Published in final edited form as:

Addiction. 2011 March ; 106(3): 507–515. doi:10.1111/j.1360-0443.2010.03299.x.

Longitudinal Association of Preference-Weighted Health-Related Quality of Life Measures and Substance Use Disorder Outcomes

Jeffrey M. Pyne, MD^{1,2}, Shanti Tripathi, MS², Michael French, PhD³, Kathryn McCollister, PhD⁴, Richard C. Rapp, MSW⁵, and Brenda M. Booth, PhD²

¹Center for Mental Healthcare and Outcomes Research, Central Arkansas Veterans Healthcare System, Little Rock, AR

²Department of Psychiatry, University of Arkansas for Medical Sciences, Little Rock, AR

³Departments of Sociology, Economics, and Epidemiology and Public Health, University of Miami, Coral Gables, FL

⁴Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL

⁵Department of Community Health, Boonshoft School of Medicine, Wright State University, Dayton, OH

Abstract

Aim—Examine the construct validity of generic preference-weighted health-related quality of life measures in a sample of patients with a substance use disorder (SUD).

Design—Longitudinal (baseline and six-month follow-up) data from a research study that evaluated interventions to improve linkage and engagement with SUD treatment.

Setting—A central intake unit that referred patients to seven SUD treatment centers in a Midwestern US metropolitan area.

Participants—495 persons with a SUD.

Measurements—Participants completed two preference-weighted measures: self-administered Quality of Well-Being scale (QWB-SA) and standard gamble weighted Medical Outcomes Study SF-12 (SF-6D). They were also administered two clinical assessments: all seven domains of the Addiction Severity Index (ASI) and a symptom checklist based on the DSM-IV. Construct validity was determined via the relationships between disease-specific SUD and generic measures.

Findings—In unadjusted analyses, the QWB-SA and SF-6D change scores were significantly correlated with six ASI subscale change scores, but not with employment status. In adjusted repeated measures analyses, 3/7 ASI subscale scores were significant predictors of QWB-SA and 5/7 ASI subscale scores were significant predictors of SF-6D. Abstinence and problematic use at follow-up were significant predictors of QWB-SA and SF-6D. Effect sizes ranged from 0.352 to 0.400 for abstinence and −0.484 to −0.585 for problematic use.

Conclusions—Generic preference-weighted health-related quality of life measures show moderate to good associations with substance-use specific measures and in certain circumstances

---

Corresponding Author: Jeffrey M. Pyne, MD, Central Arkansas Veterans Healthcare System, Department of Psychiatry – 152/NLR, 2200 Fort Roots Drive, North Little Rock, AR, 72114; phone: (501) 257-1083; fax: (501) 257-1844; jmpyne@uams.edu.

Conflict of interest: None

The authors have no conflicts of interest to report.

can be used in their stead. This study provides further support for the use of the QWB-SA and SF-6D in clinical and economic evaluations of SUD interventions.

Keywords

health-related quality of life; substance use disorder; cost-utility analysis; cost-effectiveness analysis; Quality of Well Being scale; SF-12; SF-6D

INTRODUCTION

Preference-weighted health-related quality-of-life (HRQL) measures are used to calculate quality-adjusted life years (QALYs) [1,2]. Generic QALY measures apply preference weights to symptom and functioning questions to derive a single score between death (score=0) and perfect health (score=1). Generic QALYs provide a level playing field from which to compare the effectiveness of healthcare interventions and are the recommended measure in the denominator of cost-effectiveness ratios [1,2]. However, cost per QALY analyses are rarely used to evaluate substance use disorder (SUD) interventions. Relative to disease burden analyses, cost per QALY analyses of SUD interventions are under-represented [3]. One of the key issues to address prior to more widespread use of cost per QALY analyses for SUD interventions is the construct validity of QALY measures in the context of SUD outcomes [4–6].

To our knowledge, six published studies directly examined the construct validity of a generic QALY measure using a SUD sample and widely-accepted SUD outcome measures. One longitudinal study evaluated the 5-item Euro-Qol (EQ-5D) with 52 subjects diagnosed with alcohol dependence and found small to medium effect sizes relative to changes in alcohol consumption or problem status and non-significant responsiveness using linear regression models and changes in alcohol consumption or problem status as predictors [7]. Another longitudinal study used the EQ-5D with 272 subjects diagnosed with opiate dependence and found significant differences between abstinent and non-abstinent subjects at the 16-month follow-up [8]. A third longitudinal study used the SF-36 version of the SF-6D with 165 subjects diagnosed with substance dependence and reported significant improvements for subjects meeting early remission criteria and decreased levels of alcohol consumption over a six-month follow-up [9]. Cross-sectional analyses of the baseline data presented in this paper used the self-administered Quality of Well-Being scale (QWB-SA) and SF-12 version of the SF-6D with 574 subjects referred for SUD treatment and found significant correlations with SUD symptom severity and diagnostic criteria [10]. Other cross-sectional analyses found a statistically significant decrease in the 15D (15 dimensions) (EQ-5D was non-significant) for heavy drinkers versus abstainers but the 15D result was not clinically meaningful [11]. Another study found only very high risk drinking to significantly reduce the EQ-5D compared to compared to abstainers or lower risk drinking [12].

