Health Utility Measurement with Index Scores for People Living with HIV/AIDS Under Combined Antiretroviral Therapy: A Comparison of EQ-5D-5L and SF-6D

Xiaowen Wang

Yunnan Centers for Disease Control and Prevention

Hongbing Luo

Yunnan Center for Disease Control and Prevention

Enlong Yao

Honghe municipal center for disease control and prevention

Renhai Tang

Dehong municipal center for disease control and prevention

Wenbing Dong

Yuxi municipal center for disease control and prevention

Fuyong Liu

Zhaotong municipal center for disease control and prevention

Jun Liang

Kunming municipal center for disease control and prevention

Minyang Xiao

Yunnan center for disease control and prevention

Zuyang Zhang

Yunnan center for disease control and prevention

Jin Niu

Yunnan center for disease control and prevention

Lijun Song

Yunnan center for disease control and prevention

Liru Fu

Yunnan center for disease control and prevention

Xuehua Li

Yunnan center for disease control and prevention

Shicong Qian

Wenshan municipal center for disease control and prevention

Qing Guo

Lincang municipal center for disease control and prevention

zhizhong Song (✉ songzhizhong1206@163.com)

Yunnan center for disease control and prevention <https://orcid.org/0000-0003-4214-1090>

Research

Keywords: Health utility measurement, People living with HIV/AIDS, Combined antiretroviral therapy, EQ-5D-5L, SF-6D

Posted Date: February 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-233906/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

The EQ-5D-5L and SF-6D are two widely used generic index score measures. We compared the discriminative validity, agreement and sensitivity of EQ-5D-5L and SF-6D utility scores in people living with HIV/AIDS (PLWHIV).

Methods

We conducted a cross-sectional survey among PLWHIV aged more than 18 years old in 9 municipalities in Yunnan Province, China. A convenience sample was enrolled. We administered the SF-12 and EQ-5D-5L to measure health-related quality of life (QALY). The utility index of the SF-6D was derived from the SF-12. The covariate data included demographic components, clinical components and social-psychology components. To evaluate the homogeneity of the EQ-5D-5L and SF-6D, intraclass correlation coefficients (ICCs), scatter plots and Bland-Altman plots were computed and drawn. To evaluate the capacity to discriminate between different categories of clinical components, social support and anxiety and depression status, mean and median scores were calculated and compared using one-way ANOVA and the Kruskal-Wallis test, respectively. The effect size was defined as the difference of each of the characteristics and was computed using Z/N. We also used receiver operating characteristic (ROC) curves to compare the discriminative properties and sensitivity of the econometric index.

Results

A total of 1,797 respondents, with a mean age of 45.6 ± 11.7 years (range 18 to 80), was interviewed. The distribution of EQ-5D-5L scores skewed towards full health with a skewness of -3.316. The distribution of SF-6D scores was almost centered around its mean, and the skewness was 0.084. The effect size was smaller for the EQ-5D-5L than for the SF-6D across the social support, anxiety and depression subgroups. The overall correlation between EQ-5D-5L and SF-6D index scores was 0.46 ($P < 0.001$). An ICC of 0.59 between the EQ-5D-5L and SF-6D meant a moderate correlation and indicated general agreement. The Bland-Altman plot displayed the same results as the scatter plot. The ROC curve showed that the AUC for the SF-6D was 0.776 (95% CI: 0.757, 0.796) and that for the EQ-5D-5L was 0.732 (95% CI: 0.712, 0.752) by the PCS-12, and it was 0.782 (95% CI: 0.763, 0.802) for the SF-6D and 0.690 (95% CI: 0.669, 0.711) for the EQ-5D-5L by the MCS-12.

Discussion

Our study demonstrated evidence of the performance of EQ-5D-5L and SF-6D index scores to measure health utility in people living with HIV/AIDS. Both have shown discriminative capacity and validity in measuring health status. However, there were significant differences in their performance. Users need to pay more attention to the characteristics of the target population. HIV/AIDS has transformed from being a terminal illness to being a chronic disease. We preferred to apply the SF-6D to measure the health utility of PLWHIV during the cART period.

Conclusion

Our study has demonstrated evidence for instrument choice and preference measurements in PLWHIV under cART. The differences between the measures could generate different health utilities for the same sample population, which is critical for cost-utility analyses that guide resource allocation and decision making.

Background

Although the HIV epidemic in China has a low national prevalence of 0.037%^[1], approximately 1.05 million people living with HIV (PLWHIV) reside in China, and 86.6% of them had accepted combination antiretroviral therapy (cART) by the end of 2020^[2]. Yunnan Province has reported one of the highest prevalences among various regions in China. Yunnan has reported 0.13 million PLWHIV. A recent report demonstrated that 90% of PLWHIV in Yunnan Province know their status, 90% of them have received cART, and 90% of those on cART had an undetectable viral load^[3].

Health-related quality of life (HRQoL) is an individual or group's perceived physical and mental health and multiple factors associated with it^[4]. Attributed to the accessibility and success of cART, HIV infection has been transformed into a chronic and manageable condition^[5]. It is estimated that the life expectancy of PLWHIV at 35 years old in Shanghai, China, is more than 40 years old^[6], nearly approaching the life expectancy of normal people in Yunnan Province in 2019 (75.1 years)^[7]. The quality of life for PLWHIV has become a significant area of HIV/AIDS research. Patient-level HRQoL is an important indicator for improving the health level of PLWHIV and for supporting prevention program design. A preference-based measure of health, also called health utility, generates a utility weight to calculate quality-adjusted life years (QALYs)^[8], which is widely used in economic evaluations for resource allocation.

Instruments for assessing HRQoL are increasingly used in surveys for PLWHIV, including generic and disease-specific instruments. Generic instruments can generally be classified as psychometric profile measures and econometric index measures. Psychometric measures are usually used to generate different health dimension scores (profiles). Economic measures can provide a single globe index, such as the societal preferences for health status (utility) that we used to calculate HRQoL for cost-utility analyses^[5, 9]. Despite the availability of different instruments, these instruments have different characteristics and different abilities to detect the effect of disease conditions on overall health^[9]. Not surprisingly, the results of disease burden calculation and cost-utility analyses might be dependent on the measure used. Thus far, with the development and pluralization of HIV/AIDS prevention strategies, public health policy can be based on evaluation results to ensure valid resource allocation, which has become an ongoing and great challenge. Discrepancy analysis and comparisons between instruments are considered necessary for HRQoL measurements.

