A Comparative Study of CQ-11D, SF-6Dv1, and EQ-5D-3L Outcomes in a Chronic Disease Population: Postprint

Authors: Wang Wei, Xie Shitong, Zhou Jiameng, Pan Jie, Zhu Wentao, Zhu Wentao

Date: 2022-11-23T00:00:00+00:00

Abstract

Background The rising prevalence of chronic diseases in China has made health-related quality of life a critical indicator for evaluating prevention and treatment outcomes, underscoring the importance of selecting appropriate measurement instruments for health-related quality of life in chronic disease patients.

Objective To compare the scores of three generic health utility instruments—the Chinese Quality of Life Instrument (CQ-11D), the Short Form Six-Dimension version 1 (SF-6Dv1), and the EuroQol Five-Dimension Three-Level (EQ-5D-3L)—in chronic disease patients, and to assess the validity, consistency, and correlation of these three instruments in measuring health utility values in chronic disease patients.

Methods Using quota sampling based on the gender and age distribution of the Chinese population, face-to-face interviews were conducted to collect basic demographic information and sequential self-reported responses to the three instruments (CQ-11D, SF-6Dv1, and EQ-5D-3L). Chronic disease patients were identified from the sample for analysis of their utility value measurements. Bland-Altman plots and Intraclass Correlation Coefficient (ICC) were employed to assess consistency and correlation among the three instruments, while Kruskal-Wallis tests were used for univariate analysis of gender and age.

Results A total of 692 chronic disease patients were included in the study. The mean (standard deviation) utility values for CQ-11D, SF-6Dv1, and EQ-5D-3L were 0.855 (0.168), 0.793 (0.132), and 0.876 (0.136), respectively. The three most prevalent chronic conditions were hypertension (N=275), arthritis (N=128), and hyperlipidemia (N=124). Histograms revealed a pronounced ceiling effect for EQ-5D-3L, while CQ-11D showed a broader distribution range. Bland-Altman plots demonstrated relatively good agreement among the three

instruments, with approximately 95% of points falling within the 95% limits of agreement (LOA). CQ-11D showed the highest correlation with the Visual Analogue Scale (VAS). The ICC range among the three instruments was 0.528-0.625, indicating moderate correlation; both Bland-Altman plots and ICC analysis indicated that SF-6Dv1 and EQ-5D-3L had the best agreement. Subgroup analyses of the three different chronic conditions yielded similar results to those of the overall sample. Univariate analysis revealed that male patients had higher utility values than female patients among those with hypertension and hyperlipidemia; among hyperlipidemia patients, those aged 60 years and above had the lowest utility values.

Conclusion In chronic disease populations, the EQ-5D-3L instrument exhibited a pronounced ceiling effect. CQ-11D demonstrated a broader distribution range of utility values for the three chronic conditions without observable ceiling effects and showed the highest correlation with VAS, making it relatively more suitable for measuring health-related quality of life in Chinese chronic disease populations.

Full Text

Comparison of the CQ-11D, SF-6Dv1, and EQ-5D-3L Scales in a Chronic Disease Population

WANG Wei[12], **XIE Shitong**[34], **ZHOU Jiameng**[12], **PAN Jie**[12], **ZHU Wentao**[12]*

[1]School of Management, Beijing University of Chinese Medicine, Beijing 100029, China
[2]Institute of Pharmacoeconomic Evaluation of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100029, China
[3]Department of Health Evidence, Methods and Effects, McMaster University, Hamilton L8S4L8, Canada
[4]College of Pharmaceutical Science and Technology, Tianjin University, Tianjin 300072, China

Corresponding author: ZHU Wentao, professor, doctoral supervisor; E-mail: wentao67@126.com

Abstract

Background The number of patients with chronic diseases is increasing annually in China. Health-related quality of life (HRQoL) represents an important indicator for evaluating the effectiveness of chronic disease prevention and treatment, making the selection of appropriate HRQoL measurement tools for chronic disease patients highly significant.

