Medicine®

OPEN

Health utility measurement for people living with HIV/AIDS under combined antiretroviral therapy
A comparison of EQ-5D-5L and SF-6D

Xiaowen Wang, PhD[a], Hongbing Luo, BS[a], Enlong Yao, BS[b], Renhai Tang, BS[c], Wenbing Dong, BS[d], Fuyong Liu, BS[e], Jun Liang, MA[f], Minyang Xiao, MA[a], Zuyang Zhang, MA[a], Jin Niu, MA[a], Lijun Song, MA[a], Liru Fu, MA[a], Xuehua Li, MA[a], Shicong Qian, BS[g], Qing Guo, BS[h], Zhizhong Song, PhD[a],* iD

Abstract

We compared the discriminative validity, agreement and sensitivity of EQ-5D-5L and SF-6D utility scores in people living with HIV/AIDS (PLWHIV). 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. The utility index of the SF-6D was derived from the SF-12. We calculate correlation coefficients to evaluate the relationship and agreement of 2 instruments. To evaluate the homogeneity of the EQ-5D-5L and SF-6D, intraclass correlation coefficients, scatter plots, and Bland–Altman plots were computed and drawn. We also used receiver operating characteristic curves to compare the discriminative properties and sensitivity of the econometric index. A total of 1797 respondents, with a mean age of 45.6 ± 11.7 years, was interviewed. The distribution of EQ-5D-5L scores skewed towards full health with a skewness of −3.316. The overall correlation between EQ-5D-5L and SF-6D index scores was 0.46 ($P < .001$). The association of the 2 scales appeared stronger at the upper end. An intraclass correlation coefficient 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 receiver operating characteristic 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. 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. There were significant differences in their performance. We preferred to apply the SF-6D to measure the health utility of PLWHIV during the combined antiretroviral therapy period. Our study has demonstrated evidence for instrument choice and preference measurements in PLWHIV under combined antiretroviral therapy.

Abbreviations: cART = combined antiretroviral therapy, HRQoL = health-related quality of life, ICCs = intraclass correlation coefficients, MCS = mental functional scores, PLWHIV = people living with HIV, QALY = quality of life year, ROC = receiver operating characteristic, SG = standard gamble, TTO = time trade-off, UI = utility index.

Keywords: combined antiretroviral therapy, EQ-5D-5L, health utility measurement, people living with HIV/AIDS, SF-6D

1. Introduction

Although the HIV epidemic in China is defined as the low national prevalence with 0.037%, approximately 0.98 million people living with HIV (PLWHIV) reside in China and 86.6% of them are accepting combination antiretroviral therapy (cART) by the end of 2019.[1] Yunnan province is one the regions in China reporting the highest prevalence. Yunnan has an estimated number of 0.11 million PLWHIV and cART is free of charge to all of them. By the progress to combat HIV prevalence, in 2020, Yunnan province has achieved the target of 3 90-90-90 on time, 90% of PLWHIV in Yunnan province has known their status, 90% of them has received cART, and 90% of those on cART had an undetectable viral load.[2]

Health-related quality of life (HRQoL) is an individual's or a group's perceived physical and mental health with multiple factors associated with it.[3] Attributed to the access and success of cART, HIV infection has been transformed into a chronic and manageable condition.[4] It's estimated that the expectance life of PLWHIV at their 35 years old in Shanghai, China has research more than 40 years.[5] The quality of life for PLWHIV has become a significant area of AIDS research. Patient-level HRQoL is important for the PLWHIV to monitor and support the intervention programs designing. Preference-based measure of health also called health utility, which generated a utility weight to calculate quality-adjusted life years (QALYs),[6] which is widely used in economic evaluations for the resource allocation.

Many instruments for the HRQoL measurements are increasingly used in the surveys for the PLWHIV, including generic and disease-specific instruments. Generic instruments generally could be classified as psychometric profile measures and econometric index measures.[6] Psychometric measures are usually used to generate different health dimension scores (profile). Otherwise economic measures usually need to provide a single globe index, which finally acquired a preference for health status (utility) to calculate QALYs for the cost-effectiveness analysis. Despite the availability of different instruments, these instruments have different characteristics so as to different ability to detect the disease conditions on overall health.[7] Not surprisingly, the results of disease burden calculation and cost-effectiveness analysis might be dependent on the measure that we used. So far, by the development and pluralizing of the HIV/AIDS prevention strategies, policy making for the public health agencies based on the evaluation results to make the valid resource allocation has become an ongoing and great challenge. However, different instruments could form different measurements.

The EQ-5D (EuroQol 5-Dimension) has been widely used in the population survey because of the low respondent burden. It defined the total health by 5 dimensions: mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression.[8] The primary version of EQ-5D allows respondents indicating the degree of impairment on each dimension by 3 levels: no problems, some problems and extreme problems. A new version of EQ-5D-5L includes 5 levels to indicate the degree of severity for each dimension: no problems, slight problems, moderate problems, severe problems, and extreme problems.[9] Whatever EQ-5D-3L or EQ-5D-5L, it could generate a preference-based index. The SF-12 is the abbreviate version of SF-36 (the Short Form-36 Health Survey) with 12 items, which finally provides 2 component summary scores related to physical and mental respectively. The SF-6D (the Short Form-6 Dimension) is an econometric index derived from the preference value system of SF-36 and then SF-12.[10] Considering that the SF-12 could generate both profile and index, it may provide a useful alternative to the EQ-5D.