We are aware of only four published economic studies reporting prospective generic QALY data to evaluate SUD interventions. One study used the EQ-5D with alcohol abuse or dependence patients and found no significant QALY difference between two active treatment conditions [13]. Another EQ-5D study with heroin dependent patients found significantly greater QALY gains for those patients co-prescribed heroin and methadone versus methadone alone [14]. A study employing the Assessment of Quality of Life (AQOL) in persons with heroin dependence found no significant differences in QALYs between methadone and buprenorphine maintenance [15]. Another study found no significant differences in SF-6D scores for alcohol-dependent patients receiving cognitive behavioral therapy (CBT) alone versus CBT plus naltrexone [16]. Based on our review, the remaining

SUD studies reporting cost per QALY analyses used simulation modeling methods, and most evaluated the cost-utility of tobacco prevention and treatment programs [10].

To inform the issue of construct validity for generic QALY measures and SUD outcomes, we used data from a recent study comparing motivational interviewing, strength-based case management, and usual care to improve linkage and engagement with SUD treatment. The generic QALY measures were the QWB-SA and the SF-6D (derived from the Medical Outcomes Study SF-12). We operationalized construct validity as the relationship between disease-specific symptom measures and generic preference-weighted health-related quality of life measures based on the Wilson and Cleary patient outcomes model [17]. This paper extends the previously published cross-sectional findings [10] to include longitudinal validation data. Specifically, we examine: 1) the correlation between the longitudinal generic QALY measures and SUD diagnostic and symptom severity and 2) the effect sizes of generic QALY measures relative to standard SUD outcomes.

METHODS

Design

Data for this study were collected as part of a National Institute on Drug Abuse funded clinical trial “Reducing Barriers to Drug Abuse Treatment Services” (R01 DA15690). Detailed descriptions of the parent study have been reported previously [10,18]. Briefly, the clinical trial tested the effectiveness of two interventions (motivational interviewing and strength-based case management) compared to usual care vis-à-vis improving linkage with and engagement in substance abuse treatment. These interventions occurred in a centralized intake unit (CIU) facility in a medium-sized Midwestern metropolitan area of the United States. The CIU was the county’s only point of entry for all uninsured individuals seeking treatment for SUD and mental health problems.

Subjects were eligible for the clinical trial if they met the following criteria: 1) received a recent assessment and referral at the CIU; 2) were at least 18 years of age; 3) received a clinical diagnosis of substance abuse or dependence using criteria from the DSM-IV [19] (subjects with a clinical diagnosis of alcohol abuse or dependence only were ineligible); 4) had no clinical diagnosis of schizophrenia or any other psychotic disorder; and 5) were referred to either residential, drug free outpatient or methadone maintenance treatment. The Wright State University Institutional Review Board approved the research protocol.

Subjects

A total of 495 subjects were included in the longitudinal analysis. The mean age was 32.9 (SD=9.7) years, 61.2% (303/495) were male, 51.1% (253/495) were Caucasian, 20.8% (103/495) were married or cohabitating, and 54.3% (269/495) did not graduate from high school. A total of 574 subjects completed baseline measures; 79 did not complete six-month follow-up interviews. No significant differences emerged between completers and non-completers for baseline QWB-SA, SF-6D, ethnicity, or education. Six-month completers were significantly older (32.9 versus 30.1, $p=0.02$), more likely to be female (33.5% versus 22.8%, $p=0.006$), and more likely to be married or cohabitating (20.8% versus 11.4%, $p=0.05$) than six-month non-completers.

Measures

At baseline and follow-up, all subjects completed an interviewer-administered survey including a substance abuse severity questionnaire (Addiction Severity Index) and two health-related quality of life questionnaires: QWB-SA and standard gamble weighted Medical Outcomes Study SF-12 (SF-6D). Also, extensive SUD data such as specific

symptoms of substance abuse and substance dependence as defined by the DSM-IV were collected [19]. Although the QWB-SA and SF-12 are usually self-administered, interviewers read the participants all instruments to account for varying reading levels and minimize missing data.

Addiction Severity Index (ASI)—The ASI is designed to evaluate the nature and severity of problems associated with SUDs [20]. The ASI focuses on seven functional domains: medical, employment, alcohol use, drug use, legal, family/social, and psychiatric. We report mean composite scores, which range from 0 (no problem) to 1.0 (extreme problem), in each domain for the past 30 days. Thus, a higher ASI score indicates greater severity within a given domain.

