

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

A comparison of the EQ-5D-5L and the SF-6Dv2 among patients with hypertrophic obstructive cardiomyopathy

Short title: EQ-5D-5L and the SF-6Dv2 in hypertrophic obstructive cardiomyopathy

Xiaoyu Wang¹, Haiqiang Sang^{1*}, Rui Meng², Xin Su¹, Peng Liu¹

¹Department of Cardiovascular Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou University, Zhengzhou, Henan 450052, China

²School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, Jiangsu 210009, China

*Corresponding author

Email: fccsanghq@zzu.edu.cn (HS)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

26 Abstract

27 **Background:** Hypertrophic obstructive cardiomyopathy (HOCM) is a serious and
28 complex chronic disease severely affecting patients' daily lives and health-related quality of
29 life (HRQoL). The psychometric performance of generic preference-based instruments has
30 not been compared in Chinese patients with HOCM. We aimed to identify an applicable
31 vehicle to determine HRQoL and explore the psychometric properties of SF-6Dv2 and EQ-
32 5D-5L in adults with HOCM. The interchangeability of the tools in cost-utility analysis was
33 also investigated. **Methods:** We collected data from 131 patients with HOCM from the First
34 Affiliated Hospital of Zhengzhou University, China. Assessments were performed on the day
35 of admission and three and six months after discharge using SF-6Dv2 and EQ-5D-5L. The
36 responses were converted to utility values using the corresponding Chinese value sets. The
37 tool distributions were explored, and the floor and ceiling effects were analyzed. The
38 agreement was assessed using intra-class correlation coefficients (ICCs) and Bland–Altman
39 plots. Convergent validity was tested using Spearman's rank correlation coefficient. Known
40 group validity was measured across various clinical and sociodemographic indicators using
41 relative efficiency (RE) statistics. **Results:** The mean utility scores for SF-6Dv2 and EQ-5D-
42 5L at baseline and 6-month follow-up were 0.61, 0.62, 0.736, and 0.797, respectively. The
43 EQ-5D-5L and SF-6Dv2 distribution scores showed no normality. EQ-5D-5L was more
44 sensitive to changes over time and showed a moderate to good correlation with the SF-6Dv2
45 (ICC: 0.598–0.862). The instruments' agreement and convergent validity worsened in
46 patients with a higher New York Heart Association (NYHA) cardiac function classification
47 and lower general health status. SF-6Dv2 showed higher relative efficiency statistics and a
48 greater ability to distinguish external health status. **Conclusions:** The measured results can be
49 used for future cost-utility analyses. SF-6Dv2 and EQ-5D-5L presented different results and

50 should not be used interchangeably. SF-6Dv2 is optimal for detecting differences between
51 subgroups with various health states.

52

53 **Keywords:** SF-6Dv2, EQ-5D-5L, HOCM, Health-related quality of life

54

55

56 **Introduction**

57 Hypertrophic cardiomyopathy (HCM), distinguished by left ventricular myocardial
58 hypertrophy, is a class of cardiac disorders caused by genetic factors. It often has an
59 autosomal inheritance pattern due to variants in the genes encoding myosin (or myosin-
60 associated genes) or may have an unknown genetic etiology [1]. The main feature of
61 hypertrophic obstructive cardiomyopathy (HOCM) is a left ventricular outflow tract pressure
62 difference of at least 30 mmHg at rest or with provocation. It accounts for approximately
63 70% of all HCM patients [2]. Currently, the prevalence of HCM in China is 0.076%, and the
64 mortality rate of HCM patients is as high as 3.38% [3]. HOCM is the main cause of sudden
65 cardiac death in adolescents and athletes and is closely related to heart failure and stroke in
66 older patients [4,5]. Typical symptoms include exertional dyspnea and weakness, chest pain,
67 and palpitations, which tremendously impact patients' physical health and health-related
68 quality of life (HRQoL) [6]. Parallel to this, HOCM is closely related to considerable
69 expenditures on healthcare and imposes an enormous financial burden on global health
70 budgets [7].

71 In order to provide evidence support for decision-makers to allocate limited resources
72 among competing healthcare programs appropriately, cost-utility analysis (CUA) is
73 incrementally being used. Quality-adjusted life years (QALYs) are health outcome indicators
74 in CUA, which combines years of survival with utility values [8]. Utility scores are usually

75 expressed as a value from 0 to 1, where 1 represents perfect health and 0 represents death;
76 they can also be negative, reflecting a disease situation worse than death. A series of
77 universal health scales (multi-attribute utility instruments [MAUIs]) have been developed
78 internationally to measure health utility values. Among the multitude of generic preference-
79 based scales, the EuroQoL Five-Dimension (EQ-5D) from the European Society of Quality
80 of Life [EuroQoL Group] along with the Short Form Six-Dimension (SF-6D) obtained based
81 on the Short Form 36-item [SF-36] are among the most popular [9].

82 EQ-5D, developed by the EuroQol Group, is the most commonly used MAUIs
83 worldwide. Compared to the original version of EQ-5D-3L, which contains only three
84 response levels for each dimension, EQ-5D-5L offers five levels per dimension and lower
85 ceiling effects when conducting a survey among the population [10]. The SF-6D scale is
86 second only to the EQ-5D. Moreover, because the SF-6D scale dimension level setting is
87 more abundant and sensitive, it is more suitable for chronic diseases where the clinical
88 symptoms are not obvious. Compared with SF-6D, SF-6Dv2 has a much broader scoring
89 range and addresses the problem of descriptive systems [11]. Differences in the structure and
90 valuation of different instruments may give rise to various estimates of the identical “health
91 state” for the same individual, which results in discrepancies in utility and healthcare-related
92 decision-making. Therefore, it is crucial to measure the utility of HOcm patients using these
93 two well-known instruments and to compare the psychometric properties of measurements
94 and interchangeability.

95 Comparisons of the properties of SF-6D and EQ-5D measurements have already been
96 validated for several diseases, including depressive symptoms and fibromyalgia [12,13].
97 However, no studies have compared health utility values based on Chinese HOcm patients
98 using EQ-5D-5L and SF-6Dv2. Therefore, we aimed to estimate the utility of Chinese
99 patients with HOcm using EQ-5D-5L and SF-6Dv2. Furthermore, we evaluated whether

100 EQ-5D-5L and SF-6Dv2 revealed similar empirical estimates of health-related utility and
101 interchangeability in patients with HOVM.