Several studies have compared the EQ-5D with SF-12 or SF-6D in the usual population and different disease groups.[6,7,11–14] Some studies reported that the different instruments generated widely differing HRQoL scores for the same patients groups.

Some studies supported the usage of SF-6D, which had the lower floor and ceiling effect and could better detect the different stages of the disease. Some studies considered that EQ-5D could be recommended using in the severe conditions and SF-6D using in the mild conditions. Nevertheless one review study recommended that the EQ-5D had a stronger discriminative power in the HRQoL measurement for the patients with pain. In all, these studies highlighted the variation in the results generated from the different instruments, suggesting the combination of EQ-5D and SF-12, which could provide a relatively broad coverage for various purposes. Otherwise one study found that the SF-6D derived from SF-36 was more discriminative than that derived from SF-12.[15] So the weight placed on health utility for economic evaluation purpose depended on the selection of a preferred instrument.

HRQoL has been believed to be a dynamic and relative concept in the HIV/AIDS prevention under the cART era. Stigma remained a significant challenge for the most PLWHIV, social circumstance may influence the daily life negatively.[16,17] Otherwise the scale of HIV/AIDS problems and limited resources for prevention make the resource allocation critical for the decision making. The preferred instrument chose is argent needed for the more accurate results. To be known to ourselves, few studies available have compared EQ-5D with SF-12 profile and SF-6D index in terms of their power to distinguish health status between groups defined by socioeconomic characteristics, CD4 counts, antiviral therapy modes, psychometric levels, and social supports in the PLWHIV as a whole.

The aim of this study was to evaluate the accuracy and applicability of EQ-5D-5L, SF-6D, and SF-12 version 2 for use in the PLWHIV management by comparing the discriminability and finally form the health utility weight, which may be useful for the resource allocation to prevent HIV/AIDS epidemic.

2. Methods

2.1. Study design

We conducted a cross-sectional survey among PLWHIV aged more than 18 years old in the 10 municipals in Yunnan province from October, 2019 to May, 2020. A convenience sample including 1797 participants was enrolled. The investigators with the strict training from local CDCs and social organizations implemented the investigation face to face. The study was approved by the ethics research committee of Yunnan Centers for Disease Control and Prevention, China(No.YNCDC/QR-KJB-2021-003).

2.2. HRQoL assessment

We administrated SF-12 (12-item Short Form Health Survey), which is the shortened version of SF-36 (36-item Short Form Health Survey) and could explain at least 90% accuracy of SF-36.[8] The SF-12 consists of eight domains to generate finally 2 separate summary scores, physical functional scores (PCS) and mental functional scores (MCS) ranging from 0 to 100. The Cronbach's α was 0.434.

We also administrated EQ-5D-5L (EuroQoL 5-dimensions) as the measurement simultaneously. The EQ-5D-5L comprised 2 components: utility index (UI) and the EQ-VAS. We calculated UI by the respondents scoring from the 5 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." All the responses were combined to form a 5-digital number to describe the health status. All of these were converted to UI based on the EQ-5D-5L value set for Chinese (see Table 1). UI ranged from (the worst possible health status) to 1 (the best possible health status).[18] The Cronbach's α was 0.813.

Table 1

Chinese value set of 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	-

2.3. Covariate data collection

The covariate data in the study included 3 part, demographic component, clinical component, and social–psychology component. All the participants completed a demographic questionnaire designed by the study's staff. Social support was assessed by Social Support Rating Scale designed by Xiao Shuiyuan in 1986 primarily for Chinese population.[19] The reliability is with a Cronbach's α of 0.684. The anxiety and depression was assessed by Chinese version of the Hospital Anxiety and Depression Scale.[19] The reliability of Hospital Anxiety and Depression Scale in our study is with a Cronbach's Cronbach's α of 0.237.

3. Statistical analysis

3.1. EQ-5D-5L scoring

The EQ-5D-5L could define the 3125 possible health states by the different combinations.[18] We adept the Chinese population-based preference trade-off time to transform the measures into UI (Table 1), thereby producing a single preference-based index ranging from –0.391 to 1.000. Scores for the 2 component summaries (physical and mental component summaries, PCS-12 and MSC-12) were calculated using the 2nd edition of standard US instrument scoring algorithms (SF-12v2). The details could be reviewed in this professional manuscript.[8]

3.2. SF-6D scoring

The UI of SF-6D can be derived from the SF-36 or SF-12 although the number of items included differed, 11 items and 7 items, respectively. We used the value set from SF-12 and Consider the value set from UK general population developed by Brazier et al (Fig. 1 and Table 2, which demonstrated the value set and the transfer progress of the SF-6D).[20] UI of SF-6D could be calculated as U = 1 + coefficient of different dimension + adjustment coefficient (Most = –0.085). If any dimension chose the most serious level, the final result should subtract the Most.[21]

