



Published in final edited form as:

Crit Care Med. 2010 March ; 38(3): 855–860. doi:10.1097/CCM.0b013e3181cd10c7.

Baseline Quality of Life before Intensive Care: A Comparison of Patient versus Proxy Responses

Jeneen M. Gifford, M.D., Nadia Husain, M.H.S., Victor D. Dinglas, B.S., Elizabeth Colantuoni, Ph.D., and Dale M. Needham, M.D., Ph.D.

Department of Internal Medicine, Johns Hopkins Bayview Medical Center, Baltimore, MD (JMG); Department of Biostatistics, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD (NH, EC); Division of Pulmonary and Critical Care Medicine (VD, DMN); Department of Anesthesiology and Critical Care Medicine (EC), and Department of Physical Medicine and Rehabilitation (DMN), Johns Hopkins University, Baltimore, MD

Abstract

Objective—To compare acute lung injury (ALI) patients' self-reported, retrospective baseline quality of life (QOL) before their intensive care hospitalization with population norms and retrospective proxy estimates.

Design—Prospective cohort study using the Short Form 36 (SF-36) QOL survey.

Setting—13 intensive care units at 4 teaching hospitals in Baltimore, MD, USA.

Patients—136 ALI survivors and their designated proxies.

Interventions—Both patients and proxies were asked to estimate patient baseline QOL before hospital admission using the SF-36 survey.

Measurements and Main Results—Compared to population norms, QOL scores were lower in ALI patients across all 8 domains, but the difference was significantly greater than the minimum clinically important difference in only 2 of 8 domains (Physical Role and General Health). The mean paired difference between patient versus proxy responses revealed no clinically important difference. However, kappa statistics demonstrated only fair to moderate agreement for all domains. Bland-Altman analysis revealed that for all domains, proxies tended to overestimate QOL when patient scores were low and underestimate QOL when patient scores were high.

Conclusion—Retrospective assessment of QOL prior to hospitalization revealed that ALI patients were consistently lower than population norms, but the magnitude of this difference may not be clinically important. Proxy assessments had only fair to moderate agreement with patient assessments. Across all 8 SF-36 QOL domains, proxy responses represented an attenuation of patient QOL estimates.

Keywords

Critical care; Quality of life; Acute lung injury; Proxy; Respiratory distress syndrome; adult; Health status

Introduction

Quality of life (QOL) is an important outcome measure in evaluating survivors of critical illness.(1–6) ICU survivors have impaired QOL versus age- and sex-matched population norms, and these impairments can persist for months to years.(1;2;7–11) Among critically ill patients, those with acute lung injury/acute respiratory distress syndrome (ALI) may be especially likely to have poor QOL, given their high severity of illness, prolonged ICU stay, and frequent physical and mental health morbidities.(1;2;12)

Baseline QOL prior to hospitalization for critical illness is an important consideration when analyzing the subsequent effects of critical illness and ICU treatment on long-term QOL. In addition, baseline QOL is useful in predicting mortality and assisting decision-making regarding the use of life-sustaining treatments in the ICU.(13) Because critical illness is frequently sudden and unexpected, baseline QOL generally cannot be prospectively measured in most ICU patients. In these circumstances, proxies may estimate patient QOL, or ICU survivors may be asked to retrospectively assess their QOL. Both of these QOL assessment methods are subject to bias.

It is unclear whether proxy estimates of patient QOL are accurate. In one study of ALI patients, the average difference between paired patient and proxy responses was insignificant in 5 of 8 domains of the Short Form-36 (SF-36) QOL survey, but patient-proxy agreement was only fair (kappa statistic = 0.30 – 0.40).(14) However, other studies have concluded that proxy estimates are an acceptable substitute for patients' self-reported QOL in several circumstances, including during chronic illness(15), prior to hospital admission (16–19), or after ICU discharge.(20)

Given these conflicting results, our goal was to further evaluate patient versus proxy assessments of baseline QOL in ALI survivors. Our study has two specific objectives: (1) to compare baseline QOL of ALI survivors to age- and sex-matched population norms, and (2) to evaluate the agreement of proxy versus patient estimates of baseline QOL using the Short Form General Health Survey (SF-36).

Materials and Methods

Study Design and Participants

Data for this analysis was obtained from the Improving Care of Acute Lung Injury Patients (ICAP) study, an ongoing prospective cohort study.(21) The ICAP study population was consecutively enrolled from 13 intensive care units at four teaching hospitals in Baltimore, USA. Eligible patients were ≥ 18 years old and mechanically ventilated with ALI as defined by the American-European Consensus Conference criteria.(22) Relevant exclusion criteria evaluated at the time of ALI diagnosis included preexistence of: 1) comorbid disease with a life expectancy of <6 months; 2) any pre-existing communication or language barrier; 3) known pre-existing cognitive impairment; and 4) no fixed address. The study was approved by the institutional review boards (IRB) of the Johns Hopkins University and all participating institutions.