The EQ-5D (EuroQol 5-Dimension) has been widely used in population surveys because of the low respondent burden. It defined total health by five dimensions: mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression^[9]. The primary version of the EQ-5D allows respondents to indicate the degree of impairment in each dimension by three levels: no problems, some problems and extreme problems. A new version of the EQ-5D-5L includes five levels to indicate the degree of severity for each dimension: no problems, slight problems, moderate problems, severe problems and extreme problems^[5]. Regardless of whether the EQ-5D-3L or EQ-5D-5L is used, a preference-based index can be generated. The SF-12 is the abbreviated version of the SF-36 (the Short Form-36 Health Survey) with 12 items, and it provides two component summary scores related to physical and mental health^[10]. The SF-6D (the Short Form-6 Dimension) is an econometric index derived from the preference value system of the SF-36 and SF-12^[11]. The SF-6D may provide a useful alternative to the EQ-5D. It has shown values comparable to those from the EQ-5D.

Several studies have compared the EQ-5D with the SF-12 or SF-6D in the general population and different disease groups^[8–13]. Some studies reported that the different instruments generated widely differing HRQoL scores for the same patient groups. Some studies supported the usage of the SF-6D, which had lower floor and ceiling effects and could better detect the different stages of the disease. Some studies suggested that the EQ-5D could be recommended for use in severe conditions and that the SF-6D could be recommended for use in mild conditions. Together, these studies highlighted the variation in the results generated from the different instruments.

Currently, HRQoL is believed to be a dynamic and relative concept for PLWHIV in the cART era^[14]. Otherwise, a considerable number of people living with HIV/AIDS and limited resources for prevention and therapy make resource allocation critical for decision making. A preferred instrument choice is urgently needed for more accurate results. Few studies have compared the EQ-5D-5L with the SF-6D in terms of their power to distinguish health status in PLWHIV. The main purpose of our study was to describe health state index scores with EQ-5D-5L and SF-6D scores in PLWHIV to evaluate the relationship, accuracy and applicability of the two measures.

Methods

Study design

We conducted a cross-sectional survey among PLWHIV aged more than 18 years old in the 9 municipalities in Yunnan Province from October 2019 to May 2020. A convenience sample including 1,797 participants was enrolled. Investigators with strict training from local CDCs (Center for disease control and prevention) and social organizations implemented the investigation face to face.

HRQoL assessment

We administered the 12-item Short Form Health Survey (SF-12), which is the shortened version of the 36-item Short Form Health Survey (SF-36) and can explain at least 90% of the accuracy of the SF-36. The SF-12 consists of eight domains and generates two separate summary scores, physical functional scores (PCS) and mental function scores (MCS), ranging from 0 to 100. Higher scores indicated better HRQoL^[15,16]. The Cronbach's α was 0.434.

We also administered the EQ-5D-5L (EuroQoL 5-dimensions) simultaneously. The EQ-5D-5L comprised two components: the utility index (UI) and the EQ-VAS (visual analog scale, VAS). We calculated UI by the respondents scores from the five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). For each dimension, respondents were asked to mark from 1 “no problems” to 5 “extreme problems”^[17]. All the responses were combined to form a five-digit number to describe the health status; for example, 11111 represents no problems in any of the five dimensions. All of these were converted to a UI based on the EQ-5D-5L validation set for Chinese individuals (see Table 1)^[18]. The UI ranged from 0 (the worst possible health status) to 1 (the best possible health status). At the same time, all the participants were asked to complete the EQ-VAS. They recorded their self-rated health status on a vertical VAS (visual analog scale) with the end point “the worse health you can image” at the bottom (“0”) and “the best health you can image” at the top (“100”). A higher HRQoL was associated with a higher UI and EQ-VAS score^[18]. Cronbach's α was 0.813.

Covariate data collection

The data for the covariates included in the study included three parts: a demographic component, a clinical component and a social-psychology component. All the participants completed a demographic questionnaire designed by the study staff; it included information on age, marital status, education, ethnicity, and household income per year. The clinical components included time since HIV diagnosis, infectious status at diagnosis (HIV infection; AIDS), CD4 count level at diagnosis, self-reported mode of HIV transmission, and time when ART was begun. We also recorded the number of ARTs to track the recent CD4 counts and ART regimens from the electronic follow-up system. Social support was assessed by the Social Support Rating Scale (SSRS) designed by Xiao Shuiyuan in 1986, primarily for the Chinese population^[20]. It comprised ten items to examine three dimensions: objective social support, subjective social support and support utilization. The final score was obtained by averaging all three dimension scores. A higher score indicated higher perceived social support. The reliability has a Cronbach's α of 0.684. Anxiety and depression were assessed by the Chinese version of the Hospital Anxiety and Depression Scale (HADS), which comprises an anxiety and depression subscale with seven items for each condition. Each item was rated on a four-point Likert-type scale from 0 to 3. The maximum and minimum score for each ranged from 0 to 21. A participant with a single score ≥ 8 or a total score ≥ 13 was defined as pathological. A higher score demonstrated more serious anxiety and/or depression symptoms^[21]. The reliability of the HADS in our study had a Cronbach's α of 0.237.

Statistical analysis

EQ-5D-5L scoring

The EQ-5D-5L can define 3,125 possible health states by the different answer combinations. We adopted the Chinese population-based preference trade-off time (TTO) (Table 1) to transform the measures into a UI, thereby producing a single preference-based index ranging from -0.391 to 1.000, where 0 was equal to death and -0.391 meant worse than death. The Chinese validation set is shown in Table 1. For example, when we calculated a combination of “21145”, the UI equaled $1-0.066-0-0-0.252-0.258=0.424$ [17].

SF-12 scoring

Scores for the two component summaries (physical and mental component summaries, PCS-12 and MSC-12) were calculated using the 2nd edition of the standard US instrument scoring algorithms (SF-12v2) [15].