3.3. Data analysis

We described the sample characteristics by calculating the number of individuals and the percentage in each category group. We also computed the descriptive statistics, including mean, standard deviation (SD), 95% confidence intervals (CIs), median, interquartile range (IQR), minimum and maximum for EQ-5D-5L, and SF-5D index scores. Floor effects means the proportion responding "extreme problem" on each of the dimensions (self-reported health status "55555" for EQ-5D-5L and 345555 for the SF-6D) and ceiling effects means the proportion responding "no problem" on each of the dimensions (self-reported health status "11111" for EQ-5D-5L and 111111 for the SF-6D). If the distribution of index scores were highly skewed, differences between EQ-5D-5L and SF-6D index scores were examined by Wilcoxon signed-rank test. Mann–Whitney U test was used to compare index scores across participants' characteristic for 2 groups and Kruskal–Wallis test for more than 2 groups.

We calculate correlation coefficients to evaluate the relationship and agreement of 2 instruments. To evaluate the homogeneity of EQ-5D-3L and SF-6D, intraclass correlation coefficient (ICC) and Bland–Altman plot was computed and drawn. $ICC = 1$ meant the completely correlation. $0.7 \leq ICC \leq 0.9$ meant the strongly correlation. $0.4 \leq ICC \leq 0.69$ meant the moderately correlation.[22] $0.1 \leq ICC \leq 0.39$ meant the slightly correlation. $ICC = 0$ meant no correlation. Bland-Altman plot was a method that was developed to compare 2 measurements for the same variable. Generally the plots that located in the 95% interval of limits of agreement interval should occupy 95% and simultaneously could not exceed the professional scope.[23]

We defined the relevant difference of the instrument as showing different effect size for the same group category, which also could be explained the disagreement on the amount of health burden. In the study, the discriminative properties of the econometric index also compared using receiver operating characteristic (ROC) curves. We used the SF-12 component summaries as the external indicators to analyze the performance of EQ-5D-5L and SF-12. We set the external indicator as the dichotomized variables using the median cutoff points of PCS-12 and MCS-12. The largest area under the ROC curve (AUC) for the utility measurement demonstrated the most sensitive to detect differences in the external indicators. F-ratio of the significance test for the AUC was referenced to 1.0 for the EQ-5D-5L index. If a value was over 1.0, we would consider that the SF-6D index was efficient than the EQ-5D-5L at detecting differences between this group category.

All the data analysis used STATA version 14.0 (StataCorp LLC, College Station, TX).

4. Results

4.1. Research field and subjects

A total number of 1797 respondents were interviewed, with a mean age of 45.6 ± 11.7 years, ranging from 16 to 80. 68.1% respondents were of Han nationality, while others were from minority ethnic groups, including Yi, Dai, Zhuang, Jinpo, Lisu, and Bai. 58.7% respondents declared themselves divorced or separated. 69.5% respondents had the less than 9 years compulsory education. 53.6% respondents' occupation was farmers and migrant workers. The average yearly income per capital of their households was 10,871yuan in 2020. As for HIV-related characteristics, more than two-thirds of respondents were in HIV stage (70.4%), 28.3% were in AIDS stage and only 1.2% were unclear when they were diagnosed the first time. A large proportion of the sample (68.8%) were those patients with heterosexual transmission, 20.3% reported who had the history of intravenous drug use. There were 98.7% patients taking ART, among those, 59.2% have been treated for more than 4
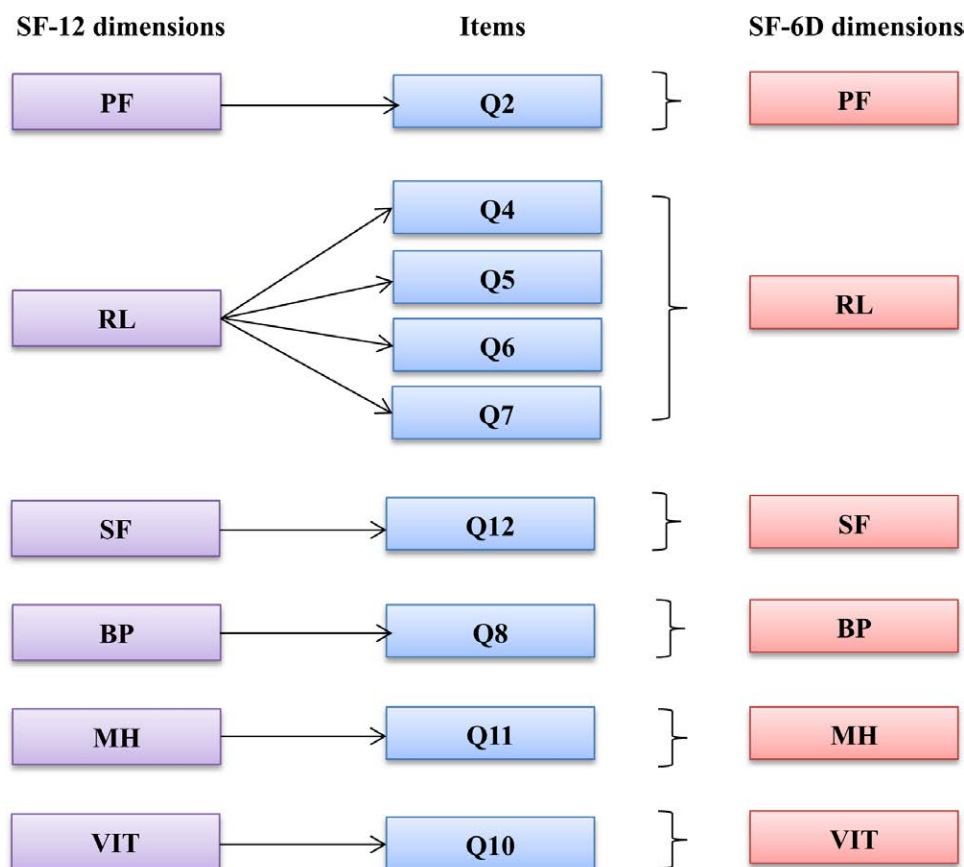
Figure 1. The relationship between SF-12 and SF-6D.