Before hospital discharge, consenting ALI survivors provided the name and contact information for their closest proxies. For this study, both patients and proxies were asked to estimate the patient's baseline QOL, defined as QOL just before the onset of the illness that resulted in the hospitalization with ALI. This QOL assessment was completed by both patient and proxy as soon as possible after patients regained capacity and provided consent for the ICAP study.(23), for patients and via phone interview for proxies who were generally

not available in hospital after patient consent. Proxies were explicitly instructed to respond using the patient's perspective of QOL as previously described.(24;25)

The SF-36 QOL instrument has 36 questions evaluating eight separate domains: Physical Functioning, Physical Role, Bodily Pain, General Health, Vitality, Social Functioning, Emotional Role, and Mental Health. The responses for each domain are scored and transformed to a 0–100 scale, with higher scores reflecting better QOL. In our study, when a domain could not be scored for either a patient or proxy (e.g. incomplete survey response), that specific domain was omitted from analysis. The SF-36 has been validated in survivors of critical illness(1;2;26) and frequently used to retrospectively assess baseline QOL in ICU survivors.(16;18;27;28) Similar to prior studies, the wording for each SF-36 item was revised slightly to evaluate baseline QOL.(29)

Statistical Analysis

For each SF-36 domain, the following statistical analyses were performed: (1) mean difference between the patient score and the associated age- and sex-matched U.S. population norm, and (2) mean difference between each patient-proxy pair. For the patient-population and patient-proxy differences, unpaired and paired t-tests were used, respectively, to test for a statistically significant difference from zero and from the minimal clinically important difference (MCID) for each domain, as estimated for patients with chronic pulmonary disease.(30;31)

Agreement between patient and proxy responses was measured using the Cohen's kappa statistic (unweighted and weighted). The kappa statistic can range from -1 (complete disagreement), to 0 (no agreement), to $+1$ (perfect agreement). For these statistics, we treated each SF-36 domain score as an ordinal variable.(32;33) For the weighted kappa, weights were assigned using a standard method for linear weighting proposed by Cicchetti and Allison.(34) This approach assigns a weight of 1 for perfect agreement and 0 for the largest possible disagreement, with weights of all other misclassifications determined linearly. Two sets of kappa and weighted kappa statistics were calculated. For the first set, perfect agreement was defined as a difference between the patient and proxy score of less than or equal to the state change (i.e. the smallest possible change in score based on a 1 unit change in response for 1 question within the domain) for the particular SF-36 domain. For the second set, perfect agreement was defined as a difference between the patient and proxy score of less than the MCID for a given domain. Based on the kappa statistic, patient-proxy agreement was qualitatively described according to recommendations from Landis and Koch: poor ($\kappa < 0$), slight ($\kappa 0 - 0.2$), fair ($\kappa 0.21 - 0.4$), moderate ($\kappa 0.41 - 0.60$), substantial ($\kappa 0.61 - 0.8$), or almost perfect ($\kappa 0.81$). (35)

In addition, Bland-Altman (B-A) plots were used to explore the relationship between the differences in the patient and proxy responses as a function of the patient response.(36) A traditional B-A plot would display the average of the patient and proxy responses along the horizontal axis. However, for this analysis, it was assumed that the patient response is measured without error, so that the patient response is most reflective of the true underlying quality of life and most appropriate for the x-axis. Linear regression models, evaluating both linear and quadratic relationships, were used to estimate the mean difference in patient and proxy responses as a function of the patient response.

For all analyses, $p < 0.05$ was considered statistically significant. All data were stored and analyzed using STATA version 10.0 (College Station, TX), except kappa statistics which were calculated using R software (University of Auckland, New Zealand).

Results

Within the ICAP study, 193 participants were potentially eligible for this patient-proxy QOL analysis. Of these potential participants, 47 were not eligible for the following reasons: no consent (primarily due to patient lack of capacity secondary to cognitive impairment, $n = 30$), no proxy available ($n=10$), or death or hospice care prior to completion of surveys ($n = 7$). Ten otherwise eligible participants were excluded because the patient and/or proxy declined to complete the survey ($n = 10$). Hence, the SF-36 data were analyzed for 136 patient-proxy pairs. Of 1088 possible patient-proxy domain scores (8 domains for each of the 136 patient-proxy pairs), only 18 (1.7%) could not be scored and reported due to missing data. Table 1 describes baseline characteristics of the patients included in this study.