Objective This study compares contemporaneous utility scores from three generic health utility scales—the Chinese Quality of Life Assessment Scale of

Traditional Chinese Medicine (CQ-11D), the Six-Dimensional Health Survey Short Form version 1 (SF-6Dv1), and the Three-Level Five-Dimensional Health Scale (EQ-5D-3L)—in chronic disease patients, and evaluates the validity, consistency, and correlation of these three instruments for measuring health utility values.

Methods Quota sampling was conducted according to the sex and age distribution of the Chinese population. Basic respondent information was collected through face-to-face surveys, followed by sequential collection of self-reported results from the CQ-11D, SF-6Dv1, and EQ-5D-3L scales. Patients with chronic diseases were selected from this sample, and their utility values from the three scales were analyzed. Bland-Altman plots and Intraclass Correlation Coefficients (ICC) were used to analyze consistency and correlation among the three scales, while Kruskal-Wallis tests performed univariate analyses by sex and age.

Results A total of 692 chronic disease patients were included. The mean (standard deviation) utility values were 0.855 (0.168) for CQ-11D, 0.793 (0.132) for SF-6Dv1, and 0.876 (0.136) for EQ-5D-3L. The three most prevalent chronic conditions were hypertension (N=275), arthritis (N=128), and hyperlipidemia (N=124). Histograms revealed a pronounced ceiling effect for EQ-5D-3L, while CQ-11D showed a broader distribution range. Bland-Altman plots demonstrated relatively good consistency among the three scales, with approximately 95% of points falling within the 95% limits of agreement (LOA). CQ-11D showed the highest correlation with Visual Analogue Scale (VAS) scores. ICC values among the three scales ranged from 0.528 to 0.625, indicating moderate correlation, with SF-6Dv1 and EQ-5D-3L demonstrating the best consistency. Subgroup analyses for the three chronic diseases yielded similar results to the overall sample. Univariate analysis showed that among hypertension and hyperlipidemia patients, males had higher utility values than females; among hyperlipidemia patients, those aged 60 and above had the lowest utility values.

Conclusion In chronic disease populations, the EQ-5D-3L exhibits a notable ceiling effect. The CQ-11D demonstrates a broader utility value distribution without ceiling effects and shows the highest correlation with VAS, making it relatively more suitable for measuring HRQoL in Chinese chronic disease populations.

Keywords Chronic disease; Health utility scores; CQ-11D; SF-6Dv1; EQ-5D-3L

Introduction

As population aging becomes increasingly prominent, chronic non-communicable diseases have gradually emerged as the primary factor affecting residents' health. Data from China' s Health Statistics Yearbook show that the prevalence of chronic diseases has risen substantially in recent years, increasing from 24.5% in

2013 to 34.3% in 2018. Chronic disease patients experience long disease courses with complex etiologies. In response to these characteristics, patient-reported outcomes (PROs) such as health-related quality of life (HRQoL) have become important complementary indicators for evaluating chronic disease prevention and treatment effectiveness alongside objective clinical measures.

Accurate measurement of health utility values is essential for calculating quality-adjusted life years (QALYs) and conducting pharmacoeconomic evaluations. Health utility values represent preferences for specific health states, typically ranging from 0 to 1, where 0 denotes death and 1 represents perfect health; negative values are permitted to indicate health states worse than death. Currently, the European Quality of Life Five-Dimensional Questionnaire (EQ-5D) and the Six-Dimensional Health Survey Short Form (SF-6D) are the most widely used health utility measurement scales globally, and both are recommended in the *China Guidelines for Pharmacoeconomic Evaluations (2020 Edition)*. However, recent studies indicate that the three-level version of EQ-5D (EQ-5D-3L) exhibits a ceiling effect when measuring HRQoL, showing poor sensitivity to minor health state changes, while the SF-6D demonstrates a floor effect in certain dimensions. More importantly, although both EQ-5D-3L and SF-6D have demonstrated good reliability and validity across various chronic disease populations, they were developed based on Western populations and may not accurately reflect health preferences and characteristics of Chinese populations.

