstract: 2 were excluded because they concerned a thesis or book, 5 because they were review articles, 11 because quality of life of a different patient group was reported, 7 because quality of life of renal donors was reported, 2 because patients had predialysis renal insufficiency, 3 because quality of life of caregivers was reported, 2 because the sample included less than 10 patients, and 3 because no mean SF-36 scores were reported. Of the remaining 157 articles, we excluded 105 on the basis of the full text, for the following reasons: other quality-of-life measurement techniques were used (N = 2), sample size was not reported (N = 4) or less than 10 patients were analyzed (N = 1), mean SF-36 scores were not reported for all eight dimensions (N = 66), patients with predialysis renal insufficiency were included (N = 3), scores among combined groups were reported (e.g., HD and PD patients, or HD, PD, and RTx patients; N = 12), and plausible overlap with another included article (N = 17). Exclusion of these studies ensured included studies to be of good quality and all included studies were of at least level 2b evidence according to the Oxford Center for Evidence-Based Medicine classification [18]. One intervention study was initiated directly after RTx [19] and did not report absolute scores at a well-defined follow-up point, therefore we used the baseline measurements from this study.

In our meta-analysis, we included 52 studies, that reported on the quality of life of 92 groups of patients on RRT measured with the SF-36 [1,8,12,19–67]; quality of life was reported for 44 HD groups (30,372 patients), 20 PD patient groups (3262 patients), and 28 RTx groups (2948 patients) (Tables of SF-36 scores and selected study and patient characteristics for all groups from the individual studies are available online at http://www.ispor.org/valueinhealth_index.asp). Tests for homogeneity were statistically significant for all variables, meaning that the null hypothesis of homogeneity was rejected. Therefore, using random-effects rather than fixed-effects models appears to be justified. Mean age, computed using random-effects models, was not significantly different for HD (55.8 years) compared with PD (52.9 years) (P = 0.085) patients but RTx recipients (mean age = 43.7 years) were significantly younger than dialysis patients (P < 0.001) (Table 1). The majority of patients were male and there were no statistically significant differences in sex distribution among the three treatment groups (P > 0.039). Prevalence of diabetes was 24% among HD, 17% among PD, and 7% among RTx patients, with a significant difference for the HD to RTx patient comparison (P < 0.001). This prevalence might be lower than that reported by the United States Renal Data System, because most studies are from other countries than the United States and in general, prevalence of diabetes is known to be lower in Europe and Asia. The mean treatment time was 44.1 months for HD patients, 24.3 months for PD, and 63.8 months for RTx recipients. Comparing all treatment groups, PD and RTx patients had a significantly different mean time on treatment (P = 0.001).

In general, SF-36-dimension scores were significantly lower for HD and PD compared with RTx patients (Fig. 1) (P < 0.01), except for the Mental Health dimension for which PD scores were not significantly different from scores of RTx recipients (P = 0.019). Scores of HD compared with PD patients were not statistically significantly different (P > 0.055).

We found 23 studies that reported the percentage of patients with diabetes mellitus in 47 patient groups [8,21,22,24–26,30–33,37,40,44,46,47,51,53,56,58,61–63,66]. The random-effects means of the SF-36 scores computed from these studies (Table 2) were very similar to those computed from all studies (Fig. 1). Significance of the differences between treatment groups was also similar, except for the Bodily Pain

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographic and clinical characteristics of renal replacement therapy patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hemodialysis</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>43</td>
</tr>
<tr>
<td>Proportion male</td>
<td>37</td>
</tr>
<tr>
<td>Proportion with diabetes</td>
<td>24</td>
</tr>
<tr>
<td>Mean time on treatment (months)</td>
<td>32</td>
</tr>
</tbody>
</table>

*Mean based on random-effects model.
CI, confidence interval; N, number of groups.
dimension (no significant difference in score between PD and RTx patients) and the Mental Health dimension (no significant difference between HD and RTx patients). The point differences of the scores of the dialysis groups compared with the RTx group varied from 2 to 32.

