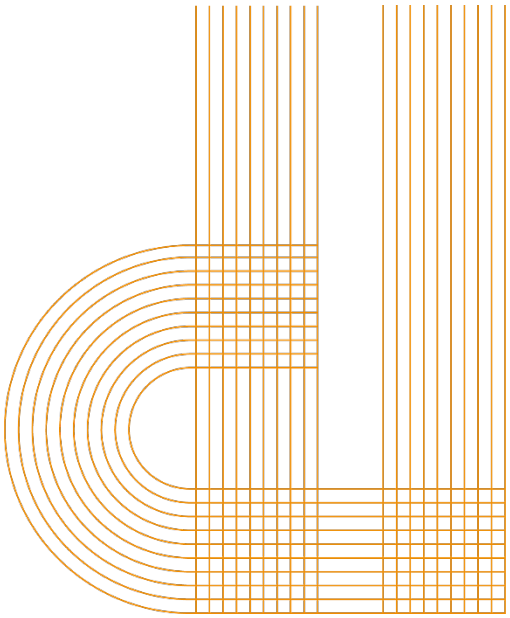
Universidade Vigo

Departamento de Economía Aplicada
https://webx06.webs.uvigo.es

Working Paper: 24/04. December 2024

Comparison of measures of health-related quality of life in patients with alcohol misuse: Mapping of two alcohol-specific instruments to three preference-based instruments



Eva Rodríguez-Míguez

Jacinto Mosquera

Coordinator: Olga Alonso-Villar
ovillar@uvigo.es

Comparison of measures of health-related quality of life in patients with alcohol misuse: mapping of two alcohol-specific instruments to three preference-based instruments*

Eva Rodríguez-Míguez[1] and Jacinto Mosquera[2]

Abstract

Problems associated with alcohol use disorder (AUD) can be quantified using generic and specific measures. The aim of this study is to compare different preference-based instruments for measuring health-related quality of life (HRQoL) in patients with AUD and to examine their relationship with alcohol-specific measures used in clinical practice. Patients with AUD were recruited from a Spanish alcoholism unit. We administered the EuroQoL-5Dimension-5levels (EQ-5D), ShortForm-6Dimension (SF-6D), AlcoholQuality-of-life-4Dimension (AlcQ-4D), Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and Alcohol Use Disorders Identification Test (AUDIT) instruments at baseline and 12 months later. Differences in HRQoL instrument scores were estimated and their ability to discriminate between known clinical severity groups was analysed. Several mapping functions were tested to transform clinical scores (AUDIT or DSM-5) into HRQoL scores (EQ-5D, SF-6D or AlcQ-4D). The results show that HRQoL scores are sensitive to the instrument used. Mean utility scores are always highest with EQ-5D, followed by SF-6D and AlcQ-4D. All HRQoL instruments discriminate between clinical severity groups defined by DSM-5 or AUDIT. Although several mapping functions were estimated for each pair of clinical vs preference-based instruments, the model using the total score of the clinical measures as the independent variable was selected for all of them. The results suggest that clinical measures used in the field of AUD could be adapted for use in economic evaluation. However, the incremental cost-utility ratio of AUD programmes, and hence the policy decisions derived from it, may depend on the HRQoL instrument used.

JEL Classification: D61, I18, I12, C12

Key words: alcohol use disorder, HRQoL, EQ-5D, SF-6D, AlcQ-4D, AUDIT, DSM-5, mapping

---

* This study was funded by Spanish Ministry of Science, Innovation and Universities [RTI2018-099403-B-I00]. The authors would like to thank the staff of ASVIDAL (Alcoholism Treatment Unit of Vigo) for their disinterested participation in the management of the interviews.

[1] Correspondence address: University of Vigo. Vigo (Spain); Instituto de Investigación Sanitaria Galicia Sur (IISGS); ECOBAS. ORCID: 0000-0002-7287-7894 emiguez@uvigo.es

[2] Galician Health Service (SERGAS), Vigo (Spain); Instituto de Investigación Sanitaria Galicia Sur (IISGS). ORCID: 0000-0002-7802-1739

1. Introduction

Alcohol consumption is a serious public health problem in large parts of the world. Alcohol is consumed by more than half of the population in three World Health Organization (WHO) regions: the Americas, Europe, and the Western Pacific. Alcohol use has been linked to more than 200 health conditions ranging from liver diseases, road injuries, and violence to cancers, cardiovascular diseases, suicides, tuberculosis, and HIV/AIDS. Of course, these conditions have a major impact on both mortality and morbidity. It is estimated that alcohol abuse causes 5.3% of deaths worldwide, mostly in the young population; for those between 20 and 39 years old, 13.5% of all deaths are attributable to alcohol. With regard to morbidity, the risk factor of alcohol use contributes 5.1% to the global burden of disease and injury as calculated in terms of disability-adjusted life years (DALYs); it is the ninth leading risk factor globally and the leading risk factor among males aged 15–49 years [1]. This state of affairs poses a huge problem at both the individual and the societal level. Therefore, public health priorities should include the early detection and diagnosis of alcohol use disorder (AUD), comprehensive assessment of its consequences, and analysis of treatment effectiveness.

In the clinical setting, problems related to alcohol use are assessed by clinical interview and by means of questionnaires and widely used diagnostic criteria. The most popular questionnaire is the Alcohol Use Disorders Identification Test (AUDIT) proposed by the WHO [2] and validated in Spanish [3]. This questionnaire enables the *detection* of risky consumption and alcohol use disorder by means of 10 questions. The *diagnosis* of AUD usually references criteria of the American Psychiatric Association, whose *Diagnostic and Statistical Manual of Mental Disorders DSM-5* [4] lists 11 criteria; a diagnosis of AUD is indicated by the presence of at least two criteria for a duration of at least 12 months. A recent revision of this manual (DSM-5 TR) [5] did not introduce any changes in this diagnostic category.

These instruments, of unquestionable usefulness for clinical management, have a crucial limitation in that they do not incorporate a measurable approximation of the *intangible* consequences of this pathology – in essence, health-related quality of life (HRQoL). Estimates of HRQoL have become extremely relevant to the management of resources and to the assessment of clinical outcomes. In a meeting of experts convened to define main indicators of the evolution in “brief” alcohol interventions, it was mentioned that "an alcohol brief interventions may not be considered useful if it did not influence quality of life", with HRQoL being the most relevant item of this dimension [6]. Another limitation of clinical measurements is that they do not allow comparison of the impact of alcohol consumption and those of other pathologies or risk factors.

Several instruments, both generic and specific, are available to measure HRQoL in relation to drinking behavior, alcohol use disorder, and treatment outcomes [7-9]. The most widely used generic instruments are the Short Form 36 (SF-36) – or its reduced versions, such as the Short Form 12 (SF-12) or the Short Form–6 Dimension (SF-6D) [10-21] – and either the 3- or 5-level version of the EuroQoL–5 Dimension

(EQ-5D) [22-26]. The cited studies support the suitability of these instruments for measuring not only how alcohol dependence affects HRQoL but also the effects of different treatment interventions. That said, other studies describe the limitations of these instruments in terms of their capacity to discriminate adequately between the health states of patients with alcohol dependence [27, 28] or to measure treatment effects [8, 29]. An additional advantage of the SF-6D and the EQ-5D is that they make it possible to quantify changes in HRQoL in terms of quality-adjusted life years (QALYs). A QALY *is* a unified and exportable indicator that allows for comparative studies with other pathologies and whose psychometric properties enable its use in economic evaluation studies.