102

103 **Material and methods**

104 **EQ-5D-5L**

105 The EQ-5D-5L consists of two parts: a questionnaire section and the EuroQoL-visual
106 analog scale (EQ-VAS). The questionnaire contained five dimensions, each with five severity
107 levels: no problem, mild, moderate, severe, and very severe [14]. Thus, 3125 ($5 \times 5 \times 5 \times 5 \times$
108 5×5) different health states described by the EQ-5D-5L questionnaire can be calculated, of
109 which 11111 are the best health states (perfectly healthy) and 55555 are the worst state of
110 health. The EQ-5D-5L utility value integral system converts a patient's health status into
111 health utility value. A Time-trade off (TTO) approach was employed to develop Chinese EQ-
112 5D-5L utility values ranging from $-0.391(55555)$ to $1(11111)$. Furthermore, the EQ-VAS
113 indicates the patient's subjective rating of self-reported health state on the day being
114 interviewed, with a vertical length of 20 cm, starting at 100 (the best health imaginable) and
115 ending at 0 (the worst health imaginable) [15]. The value of the EQ-VAS was converted to 0-
116 1 to facilitate comparison.

117

118 **SF-6Dv2**

119 The SF-6Dv2 serves a broader range of dimensions than the EQ-5D-5L, of which
120 only ten items in the SF-36 have been reclassified into six items, each dimension
121 corresponding to one item [16]. All the dimensions except pain domains (six levels) had five
122 levels, yielding a total of 18,750 ($= 5 \times 5 \times 5 \times 6 \times 5 \times 5 \times 5$) health conditions [17]. We use

123 the set of Chinese tariff values developed for SF-6Dv2 by the TTO method, which has a
124 range of values from -0.277 (555,655) to 1 (111,111).

125

126 **Sampling and data collection**

127 Participants were patients with HOCM who were hospitalized at the First Affiliated
128 Hospital of Zhengzhou University between 1 June 2021 and 31 October 2022. Patient
129 demographic characteristics and disease-related data, including age, sex, work and marital
130 status, educational background, body mass index (BMI), and left ventricular ejection fraction
131 (LVEF), were completed by the patient or extracted from the inpatient electronic medical
132 records.

133 The inclusion criteria were supposed to be in accordance with the following: 1)
134 consistent with the diagnosis of hypertrophic obstructive cardiomyopathy, and 2) the
135 completed scale contained no missing data. Patients were excluded if they 1) were under 18
136 years old, 2) refused to sign the informed consent form, or 3) had other serious illnesses or
137 psychological disorders that prevented them from understanding the contents of the
138 questionnaire. Eligible respondents were interviewed face-to-face on the day of admission by
139 a research assistant who underwent a 1-day uniform professional training before the study
140 started. After filling in the questionnaire, a second research assistant verified that the
141 questionnaire was complete and that no items were missing. Following discharge, the patient
142 was contacted by phone three and six months later. Upon completion of the questionnaire, the
143 patients uploaded and returned the results.

144 The study protocol was approved by the University of Zhengzhou Ethics Committee
145 (registration number 2021-KY-0943-002), and all participants provided written informed
146 consent before participating in the research. This study adhered to the tenets of the 1975
147 Declaration of Helsinki (Tokyo revision, 2004).

148

149 **Statistics analysis**

150 The baseline demographic statistics described continuous and categorical variables
151 using mean, standard deviation, and frequency, respectively. Overall quality of life (QOL)
152 indices were calculated and compared for the SF-6Dv2 and EQ-5D-5L questionnaires at three
153 time points: on the day of admission and three and six months after discharge. A histogram
154 represented the distributions of the two instruments. Normality tests were conducted using
155 the skewness and kurtosis values. Because of the highly skewed distribution of the index
156 scores, the Wilcoxon signed-rank test for differences between the index scores of the two
157 measurements was used. In contrast, the Kruskal–Wallis test was used for comparisons
158 between the index scores of the participants’ characteristic groups. One-way ANOVA was
159 applied to measure changes in New York Heart Association (NYHA) cardiac function classes
160 during follow-up. Ceiling and floor effects were measured by computing the percentage of
161 respondents who received the highest or lowest scores in each field. The minimally important
162 difference (MID) is often applied to explain patient-approved minimum clinical effectiveness
163 questionnaire score changes reflecting HRQoL. The treatment is clinically significant when
164 the score change reaches the MID threshold. In a previous study, the average MID values of
165 the SF-6Dv2 and EQ-5D-5L were 0.041 and 0.074, respectively [18]. Intraclass correlation
166 coefficients (ICCs) were calculated for the SF-6Dv2 and EQ-5D-5L instruments to assess
167 agreement, which was reported as a two-way random and absolute agreement with a single
168 metric model: ICC lies within 0.5-0.75 for moderate, 0.75-0.9 for good and more than 0.9 for
169 excellent. Bland–Altman plots were used to explore the degree of consistency between the
170 two utility scores [19]. Convergent validity is the degree to which the results correlate when
171 different measures are utilized to measure similar characteristics, such as anxiety/depression
172 on the EQ-5D-5L and mental health on the SF-6Dv2. Convergent validity was evaluated by

173 calculating Spearman's correlation (ρ) coefficient between the domains of the two
174 instruments. The strength of the Spearman correlation was rated as weak for less than 0.3,
175 strong for more than 0.5, and moderate for in-between [20]. Known-group validity refers to
176 the ability of two instruments to distinguish between each respondent's characteristic group
177 by calculating the average utility value for each measure and comparing them [21]. We
178 examined how well SF-6Dv2 and EQ-5D-5L detected differences in the indices of external
179 health based on the relative efficiency (RE) statistic. The RE is defined as the ratio of the
180 square of the t-statistic of the comparison tool to the reference tool [22]. A value > 1.0
181 represents, when determining differences in health-related external indicators, EQ-5D-5L
182 performs better than SF-6Dv2, while a value < 1.0 illustrates that the EQ-5D-5L performs
183 less efficiently than SF-6Dv2.