SF-6D scoring

The UI of the SF-6D can be derived from the SF-36 or SF-12, although the number of items included differs: 11 items and 7 items, respectively. We used the validation set from the SF-12 and considered the validation set from the UK general population developed by Brazier et al. (see Table 2 and Figure 1). The UI of the SF-6D can be calculated as $U=1+\text{coefficient of different dimensions}+\text{adjustment coefficient (Most}=-0.085)$. For example, when the combination was “223123” for the SF-6D, the UI equaled $1-0.012-0.068-0.069-0-0.004-0.065-0.085=0.697$. If the most serious level was chosen for any dimension, the Most was subtracted from the final result [22, 23].

Data analysis

We described the sample characteristics by calculating the number of individuals and the percentage in each category group. We also computed descriptive statistics, including the mean, standard deviation (SD), 95% confidence intervals (CIs), median, interquartile range (IQR), and the minimum and maximum EQ-5D-5L and SF-5D index scores. Ceiling and floor effects, which were defined as the proportion of respondents with the best (11111 for the EQ-5D-5L and 111111 for the SF-6D) and the worst (55555 for the EQ-5D-5L and 345555 for the SF-6D) possible theoretical scores, respectively, were calculated for the EQ-5D-5L and SF-6D. If the distribution of index scores was highly skewed, differences between the EQ-5D-5L and SF-6D index scores were examined by Wilcoxon signed-rank test. The Mann-Whitney test was used to compare index scores across participants' characteristics for two groups, and the Kruskal-Wallis test was used for more than two groups.

To evaluate the homogeneity of the EQ-5D-5L and SF-6D, intraclass correlation coefficients (ICCs), scatter plots and Bland-Altman plots were computed and drawn. $ICC=1$ indicated complete correlation; $0.7 \leq ICC \leq 0.9$ indicated a strong correlation; $0.4 \leq ICC \leq 0.69$ indicated a moderate correlation; $0.1 \leq ICC \leq 0.39$ indicated a slight correlation; and $ICC=0$ indicated no correlation [24]. A Bland-Altman plot was developed to compare two measurements of the same variable. Generally, the plots located in the 95% LOA (interval of the limit of agreement) interval occupy 95% and simultaneously cannot exceed the professional scope [25].

To evaluate the capacity to discriminate between different categories of clinical components, social support and anxiety and depression statuses, the mean scores and median scores were calculated and compared using one-way ANOVA and Kruskal-Wallis tests, respectively. Effect size was defined as the difference of each of the characteristics, which was computed using Z/N , where Z was the Mann-Whitney Z statistic, and N was the total sample size [26]. For variables with more than two categories, the effect size between extreme groups was calculated. An effect size of 0.2 demonstrated a small effect size, 0.5 demonstrated a moderate effect size, and 0.8 demonstrated a large effect size. The effect size could explain the difference in the discriminative capacity of the studied measurements. The greater the effect size was, the stronger the discriminative capacity. We defined the relevant difference of the instrument as showing different effect sizes for the same group category, which could also be explained by the disagreement on the amount of health burden. In this study, the discriminative properties of the econometric index were also compared using receiver operating characteristic (ROC) curves. We used the SF-12 component summaries as external indicators of the performance of the EQ-5D-5L and SF-12. We set the external indicator as a dichotomized variables using the median cutoff points of the PCS-12 and MCS-12. The largest area under the ROC curve (AUC) for the utility measurement demonstrated the most sensitivity to detect differences in the external indicators. The F-ratio of the significance test for the AUC was referenced to 1.0 for the EQ-5D-5L index. If a value was greater than 1.0, we considered the SF-6D index to be more efficient than the EQ-5D-5L at detecting differences between the categories [27].

Results

Research field and subjects

A total of 1,797 respondents, with a mean age of 45.6 ± 11.7 years (range 18 to 80), were interviewed. A total of 68.1% of respondents were of Han nationality, while others were from minority ethnic groups, including Yi, Dai, Zhuang, Jinpo, Lisu, and Bai. A total of 58.7% of respondents declared themselves unmarried, divorced or separated. A total of 69.5% of respondents had less than 9 years of compulsory education. A total of 53.6% of respondents' occupations were farming and migrant work. The average yearly income per capita of their households was 10,871 yuan in 2019. Regarding HIV-related clinical characteristics,

more than two-thirds of respondents were in the HIV stage (70.4%), 28.3% were in the AIDS stage, and only 1.2% were unclear about when they were diagnosed the first time. A large proportion of the sample (68.8%) obtained HIV via heterosexual transmission, and 20.3% reported a history of intravenous drug use (IDU). A total of 98.7% of patients took ART; among those, 59.2% had been treated for more than four years. The majority of patients had sustained high CD4 cell counts, and 48.5% had cell counts of more than 500 cells/ μ l. Table 3 shows the details of the sociodemographic and clinical characteristics of all the respondents.

Descriptive statistics for the EQ-5D-5L and SF-6D

The mean EQ-5D-5L index score was 0.896 ± 0.150 (median 0.942, IQR 0.115). The distribution of EQ-5D-5L scores skewed towards full health, with a skewness of -3.316. The index score ranged from -0.391 to 1.000. The percentage of respondents ranked at the floor and ceiling were 0.1% (n=2) and 33.0% (n=593), respectively (Figure 2). The mean SF-6D index score was 0.772 ± 0.137 (median 0.762, IQR 0.241). The distribution of SF-6D scores was almost centered around its mean, and the skewness was 0.084, with a range from 0.374 to 1.000. The percentages of respondents ranked at the floor and ceiling were 0.06% (n=1) and 6.5% (n=116), respectively (Figure 3). The mean SF-6D index score for respondents with the best health state on the EQ-5D-5L descriptive system (11111) was 0.862, and those with the worst health state (55555) had a mean SF-6D index score of 0.797. Conversely, the mean EQ-5D-5L index score for those with the best SF-6D health state (355151) was 0.990, and for the one respondent with the worst health state (111515), it was 0.364. Overall, the mean EQ-5D-5L index scores exceeded the mean SF-6D index scores by 0.124, and the difference between the medians was 0.22. The difference between the EQ-5D-5L and SF-6D index scores was significant for the entire sample and for some examined sociodemographic and infectious status subgroups ($P < 0.05$) (Table 4). Both the EQ-5D-5L and SF-6D index scores were significantly different across groups of age, race/ethnicity, education level, occupation, household income per year, transmission mode and duration of ART. EQ-5D-5L index scores were significantly different across initial infectious status. SF-6D index scores were significantly different across the most recent CD4 counts.