Medical Outcomes Study SF-6D—The SF-6D is a preference-weighted version of the SF-12. Brazier and colleagues developed methods for converting SF-36 and SF-12 data into quality-adjusted health index scores using preference-weighted methods [21,22]. Preference weight conversion formulas are based on visual analogue scale (VAS) [22] and standard gamble (SG) [21] methods. We used the SG conversion formula because it is more consistent with expected utility theory than the VAS [23]. The SG method for obtaining preference weights includes a choice between two options. One option is to remain in a given health state and the other option consists of accepting a treatment that will result in varying probabilities of death or perfect health. The probabilities for death and perfect health are systematically altered until the respondent is indifferent between the two options. The preference weight for the health state is equal to the probability of perfect health at the point of respondent indifference.

The six dimensions of the SF-6D are physical functioning, role limitations due to physical health or emotional problems, social functioning, pain, mental health, and vitality. The SG preference weights were derived from a general population sample of 611 subjects. The SG preference weight conversion formula transforms SF-12 data into an overall preference-weighted index score that varies from 0 (death) to 1.0 (perfect health), with higher score indicating better health. The range of the SF-6D for living participants is 0.345 to 1.0, where 0.345 is the lowest score possible for a responsive subject [24].

QWB-SA—The QWB-SA (self-administered version of the QWB) was specifically designed for cost-utility analyses, and it results in a preference-weighted index score between 0 (death) and 1.0 (perfect health) [21,22]. A variety of reliability and validity studies have been conducted using the QWB-SA [25–31]. The range for the QWB-SA in persons who are alive and responsive is 0.093 to 1.0 [28].

The QWB-SA is composed of five parts and four subscales and covers the past three days. Part I (symptom/problem subscale) asks about acute and chronic symptoms. Part II asks about self-care (e.g., hospitalization and the need for assistance with self-care). Part III (mobility subscale) asks about mobility (e.g., use of public transportation or driving). Part IV (physical activity subscale) asks about physical functioning (e.g., walking, confinement to a bed or chair). Part V asks about performance of usual activity (e.g., work, school, or housework). Parts II and V are combined to form the social activity subscale.

Abstinence and Problematic Use—Abstinence was defined as no use of 15 categories of alcohol or street drugs in the past 30 days. Non-abstinence was defined as using at least one of these substances in the past 30 days. By these definitions, 44/495 subjects (9%) were abstinent at baseline and 224/495 subjects (45%) were abstinent at the six-month followup. Problematic use was defined as meeting lifetime dependence criteria and reporting either alcohol and/or drug-related problems in the past 30 days. Based on this definition, 340/453

(75%) experienced problematic use at baseline and 172/453 (38%) experienced problematic use at the six-month followup.

Statistical analyses

Longitudinal interview data are presented for 495 subjects who completed both a baseline and six-month follow-up assessment. Analyses were conducted using SAS version 9.0. Since not all of the ASI change scores were normally distributed, we computed nonparametric Spearman correlation coefficients to describe the unadjusted relationships between changes in ASI subscale scores and changes in QWB-SA and SF-6D scores, as well as t-tests or Wilcoxon rank-sum tests to compare abstinence and problem use at six-months to changes in ASI scores (Table 2). We used a repeated measures design with general linear mixed models (PROC MIXED in SAS) because this method controls for within-subject correlation at baseline and follow-up. These models controlled for the following demographic variables: age, gender, race, marital status, and education. Standard fixed-effects models using first differencing of the dependent and independent variables produced results that were similar to the general linear mixed model results (results available from the authors upon request).

Table 3 presents the relationships between ASI subscale scores and QWB-SA and SF-6D scores over time using the full sample (N=495). Table 4 reports the estimated effects of being abstinent at the six-month follow-up for subjects who were not abstinent at baseline (N=451). Finally, Table 5 shows the relationships between problematic substance use, QWB-SA scores, and SF-6D scores at the six-month follow-up for subjects meeting lifetime dependence at baseline (N=453).

Effect sizes for abstinence and problematic use were calculated using Cohen's *d* [32]. The abstinence effect size calculation included the subgroup of subjects who were not abstinent at baseline (N=451), and the problematic use calculation included those subjects who met lifetime dependence criteria and problematic substance use in the past 30 days at baseline (N=340). The effect size calculations used unadjusted QWB-SA and SF-6D change scores in the numerator and the square root of mean square error for the respective change score in the denominator.

RESULTS

Table 1 describes baseline, six-month follow-up, and change scores for the ASI subscales, QWB-SA, and SF-6D scores. As shown in Table 2, both the QWB-SA and SF-6D unadjusted change scores were significantly correlated with six out of seven ASI change scores. The change scores for neither the QWB-SA nor SF-6D were significantly correlated with changes in the ASI employment score. Abstinence at the six-month follow-up was significantly associated with five of seven ASI change scores, and problem use at six-months was significantly associated with six of seven ASI change scores. Again, neither abstinence nor problem use were significantly correlated with changes in ASI employment score. As expected, the QWB-SA and SF-6D change scores were significantly correlated with each other ($r=0.351$, $p<0.0001$).