184 Statistical analyses were performed using the SPSS, IBM Statistics software, version
185 22.0. In all statistical tests, p-values less than 0.01 were assumed to be statistically
186 significant.

187

188 **Results**

189 **Characteristics of the study sample**

190 Four of the 135 participants died during the follow-up period. As their questionnaires
191 were incomplete, they were excluded from the sample. Consequently, 131 patients with
192 HOcm were included in this study. Participants' sociodemographic and clinical
193 characteristics at baseline are displayed in Tables 1 and 2, who were 53.4 years old, and most
194 patients with HOcm were male, accounting for 55.7%, and 91.6% of patients were married.
195 The mean body mass index (BMI) was 25.8 kg/m². NYHA functional class III was present in
196 53.4% of the patients, and the LVEF ranged from 29% to 79%, with an average of 64.4%.

197 We observed statistically significant differences in the NYHA cardiac function classification
 198 on the day of admission and at three and six months after discharge ($p < 0.01$).

199

200 **Table 1. Baseline characteristics**

Characteristics	Total (n=131)
Age, mean (SD)	53.4 (12.1)
BMI, mean (SD)	25.8 (3.2)
EF%, mean (SD)	64.4(5.66)
EQ-5D-5L utility, mean (SD)	0.62(0.28)
SF-6Dv2 utility, mean (SD)	0.61(0.19)
EQ-VAS score, mean (SD)	0.62(0.18)

201 SD: Standard deviation BMI: Body Mass Index; EF: Ejection Fraction; EQ-5D-5L: 5-level
 202 EuroQol 5-dimension; SF-6Dv2: Short Form Six-Dimension (version 2); EQ-VAS: EuroQol-
 203 Visual Analog Scale (EQ-VAS)

204

205 **Table 2. Known-groups validity of SF-6Dv2 and EQ-5D-5L**

	Ratio%	EQ-5D-5L		p	SF-6Dv2		p	RE*
		Mean	SD		Mean	SD		
Gender				0.88			0.10	0.01
Female	44.3	0.623	0.277		0.583	0.206		
Male	55.7	0.615	0.290		0.639	0.166		
Annual household income				<0.01			<0.01	0.77
<5K	24.4	0.450	0.299		0.490	0.178		
5K-1W	10.7	0.706	0.251		0.569	0.188		

1W-5W	25.2	0.674	0.282	0.652	0.153		
5W-10W	19.8	0.661	0.253	0.679	0.178		
10W-20W	9.9	0.752	0.189	0.716	0.210		
>20W	9.9	0.579	0.266	0.641	0.118		
Employment status				0.10		<0.0	0.34
Employed	32.8	0.690	0.254	0.694	0.156		
Retired	9.9	0.509	0.328	0.611	0.187		
Unemployed	57.3	0.596	0.285	0.569	0.189		
Education				0.98		0.11	0.03
Primary or below	29.0	0.602	0.255	0.552	0.180		
Junior high school	41.2	0.624	0.286	0.629	0.182		
Senior high school	18.3	0.625	0.341	0.654	0.194		
College or above	11.5	0.627	0.270	0.654	0.182		
Marital status				0.90		0.86	0.76
Divorced	1.5	0.763	0.254	0.554	0.409		
Married	91.6	0.619	0.285	0.619	0.185		
Unmarried	3.1	0.554	0.407	0.564	0.195		
Widow	3.8	0.605	0.189	0.552	0.178		

NYHA				<0.01		<0.0	1.23
functional						1	
class							
I	7.6	0.868	0.089		0.848	0.091	
II	22.9	0.784	0.135		0.688	0.186	
III	53.4	0.608	0.247		0.588	0.171	
IV	16	0.296	0.315		0.484	0.121	
Comorbidities				0.87		0.72	0.22
>2	51.1	0.623	0.304		0.620	0.185	
0-2	48.9	0.614	0.264		0.608	0.188	
Current				0.03		0.64	1.39
smoking							
Yes	17.6	0.722	0.233		0.676	0.165	
NO	82.4	0.596	0.289		0.601	0.189	

206 *Relative efficiency value of the F-statistics of one-way ANOVA (F-statistics EQ-5D-5L/ F-
 207 statistics SF-6Dv2). NYHA, New York Heart Association; EQ-5D-5L, 5-level EuroQol 5-
 208 dimension; SF-6Dv2, Short Form Six-Dimension (version 2)

209

210 Comparison

211 As shown in Fig 1, the scores from the baseline and follow-up measurements were not
 212 normally distributed. The distribution of EQ-5D-5L as a starting point and follow-up was
 213 highly left-skewed compared to SF-6Dv2. The most frequently distributed combinations of
 214 SF-6Dv2 and EQ-5D-5L at baseline were 421111 and 11121. The mean utility scores of SF-
 215 6Dv2 and EQ-5D-5L at baseline were 0.61 and 0.62, respectively. The minimum and
 216 maximum values generated from the EQ-5D-5L were -0.182 and 1.0, respectively, while the

217 SF-6Dv2 measured -0.277 and 0.962 at its minimum and maximum, respectively. In EQ-5D-
218 5L, a low ceiling effect of 3 (2.3%) was observed. The EQ-VAS measured two (1.5%)
219 participants in optimum health. No instruments produced a floor effect. Over a 3-month
220 follow-up, the mean utility of EQ-5D-5L and SF-6Dv2 was 0.849 (range: -0.391 to 1.0) and
221 0.769 (range: -0.277 to 1.0), respectively ($p = 0.000$). The floor and ceiling effects of EQ-
222 5D-5L were 2 (1.5%) and 36 (27.5%), respectively, while those of SF-6Dv2 were 2 (1.5%)
223 and 9 (6.9%), respectively. According to the EQ-VAS score, eleven individuals (8.4 %)
224 obtained the best possible health. The 6-month follow-up revealed the following: the mean
225 utility values of EQ-5D-5L and SF-6Dv2 were 0.797 and 0.736 ($p = 0.000$), respectively. The
226 range of values of the two measurements was the same as for the follow-up in March. The
227 floor and ceiling effects of EQ-5D-5L were 1 (0.8%) and 35 (26.7%), respectively, whereas
228 those of SF-6Dv2 were 1 (0.8%) and 8 (6.1%), respectively; and 14 (10.7%) people achieved
229 the highest EQ-VAS scores.