years. The majority of patients remained high CD4 cell counts, 48.5% had more than 500 cells/µL. Table 3 showed the details of the social-demographic and clinical characteristic of all the respondents.

4.2. Descriptive statistics of 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 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 floor and ceiling effects of 0.1% (n = 2) and 33.0% (n = 593) respectively (see Fig. 2). The mean SF-6D index score was 0.772 ± 0.137 (median 0.762, IQR 0.241). The distribution of EQ-5D-5L almost centered around its mean, the skewness was 0.084, with a range from 0.374 to 1.000. The percentage of respondents with the floor and ceiling effects of 0.06% (n = 1) and 6.5% (n = 116), respectively (see Supplemental Content Fig. 2b). The mean SF-6D index score for respondents with the best health state on 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 only one respondent with the worst health state (111515) was 0.364. In the whole, the mean EQ-5D-5L index scores exceeded the mean SF-6D index scores by 0.124, the difference between median was 0.22. The difference between EQ-5D-5L and SF-6D index scores was significant for the entire sample and for some examined sociodemographic and infectious statue's subgroups (Table 4).

Both EQ-5D-5L and SF-6D index scores were significantly difference across groups of age, race/ethnicity, education level, occupation, household income per year, transmission model,

and duration of ART. EQ-5D-5L index scores were significantly difference across initial infectious status. SF-6D index scores were significantly difference across the most recent CD4counts.

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

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

4.4. Relationship between EQ-5D-5L and SF-6D

The overall correlation between EQ-5D-5L and SF-6D index scores was observed as 0.46 ($P < .001$). The association of the 2 scales appeared stronger at the upper end. We also observed a degree of dispersion that the very low EQ-5D-5L scores were associated with the very high scores on the SF-6D. And

Table 2

UK value set of SF-6D and SF-12 health status.

Dimension	SF-12 items and options	SF-6D dimension	e
PF	Q2 = 3	PF1	0
	Q2 = 2	PF2	-0.012
	Q2 = 1	PF3	-0.040
RL	Q4 or Q7 = 5	RL1	0
	Q5 = (1 or 2 or 3 or 4)	RL2	-0.068
	Q6 = (1 or 2 or 3 or 4)	RL3	-0.061
	Q5 and Q6 = (1 or 2 or 3 or 4)	RL4	-0.054
SF	Q12 = 5	SF1	0
	Q12 = 4	SF2	-0.059
	Q12 = 3	SF3	-0.069
	Q12 = 2	SF2	-0.078
	Q12 = 1	SF1	-0.093
BP	Q8 = 1	BP1	0
	Q8 = 2	BP2	-0.004
	Q8 = 3	BP3	-0.039
	Q8 = 4	BP4	-0.076
	Q8 = 5	BP5	-0.140
MH	Q11 = 5	MH1	0
	Q11 = 4	MH2	-0.004
	Q11 = 3	MH3	-0.039
	Q11 = 2	MH4	-0.076
	Q11 = 1	MH5	-0.140
VIT	Q10 = 1	VIT1	0
	Q10 = 2	VIT2	-0.097
	Q10 = 3	VIT3	-0.065
	Q10 = 4	VIT4	-0.059
	Q10 = 5	VIT5	-0.103

Table 3

Characteristics of study sample.

Characteristic	Distribution(%)	
Age (yr)	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/cohabitating	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)	<5000 yuan	592 (32.9)
	5000 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 (yr)	≤1	257 (14.3)
	1–2	148 (8.2)
	2–4	289 (16.0)
	≥4	1078 (59.9)
	Not yet	25 (1.3)
The most recent CD4 counts	CD4 ≥ 500 cells/μL	876 (48.7)
	350–500 cells/μL	359 (19.9)
	200–350	299 (16.6)
	<200	158 (8.7)
	Unclear	105 (5.8)

conversely, the very high EQ-5D-5L index scores were associated with the very low scores on the SF-6D (Fig. 3). An ICC of 0.59 between EQ-5D-5L and SF-6D meant a moderately correlation and indicated a good agreement. The Bland–Altman plot displayed the same results to the scatter plot with a mean difference between the EQ-5D-5L and SF-6D index scores of 0.124. 3% 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 liner trend, which meant the more obvious difference between EQ-5D-5L and SF-6D index scores existed in the observations with a good or a weak health status, while the observation with a general health status, the homogeneity between the 2 scales seemed good (Fig. 4).