Table 3 presents the multivariate relationships between ASI subscale scores, QWB-SA, and SF-6D scores. In both specifications, all of the ASI domain coefficients were negative; indicating that increased ASI scores (worsening SUD symptoms) was associated with decreased preference-weighted health-related quality of life. Three of seven ASI subscale scores (medical, drug use, and psychiatric) were significant predictors in the QWB-SA equation. Five of seven ASI subscale scores (medical, employment, drug use, family/social, and psychiatric) were significant predictors in the SF-6D equation.

The results in Table 4 show that, among subjects who were not abstinent at baseline (N=451), being abstinent at the six-month follow-up was associated with an increase of 0.057 on the QWB-SA and 0.046 on the SF-6D. Among subjects who met criteria for lifetime dependence and problematic use at baseline (N=340), continued problematic use six months later was associated with a decrease of 0.060 on the QWB-SA and 0.064 on the SF-6D (see Table 5).

The effect size calculations used Cohen's *d* method as described above. For the abstinence effect size calculation, we included only subjects who were not abstinent at baseline. The mean QWB-SA change for subjects who were non-abstinent at the six-month follow-up was 0.015, and 0.073 for subjects who were abstinent at the follow-up. The square root of the mean square error for QWB-SA change scores was 0.165. Therefore, the QWB-SA effect size for abstinence at the six-month follow-up was 0.352 $([0.073-0.015]/0.165)$. Similarly, the SF-6D effect size for abstinence at the six-month follow-up was 0.400 $([0.093-0.041]/0.130)$. The effect size for problematic substance use was -0.484 $([0.003-0.081]/0.161)$ using the QWB-SA and -0.585 $([0.034-0.106]/0.161)$ using the SF-6D.

Comparison with cross-sectional baseline data

The unadjusted change score correlations presented here were similar to the cross-sectional results [10] except that the number of significant ASI subscales increased from four out of seven using cross-sectional SF-6D data to six out of seven using longitudinal SF-6D data. Spearman correlations were used for both the cross-sectional and longitudinal data analyses. The multivariate longitudinal results were similar to the multivariate cross-sectional results [10] except that ASI employment subscale became significant in the SF-6D specification and non-significant in the QWB-SA specification when analyzing longitudinal data. Thus, moving from the cross-sectional to longitudinal analyses resulted in the number of significant ASI subscales decreasing for the QWB-SA (4 to 3) and increasing for the SF-6D (4 to 5).

DISCUSSION

In unadjusted longitudinal analyses, the relationships between ASI subscale scores and the QWB-SA, SF-6D, abstinence, and problematic use were always in the expected direction and usually statistically significant. The absence of a significant relationship between abstinence at the six-month follow-up and change in the ASI medical subscale was most likely due to a sizable minority of patients who were abstinent because of serious physical health problems, and the severity of these physical health problems did not change over a six-month time frame [33].

In moving from unadjusted change scores (Table 2) to multivariate repeated measures analyses (Table 3), the number of statistically significant ASI subscales in relation to the QWB-SA and SF-6D decreased slightly. This finding was most likely caused by overlapping constructs associated with the ASI subscales. For example, change in the ASI alcohol use subscale was significantly correlated with change in the ASI drug use, family/social problems, and psychiatric problem subscales (range of $r=0.10$ to 0.31). Moreover change in the ASI legal problem subscale was significantly correlated with change in the ASI drug use, family/social problems, psychiatric problem, and employment status subscales (range of $r=0.11$ to 0.21).

The number of significant ASI subscales in the multivariate equations differed between the QWB-SA (3/7) and SF-6D (5/7). This suggests that the SF-6D may be capturing a wider range of SUD outcomes than the QWB-SA. Furthermore, the ASI legal problems subscale

was not a significant predictor in cross-sectional or longitudinal multivariate analyses, suggesting that the QWB-SA and SF-6D are not sensitive to current legal problems.

The magnitudes of the estimated longitudinal relationships between health-related quality of life scores and abstinence or problematic use were in a clinically important range for both the QWB-SA and the SF-6D. For example, the original interviewer-administered QWB validation studies found that health state descriptions with preference weights less than 0.03 units apart on a 0.0 (death) to 1.0 (perfect health) continuum could not be reliably rated as “different” [34]. More recently, Kaplan confirmed the clinically important difference for the interviewer-administered QWB to be at least 0.030 in patients with chronic obstructive pulmonary disease [35]. Similarly, Walters and Brazier determined the mean minimally clinically-important difference for the SF-6D to be 0.041 from eight longitudinal studies that included patients from eleven chronic physical illness groups [36]. Thus, the magnitudes of the effect sizes reported here were in the small to medium range [32].