230

231 **Fig 1 Distribution of EQ-5D-5L and SF-6Dv2.** (A) Distribution of EQ-5D-5L at baseline;
232 (B) Distribution of EQ-5D-5L at 3-month follow-up; (C) Distribution of EQ-5D-5L at 6-
233 month follow-up; (D) Distribution of EQ-5D-5L at baseline; (E) Distribution of EQ-5D-5L at
234 3-month follow-up; (F) Distribution of EQ-5D-5L at the 6-month follow-up. EQ-5D-5L: 5-
235 level EuroQol 5-dimension. SF-6Dv2: Short-Form Six-Dimensions (Version 2)

236

237 **Agreement**

238 The overall agreement for the SF-6Dv2 and EQ-5D-5L in utility scores at baseline
239 was moderate ($ICC = 0.598$). The ICCs at the 3- and 6-month follow-ups indicate good
240 agreement when comparing EQ-5D-5L and SF-6Dv2 index scores (S1 Table). The Bland–
241 Altman plots in Fig 2a show an average difference of 0.004 , with a wide range of agreement

242 from -0.418 to 0.427 between the EQ-5D-5L and SF-6Dv2 index scores. Consequently, the
243 EQ-5D-5L measurement was 35, which was 9% less or 58% larger than the measurement by
244 SF-6Dv2 for 95% of individuals. SF-6Dv2 and EQ-5D-5L had a large discrepancy for lower
245 utility values, with high variation between the two instruments depending on the patient's
246 health status. For better health, EQ-5D-5L produced higher scores, while for poorer health,
247 SF-6Dv2 yielded higher scores. The 3-month follow-up period is shown in Fig 2b. The
248 indicator scores for EQ-5D-5L exceeded those for SF-6Dv2, with an average difference of
249 0.080 in 77.1% of observations. At the 6-month follow-up, as shown in Fig 2c, more than
250 70.2% of the EQ-5D-5L utility scores outperformed the SF-6Dv2 utility scores, with a mean
251 difference of 0.061 . The limits of the agreement ranged from -0.208 to 0.330 .

252

253 **Fig 2. Bland–Altman plots of the difference in utility values of EQ-5D-5L and SF-6Dv2.**

254 (A) At baseline; (B) at the 3-month follow-up; and (C) at the 6-month follow-up. EQ-5D-5L:
255 5-level EuroQol 5-dimension. SF-6Dv2: Short Form Six-Dimension (version 2)

256

257 A mean difference at the baseline of 0.082 was found between SF-6Dv2 and EQ-5D-
258 5L, which is lower than the recommended MID of 0.041 for SF-6D and 0.074 for EQ-5D. Fig
259 2c shows a mean difference of 0.080 between SF-6Dv2 and EQ-5D-5L, similar to the MID of
260 EQ-5D but greater than the MID of SF-6D. The result of SF-6Dv2 versus EQ-5D-5L was
261 0.061 , which was lower than the MID of EQ-5D; however, it was greater than the MID of
262 SF-6D. It is possible that these instruments cannot be interchanged because of their clinical
263 significance.

264

265 **Convergent validity**

266 The correlations between the SF-6Dv2 and EQ-5D-5L dimensions based on the self-
 267 reported health of the patients are presented in Table 3. Except for the dimension of self-care
 268 (EQ-5D-5L), most correlations between SF-6Dv2 and EQ-5D-5L at baseline were moderate
 269 to weak. Indeed, the highest correlation at baseline and 3-month follow-up between the
 270 Anxiety/Depression (EQ-5D-5L) and Mental Health (SF-6Dv2) dimensions was 0.86 and
 271 0.85, followed closely by the Role Limitation (SF-6Dv2) and Usual Activities (EQ-5D-5L)
 272 dimensions. Stronger correlations were observed at the 6-month follow-up, at which point
 273 convergent validity between similar dimensions was evident: Self-Care and Physical
 274 Function (0.41) and Self-Care and Role Limitation (0.41). The correlation for each dimension
 275 tended to be moderate to strong. Overall, the SF-6Dv2 and EQ-5D-5L utilities were highly
 276 correlated at baseline ($\rho = 0.709$), 3-month follow-up ($\rho = 0.772$), and 6-month follow-up
 277 ($\rho = 0.792$).

278

279 **Table 3. Spearman correlations among the EQ-5D-5L and SF-6Dv2 domains**

EQ-5D-5L	SF-6Dv2					
	Physical function	Role limitation	Social function	Pain	Mental health	Vitality
At baseline						
Mobility	0.53**	0.52**	0.49**	-0.04	0.25**	0.29**
Self-care	0.19*	0.19*	0.24**	0.10	-0.01	0.12
Usual activities	0.66**	0.71**	0.57**	0.04	0.40**	0.45**
Pain/discomfort	0.45**	0.44**	0.32**	0.22*	0.32**	0.38**
Anxiety/Depressed	0.34**	0.35**	0.25**	0.28**	0.86**	0.45**
At 3-month follow-up						
Mobility	0.58**	0.57**	0.60**	0.16	0.29**	0.30**

Self-care	0.32**	0.29**	0.37**	0.37**	0.32**	0.26**
Usual activities	0.59**	0.60**	0.52**	0.18*	0.31**	0.32**
Pain/discomfort	0.40**	0.45**	0.26**	0.45**	0.39**	0.46**
Anxiety/Depressed	0.20*	0.34**	0.20*	0.26**	0.85**	0.46**
<hr/>						
At 6-month follow-up						
Mobility	0.63**	0.56**	0.49**	0.43**	0.24**	0.50**
Self-care	0.41**	0.41**	0.49**	0.24**	0.28**	0.36**
Usual activities	0.68**	0.60**	0.43**	0.28**	0.36**	0.49**
Pain/discomfort	0.63**	0.58**	0.41**	0.58**	0.42**	0.66**
Anxiety/Depressed	0.52**	0.51**	0.49**	0.26**	0.59**	0.54**