The mean paired difference for the patient-population comparison demonstrated significantly greater population norms for all SF-36 domains except for Vitality where this difference did not reach statistical significance ($p=0.12$) (Table 2). For 2 of the 8 domains (Physical Role and General Health) the patient-population difference was significantly greater than the corresponding MCID.

The mean paired difference for the patient-proxy comparison demonstrated significantly greater patient values for all domains except Emotional Role where this difference did not reach statistical significance ($p=0.14$) (Table 3). General Health was the only domain where the patient-proxy difference was greater in magnitude than the corresponding MCID, but this difference was not significant ($p = 0.29$). All but one patient-proxy pair had at least one domain where the difference in scores was greater than the MCID (median = 5, IQR = 3). When differences in survey timing for patients and proxies were analyzed, patient-proxy pairs with one or more domains with a score difference ≥ 60 did not differ significantly from patient-proxy pairs with all domain score differences < 60 (Wilcoxon rank sum test, $p=0.14$).

The weighted kappa statistic, with perfect agreement defined as both a difference of less than or equal to a “state change” or a difference less than the MCID, revealed similar results with patient-proxy agreement being “fair” for 7 of the 8 domains (range: 0.32 – 0.43) (Table 4). The unweighted kappa, as expected, demonstrated a lower level of agreement with generally “slight” agreement between the patient-proxy pairs.

For all 8 SF-36 domains, the B-A analysis revealed that for both relatively low and high patient scores, the proxy estimate tended to attenuate toward a more moderate score. The General Health domain demonstrates this pattern of attenuation, especially with higher patient scores (Figure 1).

Discussion

This prospective cohort study of 136 ALI patients compared patients’ retrospective baseline SF-36 QOL estimates with age- and sex-matched population norms and retrospective proxy estimates. The ALI patient cohort demonstrated lower baseline QOL scores than population norms for all 8 SF-36 domains, but this difference was both clinically important and statistically significant for only the Physical Role and General Health domains. Proxy estimates had only fair agreement with patients, with proxy results consistently demonstrating an attenuation of patient scores for all domains.

A small number of studies have compared baseline QOL in ICU survivors with population norms. In a systematic review of QOL in ARDS patients,(2) three studies specifically evaluated baseline QOL using the SF-36 survey.(28;29;37) All three studies similarly demonstrated that when compared to population norms, ICU survivors had lower mean

baseline SF-36 scores for all domains. However, one additional recent study of 46 ALI patients demonstrated that for all domains except General Health, patients had SF-36 QOL scores that were the same or better than population norms.(14) Two studies of general ICU patient populations used proxy responses to estimate baseline QOL, with one study finding a trend toward lower SF-36 scores vs. population norms for all domains,(27) and the other finding a similar trend in 6 of 8 domains.(8) In our study, when compared to population norms, patients had lower average SF-36 scores for all domains. However, only the Physical Role and General Health domains had differences that were both statistically significant and greater than the estimated MCID. The prior studies did not compare the magnitude of the patient-population difference to an MCID estimate.

Several factors may contribute to the diversity in these results, including varying sources for the baseline QOL estimate (i.e. patients, proxies, or mixed patient and proxy), differences in the presence of comorbidities between study cohorts, and insufficient normative data for patients >65 years old in some countries. Compared to the most recent ALI study cohort which reported baseline QOL scores that were the same or better than population norms, our cohort had a similar median patient age and APACHE II severity of illness. However, the burden of patients' baseline comorbidity was likely greater in our study since the majority of patients were recruited from hospitals serving inner city patients and the prevalence of documented preexisting pulmonary disease was two-fold greater (23% vs. 11%).(10;14)

Existing studies also demonstrate variable results for agreement between patient and proxy estimates of baseline QOL in ICU survivors. A recent study of ALI patients reported findings similar to our study with patient-proxy agreement that was only fair despite a mean paired difference of less than the MCID in 7 of 8 SF-36 domains.(14) However, studies of other ICU populations have reported fair to excellent agreement using the SF-36 and other QOL instruments.(16;18;19;38) Furthermore, one review article of 23 studies of non-ICU patients with chronic illness concluded that proxy estimates are reasonably accurate and that substantial discrepancies are rare.