Comparison of the SF-12 scores across different dimensions of the EQ-5D-5L and SF-6D

Both the PCS-12 and MCS-12 scores indicated significant differences across all EQ-5D-5L dimensions, with differences from 0.10 to 0.29 for the EQ-5D-5L dimensions for the PCS-12 and 0.05 to 0.28 for the EQ-5D-5L dimensions for the MCS-12. The relationship among mobility, self-care, usual activities, pain/discomfort dimension and PCS-12 and the relationship between the anxiety/depression dimension and MCS-12 were stronger. The relationship between the less comparable dimensions and component scores was weaker (Table 4). Both the PCS-12 and MCS-12 scores indicated significant differences across all SF-6D dimensions, with differences from 0.04 to 0.61 for the SF-6D dimensions for the PCS-12 and 0.02 to 0.52 for the SF-6D dimensions for the MCS-12. The relationship between the physical function, role limitation, vitality and social function dimensions and the PCS-12 and the relationship between bodily pain and the mental health dimension and the MCS-12 were stronger. The relationship between the less comparable dimensions and component scores was weaker (Table 6). The effect size was smaller for the EQ-5D-5L than the SF-6D across the social support, anxiety and depression subgroups (Table 7).

Relationship between the EQ-5D-5L and SF-6D

The overall correlation between EQ-5D-5L and SF-6D index scores was 0.46 ($P < 0.001$). The association of the two scales appeared stronger at the upper end. We also observed a degree of dispersion in which very low EQ-5D-5L scores were associated with very high scores on the SF-6D. Conversely, the very high EQ-5D-5L index scores were associated with very low scores on the SF-6D (Figure 4). An ICC of 0.59 between the EQ-5D-5L and SF-6D meant a moderate correlation and indicated general agreement. The Bland-Altman plot displayed the same results as the scatter plot, with a mean difference between the EQ-5D-5L and SF-6D index scores of 0.124. Three percent of observations were outside the 95% limits of agreement (-0.170, 0.418), which indicated an overall acceptable agreement. However, the agreement seemed weaker at the lower end of the scale, with the majority of the observations outside the limits of agreement lines. The distribution of the scatter showed a linear trend, which meant that the more obvious difference between EQ-5D-5L and SF-6D index scores existed in the observations with a good or weak health status, while the observations with a general health status and the homogeneity between the two scales seemed good (Figure 5).

Sensitivity of EQ-5D-5L and SF-6D index scores

We set the PCS-12 and MCS-12 as the gold standard for measuring health status and used the median of the PCS-12 and MCS-12. The index scores measured by the EQ-5D-5L and SF-6D were divided into two categories. The ROC analysis results were as following (Figure 6 and Figure 7): the AUC for the SF-6D was 0.776(95% CI: 0.757,0.796), while that for the EQ-5D-5L was 0.732(95% CI: 0.712, 0.752), according to the PCS-12, and that for the SF-6D was 0.782(95% CI:0.763, 0.802), while that for the EQ-5D-5L was 0.690(95% CI:0.669, 0.711), according to the MCS-12. Both AUC differences for the two groups were significant ($P < 0.05$). The AUCs for both the SF-6D and EQ-5D-5L were more than 0.5, and the F-ratios were 1.06 and 1.13, respectively, with the EQ-5D-5L as the reference, which revealed the good ability to discriminate health statuses. The SF-6D seemed more sensitive than the EQ-5D-5L in discriminating health status defined by the SF-12.

Discussion

Our study demonstrated evidence of the performance of EQ-5D-5L and SF-6D index scores in measuring health utility in people living with HIV/AIDS, showing a moderate correlation between the two measurements. Both have shown discriminative capacity and validity in measuring the health status of PLWHIV. However, some considerable overlaps existed in the two measurements, and there were significant differences in their performance, which were in accordance with the results reported in previous studies about the differences between the EQ-5D and SF-6D in the general population and several patient groups, such as rural residents in China and patients with Pompe disease, diabetes, mental health, chronic low back pain, stroke and breast cancer^[11,17,25,26,28-30].

In our study, for the mean and median EQ-5D-5L and SF-6D scores, when assessing the same sample of people living with HIV/AIDS, the EQ-5D-5L values exceeded the SF-6D values regardless of whether the whole sample or any of the subgroups was being considered, with a mean difference of 0.124 and a median difference of 0.180. This result was consistent with some studies that proved the difference in the value [11, 17, 25, 26, 28–30]. The ICC for the whole sample was 0.59, which meant a moderate correlation. We could consider this an acceptable but not very good level of agreement for the two measurements, especially at the more serious and mild ends of the scales. The two kinds of plots also revealed the details of the marked differences between the two measures. The lack of agreement highlighted the importance of considering the reasons behind the differences to assess the suitability of the instruments within a population of PLWHIV, which is important to health technology assessment and policy making. Some previous studies have explored the reasons for the differences in the EQ-5D and SF-6D to measure health utility. We mainly discussed on the following three points. First, valuation methods were considered to explain the difference. The EQ-5D-5L is based on the time-trade-off (TTO) method, whereas the SF-6D made use of the standard gamble (SG) technique [26]. Previous studies have shown that SG technique produced higher values than the TTO method [11, 17, 29], and crossover occurred in one study in which TTO values for milder states were higher than the SG values [8, 25, 30]. Our study was in accordance with this result. HIV/AIDS has transformed into a chronic disease. With scaled ART, PLWHIV can maintain good physical health. We considered that it had milder states than some diseases with disability.