280 SD: Standard deviation RE: relative efficiency **p< 0.01 (two-tailed) EQ-5D-5L: 5-level

281 EuroQol 5-dimension; SF-6Dv2: Short Form Six-Dimension (version 2)

282

283 External validity

284 SF-6Dv2 and EQ-5D-5L can distinguish between known group validity (Table 2). As
 285 measured by the SF-6Dv2 and EQ-5D-5L, participants who were less educated, unmarried,
 286 retired, or unemployed, had comorbidities ≤ 2 , and higher NYHA classification tended to
 287 have lower average utility. Additionally, EQ-5D-5L was considerably more efficient
 288 concerning NYHA functional class (RE = 1.23) and current smoking (RE = 1.39). At the
 289 same time, SF-6Dv2 performed better concerning sex (RE = 0.01), annual household income
 290 (RE = 0.77), employment status (RE = 0.34), education (RE = 0.03), marital status (RE =
 291 0.76), and comorbidities (RE = 0.22).

292

293 Discussion

294 This study compared the changes over time between the generic health status
295 questionnaires SF-6Dv2 and EQ-5D-5L in patients with HOCM. To the best of our
296 knowledge, this is the first study to compare EQ-5D-5L and SF-6Dv2 instruments in patients
297 with HOCM. A 6-month follow-up allowed us to investigate the sensitivity of the two scales
298 to changes in health status. As reported above, SF-6Dv2 and EQ-5D-5L displayed some
299 parallels with respect to the ability to detect trends in the changes in health utility values.
300 However, the absolute amounts of measured utility values were not identical. The two scales
301 had similar average utility values at baseline. However, at the subsequent follow-up visit,
302 when the physical function was better than SF-6Dv2, EQ-5D-5L generally showed higher
303 utility in the same patient. These results were consistent with a recent study comparing the
304 two instruments [23,24]. A previous study comparing EQ-5D-5L and SF-6Dv2 in breast
305 cancer patients indicated that the utility value of EQ-5D-5L exceeded that of SF-6Dv2, which
306 also agreed with our findings [25]. The percentage change from the baseline in the EQ-5D-5L
307 questionnaire was higher. It is reasonable to assume that EQ-5D-5L is more responsive to
308 change than SF-6Dv2. This might be partly attributed to the different dimensions covered by
309 the two scales and the different utility ranges corresponding to each measure. The ceiling
310 effect of EQ-5D-5L was evident as the NYHA cardiac function class improved after
311 discharge. Our follow-up results showed a higher ceiling effect [26]. The disease's morbidity
312 level is thought to be a possible factor influencing the ceiling effect in EQ-5D [27].
313 Numerous previous studies have proven the ceiling effect of EQ-5D-5L [28,29]. These results
314 suggest that an instrument with higher ceiling effects is ineffective in differentiating
315 relatively better health conditions. However, neither EQ-5D-5L nor SF-6Dv2 showed floor
316 effects. As an explanation, most patients received regular treatment on the day of admission
317 accompanied by their family members, which led to improvements in the physical
318 impairment and mental symptoms caused by the disease. In contrast, the SF-6Dv2 involved

319 the period over the previous four weeks and presumably had a recall bias that prevented
320 patients from accurately perceiving physical discomfort and mental impairment within the
321 specified time of the instrument. After the patient was discharged from the hospital at the end
322 of the intervention, due to home care and regular re-examination of chronic diseases, the
323 assessment of the patient’s worst self-health status was further reduced. There is no
324 consensus on a methodology for comparing the utility scores of different MAUIs. Owing to
325 the lack of a distinct reference standard for basic preference measures assessing health
326 outcomes, we compared the combination of indicators concerned with correlation,
327 consistency, convergent validity, and known group validity.

328 The ICC between the utility scores of the two tools can be observed to be moderate to
329 good, similar to previous findings [30,31]. Nevertheless, the agreement of the instruments
330 tended to worsen in patients with a higher NYHA cardiac function classification and poorer
331 disease, that is, patients with lower general health status. Similar findings have been found in
332 other studies, with low consistency between the EQ-5D-5L and SF-6Dv2 in the lower health
333 state and vice versa [32,33]. In addition, the Bland–Altman plots showed that the 95% levels
334 of agreement for both SF-6Dv2 and EQ-5D-5L utilities were higher than the recommended
335 threshold of minimal clinically important difference, implying that clinically significant
336 differences existed between utilities; therefore, the specific tools could not be regarded as
337 interchangeable. Thus, the findings may vary depending on the scale used.

338 Surprisingly, the “pain” domain at baseline was an exception ($\rho = 0.22$), which is
339 in contrast to the results of other studies that concluded that the correlation was strongest in
340 the area of pain [34,35]. An important consideration for inconsistencies is that pain is not the
341 most prominent symptom of HOCM. As expected, the two questionnaires correlated well in
342 similar domains with improved health status, especially six months after discharge. Hence,
343 when participants possessed better physical function, the construct validity was supported by

344 a high degree of convergence. In line with this, the phenomenon frequently appears in the
345 academic literature on the topic of this study. As described below, when describing
346 dimensions across comparable tools or representing the same concept owing to their names,
347 there is a risk that it may lead to inaccurate conclusions [36]. When describing dimensions
348 across instruments as similar or representing the same concept based on their names, there is
349 a risk that it may lead to incorrect conclusions.

350 We validated the known group validity of both SF-6Dv2 and EQ-5D-5L. The RE data
351 showed that the SF-6Dv2 had higher discriminatory efficiency than the EQ-5D-5L version
352 for sociodemographic characteristics other than the NYHA classification and whether they
353 are currently smoking. SF-6Dv2 displayed better construct validity than EQ-5D-5L in a
354 similar study [9]. The reason for the greater efficiency level of SF-6Dv2 in recognizing
355 health-related external indicators may be associated with the features of a more detailed
356 system of description, such as physical function and role limitation dimensions [37].
357 However, further research is required to understand these underlying reasons better.