Few studies have been able to identify factors underlying these varied results regarding patient-proxy agreement. Of the five studies previously discussed, only two used the SF-36 (14;18). Hence, the difference in the results may be due to different QOL instruments. There are no studies directly comparing the level of patient-proxy agreement using different QOL instruments. The retrospective nature of baseline QOL assessments introduces potential recall bias from the patient's ICU stay. For example, in one study, patient and proxy QOL estimates differed significantly in six domains at hospital discharge, but agreement improved at 6 months, suggesting that temporal proximity to the acute hospital illness may have an effect on agreement.(20) However, existing research demonstrates that severity of illness, type of admission, patient education level, and the nature of the patient-proxy relationship do not affect patient-proxy agreement for QOL prior to ICU admission. (14;16;18;20;38)

Proxy attenuation of patient QOL estimates, demonstrated by the B-A analyses in our study, is a novel finding within ICU QOL research. This relationship was present for all 8 SF-36 domains and illustrates the limitations of the mean paired difference statistic in comparing patient-proxy estimates. Despite this finding being present across all SF-36 domains, the magnitude of differences within patient-proxy pairs markedly varied by domain such that pairs did not consistently have particularly large or small differences across all domains. More investigation is needed to understand this relationship and its potential implications on proxy decision-making for critically ill patients.

Our study has several potential limitations. First, the use of an MCID based on chronic pulmonary diseases may not be appropriate for ALI patients. Moreover, the MCID was determined via an expert consensus panel since no estimate based on patient report is available for ALI or general ICU patients. Second, our study does not include the necessary data to permit reporting on the nature of the relationship of the proxy to the patient or the precise timing differences between completion of the SF-36 survey within the patient-proxy pairs. However, the closest available proxy (designated by the patient) was used in this study, and prior research has demonstrated that the nature of this relationship does not significantly affect patient-proxy agreement.(16;20;38) In addition, analysis of the available data within the study suggests that the magnitude of patient-proxy differences was not associated with the estimated timing differences. Third, although the number of capable patients and/or proxies declining to participate in the QOL survey was low (5%), similar to other related studies,(16;38) approximately 15% of patients could not complete the baseline SF-36 survey due to cognitive impairment, as commonly observed during the early post-ICU in-patient period.(39) Despite this, our final sample size makes an important contribution since it is 2- to 3-fold larger than prior studies of patient-proxy SF-36 QOL comparisons. (14;18) Finally, in our study, the SF-36 survey was administered in-person for patients and via phone for proxies, a difference which may introduce a response-mode bias.(40) We attempted to minimize this bias by administering the survey via an interviewer (rather than self-administered) for both patients and proxies. Moreover, we felt that selecting a single mode of administration (i.e. changing to in-person for proxies or to phone after hospital discharge for patients) would have resulted in a greater bias due to non-response.

Conclusions

Our retrospective assessment of QOL prior to hospitalization revealed that ALI patients were consistently lower than population norms, but the magnitude of this difference may not be clinically important. Furthermore, proxy assessments had only fair to moderate agreement with patient assessments, largely due to a proxy attenuation of patient responses in all 8 SF-36 domains.

Acknowledgments

This research is supported by the National Institutes of Health (Acute Lung Injury SCCOR Grant # P050 HL 73994). Dr. Needham is supported by a Clinician-Scientist Award from the Canadian Institutes of Health Research. The funding bodies had no role in the study design, manuscript writing or decision to submit the manuscript for publication.

Reference List

1. Dowdy DW, Eid MP, Dennison CR, Mendez-Tellez PA, Herridge MS, Guallar E, Pronovost PJ, Needham DM. Quality of life after acute respiratory distress syndrome: a meta-analysis. *Intensive Care Med.* 2006 Aug; 32(8):1115–24. [PubMed: 16783553]
2. Dowdy DW, Eid MP, Sedrakyan A, Mendez-Tellez PA, Pronovost PJ, Herridge MS, Needham DM. Quality of life in adult survivors of critical illness: a systematic review of the literature. *Intensive Care Med.* 2005 May; 31(5):611–20.
3. Hofhuis JG, van Stel HF, Schrijvers AJ, Rommes JH, Bakker J, Spronk PE. Conceptual issues specifically related to health-related quality of life in critically ill patients. *Crit Care.* 2009 Feb 19.13(1):118. [PubMed: 19239721]
4. Cuthbertson BH, Rattray J, Johnston M, Wildsmith JA, Wilson E, Hernandez R, Ramsey C, Hull AM, Norrie J, Campbell M. A pragmatic randomised, controlled trial of intensive care follow up programmes in improving longer-term outcomes from critical illness. The PRACTICAL study. *BMC Health Serv Res.* 2007; 7:116. [PubMed: 17645791]