The CQ-11D utility scale was recently developed based on Chinese population health preferences and grounded in Traditional Chinese Medicine (TCM) theory and health concepts, aiming to provide a more culturally appropriate and accurate measurement tool for Chinese populations. Given the substantial differences in dimensional structure between CQ-11D and the widely used SF-6Dv1 and EQ-5D-3L scales, measurement results may differ when applied to the same population. However, the existence, magnitude, and direction of these differences remain unclear. Therefore, comparing the utility value measurements from these three scales has important practical significance. This study simultaneously measured utility values in Chinese chronic disease patients using CQ-11D, SF-6Dv1, and EQ-5D-3L, compared their distributions, consistency, and correlation, and examined differences through univariate analysis to identify the most suitable scale for utility measurement in chronic disease populations, providing references for future HRQoL research and tools for pharmacoeconomic evaluation.

1.1 Study Population

Data for this study were derived from the research project constructing a Chinese population utility scoring system for CQ-11D, with data collection conducted from February to April 2021. During the initial phase of generating discrete choice experiment questionnaires, the Balanced Overlap method in Lighthouse Studio 9.9.2 software was used to extract balanced experimental samples from all possible health states generated by the CQ-11D scale. Based

on expert consultation, 700 pairs of health states were selected and allocated to 70 questionnaire sets. According to recommendations for discrete choice experiment research, each questionnaire set should have more than 20 respondents to ensure relatively reliable and robust model estimation. Therefore, the CQ-11D Chinese population utility scoring system construction study required a sample size greater than 1,400. Based on budget constraints, the effective sample size was expanded to 2,400, yielding approximately 34 respondents per discrete choice questionnaire to ensure adequate study quality.

The survey targeted Chinese adults aged 18 and above. Using quota sampling according to the sex and age distribution of the Chinese population from the *China Statistical Yearbook 2019*, the survey was conducted across seven geographic regions: North China, Northeast China, East China, Central China, Southwest China, Northwest China, and South China. Each region selected 2-6 representative provinces, autonomous regions, or municipalities directly under the central government (28 total), from which survey locations were chosen. Surveyors recruited respondents through convenience sampling in public areas (streets, communities, schools, etc.) within their jurisdiction. With respondents' consent, surveyors audio-recorded the interview process. This study was approved by the Ethics Committee of Beijing University of Chinese Medicine (Approval No. 2021BZYLL03012).

Inclusion criteria for survey participants were: (1) age $ $18 years; (2) Chinese citizens with Chinese nationality; (3) continuous residence in mainland China for the past five years; and (4) agreement to participate. Exclusion criteria were: (1) hearing, speaking, reading, or writing disabilities, or inability to understand survey content; and (2) abnormal mental status. The study population for this analysis comprised respondents who self-reported having physician-diagnosed chronic diseases, with chronic disease categories determined based on reference literature.

1.2 Survey Methods and Content

The survey employed face-to-face questionnaire administration. Surveyors collected sociodemographic information (sex, age, ethnicity, education level, household registration type, employment status, etc.) and chronic disease status, followed by sequential collection of self-reported results from the CQ-11D, EQ-5D-3L, and SF-6Dv1 scales. To ensure survey quality, high-quality surveyors were recruited and received unified training, with a *Surveyor Training Manual* distributed for study. The manual included specific operational procedures, questionnaire completion examples, and quality control requirements. After survey completion, team leaders and project members verified data and reviewed audio recordings, excluding questionnaires with missing data or those not administered according to the training manual to ensure data quality.

1.3 Measurement Instruments and Scoring Methods

The CQ-11D comprises 11 dimensions: mobility and self-care, appetite, bowel movements, sleep quality, mental vitality (energy, spirit, concentration), dizziness (subjective vertigo, mild cases resolving with eye closure, severe cases with sensation of self or surroundings spinning, inability to stand), palpitations (subjective awareness of heart beating), pain, fatigue/weakness, irritability, and anxiety (worry, urgency, tension, concern, unease) or depression (discouragement, loss of interest in activities, lack of pleasure, low spirits). Each dimension has four levels describing different severity levels, capable of describing 4,194,304 health states. Utility values were calculated using the utility scoring system developed by Zhu Wentao et al. based on a Chinese general population discrete choice experiment with time trade-off (DCETTO) design, with a measurable utility range of -0.868 (44444444444) to 1 (11111111111).