Second, in our study, both the EQ-5D-5L and SF-6D performed better in monitoring changes in social and psychology aspects than physical aspects for people living with HIV/AIDS; among these, the SF-6D appeared to detect more changes and had larger effect sizes than the EQ-5D-5L. This result is somewhat surprising in that the richer descriptive system of the SF-6D might make it easier to identify changes in psychological aspects, which are often smaller and more unnoticeable than physical aspects. Based on the ROC curves and AUCs, both measures revealed good ability in discriminating health status, and the SF-6D seemed more sensitive in discriminating health status. One previous study demonstrated that the difference in SE was inherently driven by the smaller SD of the SF-6D, which was a consequence of the narrower range of the index scores [26]. We considered that the reason lies in the discrepancies in the descriptive systems' contents. In a given sample population, all participants should complete the two measures simultaneously, whereby their health status would be described by the EQ-5D-5L, which includes the five areas of mobility, self-care, usual activities, pain/discomfort and anxiety/depression, while the areas of physical functioning, role limitations, bodily pain, vitality, social functioning and mental health are obtained from the SF-6D. Different descriptive contents defined the application and appropriateness. The EQ-5D-5L emphasizes the physical aspect of health more, while the SF-6D emphasizes mental health and social adaptation more. With combined antiretroviral therapy greatly improving the survival of people living with HIV/AIDS, HIV/AIDS has transformed from being a terminal illness to being a chronic disease. A rising challenge for this population is full health, which requires more consideration be given to mental health and family and society rehabilitation. Therefore, the results implied that researchers have to choose between the two instruments based on the appropriateness of the descriptive system for the severity of potential problems the patient group may encounter. From this perspective, we preferred to apply the SF-6D to measure health utility in PLWHIV during the cART.

Third, we also considered that the various scoring algorithms contributed partly to the discrepancy of the two measures [19, 31]. The validation algorithms for the EQ-5D-5L and SF-6D are presented in Tables 1 and 2 in the Methods section. There were two different kinds of algorithms that had an effect on the index score generation. For the same health status, the different scoring algorithms assigned different index scores; the worst health status measured by the EQ-5D-5L was -0.391 (worse than death), while the SF-6D index score was 0.331 . These variations resulted in different descriptive systems and different theories of scoring systems from which to choose. One previous study proved that the interpretation of the constant terms and the interaction items were the two key factors [8]. The SF-6D interpreted the constant as an expected value that was equal to one, whereas the EQ-5D interpreted the difference between the constant and one as 'any move away from full health'. For the interaction effects, the SF-6D had a simple dummy named 'MOST', which meant that the value 1 subtracted MOST if any dimension was at the 'most severe' level. The EQ-5D had a dummy named N3, which was similar to 'MOST'. MOST had a coefficient of -0.085 , while N3 had a coefficient of 0 in the Chinese validation set for the EQ-5D-5L.

In addition, the preferences of the source population may also be a possible reason for the difference. The EQ-5D-5L values reflected Chinese patient preferences, while the SF-6D values reflected UK patient preferences.

Based on these factors, users needed to pay more attention to the characteristics of the target population. We can summarize some principles for making selections. First, for the general population or a mild patient population with generally good health, the EQ-5D-5L and SF-6D were likely to perform similarly, but for a sicker population, the performance of the two measures seemed different. Second, for a patient population with greatly impacted mental health and mild or minimally impacted physical health, we suggested selection of the SF-6D; such patients could include those with mental health problems, HIV/AIDS, or early stage breast cancer and patients in the controlled disease period. Otherwise, for a patient population with greatly impacted physical health, we suggested selection of the EQ-5D-5L; such populations could include patients with disease loss capacity and patients in the advanced disease period. Third, we should also consider the availability of the scoring algorithm, the origin of the population used for the validation set, the extent of change in health status and resource allocation, when using cost-utility analysis to inform local decisions.

There were some limitations in our study. First, the results are limited to our sample population of people living with HIV/AIDS who had good ART, and thus, these results may not be generalizable to all people living with HIV/AIDS, including patients with failed ART. Second, we used the SF-12 as the gold standard to establish the comparisons; however, the results of the SF-6D are derived from the SF-12, which could generate bias for the results to some extent. Third, we constructed a cross-sectional study and could not capture the responsiveness of the two measures. Fourth, depressive and anxiety symptoms were measured based on self-reports, which could over- or underestimate these symptoms.

Conclusion

Despite these limitations, our study has demonstrated evidence for instrument choice and preference measurements in PLWHIV under cART. Both EQ-5D-5L and SF-6D have shown discriminative capacity and validity in measuring health status. However, there were significant differences in their performance. The

differences between the measures could generate different health utilities for the same sample population, which is critical for cost-utility analyses that guide resource allocation and decision making.

Abbreviations

cART

Combined antiretroviral therapy

ICCs

Intraclass correlation coefficients

PLWHIV

People living with HIV

QALY

Quality of life year

ROC

Receiver operating characteristic

SG

Standard gamble

TTO

Time-trade-off

UI

Utility Index

Declarations

Acknowledgement

We thank all the investigators of our study and all the individuals in the study for their participation.

Funding

Our study was supported by National Nature Science Foundation of China (No. 71904166), Yunnan high-level medical cultivation programme (No.H-2018103) and 13th Five-year National S&T Major Project for Comprehensive Pilots (No. 2018ZX10715006).

Availability of data and materials

We used STATA version 14.0(StataCorp LLC, College Station, TX) to perform statistical analysis. The data will not be shared because the raw data included the individual's information, and the information of people living with HIV/AIDS must be kept confidential.

Author's contribution

All authors were involved in the study's conception and design, as well as data collection, sorting, analysis and interpretation. All authors critically reviewed the report for important intellectual and practical content. WX did the study design, statistical analysis and manuscript writing. LH, YEN, DW, LJ, LF, QS, GQ, HL, ZJ, XM, ZZ, NJ, FL, LX and SL did the field investigation organization. SZ supervised the study.

Ethics approval and consent to participate

All patients were informed about the study objective, and they were assured of confidentiality. They were asked to indicate their agreement and understanding with a signed informed consent form before the investigation. The study was approved by the ethics research committee of Yunnan Centers for Disease Control and Prevention, China.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

References

1. Yogesh SP, Guo N, Gabriel T, et al. Improving access to antiretrovirals in China: economic analyses of dolutegravir in HIV-1 patients[J]. Cost Effectiveness Resource Allocation. 2019;17:26–34.
2. GmW.cn. It reported that 1.05 million people living with HIV/AIDS resided in China [EB/OL]. <https://xw.qq.com/cmsid/20201201a03yiz00>. 2020-12-1/2021-1-29.