5. Boyle M, Murgo M, Adamson H, Gill J, Elliott D, Crawford M. The effect of chronic pain on health related quality of life amongst intensive care survivors. *Aust Crit Care*. 2004 Aug; 17(3):104–13. [PubMed: 15493858]
6. Graf J, Wagner J, Graf C, Koch KC, Janssens U. Five-year survival, quality of life, and individual costs of 303 consecutive medical intensive care patients--a cost-utility analysis. *Crit Care Med*. 2005 Mar; 33(3):547–55. [PubMed: 15753746]
7. Chaboyer W, Elliott D. Health-related quality of life of ICU survivors: review of the literature. *Intensive Crit Care Nurs*. 2000 Apr; 16(2):88–97. [PubMed: 11868593]
8. Hofhuis JG, Spronk PE, van Stel HF, Schrijvers GJ, Rommes JH, Bakker J. The impact of critical illness on perceived health-related quality of life during ICU treatment, hospital stay, and after hospital discharge: a long-term follow-up study. *Chest*. 2008 Feb; 133(2):377–85. [PubMed: 17925419]
9. Kaarlola A, Pettila V, Kekki P. Quality of life six years after intensive care. *Intensive Care Med*. 2003 Aug; 29(8):1294–9.
10. Herridge MS, Cheung AM, Tansey CM, Matte-Martyn A, az-Granados N, Al-Saidi F, Cooper AB, Guest CB, Mazer CD, Mehta S, et al. One-year outcomes in survivors of the acute respiratory distress syndrome. *N Engl J Med*. 2003 Feb 20; 348(8):683–93. [PubMed: 12594312]
11. Cheung AM, Tansey CM, Tomlinson G, az-Granados N, Matte A, Barr A, Mehta S, Mazer CD, Guest CB, Stewart TE, et al. Two-year outcomes, health care use, and costs of survivors of acute respiratory distress syndrome. *Am J Respir Crit Care Med*. 2006 Sep 1; 174(5):538–44.
12. Davydow DS, Desai SV, Needham DM, Bienvenu OJ. Psychiatric morbidity in survivors of the acute respiratory distress syndrome: a systematic review. *Psychosom Med*. 2008 May; 70(4):512–9. [PubMed: 18434495]
13. Hofhuis JG, Spronk PE, van Stel HF, Schrijvers AJ, Bakker J. Quality of life before intensive care unit admission is a predictor of survival. *Crit Care*. 2007; 11(4):R78.
14. Scales DC, Tansey CM, Matte A, Herridge MS. Difference in reported pre-morbid health-related quality of life between ARDS survivors and their substitute decision makers. *Intensive Care Med*. 2006 Nov; 32(11):1826–31.
15. Sneeuw KC, Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *J Clin Epidemiol*. 2002 Nov; 55(11):1130–43. [PubMed: 12507678]
16. Capuzzo M, Grasselli C, Carrer S, Gritti G, Alvisi R. Quality of life before intensive care admission: agreement between patient and relative assessment. *Intensive Care Med*. 2000 Sep; 26(9):1288–95. [PubMed: 11089755]
17. Elliott D, Lazarus R, Leeder SR. Proxy respondents reliably assessed the quality of life of elective cardiac surgery patients. *J Clin Epidemiol*. 2006 Feb; 59(2):153–9. [PubMed: 16426950]
18. Hofhuis J, Hautvast JL, Schrijvers AJ, Bakker J. Quality of life on admission to the intensive care: can we query the relatives? *Intensive Care Med*. 2003 Jun; 29(6):974–9. [PubMed: 12734653]
19. Badia X, Diaz-Prieto A, Rue M, Patrick DL. Measuring health and health state preferences among critically ill patients. *Intensive Care Med*. 1996 Dec; 22(12):1379–84. [PubMed: 8986489]
20. Rogers J, Ridley S, Chrispin P, Scotton H, Lloyd D. Reliability of the next of kins' estimates of critically ill patients' quality of life. *Anaesthesia*. 1997 Dec; 52(12):1137–43. [PubMed: 9485965]
21. Needham DM, Dennison CR, Dowdy DW, Mendez-Tellez PA, Ciesla N, Desai SV, Sevransky J, Shanholtz C, Scharfstein D, Herridge MS, et al. Study protocol: The Improving Care of Acute Lung Injury Patients (ICAP) study. *Crit Care*. 2006 Feb.10(1):R9.
22. Bernard GR, Artigas A, Brigham KL, Carlet J, Falke K, Hudson L, Lamy M, LeGall JR, Morris A, Spragg R. The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med*. 1994 Mar; 149(3 Pt 1):818–24. [PubMed: 7509706]
23. Fan E, Shahid S, Kondreddi VP, Bienvenu OJ, Mendez-Tellez PA, Pronovost PJ, Needham DM. Informed consent in the critically ill: a two-step approach incorporating delirium screening. *Crit Care Med*. 2008 Jan; 36(1):94–9. [PubMed: 18090168]