The EQ-5D, developed by the EuroQoL Group, is one of the most widely used generic utility scales worldwide. It includes five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The 3-level version (EQ-5D-3L) contains three severity levels per dimension (no problems, some/moderate problems, extreme problems), generating 243 health states with a measurable utility range of -0.149 (33333) to 1 (11111). EQ-5D-3L includes a Visual Analogue Scale (EuroQol VAS) to directly measure respondents' quality of life. The EQ VAS is a 20cm vertical scale with 100 at the top and 0 at the bottom, representing the best and worst imaginable health states, respectively. EQ-5D-3L utility values were calculated using a Chinese population-based utility scoring system.

The SF-6D was developed by Brazier et al. based on the SF-12 and SF-36 health surveys, including six dimensions: physical functioning, role limitation, social functioning, pain, mental health, and vitality. Its reliability and validity have been verified across different disease areas in various countries. Each dimension has 4-6 levels, describing 18,000 health states, with the worst state being 645655 and the best state 111111. As a mainland China population-based utility scoring system for SF-6Dv1 has not yet been developed, the Hong Kong Chinese population-based utility scoring system was used.

1.4 Data Analysis

This study conducted two-part data analysis. The first part analyzed the overall chronic disease population measurements from CQ-11D, EQ-5D-3L, and SF-6Dv1, including descriptive analysis, consistency analysis, and correlation analysis. The second part focused on the three most prevalent chronic diseases (hypertension, arthritis, and hyperlipidemia), conducting more detailed analyses including descriptive analysis of utility values, ceiling/floor effects, and univariate analysis. Additionally, we examined whether the distribution characteristics, consistency, and correlation of utility values measured by the three scales differed from those in the overall chronic disease population.

Descriptive analysis used means and standard deviations to describe utility values, with histograms used to observe distribution characteristics. Ceiling/floor effects were evaluated based on literature indicating that proportions exceeding 15% at the maximum or minimum dimension or total scores constitute ceiling or floor effects. This study used the proportion of samples achieving the best/worst possible health state as the evaluation indicator for ceiling/floor effects.

Consistency analysis employed Intraclass Correlation Coefficients (ICC) and Bland-Altman plots. Bland-Altman consistency evaluation criteria followed literature guidelines stating that points within the 95% limits of agreement (LOA) on the plot should comprise approximately 95% of the total, while also considering whether LOA fall within professionally acceptable ranges. This study did not consider professional acceptability ranges. Generally, ICC values between 0.75-1 indicate high correlation, 0.40-0.75 indicate moderate correlation, and below 0.4 indicate low correlation. Kruskal-Wallis nonparametric tests performed univariate analyses of sociodemographic characteristics to assess the impact of sex and age on utility value measurements from the three scales for the three chronic diseases.

A two-sided p-value <0.05 was considered statistically significant. Descriptive and univariate analyses were completed using SAS 9.2 software, histograms and ICC calculations were performed using STATA 15.0, and Bland-Altman plots were generated using MedCalc software.

Results

2.1 Sample Characteristics

Among 2,586 participants, 88 questionnaires were excluded following quality control principles: (1) 57 incomplete questionnaires; (2) 5 questionnaires not meeting inclusion/exclusion criteria; (3) 9 questionnaires with logical inconsistencies; and (4) 17 questionnaires completed in less than 5 minutes. The final analysis included 2,498 valid questionnaires, yielding an effective response rate of 96.60%. Screening these valid questionnaires for self-reported physician-diagnosed chronic diseases resulted in 692 chronic disease patients, representing 27.70% of the sample. The mean age of chronic disease patients was 54 years, with 365 males (52.75%) and 327 females (47.25%). The proportions of patients with one, two, and three or more chronic conditions were 71.24% (493/692), 18.50% (128/692), and 10.26% (71/692), respectively. The three most prevalent chronic diseases were hypertension (39.74%, 275/692), arthritis (18.50%, 128/692), and hyperlipidemia (17.92%, 124/692). Basic information for the included chronic disease patients is presented in Table 1 .