Within the field of alcohol consumption, there are also specific instruments to measure HRQOL; these include the AQoLS [30], the AQoL9 [31], and one proposed by Rodríguez-Míguez and Mosquera [32] ––hereinafter referred to as ALCohol Quality-of-life–4 Dimension (AlcQ-4D). One advantage of these instruments is that they enable the analysis of quality-of-life dimensions that are strongly affected by AUD but that are not addressed by generic instruments. However, to the best of our knowledge, the only specific instrument that allows for estimating QALYs is the AlcQ-4D.

Despite the importance of quantifying how AUD and its treatment affect patients' quality of life, such measurements are seldom taken in routine clinical practice; instead, clinical instruments are most often used. This practice makes it impossible to use medical records to estimate changes in patients' quality of life or in subsequent economic evaluation studies. Hence it would be desirable to have *mapping functions* that could transform clinical measurements into HRQoL measurements, preferably in QALYs, that could be used for economic evaluation. Although the use of mapping functions is a second-best solution when compared to using a preference-based measure directly [33, 34], it is a solution that is gaining popularity because it allows, when preference-based measures are not available, researchers to estimate quality-of-life scores using data on other clinical measures. As far as we know, in the AUD field there is no study that describes how one would transform the scores from AUDIT or DSM-5 instruments into HRQoL scores measured in QALYs. The only exception is a study by Chavez et al. [35], who found no correlation between AUDIT-C (a reduced version of AUDIT) and the EQ-5D instrument.

The general objective of this research is to compare different instruments for measuring HRQoL in QALYs and to examine their relationship with instruments used in the clinical setting for the screening and diagnosis of alcohol use disorder. We use a sample of patients with this pathology and compare two clinical instruments (DSM-5 and AUDIT) and three instruments that allow us to quantify the impact on HRQoL in terms of QALYs: two generic instruments (EQ-5D and SF-6D) and one specific to AUD (AlcQ-4D).

More specifically, the study's objectives are as follows.

1. To analyze whether the HRQoL of patients with AUD differs depending on the instrument (EQ-5D, SF-6D, or AlcQ-4D) used to measure it; we assess differences among these instruments in measuring patients' quality of life and the gains from a treatment program.

2. To estimate the correlation between scores derived from HRQoL instruments and those obtained via AUDIT and DSM-5, the clinical questionnaires commonly used for the detection and diagnosis of this pathology.

3. To analyze the ability of the HRQoL instruments to discriminate among levels of clinical severity.

4. To estimate the relationship between HRQoL scores and clinical scores using statistical models to analyze the "exchange rates" between instruments.

2. Methods

2.1. Sample

The sample of 202 patients with AUD was obtained from a 12-month open, non-randomized, prospective study conducted during 2021–2022. The patients were recruited from a standard alcoholism treatment unit within the public health system of Galicia (a region of Spain). The patients met the following criteria: (a) were aged 18 years or more; (b) were attending their first consultation in the unit; (c) had no cognitive impairment that prevented study participation; (d) were diagnosed with AUD; and (e) had signed their informed consent. Sample size was estimated for a 95% confidence interval and 80% statistical power and to recognize as statistically significant a difference of at least 0.04 units. Note that the minimally important mean difference estimated (in the literature) for the SF-6D and EQ-5D is 0.041 and 0.074, respectively [36].

Patients were interviewed at the beginning of treatment and 12 months later. The treatment consists of psychosocial interventions, pharmacological interventions, or both, according to an individualized therapeutic plan, which may vary during the analysis period. Although 259 patients completed the questionnaire at the beginning of the study, 57 did not participate in the 12-month interview because they could not be located, died or refused to participate in the follow-up interview.

2.2. Questionnaire

The questionnaire (conducted by a social worker, psychologist, or doctor) was divided into three parts. The first part covered the patient's socioeconomic characteristics, consumption profile, level of motivation, use of other drugs, and chronic illnesses. In the second part, participants completed two

clinical questionnaires related to alcohol consumption: AUDIT (10 questions with 3–5 response levels) and DSM-5 (11 questions with 2 response levels). For the third part, respondents completed two generic HRQoL questionnaires – EQ-5D-5L (the 5-level version of the EQ-5D, hereafter simply the EQ-5D) and SF-6D (6 questions with 4–6 response levels) – and the AUD-specific AlcQ-4D (4 questions with 3 response levels). The EQ-5D and SF-6D were selected because they are the HRQoL instruments most often used to estimate QALYs; AlcQ-4D was chosen because, despite being a specific (non-generic) instrument, it allows for the estimation of QALYs and so its scores can be compared with those derived from the generic HRQoL instruments. The AlcQ-4D considers four consequences of alcohol-related disorders (family, social, physical health, and psychological health) with four severity levels.

2.3. Data analysis

The first requirement is a descriptive analysis of the sample, one that pays special attention to the comparison between the sample of patients who filled out both questionnaires and those who completed only the first. Our aim is to identify possible biases in the results. We are especially interested in analyzing whether the risk of sample attrition is concentrated in the most severe profiles (in terms of quality of life and clinical scores at baseline), which could be indicative of an upward bias in treatment efficacy.

The EQ-5D, SF-6D, and AlcQ-4D utility scores were obtained by applying the scoring algorithms estimated for the Spanish population [37, 38, 32]. These scores were calculated at baseline and also about 12 months later. The weights of the instruments can be compared because all of them quantify the quality of life in QALYs; a value of 1 corresponds to good health, 0 to a situation equivalent to death, and negative values to situations worse than death. We also calculated clinical instrument scores and constructed severity intervals. These instruments cannot be compared because they measure different constructs (screening vs. diagnosis) and use different scales. The AUDIT is measured on a 0–40 scale, where 0 corresponds to an abstainer who has never had any alcohol problems. This scale maps to WHO guidelines [39] that posit four levels of severity: abstinent or low risk (0–7), medium level of alcohol problem ('hazardous' drinking) (8–15), harmful drinking (16–19), and possible alcohol dependence (20–40). According to the American Psychiatry Association [4], the DSM-5 is measured on a 0–11 scale with three levels of severity: mild (2–3), moderate (4–5) and severe (6–11).

Paired samples of mean difference tests were used to analyze whether the HRQoL of patients with AUD differs depending on the measuring instrument (EQ-5D, SF-6D, or AlcQ-4D). Inter-instrument differences were analyzed in both the baseline and 12-month follow-up data sets. We also tested for whether the gain, or the difference between the baseline and 12-month follow-up scores, is significantly different between instruments. We remark that if differences between instruments remain constant across the utility distribution, then it should be possible to find differences between instruments at baseline and 12-month follow-up *without* any differences in the gain.