24. McPhail S, Beller E, Haines T. Two perspectives of proxy reporting of health-related quality of life using the Euroqol-5D, an investigation of agreement. *Med Care*. 2008 Nov; 46(11):1140–8. [PubMed: 18953224]
25. Pickard AS, Knight SJ. Proxy evaluation of health-related quality of life: a conceptual framework for understanding multiple proxy perspectives. *Med Care*. 2005 May; 43(5):493–9. [PubMed: 15838415]
26. Mahler DA, Mackowiak JI. Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with COPD. *Chest*. 1995 Jun; 107(6):1585–9. [PubMed: 7781351]
27. Cuthbertson BH, Scott J, Strachan M, Kilonzo M, Vale L. Quality of life before and after intensive care. *Anaesthesia*. 2005 Apr; 60(4):332–9. [PubMed: 15766335]
28. Wehler M, Geise A, Hadzionerovic D, Aljukic E, Reulbach U, Hahn EG, Strauss R. Health-related quality of life of patients with multiple organ dysfunction: individual changes and comparison with normative population. *Crit Care Med*. 2003 Apr; 31(4):1094–101. [PubMed: 12682478]
29. Ridley SA, Chrispin PS, Scotton H, Rogers J, Lloyd D. Changes in quality of life after intensive care: comparison with normal data. *Anaesthesia*. 1997 Mar; 52(3):195–202. [PubMed: 9124657]
30. Wyrwich KW, Fihn SD, Tierney WM, Kroenke K, Babu AN, Wolinsky FD. Clinically important changes in health-related quality of life for patients with chronic obstructive pulmonary disease: an expert consensus panel report. *J Gen Intern Med*. 2003 Mar; 18(3):196–202. [PubMed: 12648251]
31. Wyrwich KW, Tierney WM, Babu AN, Kroenke K, Wolinsky FD. A comparison of clinically important differences in health-related quality of life for patients with chronic lung disease, asthma, or heart disease. *Health Serv Res*. 2005 Apr; 40(2):577–91.
32. Walters SJ, Campbell MJ, Lall R. Design and analysis of trials with quality of life as an outcome: a practical guide. *J Biopharm Stat*. 2001; 11(3):155–76. [PubMed: 11725929]
33. Walters SJ. Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36. *Health Qual Life Outcomes*. 2004 May 25.2:26.
34. Cicchetti DV, Allison T. A new procedure for assessing reliability of scoring EEG sleep recordings. *American Journal of EEG Technology*. 1971; (11):101–9.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar; 33(1):159–74. [PubMed: 843571]
36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986 Feb 8; 1(8476):307–10. [PubMed: 2868172]
37. Graf J, Koch M, Dujardin R, Kersten A, Janssens U. Health-related quality of life before, 1 month after, and 9 months after intensive care in medical cardiovascular and pulmonary patients. *Crit Care Med*. 2003 Aug; 31(8):2163–9. [PubMed: 12973175]
38. Diaz-Prieto A, Gorriz MT, Badia X, Torrado H, Farrero E, Amador J, Abos R. Proxy-perceived prior health status and hospital outcome among the critically ill: is there any relationship? *Intensive Care Med*. 1998 Jul; 24(7):691–8. [PubMed: 9722039]
39. Hopkins RO, Weaver LK, Pope D, Orme JF, Bigler ED, Larson-LOHR V. Neuropsychological sequelae and impaired health status in survivors of severe acute respiratory distress syndrome. *Am J Respir Crit Care Med*. 1999 Jul; 160(1):50–6.
40. Hanmer J, Hays RD, Fryback DG. Mode of administration is important in US national estimates of health-related quality of life. *Med Care*. 2007 Dec; 45(12):1171–9.
41. Ware, JE., Jr; Kosinski, M.; Dewey, JE. How to score version 2 of the SF-36 Health Survey. Lincoln, RI: QualityMetric Incorporated; 2000.