2.2 Distribution of Utility Values in the Overall Chronic Disease Population

The mean (standard deviation) utility values for the overall chronic disease population were 0.855 (0.168) for CQ-11D, 0.793 (0.132) for SF-6Dv1, and 0.876

(0.136) for EQ-5D-3L. The utility value distributions are shown in Figure 1 [Figure 1: see original paper]. All three scales exhibited non-normal distributions, with overall utility values skewed toward higher values (negative skewness). The CQ-11D distribution was relatively broad, spanning both positive and negative values with a range of (-0.147, 1) and good continuity. The SF-6Dv1 distribution was concentrated in the region above 0.5, with a range of (0.346, 1). The EQ-5D-3L distribution showed a pronounced ceiling effect with a discontinuous pattern, ranging from (0.336, 1). Ceiling effects were 5.64% (39/692) for CQ-11D, 4.05% (28/692) for SF-6Dv1, and 41.76% (289/692) for EQ-5D-3L. No floor effects were observed for any of the three scales.

2.3 Consistency Analysis Results in the Overall Chronic Disease Population

Bland-Altman plots revealed relatively good consistency among the three scales, with approximately 95% of points falling within the LOA, though consistency varied between different scale pairs. The 95% CI for differences between EQ-5D-3L and SF-6Dv1 was (-0.14, 0.31), with 4.77% (33/692) of points beyond the 95% LOA, indicating the best consistency (LOA=0.952). The 95% CI for differences between CQ-11D and EQ-5D-3L was (-0.31, 0.27), with 5.06% (35/692) beyond the LOA, indicating good consistency (LOA=0.949). The 95% CI for differences between CQ-11D and SF-6Dv1 was (-0.23, 0.35), with 5.64% (39/692) beyond the LOA, indicating good consistency (LOA=0.944). These results are shown in Figures 2 [Figure 2: see original paper], 3 [Figure 3: see original paper], and 4 [Figure 4: see original paper].

ICC analysis showed moderate correlations among the three scales, with SF-6Dv1 and EQ-5D-3L showing the highest correlation, consistent with Bland-Altman plot results. Additionally, CQ-11D demonstrated moderate correlation with VAS, while SF-6Dv1 and EQ-5D-3L showed low correlations with VAS. Results are presented in Table 2 .

2.4 Utility Value Results for Three Specific Chronic Disease Groups

For hypertension patients, mean (SD) utility values were 0.846 (0.182) for CQ-11D, 0.796 (0.137) for SF-6Dv1, and 0.875 (0.143) for EQ-5D-3L. For hyperlipidemia patients, values were 0.852 (0.179), 0.781 (0.135), and 0.870 (0.139), respectively. For arthritis patients, values were 0.832 (0.173), 0.747 (0.127), and 0.844 (0.137), respectively. Distribution patterns were consistent with the overall chronic disease population. All three scales showed the lowest mean utility values for arthritis patients.

Ceiling effects were 5.45% (15/275), 4.36% (12/275), and 43.64% (120/275) for hypertension patients; 6.45% (8/124), 4.84% (6/124), and 39.52% (49/124) for hyperlipidemia patients; and 2.34% (3/128), 1.56% (2/128), and 27.34% (35/128) for arthritis patients, respectively. No floor effects were observed for any scale in any disease group.

Bland-Altman plots showed that consistency among the three scales in hypertension patients remained similar to the overall chronic disease population. However, in hyperlipidemia and arthritis patients, consistency was lower (LOA<0.95). ICC analysis revealed moderate correlations among the three scales in hypertension patients (ICC$ $0.6), consistent with overall results. Compared to the overall population, correlation between CQ-11D and EQ-5D-3L was higher in hyperlipidemia patients, while correlation between SF-6Dv1 and EQ-5D-3L was lower in arthritis patients. Regarding correlation with VAS, CQ-11D showed the highest correlation across all three chronic diseases, consistent with overall results.