In random-effects meta-regression analyses, we corrected for age and diabetes. The covariates did not attain statistical significance in many of the regression models. Only in the regression models estimating the Role Physical score and the Vitality score, the age covariate was associated with a $P$-value lower than 0.01. From the regression analyses and the mean age (52.9 years) and proportion with diabetes (0.16) in the total population, we computed adjusted scores (see Table 3). Compared with the unadjusted scores, the adjusted scores were higher for HD patients and lower for RTx recipients. The adjusted scores of PD patients were similar to the unadjusted scores. Thus, in general, the differences between the groups became smaller. This is also reflected in the point differences of the adjusted scores of the dialysis groups compared with the RTx group, which varied from 2 to 22 (unadjusted: 2–32). After adjustment, the significance of the differences of both dialysis groups compared with RTx recipients disappeared for the Vitality dimension and the Social Functioning dimension. In addition, the significance of the differences between HD and RTx patients disappeared on the Physical Functioning, the Role Physical, and the Bodily Pain dimensions. The nonsignificantly lower score of PD as compared with HD patients on the Role Physical dimension became statistically significant. In random-effects meta-regression analyses including time on replacement therapy, this variable did not show any effect on existing differences in SF-36 scores between the three treatment groups. In the subgroup of studies that reported age and diabetes as well as time on therapy, it did have an independent effect on the significance of three PD versus RTx comparisons. But in a meta-regression model adjusted for age and diabetes as well as time on RRT, the additional effect of time on therapy only persisted for one PD versus RTx comparison. Because we did not want to include too many covariates and because the additional effect of time on RRT was negligible, it was not included as a covariate in the full meta-regression model.

**Discussion**

The present meta-analysis corroborates the consensus that health-related quality of life differs across the different forms of RRT. Except for the Mental Health dimension, health-related quality of life as measured by the SF-36 was higher among RTx patients than among dialysis patients. SF-36 scores among HD patients compared with PD patients were not statistically significantly different. Nevertheless, meta-regression analyses revealed that some of the differences in scores between dialysis patients and RTx recipients...
Changes in SF-36 scores after adjustment were more pronounced for HD and RTx patients as compared with PD patients, since the average mean age and prevalence of diabetes in PD patients was closer to the average mean age and prevalence of diabetes across all treatment groups. This may explain why the differences between scores of PD patients and RTx patients disappeared with adjustment for only two dimensions, whereas with adjustment, the differences in scores disappeared in five dimensions for HD compared with RTx patients.

Wu and colleagues [67] report both unadjusted SF-36 scores and scores adjusted for age, sex, race, education, albumin, creatinine, hemoglobin, and comorbidity score for HD and PD patients after 1 year of dialysis treatment. Unadjusted scores of PD patients were significantly higher for the Bodily Pain dimension and lower for the Vitality dimension. After adjustment, the difference in Vitality disappeared, but the difference in Bodily Pain remained. Nevertheless, all differences were borderline significant (P-value between 0.03 and 0.05). Merkus and colleagues showed that a significantly higher unadjusted quality of life of PD patients as measured on four dimensions of the SF-36 only persisted for the Mental Health dimension after adjustment [44].

From their study among HD, PD, and RTx patients and patients receiving conservative therapy, Baiardi and colleagues [21] conclude that treatment method and age independently influenced quality of life. Dimensions most affected were Physical Functioning, Bodily Pain, General Health, and Vitality. Scores for the four dimensions corrected for mean age and hemoglobin level showed that for the Physical Functioning and Bodily Pain dimensions, patients receiving conservative treatment and RTx patients had better results than those on dialysis. Similar to our results, compared with their unadjusted scores the adjusted scores of HD patients were in general higher and the adjusted scores of RTx patients were lower. Adjusted scores of PD patients were only slightly higher than unadjusted scores. In general, differences across the treatment groups became smaller. Adjusted differences in quality of life between dialysis and RTx patients might be smaller than generally thought because RTx patients commonly have a long history including dialysis, which affects their quality of life.