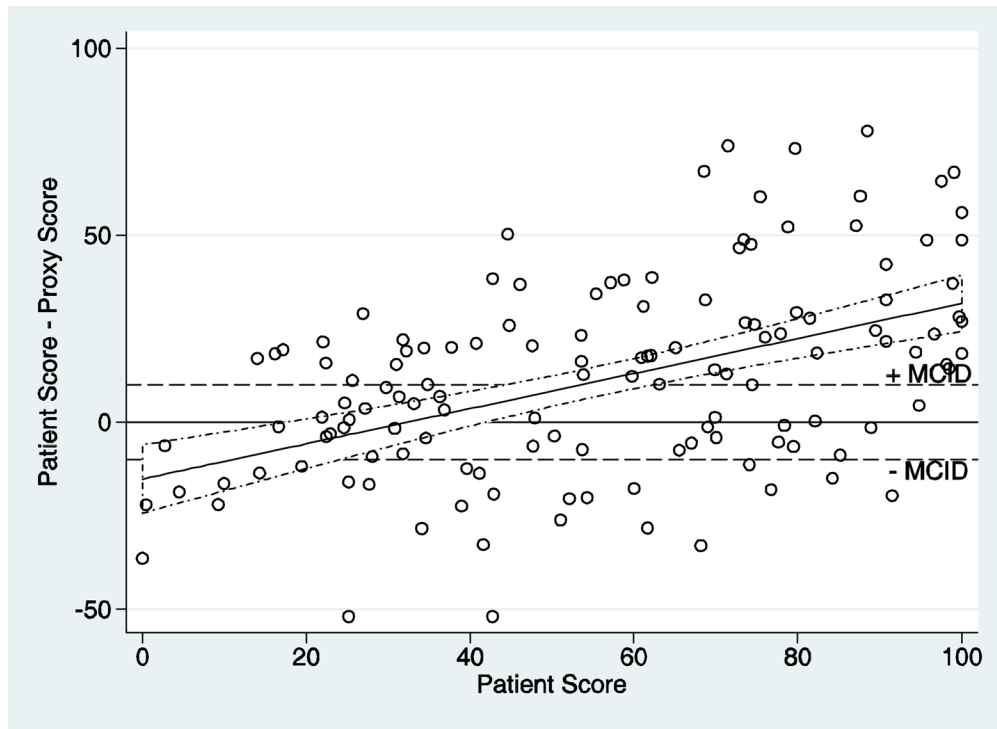


Figure 1. Bland Altman plot for the SF-36 General Health domain

This figure displays the relationship of the patient SF-36 scores to the difference between the patient and proxy scores (i.e. patient minus proxy). The area between the dashed horizontal lines represents a difference of $0 \pm$ the minimum clinically important difference (MCID) for patients with chronic pulmonary disease.(30) Each circle represents one patient-proxy pair. The fitted line represents the linear regression model used to estimate the mean difference in patient and proxy responses as a function of the patient response. The dashed lines surrounding the fitted line represents 95% confidence intervals.

Table 1

Description of study participants

Baseline characteristic	N=136
Age , median (IQR) years	49 (40–60)
Male , no. (%)	72 (53)
Race , no. (%)	
White	87 (64)
African-American	47 (35)
Other	2 (1)
Preexisting pulmonary disease^I , no (%)	31 (23)
Charlson Comorbidity score , (median, IQR)	1.0 (0.0–3.0)
APACHE II score , median (IQR)	22 (17–27)
ICU Admission Diagnosis , no. (%)	
Respiratory, including pneumonia	78 (57)
Gastrointestinal	17 (13)
Infectious disease	9 (7)
Trauma	6 (4)
Cardiovascular	4 (3)
Other	22 (16)
ICU Type , no. (%)	
Medical	107 (79)
Surgical	19 (14)
Trauma	10 (7)

Abbreviations: ALI – acute lung injury; APACHE II – Acute Physiology and Chronic Health Evaluation II; IQR – inter-quartile range

^I Includes the documented diagnosis of any obstructive or restrictive lung disease, asthma, dyspnea with moderate or less than moderate activity associated with lung pathology, known use of home oxygen, chronic hypoxia or hypercapnea, severe pulmonary hypertension, and history of lung transplant.

Table 2

Patient versus population norms for baseline quality of life

Domain	N ^a	Mean Population norm ^b	Mean Patient estimate	Mean paired difference [95% CI] (population - patient)	P-value for difference >0 ^c	MCID (30)	P-value for difference MCID ^d
Physical Function	136	82.1	69.6	12.5 [7.0, 18.0]	<0.001	10.0	0.190
Physical Role	135	81.9	64.0	17.9 [11.9, 23.9]	<0.001	12.5	0.038
Bodily Pain	133	70.1	62.8	7.3 [1.4, 13.2]	0.008	10.0	0.820
General Health	133	70.3	56.4	13.9 [9.2, 18.7]	<0.001	10.0	0.049
Vitality	134	58.9	56.4	2.6 [-1.6, 6.7]	0.120	12.5	>0.99
Social Function	134	84.2	68.5	15.7 [10.2, 21.3]	<0.001	12.5	0.130
Emotional Role	132	87.6	75.6	12.0 [6.3, 17.8]	<0.001	8.3	0.099
Mental Health	133	75.5	68.8	6.7 [2.6, 10.9]	0.001	10.0	0.940