Comparing HRQoL scores with clinical scores required different analyses of the data pool (baseline and follow-up). First, we calculated the Pearson correlation coefficient between all the instruments. Second, we used the mean difference test to assess the discriminatory capacity of our HRQoL instruments to identify severity groups predicted by the clinical instruments.

Finally, to estimate the relationship between HRQoL and clinical instruments, we performed a regression analysis in which the dependent variable was HRQoL scores (EQ-5D, SF-6D, or AlcQ-4D) and the independent variable was different specifications of the clinical instruments. For each HRQoL instrument, we estimated five models that differed in terms of the independent variable: three models using the AUDIT and two models using the DSM-5. In the case of AUDIT, the following specifications were used as independent variables: global scores (it is assumed that all dimension levels contribute to the global score with equal weight), dimension scores (it is assumed that the levels of a dimension have the same weight but that these may differ across dimensions), and item levels (it is assumed that every dimension level could have a different weight). Global and dimension scores were treated as continuous variables, and item levels were modeled as discrete dummy variables. In the case of dimension scores and item responses, models were fitted by backward regression removing non-significant dimensions ($p > .1$) and grouping two consecutive levels once it has been determined both that their coefficients are not significantly different from each other and that the signs are inconsistent. The same strategy was followed for the DSM-5, although only two models were estimated because using dimension scores and item levels produces the same model.

Random-effects regression was used in all models because we have two observations (baseline and 12-month follow-up) for each patient and so those observations were not independent. We believe that it is appropriate to work with the entire data pool because the more severe states are prevalent at baseline whereas the milder states are prevalent in the follow-up data. So, if our transformation function is to cover the broad spectrum of possible situations, it will be especially useful to work with all of the data pool. The models' goodness of fit was measured using the Akaike information criterion (AIC), for which smaller values correspond to a better model. Following Burnham and Anderson [40], we consider models separated by more than 10 points (relative to the model with the lowest AIC value) to have almost no support. For each model, we also report the estimated R^2 values and the root mean square error (RMSE).

This study was approved by the Committee on Ethics of Clinical Research in Galicia (reference code 2017/177). Signed informed consent was obtained from all participants in the study before enrollment. Statistical analyses were performed using Stata software.

3. Results

3.1. Sample description

Table 1 presents the sample's main characteristics at baseline and distinguishes between participants who completed the follow-up interview (the "traced" sample) and those who dropped out of the study. With regard to possible attrition bias, the traced sample and the dropout sample exhibit no significant differences in socioeconomic characteristics. Results from the quality-of-life and clinical instruments suggest lower severity profiles in the dropout sample, although the AlcQ-4D and AUDIT scores do not show significant differences. This lower severity is compatible with the dominant type of drinking; whereas about 50% of the dropout sample reported beer as their dominant type of alcohol consumption and 11% reported wine, the corresponding percentages were 30% and 27% in the follow-up sample.

[INSERT **Table 1** about Here]

3.2. Comparing HRQoL instruments

Table 2 presents the scores for all instruments at baseline and at the 12-month follow-up. The estimation of quality of life in patients with AUD is sensitive to the HRQoL instrument used. Paired tests revealed that, in both analysis periods, there are significant differences (at the 1% level) between all instruments except between EQ-5D (12-month) and AlcQ-4D (12-month), whose difference is significant at the 5% level ($p = 0.019$). The mean score is always significantly lower in SF-6D than in EQ-5D; the AlcQ-4D instrument yields the lowest mean scores at baseline but occupies an intermediate position in the follow-up.

[INSERT **Table 2** about Here]

Table 2 also shows that treatment has a positive effect on patients' quality of life and in reducing the severity scores of clinical instruments; all gains are statistically significant ($p < 0.001$). Among the HRQoL instruments, the AlcQ-4D estimates a significantly greater treatment effect, followed by the SF-6D and the EQ-5D. Figure 1 plots, for each RHQoL instrument, the distribution of initial scores and of gains.

[INSERT **Figure 1** about Here]

3.3. Comparing clinical and HRQoL instruments: Discriminant ability of HRQoL instruments

Table 3 reports the Pearson correlation coefficients between the different instruments for the traced sample pool of data (baseline + 12-month follow-up); all values are significant at the 1% level. The highest correlations are between the EQ-5D and SF-6D (Pearson rho = −0.84; $p < .0001$) and between DSM-5 and AUDIT (Pearson rho = −0.89; $p < .0001$). As for the correlation between quality-of-life and clinical instruments, the AlcQ-4D has the highest correlation (Pearson rho = −0.74 with both AUDIT and DSM-5; $p < .0001$); neither SF-6D nor EQ-5D correlates more than −0.60 with any of the clinical instruments.

[INSERT **Table 3** about Here]

Table 4 gives the estimated mean HRQoL for each of the severity groups as derived from the clinical instruments, which allows to analyze the discriminant ability of HRQoL instruments. All three of them can discriminate among severity levels, but the discriminating capacity of AlcQ-4D and SF-6D is greater: (a) AlcQ-4D finds significant differences at the 1% level between all groups *except* between AUDIT-level2 and AUDIT-level3, which are different at the 5% level; (b) SF-6D finds significant differences at the 1% level between all groups *except* between AUDIT-level2 and AUDIT-level3, AUDIT-level3 and AUDIT-level4, and DSM-level2 and DSM-level3, which are all significantly different at the 5% level; (c) EQ-5D does not adequately discriminate between AUDIT-level2 and AUDIT-level3 ($p = 0.405$) or between DSM-level2 and DSM-level3 ($p = 0.209$). All other comparisons are significantly different at 1% *except* AUDIT-level3 and AUDIT-level4, which are significantly different at 5%. The EQ-5D invariably produces higher scores than does the SF-6D in all severity groups, and SF-6D yields higher scores than AlcQ-4D *except* for the mildest levels of AUDIT and DSM-5.

[INSERT **Table 4** about Here]

3.4. Mapping between clinical scores onto HRQoL scores

Table 5 summarizes results from the different regressions estimated to approximate the "exchange rates" between instruments – that is, so as to estimate patients' quality of life from the scores obtained with clinical instruments in studies that did not include any HRQoL measures. In the case of AUDIT, three models are shown for each HRQoL instrument according as how the dependent variable is defined: as the total score (AUDIT-model 1), as the score of (significant) dimensions (AUDIT-model 2), or as levels of (significant) dimensions (AUDIT-model 3). In the DSM-5 case, two models per instrument are presented; because there are only two levels per dimension, using scores by dimensions or by levels per dimension amounts to the same thing. Table 5 also reports the values for each model's R^2, AIC, and RMSE.

[INSERT **Table 5** about Here]

With respect to AUDIT, models with more predictors yield (as expected) higher R^2 values. However, the RMSE – or the average deviation between the predicted HRQoL made by the model and the actual values in the data set – differs very little between models (the largest difference between any two models is 0.005). Yet the AIC, which penalizes models that use more parameters, is lowest (indicating a better fit) for EQ-5D and SF-6D in Model 1 and for AlcQ-4D in Model 3. Even so, Model 1 should not be discounted for AlcQ-4D given that the difference in AIC between it and Model 3 is in the range 4–7 [40]. In light of these considerations and for the sake of parsimony, we suggest using Model 1 when predicting HRQoL from all three instruments along with the corresponding parameters in each case. So, in the case of EQ-5D and AUDIT, for example, the mapping function selected would be: EQ-5D (estimated) = 0.91 – 0.008 × AUDIT score. In any case, it is interesting to note that dimensions 1, 5, and 6 are the ones most closely related to quality of life; together, they have as much explanatory power as the total score.