Currently, there is no generally agreed upon method for measuring preference-weighted health-related quality of life [1,37]. The QWB-SA and SF-6D measure the same construct (preference-weighted health-related quality of life), and preference weights for both instruments were obtained from general community samples. Differences between the QWB-SA and SF-6D include the health state descriptive systems. Although both measures include social and occupational functioning, the QWB-SA has a wider range and number of specific physical and mental health symptoms/problems while the SF-12 only includes depression, anxiety, and pain symptoms. In addition, the QWB-SA preference weights are based on VAS scores, and the SF-6D preference weights are based on SG scores. As mentioned above, the standard gamble method is more consistent with expected utility theory than the VAS [23]. Typically, subjects will assign lower preference scores using VAS methods than SG methods [1]. This may explain why the possible range of SF-6D scores for living patients is smaller (0.345 to 1.0) than the possible QWB-SA range (0.093 to 1.0), and why the mean QWB-SA scores were lower than the mean SF-6D scores in our sample. Another explanation for the wider range in QWB-SA scores could be that the QWB-SA health states include a greater degree of severity than the SF-12. For these reasons, the QWB-SA may be less prone to floor or ceiling effects [38]. Although probably minor for most applications, another difference between the measures is that the QWB-SA takes longer to complete (approximately 10 minutes) than the SF-6D (<5 minutes). (<http://famprevmed.ucsd.edu/hoap/index.html>)

Preference for one instrument over the other may also be based on the setting, intervention(s), and subjects. Because it can be completed somewhat more quickly, the SF-6D may be preferred when decreasing respondent burden is important. However, if an intervention includes the risk of a specific side effect(s), then the QWB-SA may be preferred because of its wider range of specific physical and mental health symptoms/problems.

Among the limitations of this study, subjects were recruited from a single Midwestern state rather than from a broader geographical area. The vast majority of subjects were also contemplating treatment, and it is unknown whether the results would change if we analyzed data from subjects with substance use disorders who were not considering treatment (i.e., pre-contemplation). Other preference weighted health-related quality of life measures (e.g., Health Utilities Index, EQ-5D, and Assessment of Quality of Life) [39–41] are available, but were not tested in this study. The preference weights used to calculate the QWB-SA and SF-6D scores were derived from general population samples in the US and UK, respectively. Studies have found no significant differences in preference weights from US and UK subjects [42]. However, it is possible that the preference weights from patients with SUDs may systematically differ from those of the general population [43,44], as none of the

health-related quality of life measures used preference weights from a SUD sample. Nevertheless, the current recommended source of preference weights for cost-utility analyses is the general population [1]. One item on the QWB-SA directly asks about substance abuse, but the SF-12 does not contain a substance abuse question. Neither the QWB-SA nor SF-12 addresses the disutility of not using alcohol or drugs. Like most studies in this area, the reliability of the data depend on the subjects' accurate recall and reporting. However, several published studies on this topic conclude that self-reported data in these settings are reliable [45].

While establishing validity is an ongoing mission, this paper provides further evidence to support the validity of both the QWB-SA and SF-6D as measures of longitudinal preference-weighted health-related quality of life in patients with SUDs. Neither instrument has a clear advantage over the other except that a larger number of ASI subscales predict the SF-6D in adjusted analyses and the SF-12 may take a shorter time to complete than the QWB-SA. In summary, these findings offer scientific justification for the use of these measures in future economic evaluations of SUD interventions.

Acknowledgments

The authors acknowledge and thank Valorie Shue for contributing to the production of this paper. We owe special gratitude to Dr. Harvey A. Siegal (deceased) who was a vital member of the research team from its inception and gracious in his support of this research project.

Funding sources: Financial support for this study was provided in part by grants from NIDA R01 DA18980, NIDA R01 DA15690, NIDA R01 DA18645, NIDA K01 DA13962, NIAAA R01 AA15695, and by a VA Research Career Development Award. The funding agreement ensured the authors' independence in designing the study, interpreting the data, and preparing and publishing manuscripts.