358 This study fills a data gap in measuring utility values in Chinese patients with HOCM
359 using EQ-5D-5L and SF-6Dv2. Its main advantage is the absence of missing data due to
360 thorough follow-up, which makes the results realistic and reliable. Another advantage of this
361 study is that the EQ-5D-5L and SF-6Dv2 value sets were calculated using the Chinese tariff,
362 which can precisely describe the characteristics of Chinese people. These results can be used
363 for future cost-utility and value analyses. The first limitation of this study was the limited
364 sample size and the single-center design. However, it should be considered that hypertrophic
365 obstructive cardiomyopathy is rare and has a relatively low prevalence. Thus, caution should
366 be exercised when making straightforward generalizations and extrapolating the results.
367 Moreover, the questionnaire was completed at baseline using face-to-face interviews and self-
368 reports at follow-up, which may have affected the final validity to a certain degree. Finally,

369 considering that only a few health-related external indicators were examined, a more
370 comprehensive assessment of the content of the health survey should be conducted in the
371 future.

372

373 **Conclusions**

374 The data from Chinese patients with HOCM suggest that the different results obtained
375 for SF-6Dv2 and EQ-5D-5L in the context of CUA were different. Researchers should decide
376 to compare instruments with one another depending on the severity of the problems typically
377 encountered in each area of the disease under study. Notably, considering the distribution
378 advantage, ceiling effect, and discriminant validity, the SF-6Dv2 is appropriate for evaluating
379 patients with HOCM.

380

381 **Acknowledgements**

382 The authors sincerely thank all the inpatients who participated in the questionnaire,
383 the assistants who helped us with the questionnaire, and the people who helped us with the
384 mapping: Chongwen Gui (Henan University of Technology), Bowen Xie (Johnson &
385 Johnson), and Jing Liu (Novartis).

386

387 **References**

- 388 1. Ommen SR, Mital S, Burke MA, Day SM, Deswal A, Elliott P, et al. 2020 AHA/ACC
389 guideline for the diagnosis and treatment of patients with hypertrophic
390 cardiomyopathy[J/OL]. *J Am Coll Cardiol*. 2020;76:e159-240.
391 DOI:10.1016/j.jacc.2020.08.045.

- 392 2. Heitner SB, Jacoby D, Lester SJ, Owens A, Wang A, Zhang D, et al. Mavacamten
393 treatment for obstructive hypertrophic cardiomyopathy: A clinical trial. *Ann Intern*
394 *Med.* 2019;170:741–8. DOI:10.7326/M18-3016.
- 395 3. Bai Y, Zheng JP, Lu F, Zhang XL, Sun CP, Guo WH, et al. Prevalence, incidence and
396 mortality of hypertrophic cardiomyopathy based on a population cohort of 21.9
397 million in China. *Sci Rep.* 2022;12:18799. DOI:10.1038/s41598-022-20042-9.
- 398 4. Maron BJ, Desai MY, Nishimura RA, Spirito P, Rakowski H, Towbin JA, et al.
399 Management of Hypertrophic Cardiomyopathy: JACC State-of-the-Art Review. *J Am*
400 *Coll Cardiol.* 2022;79(4):390-414. doi: 10.1016/j.jacc.2021.11.021.
- 401 5. Makavos G, Kairis C, Tselegkidi ME, Karamitsos T, Rigopoulos AG, Noutsias M, et
402 al. Hypertrophic cardiomyopathy: An updated review on diagnosis, prognosis, and
403 treatment. *Heart Fail Rev.* 2019;24:439-59. DOI:10.1007/s10741-019-09775-4.
- 404 6. Xie J, Wang Y, Xu Y, Fine JT, Lam J, Garrison LP. Assessing health-related quality-
405 of-life in patients with symptomatic obstructive hypertrophic cardiomyopathy: EQ-
406 5D-based utilities in the Explorer-HCM trial. *J Med Econ.* 2022;25:51-8.
407 DOI:10.1080/13696998.2021.2011301.
- 408 7. Schoonvelde SAC, Wiethoff I, Hiligsmann M, Evers SMAA, Michels M. Quality of
409 life and societal costs in hypertrophic cardiomyopathy: Protocol of the AFFECT-
410 HCM study. *Neth Heart J.* 2023 [2023-01-26].
411 <https://link.springer.com/10.1007/s12471-022-01753-0>. DOI:10.1007/s12471-022-
412 01753-0.
- 413 8. Murphy RP, Boyce CJ, Dolan P, Wood AM. Valuing the Q in QALYs: Does
414 providing patients' ratings affect population values? *Health Psychol.* 2020;39:37-45.
415 DOI:10.1037/hea0000806.

- 416 9. Abdin E, Chong SA, Seow E, Peh CX, Tan JH, Liu J, et al. A comparison of the
417 reliability and validity of SF-6D, EQ-5D and HUI3 utility measures in patients with
418 schizophrenia and patients with depression in Singapore. *Psychiatry Res.*
419 2019;274:400-8. DOI:10.1016/j.psychres.2019.02.077.
- 420 10. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L
421 value set for China. *Value Health.* 2017;20:662-9. DOI:10.1016/j.jval.2016.11.016.
- 422 11. Wu J, Xie S, He X, Chen G, Bai G, Feng D, et al. Valuation of SF-6Dv2 health states
423 in China using time trade-off and discrete-choice experiment with a duration
424 dimension. *Pharmacoeconomics.* 2021;39:521-35. DOI:10.1007/s40273-020-00997-1.
- 425 12. Franco KFM, Cabral CMN, Salvador EMES, Miyamoto GC. Comparison between
426 different health state utility instruments in patients with fibromyalgia. *Braz J Phys*
427 *Ther.* 2021;25:573-82. DOI:10.1016/j.bjpt.2021.02.006.
- 428 13. Stolz M, Albus C, Beutel ME, Deter HC, Fritzsche K, Herrmann-Lingen C, et al.
429 Assessment of health-related quality of life in individuals with depressive symptoms:
430 Validity and responsiveness of the EQ-5D-3L and the SF-6D. *Eur J Health Econ.*
431 2022 [2022-11-22]. <https://link.springer.com/10.1007/s10198-022-01543-w>.
432 DOI:10.1007/s10198-022-01543-w.
- 433 14. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development
434 and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual*
435 *Life Res.* 2011;20:1727-36. DOI:10.1007/s11136-011-9903-x.
- 436 15. Devlin N, Parkin D, Janssen B. *Methods for analysing and reporting EQ-5D*
437 *data[M/OL]. Cham: Springer International Publishing; 2020 [2022-11-16].*
438 <http://link.springer.com/10.1007/978-3-030-47622-9>. DOI:10.1007/978-3-030-47622-
439 9.