The DSM-5 instrument delivers slightly higher RMSE than AUDIT in all models, which indicates a larger deviation of the estimated values from the real values. As in the case of AUDIT, Model 1 is recommended for all HRQoL instruments. In the case of the EQ-5D, the difference in RMSE between Models 1 and 2 is very small and the AIC is the smallest in Model 1. For the SF-6D and AlcQ-4D, the AIC is higher in Model 1 than in Model 2; however, the difference is less than 6 points and the RMSE is similar.

4. Discussion

This study shows that the estimation of HRQoL in patients with AUD is sensitive to the instrument used: EQ-5D always produces the highest mean utility score, and SF-6D produces a higher score than AlcQ-4D at baseline but yields the lowest mean scores in the follow-up sample. Although all HRQoL instruments can discriminate between severity levels previously established by clinical instruments, the discriminatory ability of SF-6D and AlcQ-4D is slightly higher than that of EQ-5D. In addition, both clinical and HRQoL instruments detect a positive and significant effect of treatment. This paper also proposes algorithms that enable using the total scores provided by clinical instruments (AUDIT and DSM) to "predict" preference-based HRQoL scores (EQ-5D, SF-6D, and AlcQ-4D).

One advantage of our study is its use of a sample of patients diagnosed with AUD. The sample is important because general population studies of the relationship between alcohol abuse and HRQoL report conflicting results: some studies find a negative effect on HRQoL [41-43], but other studies report little or no effect [44-49] or even a positive effect [35, 50, 51]. These disparities may be explained, in part, by the difficulty that population-based surveys have in discriminating between AUD and "mere" heavy drinking. In addition, failure of cross-sectional cost studies to identify the long-term effects of this

pathology may mask its impact on quality of life; for example, a severe reduction in health that results from heavy drinking may lead to reduced alcohol intake, which might (misleadingly) suggest a positive relationship between low alcohol consumption and *poorer* quality of life. However, in samples consisting of diagnosed patients (as in our study), the relationship between severity of AUD and poor HRQoL is clear; see the studies reviewed by Ugochukwu et al. [52] or subsequent research [14, 15, 21, 53-56].

Although our study finds a significant negative relationship between AUD severity and HRQoL for all instruments analyzed, the results suggest that EQ-5D is less sensitive than the other instruments. Other studies have likewise found that the EQ-5D has some limitations for capturing those quality-of-life dimensions most affected by AUD [8]. At least two causes for this lower sensitivity are worth mentioning. On the one hand, the EQ-5D's descriptive system focuses mainly on physical condition, and – as Miller and Miller [57] point out – "the dominance of a health approach for addiction treatment may not be reflective of an addict's main concerns". It follows that emotional problems, lack of energy and vitality, and the impact of this pathology on social and family life may not be adequately captured by the EQ-5D, which could explain the greater sensitivity of the SF-6D and the AlcQ-4D found in our research. The analysis reveals also that an instrument that incorporates the social and family dimension will identify a greater treatment effect. The instrument we use (AlcQ-4D) is quite modest, and more research is needed on how best to incorporate those dimensions into an instrument that produces QALYs; yet our results indicate that their exclusion would lead to severely underestimating the outcomes of AUD treatment. On the other hand, the EQ-5D's relatively low sensitivity may be due to the "ceiling" effect that characterizes this instrument (i.e., its poor discrimination among mild health states). Although one of the EQ-5D-5L goals was to smooth the ceiling effect of the EQ-5D-3L, which is found also in the alcohol domain [27, 58], there is evidence that the ceiling effect persists (albeit smoothed) when the EQ-5D-5L is used [59]. Furthermore, direct comparisons between the EQ-5D-5L and the SF-6D reveal that the former has a higher ceiling effect than the latter; this difference has been demonstrated both in patient studies [60-62] and in general population studies (e.g., [63]).

With regard to clinical instruments, we have been unable to locate any studies that compare AUDIT and DSM-5 in patients diagnosed with AUD. In population-based studies, there is evidence showing a strong association between AUDIT and DSM-5 for detecting AUD [64, 65]. There is also evidence in favor of AUDIT (as compared with DSM-5) for detecting alcohol misuse among young people [66].

Our estimated mapping models perform well. The adjusted R^2 are all greater than 0.5, reaching 0.7 in some models. Overall, models mapping a generic non–preference-based onto a generic preference-based measure achieved an R^2 of more than 0.5 within sample, but the fit of functions that map from a condition-specific measure to a generic preference-based measures is usually lower [33]. The RMSE of our models ranged from 0.10 to 0.17 – in line with Brazier et al.'s [33] systematic review, which report values between 0.08 and 0.2. In the AUD field, Chavez et al. [35] considered two possible mappings: AUDIT-C scores to EQ-5D weights and to SF-6D weights. Yet no function was proposed by these authors

because neither measure (EQ-5D or SF-6D) suggested meaningful differences in HRQoL based on AUDIT-C categories. The authors suggest that generic HRQOL measures may not adequately capture differences associated with alcohol use, a very different result from what we obtained. These results might be explained by that study's use of a general population sample, its failure to identify sample members who were diagnosed with AUD (since AUDIT is based on *self*-reported alcohol consumption), and/or its use of a "reduced" version of the AUDIT (note that AUDIT-C includes only 3 items).

Our study has several limitations, most of which are related to its external validity. First, the sample size is small. Small samples are common when the population under study consists of patients diagnosed with alcohol problems; in Luquiens et al.'s review [8], less than a third of such studies featured sample sizes larger than 225 participants. Nonetheless, the small sample size here could call into question the generalizability of our results to the wider AUD patient population and may also have affected whether some of the estimated parameters were (or were not) statistically significant. Second, although the population was selected by systematically recruiting the population attending a drug treatment center, it may not be representative of the profiles of alcohol-dependent persons in other areas. Third, internal validity may also be compromised in the absence of a control group, which makes it practically impossible to determine with certainty whether changes in scores on the instruments analyzed are due to the treatment intervention or to another confounding variable. Note that intervention outcomes were established by comparing the values of the instruments at the beginning of the treatment and one year later (a control group was not established because this is the usual treatment at the center and no alternative approach was available for comparison). Finally, 22% of the initial sample dropped out of the study; these departures could affect both the internal and external validity of our results if there is any evidence of attrition bias. However, such bias is unlikely because we found few meaningful differences between the traced and drop-out samples.

The limitations just described would be of considerable relevance if our study's main objective had been to measure treatment efficacy. Yet they are of minor importance when one considers that our study's objective was instead to *compare within-sample instrument differences* for the purpose of (a) assessing patients' quality of life and its evolution, (b) evaluating their respective capacities to discriminate between severity groups established on the basis of clinical instruments, and (c) estimating a transformation function for converting clinical instrument scores into HRQoL instrument scores.