References

1. Gold, MR.; Siegel, JE.; Russell, LB.; Weinstein, MC., editors. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press, Inc; 1996.
2. Drummond, M.; McGuire, A., editors. Economic Evaluation in Health Care. Oxford: Oxford University Press, Inc; 2006.
3. Neumann PJ, Rosen AB, Greenberg D, Olchanski NV, Pande R, Chapman RH, et al. Can we better prioritize resources for cost-utility research? *Med Decis Making*. 2005; 25(4):429–436. [PubMed: 16061895]
4. Mortimer D, Segal L. Economic evaluation of interventions for problem drinking and alcohol dependence: Cost per QALY estimates. *Alcohol Alcohol*. 2005; 40(6):549–555. [PubMed: 16103034]
5. Barnett PG, Zaric GS, Brandeau ML. The cost-effectiveness of buprenorphine maintenance therapy for opiate addiction in the United States. *Addiction*. 2001; 96(9):1267–1278. [PubMed: 11672491]
6. Barnett PG. The cost-effectiveness of substance abuse treatment. *Curr Psychiatry Rep*. 1999; 1(2): 166–171. [PubMed: 11122920]
7. Gunther OH, Roick C, Angermeyer MC, König HH. Responsiveness of EQ-5D utility indices in alcohol-dependent patients. *Drug Alcohol Depend*. 2008; 92(1–3):291–295. [PubMed: 17888587]
8. De Jong CA, Roozen HG, van Rossum LG, Krabbe PF, Kerkhof AJ. High abstinence rates in heroin addicts by a new comprehensive treatment approach. *Am J Addict*. 2007; 16(2):124–130. [PubMed: 17453614]
9. Pyne JM, Booth BM, Farahati F, Tripathi S, Smith GR, Marques PR. Preference-weighted health status associated with substance use-disorders treatment. *J Stud Alcohol*. 2006; 67(3):436–444. [PubMed: 16608154]
10. Pyne JM, French M, McCollister K, Tripathi S, Rapp R, Booth B. Preference-weighted health-related quality of life measures and substance use disorder severity. *Addiction*. 2008; 103:1320–1329. [PubMed: 18397359]

11. Saarni SI, Joutsenniemi K, Koskinen S, Suvisaari J, Pirkola S, Sintonen H, et al. Alcohol consumption, abstaining, health utility, and quality of life--a general population survey in Finland. *Alcohol Alcohol*. 2008; 43(3):376–386. [PubMed: 18245136]
12. Petrie D, Doran C, Shakeshaft A, Sanson-Fisher R. The relationship between alcohol consumption and self-reported health status using the EQ5D: evidence from rural Australia. *Soc Sci Med*. 2008; 67(11):1717–1726. [PubMed: 18945533]
13. UKATT Research Team. Cost effectiveness of treatment for alcohol problems: Findings of the randomized UK Alcohol Treatment Trial (UKATT). *Br Med J*. 2005; 331:554–559.
14. Dijkgraaf MG, van der Zanden BP, de Borgie CA, Blanken P, van Ree JM, van den Brink W. Cost utility analysis of co-prescribed heroin compared with methadone maintenance treatment in heroin addicts in two randomised trials. *Br Med J*. 2005; 330(7503):1297. [PubMed: 15933353]
15. Harris AH, Gospodarevskaya E, Ritter AJ. A randomized trial of the cost effectiveness of buprenorphine as an alternative to methadone maintenance treatment for heroin dependence in a primary care setting. *Pharmacoeconomics*. 2005; 23(1):77–91. [PubMed: 15693730]
16. Walters D, Connor JP, Feeney GF, Young RM. The cost effectiveness of naltrexone added to cognitive-behavioral therapy in the treatment of alcohol dependence. *J Addict Dis*. 2009; 28(2): 137–144. [PubMed: 19340676]
17. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *J Am Med Assoc*. 1995; 273(1):59–65.
18. Rapp RC, Otto AL, Lane DT, Redko C, McGatha S, Carlson RG. Improving linkage with substance abuse treatment using brief case management and motivational interviewing. *Drug Alcohol Depend*. 2008; 94:172–182. [PubMed: 18242883]
19. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV TR). Fourth ed.. Washington, DC: American Psychiatric Association; 2000.
20. McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, et al. The Fifth Edition of the Addiction Severity Index. *J Subst Abuse Treat*. 1992; 9(3):199–213. [PubMed: 1334156]
21. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002; 21:271–292. [PubMed: 11939242]
22. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol*. 1998; 51(11):1115–1128. [PubMed: 9817129]
23. Drummond, MF.; O'Brien, B.; Stoddart, GL.; Torrance, GW. *Methods for the Economic Evaluation of Health Care Programmes*. 2nd ed.. New York: Oxford University Press, Inc; 1997.
24. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care*. 2004; 42(9):851–859. [PubMed: 15319610]
25. Kaplan RM, Sieber WJ, Ganiats TG. The Quality of Well-Being Scale: comparison of the interviewer-administered version with the self-administered questionnaire. *Psychol Health*. 1997; 12:783–791.
26. Andresen EM, Rothenberg BM, Kaplan RM. Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults. *Med Care*. 1998; 36(9): 1349–1360. [PubMed: 9749658]
27. Sieber WJ, David KM, Adams JE, Kaplan RM, Ganiats TG. Assessing the impact of migraine on health-related quality of life: An additional use of the quality of well-being scale-self-administered. *Headache*. 2000; 40(8):662–671. [PubMed: 10971663]
28. Pyne JM, Sieber WJ, David K, Kaplan RM, Rapaport MH, Williams DK. Use of the Quality of Well-Being self-administered version (QWB-SA) in assessing health-related quality of life in depression patients. *J Affect Disord*. 2003; 76:237–247. [PubMed: 12943954]
29. Pyne JM, Sullivan G, Kaplan R, Williams DK. Comparing the sensitivity of generic effectiveness measures with symptom improvement in persons with schizophrenia. *Med Care*. 2003; 41(2):208–217. [PubMed: 12555049]
30. Pyne JM, Patterson TL, Kaplan RM, Gillin JC, Koch WL, Grant I. Assessment of the quality of life of patients with major depression. *Psychiatr Serv*. 1997; 48(2):224–230. [PubMed: 9021855]
31. Patterson TL, Kaplan RM, Grant I, Semple SJ, Moscona S, Koch WL, et al. Quality of well-being in late-life psychosis. *Psychiatr Res*. 1996; 63(2–3):169–181.

32. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed.. Hillsdale, NJ: Lawrence Erlbaum Assoc; 1988.
33. Williams EC, Palfai T, Cheng DM, Samet JH, Bradley KA, Koepsell TD, et al. Physical health and drinking among medical inpatients with unhealthy alcohol use: A prospective study. *Alcohol Clin Exp Res*. 2010 May 4. eprint.
34. Kaplan RM, Bush JW, Berry CC. Health status: Types of validity and the index of well-being. *Health Serv Res*. 1976; 11(4):478–507. [PubMed: 1030700]
35. Kaplan RM. The minimally clinically important difference in generic utility-based measures. *J Chron Obstruct Pulmon Dis*. 2005; 2(1):91–97.
36. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. *Health Qual Life Outcomes*. 2003; 1:4. [PubMed: 12737635]
37. Green C, Brazier J, Deverill M. Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics*. 2000; 17(2):151–165. [PubMed: 10947338]
38. Naglie G, Tomlinson G, Tansey C, Irvine J, Ritvo P, Black SE, et al. Utility-based Quality of Life measures in Alzheimer's disease. *Qual Life Res*. 2006; 15(4):631–643. [PubMed: 16688496]
39. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions. *Health Utilities Index*. *Pharmacoeconomics*. 1995; 7(6):503–520. [PubMed: 10155336]
40. EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy*. 1990; 16:199–208. [PubMed: 10109801]
41. Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. *Qual Life Res*. 1999; 8(3):209–224. [PubMed: 10472152]
42. Patrick DL, Sittampalam Y, Somerville SM, Carter WB, Bergner M. A cross-cultural comparison of health state values. *Am J Public Health*. 1985; 75:1402–1407. [PubMed: 4061712]
43. Hadorn DC. The problem of discrimination in health care priority setting. *J Am Med Assoc*. 1992; 268:1454–1459.
44. Boyd N, Sutherland H, Heasman K, Tritchler DL, Cummings BJ. Whose utilities for decision analysis? *Med Decis Making*. 1990; 10(1):58–67. [PubMed: 2182963]
45. Smith G, Burnam MA, Mosley C, Hollenberg J, Mancino M, Grimes W. Reliability and Validity of the Substance Abuse Outcomes Module. *Psychiatr Serv*. 2006; 57(10):1452–1460. [PubMed: 17035565]

Table 1

Measures of central tendency and dispersion in ASI subscales, QWB-SA scores, and SF-6D scores at each time point (N=495)

Variable	Baseline	Six-month follow-up	Change
ASI Medical			
Mean (SD)	0.21 (0.34)	0.20 (0.32)	– 0.01 (0.40)
Median (Range)	0.00 (0 to 1)	0.00 (0 to 1)	0.00 (–1 to 1)
ASI Employment			
Mean (SD)	0.76 (0.26)	0.71 (0.30)	– 0.05 (0.24) ***
Median (Range)	0.78 (0.03 to 1)	0.75 (0.02 to 1)	0.00 (–0.91 to 0.93)
ASI Alcohol			
Mean (SD)	0.13 (0.22)	0.08 (0.20)	– 0.05 (0.21) ***
Median (Range)	0.03 (0 to 0.99)	0.00 (0 to 0.92)	0.00 (–0.94 to 0.87)
ASI Drugs			
Mean (SD)	0.20 (0.15)	0.10 (0.13)	– 0.12 (0.17) ***
Median (Range)	0.21 (0 to 0.60)	0.00 (0 to 0.56)	– 0.10 (–0.60 to 0.36)
ASI Legal			
Mean (SD)	0.21 (0.22)	0.15 (0.21)	– 0.06 (0.24) ***
Median (Range)	0.17 (0 to 0.98)	0.00 (0 to 0.97)	0.00 (–0.74 to 0.80)
ASI Family / Social			
Mean (SD)	0.20 (0.21)	0.13 (0.17)	– 0.07 (0.24) ***
Median (Range)	0.18 (0 to 0.93)	0.02 (0 to 0.76)	– 0.02 (–0.87 to 0.67)
ASI Psychiatric			
Mean (SD)	0.26 (0.25)	0.16 (0.23)	– 0.10 (0.25) ***
Median (Range)	0.25 (0 to 0.86)	0.00 (0 to 0.88)	0.00 (–0.77 to 0.82)
QWB-SA			
Mean (SD)	0.60 (0.15)	0.63 (0.17)	0.03 (0.17) ***
Median (Range)	0.61 (0.13 to 1)	0.61 (0.21 to 1)	0.01 (–0.52 to 0.65)
SF-6D			
Mean (SD)	0.71 (0.13)	0.76 (0.13)	0.05 (0.13) ***
Median (Range)	0.73 (0.37 to 1)	0.80 (0.37 to 1)	0.05 (–0.55 to 0.46)