In a meta-analysis of emotional distress and psychological well-being of HD, PD, and RTx patients, Cameron and colleagues reported comparable differences in quality of life [6]. The authors did not formally adjust for covariates influencing these differences, but did report differences in covariates among the treatment modalities. They concluded that because RTx patients are generally younger, healthier, more highly educated, and more likely to be employed, it cannot be ruled out that differences in emotional distress and psychological well-being are partly a result of differences in case mix.

To be able to assess real differences in quality of life among RRT patients, Cameron and colleagues suggested a prospective repeated-measures experimental design in which the same cohort of patients can be assessed repeatedly and at clinically significant milestones, such as when a patient switches to a different treatment modality [6]. Nevertheless, such switches are usually induced by changes in demographic or clinical variables, so analysis of data from such a study should still be adjusted for case-mix variables. Another solution the authors suggested was to limit research participants to those patients for whom any form of RRT would be equally suitable. Results, however, would not be generalizable to the entire patient population. This problem is also relevant to our meta-regression analyses, because we used the average mean age and proportion with diabetes across all treatment groups to calculate adjusted scores. Adjusted scores should therefore not be interpreted as actual scores. More interesting is the direction of the change of the scores after adjustment and the significance of the
differences of adjusted scores among the treatment groups.

There are several limitations to our study that should be mentioned. First, the results of meta-analyses always rely on completeness of available published literature. Thus, they are known to be influenced by publication bias. Nevertheless, ours is a study of mostly noncomparative studies and from the comparative studies that we included only absolute measures were extracted. So if publication bias should have affected our study, its effect can be assumed to be very small.

Second, we would have wanted to adjust for all possible case-mix differences that might influence reported health-related quality of life. Age, sex, ethnicity, socioeconomic status, education, employment, and income are considered to be independent demographic predictors of quality of life [68]. Additionally, several disease-associated factors such as primary renal disease, treatment history, anemia and comorbid disease are also known to be associated with quality of life [68]. Unfortunately, not all these covariates were consistently published in the studies and the number of studies was too small to be able to correct for many case-mix differences. Therefore, we decided to only adjust for case mix in a subset of studies and for a limited number of covariates. But even if more studies reporting all these covariates would have been available, caution is required in selecting covariates for the regression analysis. Covariates need to be prespecified to avoid data dredging [17]. Furthermore, results are easier to interpret when the covariate has a high variability across studies compared with within studies. But still, interpretation of meta-regression analyses using patient covariates is not always straightforward because of the fact that the relationship of the outcome with patient averages across trials may not be the same as the relationship of the outcome across patients within trials. This phenomenon is also referred to as aggregation bias or the ecologic fallacy [17].

Lambert and colleagues compared a meta-regression analysis using mean patient covariates to an analysis with individual patient-level data [69]. They found that although the estimates from the meta-regression analysis were not biased, there was a greater variation for the meta-regression estimates than for the estimates from the analysis on the individual patient-level data. They concluded that to investigate whether patient characteristics are related to treatment, a meta-analysis of summary data might not be apt and that an individual-patient-level data analysis will generally be necessary to uncover such relationships. We did, however, find significant changes with adjustment.

From this meta-analysis we conclude that HD and PD patients tend to have a lower quality of life than RTx recipients. Quality of life seems comparable for HD and PD patients. Adjusting the SF-36 scores for mean age and prevalence of diabetes in a meta-regression analysis showed that some differences in scores between dialysis and RTx patients can be partly explained by unequal patient selection for the different RRT modalities. This implies that although quality of life of dialysis patients is worse than that of RTx recipients, the difference between these groups might not be as big as generally thought. This is important in choosing an appropriate therapy for an individual patient. In targeting interventions aimed at improving quality of life, factors independently associated with quality of life, such as comorbidities, should also be taken into account.

Source of financial support: This project was supported by the Erasmus University Medical Center Rotterdam and ZonMw, The Netherlands Organization for Health Research and Development.

References


Supplementary material for this article can be found at: http://www.ispor.org/publications/value/VIHsupplementary.asp