Univariate analysis showed that in hypertension patients, utility values differed by sex and age, with males having higher values than females across all three scales. Maximum utility values were observed in the 30-39 age group for all three scales, while minimum values were in the $ $60 age group for CQ-11D and SF-6Dv1, and in the 18-29 age group for EQ-5D-3L. In hyperlipidemia patients, utility values differed by sex, with males showing higher values across all three scales. Maximum utility values were in the 30-39 age group for CQ-11D and EQ-5D-3L, and in the 18-29 age group for SF-6Dv1, while minimum values were in the $ $60 age group for all three scales. In arthritis patients, no statistically significant differences in utility values were observed by sex or age across any of the three scales. Results are presented in Table 3 .

Discussion

3.1 CQ-11D Demonstrates Certain Advantages in Measurement Validity

The results indicate that EQ-5D-3L measured the highest mean utility values, followed by CQ-11D, with SF-6Dv1 measuring the lowest. Regarding distribution, CQ-11D showed the broadest range for the overall chronic disease population, including some negative values, indicating its ability to reflect utility values for chronic disease patients across different health states. The SF-6Dv1 distribution was relatively continuous and uniform but concentrated in the region above 0.5. The EQ-5D-3L distribution was concentrated in higher utility regions with discontinuities. Only EQ-5D-3L exhibited a ceiling effect, consistent with previous domestic studies, likely due to its insensitivity to minor health state changes, resulting in overestimation of utility values in chronic disease patients. Previous research has indicated that SF-6Dv1 has lower sensitivity when applied to poorer health states, showing substantial floor effects. However, this study found no floor effects for any scale, possibly because chronic disease patients do not represent extremely poor health states, or because SF-6Dv1, being developed in foreign populations, may not manifest floor effects when applied to Chinese populations due to ethnic and cultural differences. Separate analyses of hypertension, hyperlipidemia, and arthritis patients yielded consistent results.

3.2 Consistency Results Indicate Differences Between CQ-11D and Other Scales

Consistency analysis revealed ICC values between 0.528-0.625 across the three scales, indicating moderate but not high correlation, with the highest correlation between SF-6Dv1 and EQ-5D-3L. Both SF-6Dv1 and EQ-5D-3L showed low correlations with VAS, while CQ-11D demonstrated moderate correlation with VAS. Bland-Altman plots showed consistency exceeding 0.95 between SF-6Dv1 and EQ-5D-3L, while consistency between CQ-11D and the other two scales was below 0.95 (0.944-0.949). Two possible explanations exist: First, CQ-11D includes more dimensions, enabling more specific health state description than SF-6Dv1 and EQ-5D-3L, which may contribute to slightly lower consistency. Second, as CQ-11D was designed specifically for Chinese populations and incorporates TCM theory and health concepts with more comprehensible items, it may better reflect patients' direct perceptions of their health status. Since VAS directly measures health status, CQ-11D shows the highest correlation with VAS among the three scales. Similar patterns were observed in hypertension, hyperlipidemia, and arthritis subgroups.

3.3 Consistency Patterns Vary Across Different Chronic Diseases

Compared to the overall chronic disease population, hypertension patients showed moderate correlations among the three scales with consistency around 0.95, suggesting the three scales may be somewhat interchangeable for measuring hypertension populations. In contrast, hyperlipidemia patients showed higher correlation between CQ-11D and EQ-5D-3L, while arthritis patients showed lower correlation between SF-6Dv1 and EQ-5D-3L. Consistency among the three scales was below 0.95 in both hyperlipidemia and arthritis patients, indicating substantial differences in utility value measurement across scales for these conditions and suggesting that scale selection should be based on specific measurement characteristics and health dimensions.

3.4 Sex and Age Influence Utility Values in Hypertension and Hyperlipidemia Populations

In hypertension and hyperlipidemia patients, all three scales consistently showed higher utility values in males than females. Across age groups, CQ-11D and SF-6Dv1 showed lowest utility values in patients aged 60 and above. These findings suggest that female and elderly patients with hypertension and hyperlipidemia warrant particular attention in treatment and daily health management.