Note: Significance of change score was determined using the non-parametric Wilcoxon rank-sum test for non-normal distributions (ASI Medical, ASI Employment, ASI Alcohol, ASI Legal, ASI Psychiatric). The other change scores were normally distributed and compared using paired t-tests,

***
p-value < 0.001

Table 2

Unadjusted associations with ASI change scores

Variable	Δ QWB-SA ¹ (N=495)	Δ SF-6D ¹ (N=495)	Abstinent at six-month follow-up ² (N=451) ³	Problematic use at six-month follow-up ² (N=340) ⁴
Δ ASI Medical	- 0.177 ***	- 0.239 ***	z-value = -1.6	z-value = 2.41 *
Δ ASI Employment	- 0.033	- 0.064	z-value = -0.50	z-value = 1.30
Δ ASI Alcohol Use	- 0.136 **	- 0.153 ***	z-value = - 6.56 ***	z-value = 6.43 ***
Δ ASI Drug Use	- 0.195 ***	- 0.307 ***	t-value = 12.8 ***	t-value = 17.62 ***
Δ ASI Legal	-0.112 *	- 0.155 ***	z-value = - 4.70 ***	z-value = 5.15 ***
Δ ASI Family / Social	- 0.109 *	- 0.272 ***	t-value = 4.00 ***	t-value = 2.01 *
Δ ASI Psychiatric	- 0.273 ***	- 0.415 ***	z-value = - 2.10 *	z-value = 1.98 *

Note:

*
p < 0.05**
p < 0.01***
p < 0.001¹ Association determined using Spearman correlation² Association determined using paired t-test (t-value) or Wilcoxon rank sum test (z-value) depending on whether the distribution of ASI change score was normal (t-value) or non-normal (z-value)³ N=451 refers to those subjects who were not abstinent at baseline⁴ N=340 refers to those subjects meeting lifetime dependence and problematic use at baseline

Table 3

Repeated measures general linear mixed model results for ASI, QWB-SA and SF-6D scores (Baseline N=495)

Variable	QWB-SA	SF-6D
Gender (Male)	0.012	0.015 *
Race (White)	– 0.0001	– 0.002
Marital (Single)	– 0.016	0.0003
Education (HS Grad or More)	0.002	– 0.006
Age	– 0.001	– 0.001 **
ASI Medical	– 0.111 ***	– 0.110 ***
ASI Employment	– 0.032	– 0.035 **
ASI Alcohol Use	– 0.006	– 0.010
ASI Drug Use	– 0.10 **	– 0.120 ***
ASI Legal	– 0.016	– 0.020
ASI Family / Social	– 0.020	– 0.057 ***
ASI Psychiatric	– 0.251 ***	– 0.234 ***
Time (six-month)	0.005	– 0.015 *
Intercept	0.766 ***	0.898 ***

Note:

*
p < 0.05**
p < 0.01***
p < 0.001

Table 4

Repeated measures general linear mixed model results for QWB-SA and SF-6D scores when subject is abstinent at the six-month follow-up (Baseline N=451)^I

Variable	QWB-SA	SF-6D
Gender (Male)	0.030 *	0.032 **
Race (White)	– 0.016	– 0.020
Marital (Single)	– 0.021	– 0.004
Education (HS Grad or More)	0.011	0.004
Age	– 0.002 ***	– 0.003 ***
6-month abstinent (yes)	0.057 ***	0.046 ***
Time (six month)	0.015	0.043 ***
Intercept	0.690 ***	0.822 ***

Note:

*
p < 0.05

**
p < 0.01

***
p < 0.001

^I
N=451 refers to those subjects who were not abstinent at baseline

Table 5

Repeated measures general linear mixed model results for QWB-SA and SF-6D scores when subject has problematic use at six-month follow-up (Baseline N=453)^I

Variable	QWB	SF-6D
Gender (Male)	0.030 *	0.032 **
Race (White)	– 0.0001	– 0.0005
Marital (Single)	– 0.023	– 0.005
Education (HS Grad or More)	0.011	0.006
Age	– 0.001 *	– 0.002 **
6-month problematic use (yes)	–0.060 ***	–0.064 ***
Time (six month)	0.020 *	0.041 ***
Intercept	0.632 ***	0.760 ***

Note:

*
p < 0.05

**
p < 0.01

***
p < 0.001

^I N=453 refers to those subjects meeting lifetime dependence criteria at baseline