- 440 16. Poder TG, Fauteux V, He J, Brazier JE. Consistency Between Three Different Ways
441 of Administering the Short Form 6 Dimension version 2. *Value Health*. 2019;22:837-
442 42. DOI:10.1016/j.jval.2018.12.012.
- 443 17. Brazier JE, Mulhern BJ, Bjorner JB, Gandek B, Rowen D, Alonso J, et al. Developing
444 a New Version of the SF-6D Health State Classification System From the SF-36v2:
445 SF-6Dv2. *Med Care New Version: SF-6Dv2*. 2020;58:557-65.
446 DOI:10.1097/MLR.0000000000001325.
- 447 18. Walters SJ, Brazier JE. Comparison of the minimally important difference for two
448 health state utility measures: EQ-5D and SF-6D. *Qual Life Res*. 2005;14:1523-32.
449 DOI:10.1007/s11136-004-7713-0.
- 450 19. Giavarina D. Understanding bland Altman analysis. *Biochem Med*. 2015;25:141-51.
451 DOI:10.11613/BM.2015.015.
- 452 20. Hazra A, Gogtay N. Biostatistics series module 6: Correlation and linear regression.
453 *Indian J Dermatol*. 2016;61:593–601. DOI:10.4103/0019-5154.193662.
- 454 21. Heslin M, Chua KC, Trevillion K, Nath S, Howard LM, Byford S. Psychometric
455 properties of the five-level EuroQoL-5 dimension and Short Form-6 dimension
456 measures of health-related quality of life in a population of pregnant women with
457 depression. *BJPsych Open*. 2019;5:e88. DOI:10.1192/bjo.2019.71.
- 458 22. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and
459 SF-6D based on hypothetical preferences in a general population. *Health Econ*.
460 2005;14:1169-89. DOI:10.1002/hec.1006.
- 461 23. Ye Z, Sun L, Wang Q. A head-to-head comparison of EQ-5D-5 L and SF-6D in
462 Chinese patients with low back pain. *Health Qual Life Outcomes*. 2019;17:57.
463 DOI:10.1186/s12955-019-1137-6.

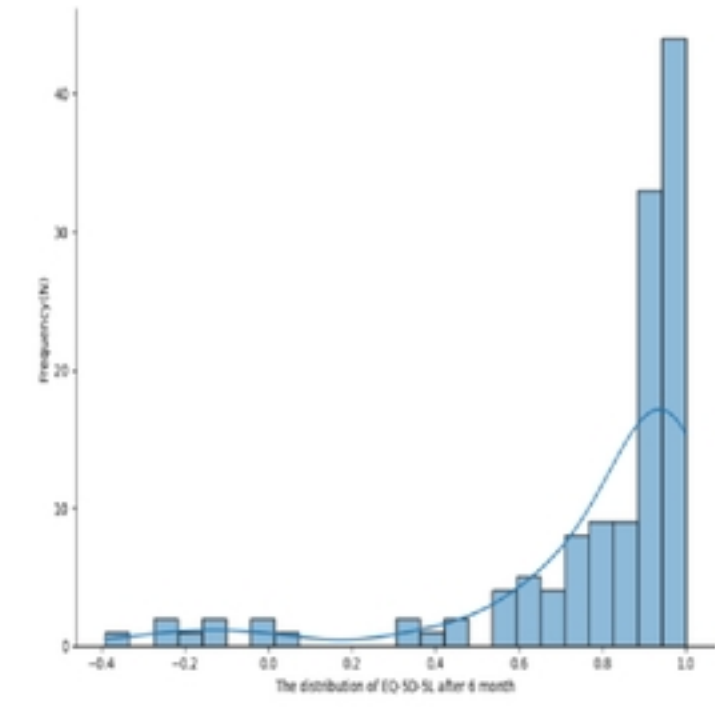
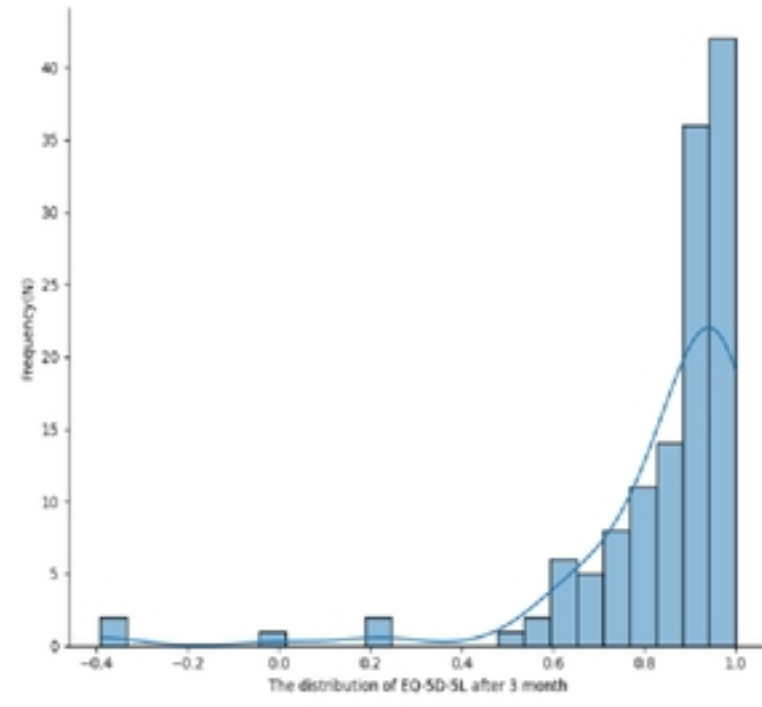
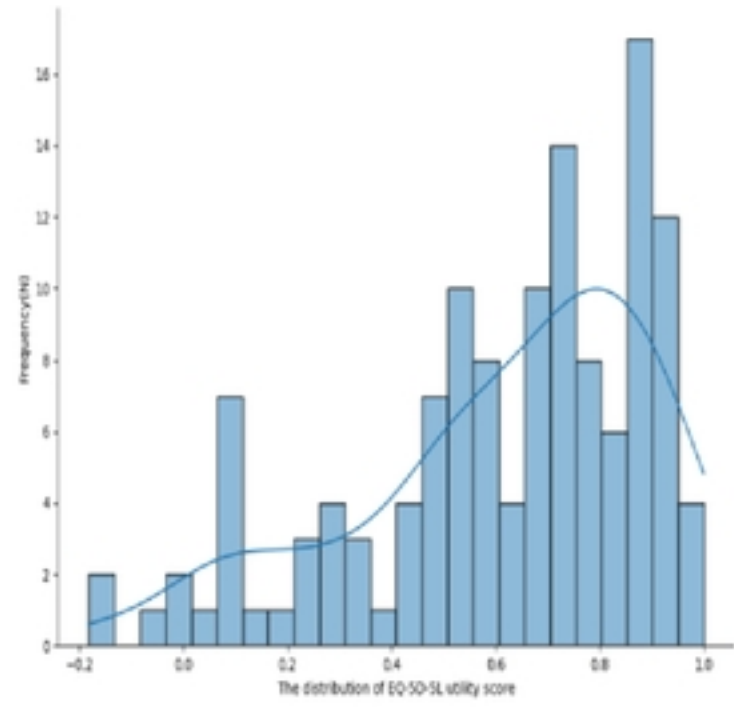
- 464 24. Bjerk M, Brovold T, Davis JC, Bergland A. Evaluating a falls prevention intervention
465 in older home care recipients: A comparison of SF-6D and EQ-5D. *Qual Life Res.*
466 2019;28:3187-95. DOI:10.1007/s11136-019-02258-x.
- 467 25. Li S, Wang M, Liu L, Chen G. Which approach is better in eliciting health state
468 utilities from breast cancer patients? Evidence from mainland China. *Eur J Cancer*
469 *Care (Engl)*. 2019;28:e12965. DOI:10.1111/ecc.12965.
- 470 26. Li N, Boonen A, van den Bergh JP, van Kuijk SMJ, Wyers CE, van Oostwaard M, et
471 al. A head-to-head comparison of EQ-5D-5L and SF-6D in Dutch patients with
472 fractures visiting a Fracture Liaison Service. *J Med Econ*. 2022;25:829-39.
473 DOI:10.1080/13696998.2022.2087409.
- 474 27. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to
475 assess their ceiling effects in the US general population. *Value Health*. 2006;9:262-71.
476 DOI:10.1111/j.1524-4733.2006.00108.x.
- 477 28. Kangwanrattanakul K. A comparison of measurement properties between UK SF-6D
478 and English EQ-5D-5L and Thai EQ-5D-5L value sets in general Thai population.
479 *Expert Rev Pharmacoecon Outcomes Res*. 2021;21:765-74.
480 DOI:10.1080/14737167.2021.1829479.
- 481 29. Eneqvist T, Nemes S, Kärrholm J, Burström K, Rolfson O. How do EQ-5D-3L and
482 EQ-5D-5L compare in a Swedish total hip replacement population? *Acta Orthop*.
483 2020;91:272-8. DOI:10.1080/17453674.2020.1746124.
- 484 30. Xu RH, Dong D, Luo N, Wong EL, Wu Y, Yu S, et al. Evaluating the psychometric
485 properties of the EQ-5D-5L and SF-6D among patients with haemophilia. *Eur J*
486 *Health Econ*. 2021;22:547-57. DOI:10.1007/s10198-021-01273-5.
- 487 31. Kularatna S, Senanayake S, Gunawardena N, Graves N. Comparison of the EQ-5D 3L
488 and the SF-6D (SF-36) contemporaneous utility scores in patients with chronic kidney