This study has several limitations. First, the number and types of chronic diseases included were limited; future research should incorporate more diverse chronic disease populations. Second, due to study design, EQ-5D-5L was not used, which might have mitigated the ceiling effect observed with EQ-5D-3L and could have influenced our ceiling effect findings. Third, disease severity was not considered; subsequent studies should stratify chronic diseases by severity stage

to analyze scale sensitivity to disease changes and performance across disease stages, providing references for appropriate scale selection and enabling precise utility measurement and improved HRQoL research in chronic disease patients. Fourth, CQ-11D items were designed from a TCM perspective, requiring surveyors to have accurate and comprehensive understanding of each item' s meaning. This creates a knowledge barrier for widespread application of CQ-11D. Surveyors should receive training on basic TCM theory and scale item interpretation before using the instrument to avoid inaccurate measurement resulting from improper understanding or explanation.

In summary, compared to EQ-5D-3L, neither CQ-11D nor SF-6Dv1 showed ceiling effects, suggesting these scales may be more sensitive for measuring minor health state changes and more suitable for slowly progressing chronic diseases, while EQ-5D-3L may be more appropriate for measuring relatively poor health states. Moreover, compared to SF-6Dv1 and EQ-5D-3L, CQ-11D demonstrated higher correlation with VAS and a broader utility value distribution, better reflecting patients' self-perceived health status across different levels. Therefore, CQ-11D may be more suitable for measuring utility values in chronic disease patients.

Author Contributions

WANG Wei and XIE Shitong conceived the research idea, analyzed the data, and wrote the manuscript. ZHOU Jiameng and PAN Jie organized and analyzed the research data and performed secondary verification of the analysis results. ZHU Wentao provided the original data and supervised and guided manuscript writing.

Conflict of Interest Statement

The authors declare no conflicts of interest.

References

[1] Chen YC, Chen H, Zhou Y, et al. Early death probability of major chronic diseases among residents in Pudong New Area, Shanghai from 2002 to 2020 [J]. *Chinese General Practice*, 202, 25(9): 1098-1104. DOI: 10.12114/j.issn.1007-9572.2021.02.074.

[2] World Health Organization. *Global status report on noncommunicable diseases 2014*[R]. Geneva: World Health Organization, 2014.

[3] Statistical information center of National Health Commission. *2021 China Health Statistics Yearbook* [M]. Beijing: China Union Medical University Press, 2021.

[4] Li XM, Wan CH, Wang GH, et al. Evaluation of quality of life in patients with chronic diseases [J]. *Chinese Journal of General Medicine*, 2007(01): 20-22.

[5] Liu XD, Wang ZX, Wu BY, et al. Based on the literature analysis of the problems China' s chronic noncommunicable diseases field [J]. *Journal of public health of China*, 2019, 35(8): 959-962. DOI: 10.11847/zgggws1118013.

[6] Sun XH, Liu KK, Jiang H, et al. Application of chronic health care model in long-term management of diabetes mellitus [J]. *Chin J general practice*, 2018, 21(31): 3878-3881. DOI: 10.12114/j.issn.1007-9572.2018.00.043.

[7] Sun LH. *Pharmaceutical economics* [M]. China Medical Science and Technology Press, 2015.

[8] Liu GG, Hu SL, Wu JH, Wu J, Dong ZH, Li HC, et al. *China guidelines for pharmacoeconomic evaluations 2020* (Chinese–English version). Beijing: China Market Press; 2020.

[9] Zhang TQ, Wu HY, Cai YF, et al. Study on quality of life status and influencing factors of Chinese general population based on EQ-5D-5L and SF-6D scales [J]. *China health service management*, 2020, 37(08): 631-634.

[10] Huang HY, Tang SL, Liu JJ. Health related quality of life and its influencing factors in patients with chronic diseases in Zhejiang, Sichuan, Hebei and Gansu provinces based on European five-dimension health Scale [J]. *Chronic disease prevention and control in China*, 2019, 27(8): 575-579+584. DOI: 10.16386/j.cjpccd.issn.1004-6194.2019.08.004.

[11] Du XD, Zhu P, Li ME, et al. Comparison of EQ-5D and SF-6D health efficacy in patients with stroke [J]. *Journal of sichuan university (medical edition)*, 2018, 49(02): 252-257. DOI: 10.13464/j.scuxbyxb.2018.02.020.

[12] Sun SP, Long JH, Zhang SZ. Reliability and validity of the Chinese version of the SF-36 scale for quality of life assessment of patients with anorectal benign diseases [J]. *Colorectal anal surgery*, 2019, 25(05): 497-505+512.