This study enables us to draw some conclusions relevant for public decision making. First, our results establish that all the HRQoL instruments used allow for discriminating between the AUD condition's levels of severity and that any of them is a reasonable choice in the AUD setting. Second, we have observed that the HRQoL – measured in QALYs– associated with alcohol use disorders and the effects of treatment intended to manage it, are sensitive to the choice of HRQoL instrument; hence the results from cost-of-illness studies or an economic evaluation can differ depending on which one is used. Because there is no "gold standard" against which instruments can be compared, no particular approach can be

considered superior. In any case, our results anticipate that the incremental cost–utility ratio will be lower – making it more likely that such treatment will receive public funding – when using the SF-6D or the AlcQ-4D than when using the EQ-5D. Third, our findings suggest that failing to incorporate family and social dimensions will lead to underestimating the treatment effect on quality of life. How best to incorporate these variables in the assessment of treatments is still an open question. Finally, given that clinical instruments do not have the psychometric properties necessary for their incorporation into economic evaluation studies, this study also proposes functions that can be used to transform clinical instrument scores into QALYs. In addition, our proposed three adjustment functions for each clinical instrument (i.e., one for each HRQoL instrument evaluated) facilitates sensitivity analyses when assessing quality of life and the effect of treatments on patients with alcohol use disorder. Mapping functions are never preferable to using a preference-based measure directly, but they do allow data obtained in the clinical setting to be transformed for use in cost-effectiveness analyses and thereby improve evidence-informed decision making in health policy.

References

1.World Health Organization. (2019). *Global status report on alcohol and health 2018*. World Health Organization.

2.Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, *88*(6), 791-804.

3.Contel, M., Gual, A. & Colom J. (1999). Test para la identificación de transtornos por uso de alcohol (AUDIT): Traducción y validación del AUDIT al catalán y castellano. *Adicciones*, *11*(4), 337-347.

4.American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders, fifth edition. Washington: American Psychiatric Association.

5.American Psychiatric Association (2022). Diagnostic and statistical manual of mental disorders, fifth edition, text revision. Washington: American Psychiatric Association.

6.Shorter, G. W., Bray, J. W., Heather, N., Berman, A. H., Giles, E. L., Clarke, M., ... & Newbury-Birch, D. (2021). The "outcome reporting in brief intervention trials: Alcohol" (orbital) core outcome set: International consensus on outcomes to measure in efficacy and effectiveness trials of alcohol brief interventions. *Journal of studies on alcohol and drugs*, *82*(5), 638-646.

7.Donovan, D., Mattson, M. E., Cisler, R. A., Longabaugh, R., & Zweben, A. (2005). Quality of life as an outcome measure in alcoholism treatment research. *Journal of Studies on Alcohol, Supplement*, (15), 119-139.

8.Luquiens, A., Reynaud, M., Falissard, B., & Aubin, H. J. (2012). Quality of life among alcohol-dependent patients: How satisfactory are the available instruments? A systematic review. *Drug and alcohol dependence*, *125*(3), 192-202.

9.Dyer, A., Böhnke, J. R., Curran, D., McGrath, K., & Toner, P. (2023). A systematic review of quality of life and health-related quality of life as outcomes in substance and behavioural addictions. *Drug and Alcohol Review, 42*(7), 1680-1700.

10.Daeppen, J.B., Krieg, M.A., Burnand, B., Yersin, B. (1998): MOS-SF-36 in evaluating health-related quality of life in alcohol-dependent patients. *American Journal of Drug and Alcohol Abuse 24*(4): 685–694

11.Dawson, D. A., Li, T. K., Chou, S. P., & Grant, B. F. (2009). Transitions in and out of alcohol use disorders: their associations with conditional changes in quality of life over a 3-year follow-up interval. *Alcohol & Alcoholism*, *44*(1), 84-92.

12.LoCastro, J. S., Youngblood, M., Cisler, R. A., Mattson, M. E., Zweben, A., Anton, R. F., & Donovan, D. M. (2009). Alcohol treatment effects on secondary nondrinking outcomes and quality of life: the COMBINE study. *Journal of studies on alcohol and drugs, 70*(2), 186-197.

13.Peltzer, K., & Pengpid, S. (2012). Alcohol use and health-related quality of life among hospital outpatients in South Africa. *Alcohol and Alcoholism*, *47*(3), 291-295.

14. Daeppen, J. B., Faouzi, M., Sanchez, N., Rahhali, N., Bineau, S., & Bertholet, N. (2014). Quality of life depends on the drinking pattern in alcohol-dependent patients. *Alcohol and Alcoholism*, *49*(4), 457-465.

15. Mosquera, J. & Rodríguez-Míguez, E. (2015). Using the SF-6D to measure the impact of alcohol dependence on health-related quality of life. *The European Journal of Health Economics*, *16*, 347-356.

16. Põlluste, K., Aart, A., Kallikorm, R., Kull, M., Kärberg, K., Müller, R., ... & Lember, M. (2016). Adverse lifestyle and health-related quality of life: gender differences in patients with and without chronic conditions. *Scandinavian journal of public health, 44*(2), 209-216.

17. Schrieks, I. C., Wei, M. Y., Rimm, E. B., Okereke, O. I., Kawachi, I., Hendriks, H. F., & Mukamal, K. J. (2016). Bidirectional associations between alcohol consumption and health-related quality of life amongst young and middle-aged women. *Journal of internal medicine*, *279*(4), 376-387.

18. Imtiaz, S., Loheswaran, G., Le Foll, B., & Rehm, J. (2018). Longitudinal alcohol consumption patterns and health-related quality of life: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug and alcohol review*, *37*(1), 48-55.

19. Favrod-Coune, T., Aebischer, G., Grondin-Giletti, F., Girod, I., Simoneau, H., & Broers, B. (2019). "Alcochoix+": controlled drinking within a structured programme–a cohort study in Switzerland. *Swiss medical weekly*, *149*(3940), w20120-w20120.

20. Lee, Y. Y., Wang, P., Abdin, E., Chang, S., Shafie, S., Sambasivam, R., ... & Subramaniam, M. (2020). Prevalence of binge drinking and its association with mental health conditions and quality of life in Singapore. *Addictive behaviors*, *100*, 106114.

21. Barbosa, C., Bray, J. W., Dowd, W. N., Barnosky, A., & Wittenberg, E. (2021). SF-6D utility scores for alcohol use disorder status and alcohol consumption risk levels in the US population. *Addiction*, *116*(5), 1034-1042.

22. Koivunen, M., Harju, S., Kauko, T., & Välimäki, M. (2017). Alcohol risk drinking, quality of life and health state among patients treated at the sobering unit in the emergency department–one year follow-up study. *International emergency nursing*, *31*, 22-29.

23. Wallhed Finn, S., Hammarberg, A., & Andreasson, S. (2018). Treatment for alcohol dependence in primary care compared to outpatient specialist treatment—a randomized controlled trial. *Alcohol and alcoholism*, *53*(4), 376-385.