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

We set the PCS-12 and MCS-12 as the golden standard to measure the health status respectively and use the median of the PCS-12, MCS-12, EQ-5D-5L, and SF-6D to divide the health status by 2 categories. The ROC listed as following (Figs. 5 and 6), which showed that the AUC of SF-6D was 0.776 (95%CI: 0.757,0.796) and
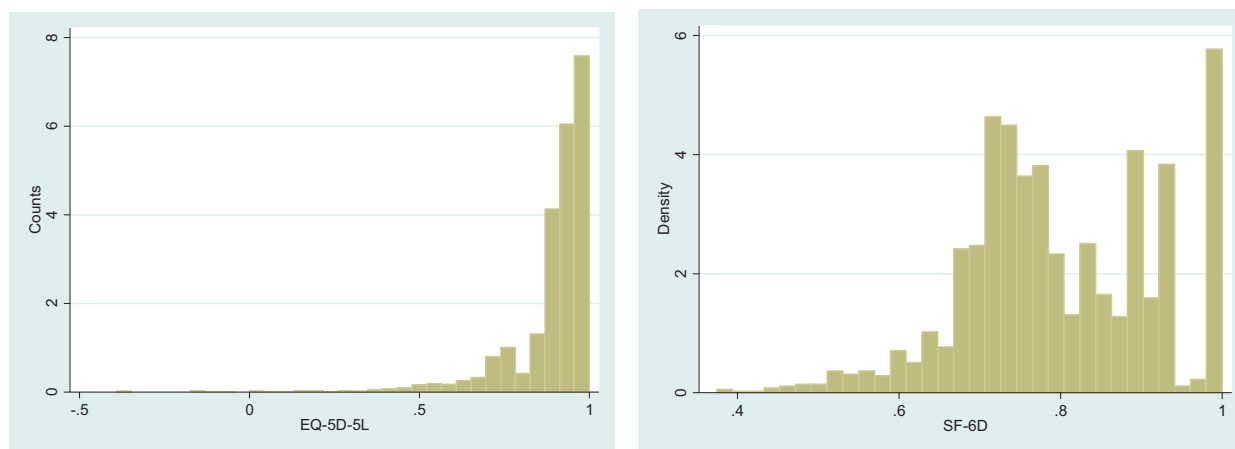
Figure 2. Distribution of EQ-5D-5L and SF-6D (a. EQ-5D-5L, b. SF-6D).

EQ-5D-5L was 0.732 (95%CI: 0.712, 0.752), respectively, by PCS-12 and 0.782 (95%CI:0.763, 0.802) for SF-6D and 0.690 (95% CI:0.669, 0.711) for EQ-5D-5L, respectively, by MCS-12. Both the difference of AUC for the 2 groups was significant ($P < .05$). Both the AUC of SF-6D and EQ-5D-5L were more than 0.5, F-ratio was 1.06 and 1.13, respectively, with EQ-5D-5L as the reference, which revealed the good value to discriminate the health status and SF-6D seemed more sensitivity than EQ-5D-5L to discriminate the health status defined by SF-12.

5. Discussion

Our study demonstrated the evidence of the performance of EQ-5D-5L and SF-6D index scores to measure the health utility in people living with HIV/AIDS, which showed the moderate correlation between the 2 measurements,[24] and both showed the discriminative capacity and validity to measure the health status for the people living with HIV. However, some considerable overlaps existed in the 2 measurements, there were significant differences in their performance, which were in accordance with the results reported in the previous studies about the preferences of EQ-5D and SF-6D in the general population and several patient groups,[25] such as diabetes, heart diseases, stoke, and chronic pain.[17,24,26,27]

In our study, for the mean and the median of EQ-5D-5L and SF-6D to measure the same samples of people living with HIV/AIDS, the EQ-5D-5L values were higher than the SF-6D values whatever in the whole samples or across all the subgroups, with a mean difference of 0.124 and a median difference of 0.180, which was consistence with some studies that proved the value difference. One study carried out across 7 patient groups found that the mean SF-6D value exceeded the EQ-5D while the median EQ-5D value exceeding the SF-6D value.[8] Whatever it implied us the importance to concern the reasons for these differences. For the interventions to help to restore the full health, the size of the difference may not impact greatly on the QALY estimation. But for the interventions only to help to achieve partial relief, the range of differences could significantly influence the estimation of QALY gains.