3. Yunnan CN. Yunnan held the briefing of AIDS epidemic and prevention progress [EB/OL].<http://m.yunnan.cn/system/2020/12/01/031153259.shtml>. 2020-12-1/2021-1-29.
4. Morad G, Nouredine E, Aicha B, et al. Assessment of quality of life in children, adolescents, and adults with celiac disease through specific questionnaires: Review [J]. *Nutrition Clinique et Métabolisme*. 2020;34:194–200.
5. Bach XT, Arto O, Long TN. Quality of life profile and psychometric properties of the EQ-5D-5L in HIV/AIDS patients [J]. *Health Quality of Life Outcomes*. 2012;10:132–9.
6. Sun GM, Gu KK, Han zhiy. Life expectancy of 917 cases with HIV/AIDS in Shanghai [J]. *Chin Prev Med*. 2019;20:32–6.
7. Yunnan.CN. The average Life expectancy of Yunnan in 2019 has approached to 75.1 [EB/OL]. <http://yn.yunnan.cn/system/2020/12/29/031207628.shtml>. 2020-12-29/2021-02-03.
8. John B, Jennifer R, Aki T, et al. A comparison of the EQ-5D and SF-6D across seven patients groups [J]. *Health Econ*. 2004;13:873–84.
9. Yuan SH, Jennifer T, A.Sarah W, et al. Mapping the medical outcomes study HIV health survey (MOS-HIV) to the EuroQoL 5 dimension(EQ-5D-3L) utility index[J]. *Health Quality of Life Outcomes*. 2019;17:83–91.
10. Wang XW, Guo GP, Zhou L, et al. Health-related quality of life in pregnant women living with HIV: a comparison of EQ-5D and SF-12 [J]. *Health Quality of Life Outcomes*. 2017;15:158–69.
11. Kanters TA, Redekop WK, Kruijshaar, et al. Comparison of EQ-5D and SF-6D utilities in Pompe disease[J]. *Oual Life Res*. 2015;24:837–44.
12. Johnson AJ, Pickard AS. Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta, Canada [J]. *Med Care*. 2000;38:115–21.
13. Wu JQ, Cao LL, Sun LH. Comparative study of EQ-5D and SF-6D in measuring health utility value over the same period [J]. *Health Economics Research*. 2020;37:42–5.
14. Vinicius CV, Liliane L, Viviane AS, et al. Oral health and health-related quality of life in HIV patients [J]. *BMC Oral Health*. 2018;18:151–6.
15. Ware JE, Kosinski M, Keller SD, et al. How to score the SF-12 physical and mental summary scale. 2nd ed[M]. Massachusetts: The Health Institute, New England Medical Center; 1995.pp. 1–97.
16. Hu H, Luan L, Yang K, et al. Psychometric validation of Chinese Health assessment questionaaire for use in rheumatoid arthritis patients in China [J].*Int J Rheum Dis* 2016; 1–6.
17. Thomas GP, Wang L, Nathalie C. EQ-5D-5L and SF-6Dv2 utility scores in people living with chronic low back pain: a survey from Quebec [J]. *BMJ Open*. 2020;10:e035722.
18. Luo N, Liu G, Li MH, et al. Estimating an EQ-5D-5L value set for China [J]. *Value in Health*. 2017;20:662–9.
19. Luo N, Li MH, Liu GG, et al. Developing the Chinese version of the new 5-level EQ-5D descriptive system: the response scaling approach [J]. *Qual Life Res*. 2013;22:885–90.
20. Wang YY, Shiu CS, Simoni JM, et al. Psychometric testing of the Chinese version of the medical outcomes study social support survey among people living with HIV/AIDS in China [J]. *Appl Nurs Res*. 2015;28:328–33.
21. Reda A. Reliability and validity of the Ethiopian version of the hospital anxiety and depression scale. (HADS) in HIV infected patients [J]. *Plos One*. 2011;6:e16049.
22. John EB, Jennifer R. The estimation of a preference-based measure of health from the SF-12 [J]. *Med Care*. 2004;42:851–9.
23. Ding SX, Ma X, Tang X, et al. Measurement and influencing factors of health utility value of middle-aged and elderly residents in rural areas [J]. *Chinese General Practice*. 2017;20:3272–6.
24. Yu HM, Luo YH, Sa J, et al. The Intraclass Correlation Coefficient and software calculation [J]. *Chinese Journal of Health Statistics*. 2011;28:497–500.
25. Du XD, Zhu P, Li M, et al. Health utility of patients with stroke measured by EQ-5D and SF-6D [J]. *J Sichuan Univ(Med Sci Edi)*. 2018;49:252–7.
26. Fatima AS, Qiu WY, Xie F, et al. Comparative performance of the EQ-5D-5L and SF-6D index scores in adults with type 2 diabetes [J]. *Qual Life Res* 2017.
27. Anik RP, Richard TL, Carlo AM, et al. The validity of the SF-12 and SF-6D instruments in people living with HIV/AIDS in Kenya [J]. *Health Quality of Life Outcomes*. 2017;15:143–51.
28. Jin H, Wang B, Gao Q, et al. Comparison between EQ-5D and SF-6D utility in rural residents of Jiangsu province, China [J]. *Plos One*. 2012;7:e41550.
29. Lamers LM, Bouwmans CAM, Van Straten A, et al. Comparison of EQ-5D and SF-6D utilities in mental health [J]. *Health Econ*. 2006;15:1229–306.
30. Mahmood Y, Safa N, Sharam G, et al. Comparison of SF-6D and EQ-5D scores in patients with breast cancer [J]. *Iran Red Crescent Med J*. 2016;18:e23556.
31. Brazier J, Robert J, Deverill M. The estimation of a preference-based measure of health from the SF-36 [J]. *Journal of Health Economics*. 2002;21:271–92.