24. Pham, T. T., & Park, B. (2019). Alcohol use disorder and health-related quality of life in Korean night-shift workers: A cross-sectional study using the KNHANES 2007-2015 data. *Plos one*, *14*(4), e0214593.

25. Eum, M. J., & Choi, M. J. (2021). Association between Exposure to Alcohol's Harm to others and health-related quality of life in korean adults: A nationwide population-based study. *International journal of environmental research and public health*, *18*(5), 2714.

26. Noto, S., Saito, S., Shiroiwa, T., & Fukuda, T. (2021). Estimation of Lifetime QALYs Based on Lifestyle Behaviors. *International Journal of Environmental Research and Public Health*, *18*(19), 9970.

27. Günther, O. H., Roick, C., Angermeyer, M. C., & König, H. H. (2008). Responsiveness of EQ-5D utility indices in alcohol-dependent patients. *Drug and alcohol dependence*, *92*(1-3), 291-295.

28. Payakachat, N., Ali, M. M., & Tilford, J. M. (2015). Can the EQ-5D detect meaningful change? A systematic review. *Pharmacoeconomics, 33*, 1137-1154.

29. Essex, H. N., White, I. R., Khadjesari, Z., Linke, S., McCambridge, J., Murray, E., ... & Godfrey, C. (2014). Quality of life among hazardous and harmful drinkers: EQ-5D over a 1-year follow-up period. *Quality of Life Research, 23*, 733-743.

30. Luquiens, A., Whalley, D., Crawford, S. R., Laramée, P., Doward, L., Price, M., ... & Aubin, H. J. (2015). Development of the Alcohol Quality of Life Scale (AQoLS): A new patient-reported outcome measure to assess health-related quality of life in alcohol use disorder. *Quality of Life Research, 24*, 1471-1481

31. Malet, L., Llorca, P. M., Beringuier, B., Lehert, P., & Falissard, B. (2006). AlQoL 9 for measuring quality of life in alcohol dependence. *Alcohol and Alcoholism, 41*(2), 181-187.

32. Rodríguez-Míguez, E., & Mosquera Nogueira, J. (2017). Measuring the impact of alcohol-related disorders on quality of life through general population preferences. *Gaceta sanitaria*, *31*, 89-94.

33. Brazier, J. E., Yang, Y., Tsuchiya, A., & Rowen, D. L. (2010). A review of studies mapping (or cross walking) non-preferencebased measures of health to generic preference-based measures. *The European journal of health economics*, *11*, 215-225.

34. Mukuria, C., Rowen, D., Harnan, S., Rawdin, A., Wong, R., Ara, R., & Brazier, J. (2019). An updated systematic review of studies mapping (or cross-walking) measures of health-related quality of life to generic preference-based measures to generate utility values. *Applied health economics and health policy, 17*, 295-313.

35. Chavez, L. J., Bradley, K., Tefft, N., Liu, C. F., Hebert, P., & Devine, B. (2016). Preference weights for the spectrum of alcohol use in the US Population. *Drug and Alcohol Dependence*, *161*, 206-213.

36. Walters, S. J., & Brazier, J. E. (2005). Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of life research*, *14*(6), 1523-1532.

37. Ramos-Goñi, J. M., Craig, B. M., Oppe, M., Ramallo-Fariña, Y., Pinto-Prades, J. L., Luo, N., & Rivero-Arias, O. (2018). Handling data quality issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. *Value in Health*, *21*(5), 596-604.

38. Abellán Perpiñán, J. M., Sánchez Martínez, F. I., Martínez Pérez, J. E., & Méndez, I. (2012). Lowering the 'floor'of the SF-6D scoring algorithm using a lottery equivalent method. *Health economics*, *21*(11), 1271-1285.

39. Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., Monteiro, M. G., & World Health Organization. (2001). AUDIT: the alcohol use disorders identification test: guidelines for use in primary health care (No. WHO/MSD/MSB/01.6 a). World Health Organization.

40. Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, *33*(2), 261-304.

41. Pisinger, C., Toft, U., Aadahl, M., Glümer, C., & Jørgensen, T. (2009). The relationship between lifestyle and self-reported health in a general population: the Inter99 study. *Preventive medicine*, *49*(5), 418-423.

42. Noronha, D. D., Martins, A. M. E. D. B. L., Dias, D. D. S., Silveira, M. F., Paula, A. M. B. D., & Haikal, D. S. A. (2016). Factors in adult health-related quality of life: a population-based study. *Ciencia & saude coletiva*, *21*, 463-474.

43. Lu, J., Yang, Y., Cui, J., Xu, W., Wu, C., Li, J., & Li, X. (2022). Alcohol use disorder and its association with quality of life and mortality in Chinese male adults: a population-based cohort study. *BMC Public Health, 22*(1), 1-11.

44. Saarni S, Suvisaari J, Sintonen H, Pirkola S, Koskinen S, Aromaa A, Lonnqvist J. (2007): Impact of psychiatric disorders on health-related quality of life: general population survey. *British journal of psychiatry 190: 326-332.*

45. Maheswaran, H., Petrou, S., Rees, K., & Stranges, S. (2013). Estimating EQ-5D utility values for major health behavioural risk factors in England. *J Epidemiol Community Health, 67*(2), 172-180.

46. Van Wilder, L., Rammant, E., Clays, E., Devleesschauwer, B., Pauwels, N., & De Smedt, D. (2019). A comprehensive catalogue of EQ-5D scores in chronic disease: results of a systematic review. *Quality of Life Research, 28*, 3153-3161.

47. Ortolá, R., García-Esquinas, E., Galán, I., & Rodríguez-Artalejo, F. (2016). Patterns of alcohol consumption and health-related quality of life in older adults. *Drug and alcohol dependence*, *159*, 166-173.

48. Martín-Fernández, J., Ariza-Cardiel, G., Polentinos-Castro, E., Sanz-Cuesta, T., Sarria-Santamera, A., & Cura-González, I. D. (2019). Explaining differences in perceived health-related quality of life: a study within the Spanish population. *Gaceta Sanitaria, 32*, 447-453.

49. Vargas-Martínez, A. M., Trapero-Bertran, M., Lima-Serrano, M., Anokye, N., Pokhrel, S., & Mora, T. (2019). Measuring the effects on quality of life and alcohol consumption of a program to reduce binge drinking in Spanish adolescents. *Drug and Alcohol Dependence*, *205*, 107597.

50. Kaplan, M. S., Huguet, N., Feeny, D., McFarland, B. H., Caetano, R., Bernier, J., ... & Ross, N. (2012). Alcohol use patterns and trajectories of health-related quality of life in middle-aged and older adults: a 14-year population-based study. *Journal of studies on alcohol and drugs*, *73*(4), 581-590.

51. Orji, C. C., Ghosh, S., Nwaobia, O. I., Ibrahim, K. R., Ibiloye, E. A., & Brown, C. M. (2021). Health behaviors and health-related quality of life among US adults aged 18–64 years. *American journal of preventive medicine*, *60*(4), 529-536.

52. Ugochukwu, C., Bagot, K. S., Delaloye, S., Pi, S., Vien, L., Garvey, T., ... & IsHak, W. W. (2013). The importance of quality of life in patients with alcohol abuse and dependence. *Harvard review of psychiatry*, *21*(1), 1-17.