The scoring system of EQ-5D-5L was based on the time-trade-off.[11] Otherwise, the SF-6D was based on the standard gamble technique.[11] The preference of the source population was also a possible reason for the difference. EQ-5D-5L values derived from Chinese preferences while SF-6D values derived from US preferences.[10,18]

The ICC was 0.59, which meant a moderate correlation.[22] We could consider an acceptable level but not very good level of agreement for the 2 measurements, especially at the more

serious and the more slight ends of the scales. Theoretically, the range of EQ-5D-5L was much wider. In fact, the distribution of SF-6D index scores approached more to the normal while EQ-5D-5L index scores skewed negatively. SF-6D demonstrated smaller ceiling effect compared to EQ-5D-5L. One study carried out in the breast cancer population found that the RL dimension of SF-6Dv2 has the highest ceiling effect (49.52%).[28] Another study carried out also in the colorectal and breast cancer population in developing counties, which demonstrated that the ceiling effect of SF-6D-v2 was 0.45% and of EQ-5D-5L was 5.84%.[29] The ceiling effect of our study was higher than the measurement in breast cancer patients. It's due to the different seriousness of the disease types. We could consider that PLWHIV under ART have got a better health condition. Otherwise it seemed that SF-6D performed better than EQ-5D-5L in terms of ceiling effects.

Both EQ-5D-5L and SF-6D performed better to monitor the changes in the social and psychology aspects than in the physical aspects for people living with HIV/AIDS.SF-6D appeared to detect more changes and had larger effect sizes compared to EQ-5D-5L. The result is somewhat surprising that the richer descriptive system of SF-6D might be easier to find the changes in the psychology aspects, which was often smaller and more unnoticeable than the physical aspects. Based on the ROC and AUC, both the 2 measures revealed the good value to discriminate the health status, SF-6D also seemed more sensitivity to discriminate the health status. One previous study demonstrated that the difference of SE were driven by the smaller SD of SF-6D inherently, which was a consequence of the narrower range of the index scores.[30]

Some previous studies have explored the reasons to lead to the differences between EQ-5D and SF-6D to measure the health utilities.[31–34] The main reason lied in the discrepancies of the descriptive systems' contents. For the same sample population, all of them should complete the 2 measures simultaneously, whereby their health status would be described by EQ-5D-5L, which included 5 contents of mobility, self-care, usual activities, pain/discomfort and anxiety/depression, while the contents of physical functioning, RLs, bodily pain, vitality, social functioning, RL, and mental health from SF-6D. Different descriptive contents defined the application and appropriateness. EQ-5D-5L emphasized more on the physical aspect of the health, while SF-6D emphasized more on the mental health and the social adaption. By the great improving of the survival for the people living with HIV/AIDS under the period of combined antiretroviral therapy, HIV/AIDS has transformed from terminal illness to "chronic disease." A rising challenge for this population was the full health, which made more consideration to the mental health and the family and society rehabilitation

Table 4

Comparison of EQ-5D-5L and SF-6D index scores 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 (yr)*				
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 Marital status	0.896 (0.150)	0.942 (0.114)	0.772 (0.138)	0.762 (0.241)
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)
5000 to 10,000yuan	0.897 (0.150)	0.942 (0.114)	0.772 (0.138)	0.762 (0.241)
10,000 to 50,000 yuan	0.896 (0.150)	0.942 (0.115)	0.772 (0.138)	0.762 (0.241)
≥ 50,000 yuan	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* (yr)				
≤ 1	0.896 (0.150)	0.942 (0.115)	0.772 (0.138)	0.761 (0.241)
1–2	0.896 (0.150)	0.942 (0.115)	0.772 (0.138)	0.759 (0.241)
2–4	0.896 (0.150)	0.942 (0.115)	0.771 (0.137)	0.759 (0.241)
≥ 4	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 CD4 counts				
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 < .001.
**P < .05.

so as to the role playing. On the perspective, our study preferred to the application of SF-6D to measure the health utility of people living with HIV/AIDS. We also considered the various scoring algorithms contributed to the discrepancy of the 2 measures. For the same health status, the different scoring algorithms assigned different index scores, the worst health status measured by EQ-5D-5L was assigned –0.391 (the worse than the death), while the SF-6D index score was 0.331. These variations resulted in the different descriptive contents and the different theories of scoring system to choose the preference.

Table 5

Mean(SD) SF-12 component scores divided by different EQ-5D-5L dimensions.

EQ-5D-5LDimension	Level	n	PCS-12			MCS-12		
			$\bar{X} \pm S$	P	η-sq*	$\bar{X} \pm S$	P	η-sq*
Mobility	1	1487	46.42 ± 8.67	.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	.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	.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	.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	.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		

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

Table 6

Mean (SD) SF-12component scores divided by different EQ-6D dimensions.

SF-6D Dimension	Level	n	PCS-12			MCS-12		
			$\bar{X} \pm S$	P	η-sq*	$\bar{X} \pm S$	P	η-sq*
Physical function	1	105	46.40 ± 8.69	.001	0.61	48.15 ± 9.63	.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	.001	0.04	48.13 ± 9.64	.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	.001	0.25	47.78 ± 9.60	.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	.001	0.54	48.13 ± 9.64	.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	.001	0.08	48.08 ± 9.63	.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	.001	0.31	48.13 ± 9.64	.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		

MCS = mental functional scores, PCS = physical functional scores.