Tables

Table 1 Chinese value set for EQ-5D-5L health status

Valuable	EQ-5D-5L
C	-
M02	0.066
M03	0.158
M04	0.287
M05	0.345
SC2	0.048
SC3	0.116
SC4	0.21
SC5	0.253
UA2	0.045
UA3	0.107
UA4	0.194
UA5	0.233
PD2	0.058
PD3	0.138
PD4	0.252
PD5	0.302
AD2	0.049
AD3	0.118
AD4	0.215
AD5	0.258
N3	-

Table 2 UK value set for SF-6D (SF-12) health status

PF			RL			SF			BP			MH	
SF-12 items and Options	SF-6D dimension	e	SF-12 items and Options	SF-6D dimension	e	SF-12 items and Options	SF-6D dimension	e	SF-12 items and Options	SF-6D dimension	e	SF-12 items and Options	SF-6D dimer
Q2=3	PF1	0	Q4 or Q7=5	RL1	0	Q12=5	SF1	0	Q8=1	BP1	0	Q11=5	MH1
Q2=2	PF2	-0.012	Q5=(1 or 2 or 3 or 4)	RL2	-0.068	Q12=4	SF2	-0.059	Q8=2	BP2	-0.004	Q11=4	MH2
Q2=1	PF3	-0.040	Q6=(1 or 2 or 3 or 4)	RL3	-0.061	Q12=3	SF3	-0.069	Q8=3	BP3	-0.039	Q11=3	MH3
-	-	-	Q5 and Q6=(1 or 2 or 3 or 4)	RL4	-0.054	Q12=2	SF2	-0.078	Q8=4	BP4	-0.076	Q11=2	MH4
-	-	-	-	-	-	Q12=1	SF1	-0.093	Q8=5	BP5	-0.140	Q11=1	MH5

Table 3 Study sample characteristics

Characteristic	Distribution(%)	
Age(Years)	16-18	17(0.9)
	18-30	112(6.2)
	30-45	702(39.0)
	45-60	766(42.6)
	≥60	200(11.1)
Race/ethnicity	Han nationality	1225(68.1)
	Yi minority	122(6.7)
	Zhuang minority	211(11.7)
	Other minority ethnic group	239(13.2)
Marital status	Separated/Divorced	1057(58.8)
	Married/Cohabiting	740(41.1)
Education level	<9 years	1308(72.7)
	≥9 years	489(27.2)
Occupation	Workers	140(7.7)
	Public officers/Staff member	114(6.3)
	Farmers	727(40.4)
	Migrant workers	239(13.2)
	Self-employed	317(17.6)
	Unemployed	260(14.4)
Household income per year(CNY)	<5,000 yuan	592(32.9)
	5,000 to 10,000yuan	559(31.1)
	10,000 to 50,000yuan	602(33.5)
	≥50,000yuan	44(2.4%)
Initial infectious status	HIV status	1266(70.4)
	AIDS status	509(28.3)
	Unclear	22(1.2)
Transmission model	Heterosexual transmission	1238(68.8)
	Homosexual transmission	107(5.9)
	Intravenous Drug Use	365(20.3)
	Mother-to-infant	23(1.2)
	Unclear	64(3.5)
Duration of ART	≤1 years	257(14.3)
	1-2 years	148(8.2)
	2-4 years	289(16.0)
	≥4 years	1078(59.9)
	Not yet	25(1.3)
The most recent CD4 counts	C≥500 cells/μl	876(48.7)
	350-500 cells/μl	359(19.9)
	200-350cells/μl	299(16.6)
	<200 cells/μl	158(8.7)
	Unclear	105(5.8)

Table 4 EQ-5D-5L and SF-6D index score comparison overall and across participant characteristics

Characteristic	EQ-5D-5L		SF-6D	
	Mean(SD)	Median(IQR)	Mean(SD)	Median(IQR)
All participants	0.896(0.004)	0.942(0.003)	0.772(0.003)	0.762(0.241)
Age(Years)*				
16-18	0.893(0.152)	0.942(0.118)	0.768(0.137)	0.749(0.237)
18-30	0.896(0.151)	0.942(0.115)	0.771(0.139)	0.758(0.241)
30-45	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
45-60	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
≥60	0.896(0.151)	0.942(0.115)	0.771(0.137)	0.759(0.241)
Race/ethnicity*				
Han nationality	0.896(0.159)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Yi nationality	0.897(0.150)	0.942(0.111)	0.772(0.138)	0.762(0.241)
Zhuang nationality	0.896(0.150)	0.885(0.115)	0.772(0.138)	0.762(0.241)
Other minority ethnic group	0.896(0.150)	0.942(0.114)	0.772(0.138)	0.762(0.241)
Marital status				
Separated/Divorced	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Married/Cohabiting	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Education level**				
<9 years	0.896(0.150)	0.897(0.115)	0.772(0.138)	0.772(0.241)
≥9 years	0.897(0.150)	0.942(0.114)	0.772(0.138)	0.762(0.241)
Occupation**				
Workers	0.897(0.150)	0.942(0.113)	0.772(0.138)	0.762(0.241)
Public officers/Staff member	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Farmers	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Migrant workers	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.759(0.241)
Self-employed	0.893(0.154)	0.942(0.118)	0.767(0.136)	0.746(0.237)
Unemployed	0.896(0.150)	0.942(0.115)	0.772(0.137)	0.762(0.241)
Household income per year(CNY)*				
<5,000yuan	0.897(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
5,000 to 10,000yuan	0.897(0.150)	0.942(0.114)	0.772(0.138)	0.762(0.241)
10,000 to 50,000yuan	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
≥ 50,000yuan	0.895(0.152)	0.942(0.115)	0.772(0.137)	0.756(0.239)
Initial infectious status				
HIV status	0.896(0.150)*	0.942(0.115)*	0.772(0.138)	0.762(0.241)
AIDS status	0.896(0.150)	0.942(0.115)	0.772(0.137)	0.762(0.241)
Unclear	0.893(0.153)	0.942(0.118)	0.769(0.136)	0.751(0.239)
Transmission model*				
Heterosexual transmission	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Homosexual transmission	0.896(0.151)	0.942(0.115)	0.772(0.138)	0.759(0.241)
Intravenous Drug Use	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Mother-to-infant	0.893(0.152)	0.942(0.118)	0.772(0.137)	0.749(0.241)
Unclear	0.896(0.151)	0.942(0.115)	0.772(0.137)	0.759(0.241)
Duration of ART*				
≤ 1 year	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.761(0.241)