53. Witkiewitz, K., Kranzler, H. R., Hallgren, K. A., O'Malley, S. S., Falk, D. E., Litten, R. Z., ... & Anton, R. F. (2018). Drinking risk level reductions associated with improvements in physical

health and quality of life among individuals with alcohol use disorder. *Alcoholism: Clinical and Experimental Research*, *42*(12), 2453-2465.

54. Patkar, P., Saldanha, D., Chaudhury, S., & Singh, I. (2019). Quality of life and disability in males with alcohol dependence syndrome. *Industrial Psychiatry Journal*, *28*(2), 262.

55. Lee, S. B., Chung, S., Seo, J. S., Jung, W. M., & Park, I. H. (2020). Socioeconomic resources and quality of life in alcohol use disorder patients: the mediating effects of social support and depression. *Substance abuse treatment, prevention, and policy*, *15*(1), 1-8.

56. Witkiewitz, K., Kranzler, H. R., Hallgren, K. A., Hasin, D. S., Aldridge, A. P., Zarkin, G. A., ... & Anton, R. F. (2021). Stability of drinking reductions and long-term functioning among patients with alcohol use disorder. Journal of General Internal Medicine, 36, 404-412.

57. Miller, P. G., & Miller, W. R. (2009). What should we be aiming for in the treatment of addiction?[Editorial]. *Addiction, 104*(5), 685–686

58. Sonntag, M., König, H. H., & Konnopka, A. (2015). The responsiveness of the EQ-5D and time trade-off scores in schizophrenia, affective disorders, and alcohol addiction. *Health and quality of life outcomes, 13*, 1-9.

59. Janssen, M. F., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., ... & Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of life research*, *22*, 1717-1727.

60. Richardson, J., Khan, M. A., Iezzi, A., & Maxwell, A. (2015). Comparing and explaining differences in the magnitude, content, and sensitivity of utilities predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D multiattribute utility instruments. *Medical Decision Making*, *35*(3), 276-291.

61. Sayah, F. A., Qiu, W., Xie, F., & Johnson, J. A. (2017). Comparative performance of the EQ-5D-5L and SF-6D index scores in adults with type 2 diabetes. *Quality of Life Research*, *26*, 2057-2066.

62. Casal, B., Rodríguez-Míguez, E., & Rivera, B. (2020). Measuring intangible cost-of-morbidity due to substance dependence: implications of using alternative preference-based instruments. *The European Journal of Health Economics*, *21*, 1039-1048.

63. Zhou, J., Xu, L., Pan, J., Wang, M., Zhou, P., Wang, W., ... & Zhu, W. (2024). A comparative study of Chinese medicine quality of life assessment scale (CQ-11D) and EQ-5D-5L and SF-6D scales based on Chinese population. *Quality of Life Research*, *33*(1), 113-122.

64. Moehring, A., Rumpf, H. J., Hapke, U., Bischof, G., John, U., & Meyer, C. (2019). Diagnostic performance of the alcohol use disorders identification test (AUDIT) in detecting DSM-5 alcohol use disorders in the general population. *Drug and alcohol dependence, 204*, 107530.

65. Källmén, H., Elgán, T. H., Wennberg, P., & Berman, A. H. (2019). Concurrent validity of the Alcohol Use Disorders Identification Test (AUDIT) in relation to Alcohol Use Disorder (AUD) severity levels according to the brief DSM-5 AUD diagnostic assessment screener. *Nordic Journal of Psychiatry 73*(7), 397-400.

66. Kuteesa, M. O., Cook, S., Weiss, H. A., Kamali, A., Weinmann, W., Seeley, J., ... & Webb, E. L. (2019). Comparing Alcohol Use Disorders Identification Test (AUDIT) with Timeline Follow

Back (TLFB), DSM-5 and Phosphatidylethanol (PEth) for the assessment of alcohol misuse among young people in Ugandan fishing communities. *Addictive behaviors reports, 10*, 100233.

Table 1. Baseline sample description: traced sample vs. drop-out sample

		Traced (n=202)	Drop-out (n=57)	*p-value*[1]
EQ-5D score (mean)		0.727	0.800	***0.010***
SF-6D score (mean)		0.558	0.666	***0.004***
AlcQ score (mean)		0.504	0.552	*0.133*
AUDIT score (mean)		21.990	22.632	*0.564*
DSM score (mean)		7.233	6.509	***0.036***
Gender (% men)		73.27	80.7	*0.253*
Age (mean)		50.322	47.667	*0.095*
Education (%)	Less than primary	3.64	8.91	*0.290*
	Primary	47.27	52.97	
	Secondary	36.36	25.25	
	University	12.73	12.87	
Type of family living (%)	Alone	26.73	21.05	*0.537*
	With a partner	35.64	36.84	
	With family of origin (no with a partner)	15.84	12.28	
	Other	21.78	29.82	
Labor status (%)	Employed	31.19	38.6	*0.208*
	Unemployed	33.66	38.6	
	Inactive	35.15	22.81	
Standard beverage units (SBU/week)		55.22	70.35	***0.043***
Dominant consumption profile (%)	Daily	78.95	66.83	*0.214*
	Weekend	5.26	8.42	
	Mixed	15.79	24.75	
Dominant type of beverage (%)	Wine	27.36	10.53	***0.010***
	Beer	29.85	50.88	
	Liquor	10.45	10.53	
	Mixed	32.34	28.07	
Level of motivation (%)	Precontemplation	5.5	3.51	*0.282*
	Contemplation	30	22.81	
	Preparation	31	26.32	
	Action	33.5	47.37	
Other drug use (%)	Tobacco	62.38	66.67	*0.553*
	Hashish	11.39	17.54	*0.218*
	Cocaine	8.42	14.04	*0.205*
	Other	0.99	3.51	*0.173*

[1] Differences between samples tests

Table 2. Mean scores of instruments

	Basal		12 months		Gain	
	Mean	SD	Mean	SD	Mean	SD
EQ-5D-5L	0.73	0.19	0.88	0.13	0.15**	0.16
SF-6D	0.56	0.25	0.78	0.18	0.23**	0.23
AlcQ-4D	0.50	0.21	0.85	0.19	0.35**	0.26
AUDIT	21.99	7.55	4.41	7.60	17.58**	9.74
DSM-5	7.23	2.19	1.52	2.89	5.71**	3.29
n	202		202			

** Significant at the 1% level.

Note: All HRQoL instruments (EQ-5D-5L, SF-6D, and AlcQ-4D) are different from each other at the 1% level, both at baseline and at 12 months *except* EQ-5D-5L (12 months) vs. AlcQ-4D (12 months), which are different at the 5% level (p = 0.0197). The AUDIT and DSM-5 instruments cannot be compared because they do not have the same measurement scale.

Table 3. Pearson correlation coefficients between instruments

	EQ-5D-5L	SF-6D	AlcQ-4D	AUDIT
EQ-5D-5L	1.000			
SF-6D	0.844	1.000		
AlcQ-4D	0.673	0.719	1.000	
AUDIT	-0.517	-0.589	-0.740	1.000
DSM-5	-0.507	-0.583	-0.738	0.885

n=404

All values are significant at the 1% level.