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

Based on these factors, we consider that the 2 measures could not have the interchangeability, which meant that we could not acquire the results of the EQ-5D-5L based on the SF-6D and vice versa. So we should take into consideration in the selection between the 2 measures. User needed to pay more attention to the characteristics of the target population. We had some
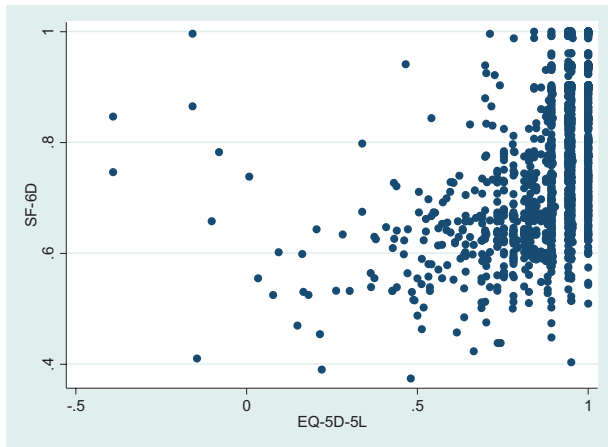
Figure 3. Scatter plot of EQ-5D-5L and SF-6D index scores.

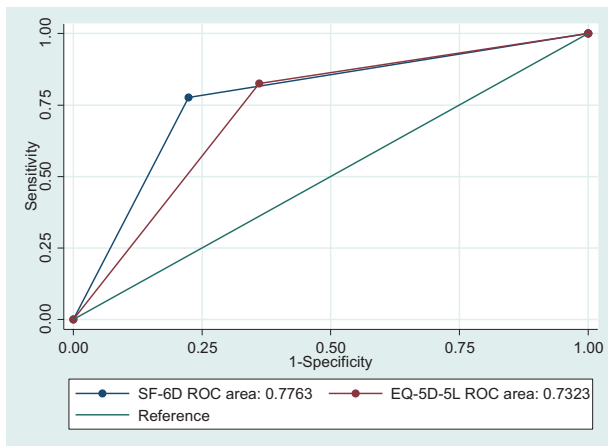**Figure 4.** Bland–Altman plot of EQ-5D-5L and SF-6D index scores.

Figure 5. ROC of EQ-5D-5L and SF-6D by PCS-12. PCS = physical functional scores.

Figure 6. ROC of EQ-5D-5L and SF-6D by MCS-12. MCS = mental functional scores.

breast cancer with early stage, and patients in the controlled disease period. Otherwise, for the patient population with a great influence by the physical health, we suggested the selection of EQ-5D-5L, such as the disease with the losing capacity, patients in the advanced disease period. Third, we should also consider the availability of the scoring algorithm, the population origin of value set, the change extent of the health status and the resource allocation when we used for the cost–utility analysis to inform the local decisions.

There were some limitations in our study. First, the results limited to our sample population of people living with HIV/AIDS who had a good therapy of ART, and thus these results may not be generalized to all people living with HIV/AIDS, including patients with failure ART. Second, we used SF-12 as the golden standard to establish the comparison and the results of SF-6D derived from the SF-12, which could generate bias for the results. Some studies have found the different performances between the SF-6D-v2 and SF-6D-v1$_{SF-36}$.[35] The differences between EQ-5D-5L and SF-6D-v2 were smaller than those for EQ-5D-5L and SF-6D-v1$_{SF-36}$ in breast cancer patient. So the difference between the SF-6D and EQ-5D-5L may be over-estimated when used the SF-6D derived from SF-36 or SF-12. We should further explore the difference between SF-6D-v2 and SF-6D-v1$_{SF-36}$ measured in PLWHIV. Third, we constructed a cross-sectional study and could not capture the responsiveness of the 2 measures. Fourth, the depressive and anxiety symptoms were measured based self-report, which could over- or under-estimate these symptoms.

Acknowledgments

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

Author contributions

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.

Data curation: Xiaowen Wang, Hongbing Luo.

Formal analysis: Xiaowen Wang.

Investigation: Hongbing Luo, Enlong Yao, Renhai Tang, Wenbing Dong, Fuyong Liu, Jun Liang, Minyang Xiao,

principles for the selection. First, for the general population or a gentle patient population with generally good health, EQ-5D-5L and SF-6D were likely to perform similarly, but for the sicker population, the performance of the 2 measures seemed different. Second, for the patient population with the disease for the mental health impacted greatly and the physical health impacted gentle or little, the sample population could keep a good physical health, we suggested the selection of SF-6D, such as, HIV/AIDS,
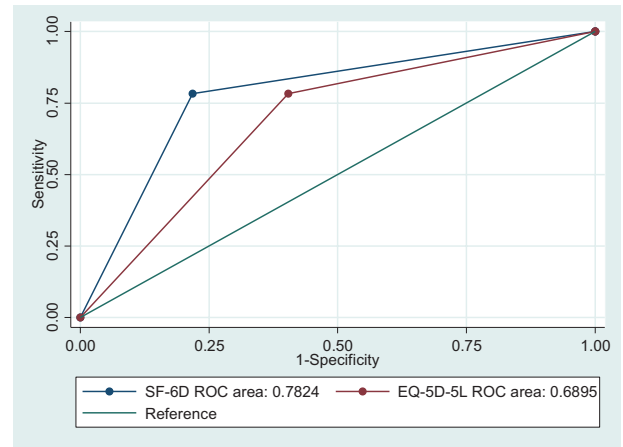
Zuyang Zhang, Jin Niu, Lijun Song, Liru Fu, Xuehua Li, Shicong Qian, Qing Guo.