1-2 years	0.896(0.150)	0.942(0.115)	0.772(0.138)	0.759(0.241)
2-4 years	0.896(0.150)	0.942(0.115)	0.771(0.137)	0.759(0.241)
≥ 4 years	0.893(0.150)	0.942(0.115)	0.772(0.138)	0.762(0.241)
Not yet	0.896(0.152)	0.942(0.115)	0.772(0.137)	0.762(0.241)
The most recent CD4counts				
CD4≥500 cells/ul	0.894(0.152)	0.942(0.115)	0.769(0.137)*	0.753(0.239)*
350-500 cells/ul	0.905(0.129)	0.942(0.107)	0.776(0.138)	0.766(0.237)
200-350 cells/ul	0.893(0.135)	0.942(0.152)	0.773(0.141)	0.764(0.247)
<200 cells/ul	0.894(0.152)	0.942(0.115)	0.767(0.136)	0.745(0.237)
Unclear	0.916(0.137)	1.000(0.107)	0.820(0.139)	0.831(0.217)

* $P<0.001$, ** $P<0.05$

Table 5 Mean(SD) SF-12component scores by EQ-5D-5L dimensions

EQ-5D-5LDimension	Level	n	PCS-12			MCS-12		
				<i>P</i>	*		<i>P</i>	*
Mobility	1	1487	46.42±8.67	0.001	0.26	48.13±9.64	0.001	0.12
	2	229	46.43±8.69			48.13±9.65		
	3	53	46.39±8.70			48.16±9.64		
	4	19	46.12±8.73			47.77±9.67		
	5	9	46.17±8.74			47.84±9.62		
Self-care	1	1687	46.42±8.67	0.001	0.10	48.13±9.64	0.001	0.05
	2	85	46.42±8.68			48.11±9.64		
	3	11	46.31±8.74			48.17±9.67		
	4	10	46.15±8.73			47.83±9.67		
	5	4	46.13±8.72			47.88±9.65		
Usual activities	1	1516	46.16±8.74	0.001	0.27	47.78±9.60	0.001	0.11
	2	227	46.22±8.75			47.97±9.67		
	3	29	46.42±8.68			48.12±9.64		
	4	18	46.42±8.68			48.12±9.64		
	5	7	46.42±8.67			48.13±9.64		
Pain/discomfort	1	945	46.13±8.77	0.001	0.29	47.81±9.60	0.001	0.11
	2	713	46.15±8.73			47.75±9.61		
	3	100	46.30±8.69			48.00±9.64		
	4	32	46.42±8.67			48.13±9.64		
	5	7	46.42±8.67			48.13±9.65		
Anxiety/depression	1	874	46.42±8.67	0.001	0.13	48.13±9.64	0.001	0.28
	2	787	46.43±8.68			48.12±9.64		
	3	101	46.32±8.69			48.04±9.63		
	4	18	46.22±8.76			47.99±9.68		
	5	17	46.19±8.72			47.88±9.61		

*=SSmodel/SStotal,which means the strength of the relationship in ANOVA without the influence of sample size.

Table 6 Mean (SD) SF-12component scores by EQ-6D dimensions

SF-6D Dimension	Level	n	PCS-12			MCS-12		
				<i>P</i>	*		<i>P</i>	*
Physical function	1	105	46.40±8.69	0.001	0.61	48.15±9.63	0.001	0.09
	2	495	46.43±8.68			48.13±9.64		
	3	1197	46.42±8.67			48.13±9.64		
Role limitation	1	51	46.36±8.68	0.001	0.04	48.13±9.64	0.001	0.02
	2	1538	46.42±8.67			48.11±9.64		
	3	178	46.16±8.73			48.17±9.67		
	4	30	46.15±8.73			47.83±9.67		
Bodily pain	1	39	46.26±8.72	0.001	0.25	47.78±9.60	0.001	0.52
	2	104	46.41±8.69			47.97±9.67		
	3	366	46.42±8.69			48.12±9.64		
	4	468	46.40±8.67			48.12±9.64		
	5	820	46.42±8.68			48.13±9.64		
Vitality	1	684	46.42±8.68	0.001	0.54	48.13±9.64	0.001	0.26
	2	443	46.42±8.69			48.13±9.65		
	3	513	46.37±8.68			48.03±9.64		
	4	109	46.42±8.69			48.15±9.64		
	5	48	46.33±8.72			48.17±9.70		
Social function	1	69	46.38±8.69	0.001	0.08	48.08±9.63	0.001	0.50
	2	145	46.37±8.69			48.03±9.65		
	3	505	46.43±8.67			48.13±9.64		
	4	762	46.42±8.68			48.13±9.64		
	5	316	46.43±8.68			48.12±9.65		
Mental health	1	422	46.42±8.68	0.001	0.31	48.13±9.64	0.001	0.50
	2	691	46.42±8.67			48.12±9.64		
	3	378	46.42±8.68			48.04±9.63		
	4	224	46.42±8.68			47.99±9.68		
	5	82	46.36±8.70			47.88±9.61		

*=SSmodel/SStotal, which means the strength of the relationship in ANOVA without the influence of sample size

Table 7 Discriminant validity of EQ-5D-5L and SF-6D index scores

Characteristic	EQ-5D-5L	SF-6D
	Median±IQR	Median±IQR
Duration of ART		
<4 years	0.942(0.115)	0.762(0.241)
≥4 years	0.942(0.115)	0.762(0.241)
D-value	0	0
ES	0.003	0.003
The most recent CD4counts		
≥350cells/ml	0.942(0.115)	0.753(0.239)
<350cells/ml	0.942(0.115)	0.745(0.237)
D-value	0	0.008
ES	0.001	0.0003
Social support		
≤33 scores	0.942(0.115)	0.762(0.241)
>33 scores	0.942(0.115)	0.762(0.241)
D-value	0	0
ES	0.0036	0.0044
Anxiety		
<8 scores	0.942(0.115)	0.760(0.241)
≥8 scores	0.942(0.115)	0.762(0.241)
D-value	0	0.002
ES	0.002	0.003
Depression		
<8 scores	0.942(0.115)	0.760(0.241)
≥8 scores	0.942(0.115)	0.762(0.241)
D-value	0	0.002
ES	0.0007	0.001

Figures

Figure 1

the item correlation between SF-12 and SF-6D

Figure 2

Page 16/21

Figure 3
Distribution of SF-6D

Figure 4

Scatter plot of EQ-5D-5L and SF-6D index scores

Figure 5

Bland-Altman plot of EQ-5D-5L and SF-6D index scores

Figure 6
ROC of EQ-5D-5L and SF-6D by PCS-12

Figure 7
ROC of EQ-5D-5L and SF-6D by MCS-12