Table 4. HRQoL values by clinical severity group

		EQ-5D		SF-6D		AlcQ-4D	
	n	Mean	SD	Mean	SD	Mean	SD
AUDIT-level 1 (A1)	168	**0.900**	0.111	**0.819**	0.139	**0.895**	0.137
AUDIT-level 2 (A2)	47	**0.808**	0.161	**0.706**	0.190	**0.673**	0.210
AUDIT-level 3 (A3)	43	**0.780**	0.152	**0.609**	0.224	**0.570**	0.225
AUDIT-level 4 (A4)	146	**0.698**	0.202	**0.509**	0.258	**0.465**	0.207
Note: All differences are significant at the 1% level except:		*A2 vs A3 (p=0.405)* *A3 vs A4 (p=0.015)*		*A2 vs A3 (p=0.029)* *A3 vs A4 (p=0.022)*		*A2 vs A3 (p=0.028)*	
DSM-level 1 (D1)	174	**0.900**	0.109	**0.820**	0.136	**0.894**	0.139
DSM-level 2 (D2)	53	**0.760**	0.206	**0.631**	0.241	**0.610**	0.201
DSM-level 3 (D3)	177	**0.722**	0.189	**0.538**	0.251	**0.488**	0.217
Note: All differences are significant at the 1% level except:		*D2 vs D3 (p=0.209)*		*D2 vs D3 (p= 0.019)*			

Table 5. Relationship between HRQoL scores and clinical scores

	Dependent variable: HRQoL score								
	EQ-5D			**SF-6D**			**AlcQ-4D**		
	Coef.	Std. Err.	p-value	Coef.	Std. Err.	p-value	Coef.	Std. Err.	p-value
AUDIT-Model 1 (Independent variable: score AUDIT)									
Total score	**-0.008**	0.001	0.000	**-0.012**	0.001	0.000	**-0.017**	0.001	0.000
cons	**0.910**	0.012	0.000	**0.833**	0.015	0.000	**0.905**	0.014	0.000
R^2 within	*0.533*			*0.576*			*0.705*		
R^2 between	*0.114*			*0.180*			*0.247*		
R^2 overall	*0.267*			*0.346*			*0.548*		
RMSE/AIC	*0.107 / -414.564*			*0.147 / -198.247*			*0.170 / -239.576*		
AUDIT-Model 2 (Independent variable: score dimensions AUDIT)									
Dim 1	**-0.028**	0.005	0.000	**-0.042**	0.006	0.000	**-0.073**	0.007	0.000
Dim 5	**-0.027**	0.007	0.000	**-0.040**	0.009	0.000	**-0.043**	0.009	0.000
Dim 6	**-0.027**	0.006	0.000	**-0.039**	0.009	0.000	**-0.035**	0.009	0.000
cons	**0.914**	0.012	0.000	**0.837**	0.015	0.000	**0.916**	0.014	0.000
R^2 within	*0.568*			*0.606*			*0.715*		
R^2 between	*0.162*			*0.203*			*0.263*		
R^2 overall	*0.311*			*0.374*			*0.554*		
RMSE/AIC	*0.103 / -437.584*			*0.143 / -215.638*			*0.170 / -241.894*		
AUDIT-Model 3 (Independent variable: levels dimensions AUDIT)									
Dim 1 (ref: level 0)									
Level 1	**-0.041**	0.026	0.111	**-0.050**	0.035	0.157	**-0.055**	0.036	0.131
Level 2	**-0.058**	0.031	0.063	**-0.074**	0.042	0.080	**-0.187**	0.044	0.000
Level 3	**-0.077**	0.023	0.001	**-0.116**	0.031	0.000	**-0.207**	0.034	0.000
Level 4	**-0.121**	0.021	0.000	**-0.175**	0.028	0.000	**-0.281**	0.029	0.000
Dim 5 (ref: level 0)									
Level 1	**-0.001**	0.024	0.987	**-0.003**	0.033	0.929	**-0.077**	0.034	0.023
Level 2	**-0.058**	0.024	0.015	**-0.112**	0.033	0.001	**-0.145**	0.033	0.000
Level 3	**-0.063**	0.024	0.009	**-0.121**	0.033	0.000	**-0.139**	0.032	0.000
Level 4	**-0.113**	0.031	0.000	**-0.136**	0.042	0.001	**-0.166**	0.041	0.000
Dim 6 (ref: level 0-2)									
Level 3	**-0.037**	0.032	0.249	**-0.089**	0.043	0.040	**-0.068**	0.044	0.123
Level 4	**-0.144**	0.029	0.000	**-0.188**	0.040	0.000	**-0.162**	0.040	0.000
cons	**0.913**	0.012	0.000	**0.834**	0.016	0.000	**0.919**	0.016	0.000
R^2 within	*0.583*			*0.619*			*0.716*		
R^2 between	*0.208*			*0.215*			*0.273*		
R^2 overall	*0.344*			*0.386*			*0.560*		
RMSE/AIC	*0.102 / -441.860*			*0.142 / -211.362*			*0.170 / -232.875*		
DSM-Model 1 (Independent variable: score DSM-5)									
Total score	**-0.024**	0.002	0.000	**-0.036**	0.002	0.000	**-0.052**	0.002	0.000
cons	**0.908**	0.012	0.000	**0.831**	0.015	0.000	**0.905**	0.014	0.000
R^2 within	*0.504*			*0.537*			*0.702*		
R^2 between	*0.115*			*0.198*			*0.249*		
R^2 overall	*0.258*			*0.340*			*0.545*		
RMSE/AIC	*0.110 / -402.811*			*0.154 / -184.514*			*0.171 / -236.597*		
DSM-Model 2 (Independent variable: dimensions DSM-5)									
Dim 1							**-0.075**	0.031	0.017
Dim 4	**-0.055**	0.018	0.003	**-0.065**	0.025	0.010	**-0.067**	0.025	0.007
Dim 6							**-0.100**	0.029	0.001
Dim 7	**-0.057**	0.018	0.001	**-0.065**	0.025	0.008	**-0.088**	0.024	0.000
Dim 8	**-0.039**	0.020	0.048	**-0.069**	0.027	0.010			
Dim 9	**-0.073**	0.018	0.000	**-0.092**	0.029	0.001	**-0.053**	0.031	0.086
Dim 10				**-0.067**	0.026	0.010	**-0.071**	0.027	0.010
Dim 11	**-0.080**	0.021	0.000	**-0.073**	0.028	0.009	**-0.096**	0.027	0.000
cons	**0.907**	0.012	0.000	**0.829**	0.015	0.000	**0.907**	0.014	0.000
R^2 within	*0.524*			*0.545*			*0.707*		
R^2 between	*0.166*			*0.203*			*0.255*		
R^2 overall	*0.297*			*0.347*			*0.552*		
RMSE/AIC	*0.108 / -414.860*			*0.154 / -180.262*			*0.171 / -231.300*		
Observations	*404*			*404*			*404*		

Figure 1. Distribution of baseline scores and treatment impact by HRQoL instrument