Methodology: Xiaowen Wang, Hongbing Luo, Enlong Yao, Renhai Tang, Wenbing Dong, Jun Liang, Zuyang Zhang, Jin Niu, Liru Fu, Shicong Qian, Zhizhong Song.

Supervision: Zhizhong Song.

Validation: Zhizhong Song.

Writing—original draft: Xiaowen Wang.

Writing—review and editing: Xiaowen Wang.

References

[1] Yogesh SP, Guo N, Gabriel T, et al. Improving access to antiretrovirals in China: economic analyses of dolutegravir in HIV-1 patients. Cost Eff Resour Alloc. 2019;17:26–34.

[2] Xinhua.net. Yunnan has achieved the goal of three 90% for HIV prevention. Available at: https://baijiahao.baidu.com/s?id=1684890263547198748&wfr=spider&for=pc [access date January 19, 2021].

[3] Yunnan CN. Yunnan held the briefing of AIDS epidemic and prevention progress [EB/OL]. Available at: http://m.yunnan.cn/system/2020/12/01/031153259.shtml [access date January 19, 2021].

[4] Morad G, Noureddine E, Aicha B, et al. Assessment of quality of life in children, adolescents, and adults with celiac disease through specific questionnaires: review. Nutr 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. Health Qual Life Outcomes. 2012;10:132–9.

[6] Sun GM, Gu KK, Zhiy H. Life expectancy of 917 cases with HIV/AIDS in Shanghai. 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]. Available at: http://yn.yunnan.cn/system/2020/12/29/031207628.shtml [access date January 19, 2021].

[8] John B, Jennifer R, Aki T, et al. A comparison of the EQ-5D and SF-6D across seven patients groups. Health Econ. 2004;13:873–84.

[9] Yuan SH, Jennifer T, Sarah AW, et al. Mapping the medical outcomes study HIV health survey (MOS-HIV) to the EuroQoL 5 dimension(EQ-5D-3L) utility index. Health Qual 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. Health Qual Life Outcomes. 2017;15:158–69.

[11] Kanters TA, Redekop WK, Kruijshaar ME, et al. Comparison of EQ-5D and SF-6D utilities in Pompe disease. 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. 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. Health Econ Res. 2020;37:42–5.

[14] Vinicius CV, Liliane L, Viviane AS, et al. Oral health and health-related quality of life in HIV patients. 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. Massachusette: The Health Institute, New England Medical Center. 1995. 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. Int J Rheum Dis 2017;20:1987–92.

[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. BMJ Open. 2020;10:e035722.

[18] Luo N, Liu G, Li MH, et al. Estimating an EQ-5D-5L value set for China. Value 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. 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. 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. PLoS One. 2011;6:e16049.

[22] Hongmei Y, Yanhong L, Jian S, et al. Intraclass correlation coefficient and software realization. Chin J Health Stat. 2011;28:497–500.

[23] John EB, Jennifer R. The estimation of a preference-based measure of health from the SF-12. Med Care. 2004;42:851–9.

[24] Yu HM, Luo YH, Sa J, et al. The intraclass correlation coefficient and software calculation. Chin J Health Stat. 2011;28:497–500.

[25] 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. Chin Gen Pract. 2017;20:3272–6.

[26] Du XD, Zhu P, Li M, et al. Health utility of patients with stroke measured by EQ-5D and SF-6D. J Sichuan Univ (Med Sci Edi). 2018;49:252–7.

[27] 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. Qual Life Res. 2017;26:2057–66.

[28] Azin N, Hossein S, Hosein A. Psychometric properties of the SF-6Dv2 in an Iranian breast cancer population. Breast Cancer. 2021;28:937–43.

[29] Mahmood Y, Azin N, Ali AS, et al. Mapping QLQ-C30 onto EQ-5D-5L and SF-6D-V2 in patients with colorectal and breast cancer from a developing country. Value Health Reg Issues. 2021;24:57–66.

[30] 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. Health Qual Life Outcomes. 2017;15:143–51.

[31] Jin H, Wang B, Gao Q, et al. Comparison between EQ-5D and SF-6D utility in rural residents of Jiangsu province, China. PLoS One. 2012;7:e41550.

[32] Lamers LM, Bouwmans CAM, Van Straten A, et al. Comparison of EQ-5D and SF-6D utilities in mental health. Health Econ. 2006;15:1229–36.

[33] Mahmood Y, Safa N, Sharam G, et al. Comparison of SF-6D and EQ-5D scores in patients with breast cancer. Iran Red Crescent Med J. 2016;18:e23556.

[34] Brazier J, Robert J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002;21:271–29.

[35] Azin N, Hossein S, Hosein A. Comparing the performance of the EQ-5D-5L with two versions of the SF-6D-v2 in patients with breast cancer. Health Service Outcomes Res Methodol. 2020;20:183–94.