Abbreviations: CI – confidence interval; MCID – minimal clinically important difference

^aWhen a domain could not be scored for either a patient or proxy, that specific domain was omitted from analysis

^bAge- and sex-matched U.S. population norm (41)

^cP-value for testing if the mean paired difference is greater than zero

^dP-value for testing if the mean paired difference is greater than or equal to the MCID (30)

Table 3

Patient versus proxy comparison for baseline quality of life

SF-36 domain	N ^a	Mean Patient estimate	Mean Proxy estimate	Mean paired difference [95% CI] (patient - proxy)	P-value for difference >0 ^b	MCID (30)	P-value for Difference MCID ^c
Physical Function	136	69.6	63.4	6.2 [0.42, 12.0]	0.036	10.0	0.903
Physical Role	135	64.0	56.7	7.3 [1.6, 13.0]	0.012	12.5	0.964
Bodily Pain	133	62.8	56.1	6.7 [0.9, 12.4]	0.023	10.0	0.875
General Health	133	56.4	45.1	11.2 [6.7, 15.8]	<0.001	10.0	0.294
Vitality	134	56.4	49.8	6.6 [2.0, 11.2]	0.005	12.5	0.994
Social Function	134	68.5	61.2	7.3 [1.2, 13.4]	0.020	12.5	0.954
Emotional Role	132	75.6	71.0	4.6 [-1.6, 10.8]	0.140	8.3	0.880
Mental Health	133	68.8	63.3	5.5 [1.5, 9.4]	0.007	10.0	0.987

Abbreviations: CI – confidence interval; MCID – minimal clinically important difference

^aWhen a domain could not be scored for either a patient or proxy, that specific domain was omitted from analysis

^bP-value for testing if the mean paired difference is greater than zero

^cP-value for testing if the mean paired difference is greater than or equal to the MCID (30)

Table 4

Agreement between patient and proxy baseline quality of life scores

SF-36 domain	State Change ^a	Kappa ^b [95% CI]	Weighted Kappa ^{b,d} [95% CI]	MCID (30)	MCID Kappa ^c [95% CI]	MCID Weighted Kappa ^{c, d} [95% CI]
Physical Function	5.0	0.16 [0.11, 0.21]	0.36 [0.24, 0.47]	10.0	0.21 [0.14, 0.29]	0.36 [0.25, 0.48]
Role Physical	6.25	0.13 [0.08, 0.18]	0.38 [0.27, 0.50]	12.5	0.17 [0.09, 0.24]	0.39 [0.27, 0.51]
Bodily Pain	10.0	0.08 [0.02, 0.14]	0.34 [0.22, 0.46]	10.0	0.08 [0.02, 0.14]	0.34 [0.22, 0.46]
General Health	5.0	0.02 [-0.01, 0.05]	0.35 [0.25, 0.45]	10.0	0.10 [0.03, 0.17]	0.37 [0.26, 0.47]
Vitality	6.25	0.07 [0.03, 0.11]	0.32 [0.23, 0.42]	12.5	0.17 [0.09, 0.25]	0.35 [0.24, 0.45]
Social Function	12.5	0.18 [0.12, 0.25]	0.34 [0.23, 0.46]	12.5	0.18 [0.12, 0.25]	0.34 [0.23, 0.46]
Role Emotional	8.3	0.16 [0.09, 0.24]	0.40 [0.28, 0.52]	8.3	0.16 [0.09, 0.23]	0.40 [0.28, 0.52]
Mental Health	5.0	0.06 [0.02, 0.10]	0.41 [0.31, 0.51]	10.0	0.17 [0.09, 0.24]	0.43 [0.33, 0.54]

Abbreviations: CI – confidence interval; COPD – chronic obstructive pulmonary disease; MCID – minimal clinically important difference; QOL – quality of life; SF-36 – Short-Form 36

^a A state change is defined as the smallest possible change in score based on a 1 unit change in response for 1 question within a domain

^b Perfect agreement is defined as having an absolute difference between the patient and proxy score of less than or equal to the state change

^c Perfect agreement is defined as having an absolute difference between the patient and proxy score of less than the MCID

^d For the weighted kappa, weights were assigned using linear weighting which assigns a weight of 1 for perfect agreement and 0 for the largest possible disagreement, with other weights of misclassifications determined linearly.(34)