[13] Ware JE Jr. SF-36 health survey update. *Spine (Phila Pa 1976)*. 2000 Dec 15;25(24):3130-9. doi: 10.1097/00007632-200012150-00008. PMID: 11124729.

[14] Stephen J. Walters and John E. Brazier. Sample Sizes for the SF-6D Preference Based Measure of Health from the SF-36: A Comparison of Two Methods[J]. *Health Services and Outcomes Research Methodology*, 2003, 4(1): 35-47.

[15] Xie ST, Chen PN, He XN, et al. Chinese journal of health policy research, 2020, 13(08): 58-68.

[16] TCM quality of life assessment scale [j] *Chinese pharmacoeconomics*, 2022, 17(05): 5-9.

[17] Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user' s guide. *Pharmacoeconomics*. 2008;26(8): 661-77. doi: 10.2165/00019053-200826080-00004. PMID: 18620460.

[18] *China Statistical Yearbook* [Z]. China Statistics Press. 2019.

[19] Wu HY. *Research on health utility value measurement* [D]. Shenyang Pharmaceutical University, 2012.

[20] WARE J E, SHERBOURNE C D. The MOS 36-item short form health survey(SF-36): I. Conceptual framework and item selection[J]. *Medical care*, 1992, 30(6): 473-483.

[21] Christiansen ASJ, Møller MLS, Kronborg C, Haugan KJ, Køber L, Højberg S, Brandes A, Graff C, Diederichsen SZ, Nielsen JB, Krieger D, Holst AG, Svendsen JH. Comparison of the three-level and the five-level versions of the EQ-5D. *Eur J Health Econ*. 2021 Jun;22(4): 621-628. doi: 10.1007/s10198-021-01279-z. Epub 2021 Mar 18. PMID: 33733344.

[22] Liu GG, Wu H, Li M, et al. Chinese time trade-off values for EQ-5D health states. *Value Health*. 2014 Jul;17(5): 597-604.

[23] BRAZIER J, ROBERTS J, DEVERILL M. The estimation of a preference based measure of health from the SF-36[J]. *Jhealth econ*, 2002, 21(2): 271–292.

[24] Liu XH, Li SP, Chen G. Evolution of the sf-6d multiattribute utility scale and its introduction [J]. *Chinese health economics*, 2019, 38(2): 8-11.

[25] Ye ZP, Zheng Y, Sun LH. Introduction and research status of six-dimensional health status classification System SF-6D[J]. *Chinese journal of health economics*, 2018, 37(02): 70-73.

[26] McGhee SM, Brazier J, Lam CL, et al. Quality-adjusted life years: population-specific measurement of the quality component. *Hong Kong Med J*. 2011 Dec;17 Suppl 6: 17-21. PMID: 22147354.

[27] McHorney C A and Tarlov A R. Individual-patient monitoring in clinical practice: are available health status surveys adequate?[J]. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, 1995, 4(4): 293-307.

[28] Sa J, Liu GF. Consistency evaluation of quantitative measurement results and application of Bland-Altman method[J]. *Chinese Health Statistics*, 2011, 28(04): 409-411+413.

[29] Li F, Lin X, Liu J. Variability of urinary biomarkers of neonicotinoid insecticides in Chinese population: Implications for human exposure assessment. *Chemosphere*. 2022 Nov;307(Pt 1): 135705. doi: 10.1016/j.chemosphere.2022.135705. Epub 2022 Jul 14. PMID: 35843435.

[30] YANG F, LAU T, LEE E, et al. Comparison of the preference-based EQ-5D-3L and SF-6D in patients with endstage renal disease (ESRD). *Eur J Health Econ*, 2015, 16(9): 1019-1026.

[31] Jin H, Wang B, Gao Q, Chao J, Wang S, Tian L, Liu P. Comparison between EQ-5D and SF-6D utility in rural residents of Jiangsu Province, China. *PLoS One*. 2012;7(7): e41550. doi: 10.1371/journal.pone.0041550. Epub 2012 Jul 27. PMID: 22848526; PMCID: PMC3407238.