- 489 disease in Sri Lanka: A cross-sectional survey. *BMJ Open*. 2019;9:e024854.
490 DOI:10.1136/bmjopen-2018-024854.
- 491 32. Sun CY, Liu Y, Zhou LR, Wang MS, Zhao XM, Huang WD, et al. Comparison of
492 EuroQol-5D-3L and short Form-6D utility scores in family caregivers of colorectal
493 cancer patients: A cross-sectional survey in China. *Front Public Health*.
494 2021;9:742332. DOI:10.3389/fpubh.2021.742332.
- 495 33. Kontodimopoulos N, Stamatopoulou E, Gazi S, Moschou D, Krikelis M, Talias MA.
496 A comparison of EQ-5D-3L, EQ-5D-5L, and SF-6D utilities of patients with
497 musculoskeletal disorders of different severity: A health-related quality of life
498 approach. *J Clin Med*. 2022;11:4097. DOI:10.3390/jcm11144097.
- 499 34. Selva-Sevilla C, Ferrara P, Gerónimo-Pardo M. Interchangeability of the EQ-5D and
500 the SF-6D, and comparison of their psychometric properties in a spinal postoperative
501 Spanish population. *Eur J Health Econ*. 2020;21:649-62. DOI:10.1007/s10198-020-
502 01161-4.
- 503 35. Thuppal S, Markwell S, Crabtree T, Hazelrigg S. Comparison between the EQ-5D-3L
504 and the SF-6D quality of life (QOL) questionnaires in patients with chronic
505 obstructive pulmonary disease (COPD) undergoing lung volume reduction surgery
506 (LVRS). *Qual Life Res*. 2019;28:1885-92. DOI:10.1007/s11136-019-02123-x.
- 507 36. Goodwin PC, Ratcliffe J, Morris J, Morrissey MC. Using the knee-specific Hughston
508 Clinic Questionnaire, EQ-5D and SF-6D following arthroscopic partial meniscectomy
509 surgery: A comparison of psychometric properties. *Qual Life Res*. 2011;20:1437-46.
510 DOI:10.1007/s11136-011-9880-0.
- 511 37. Seymour J, McNamee P, Scott A, Tinelli M. Shedding new light onto the ceiling and
512 floor? A quantile regression approach to compare EQ-5D and SF-6D responses.
513 *Health Econ*. 2010;19:683–96.

514

515 **Supporting information Caption**

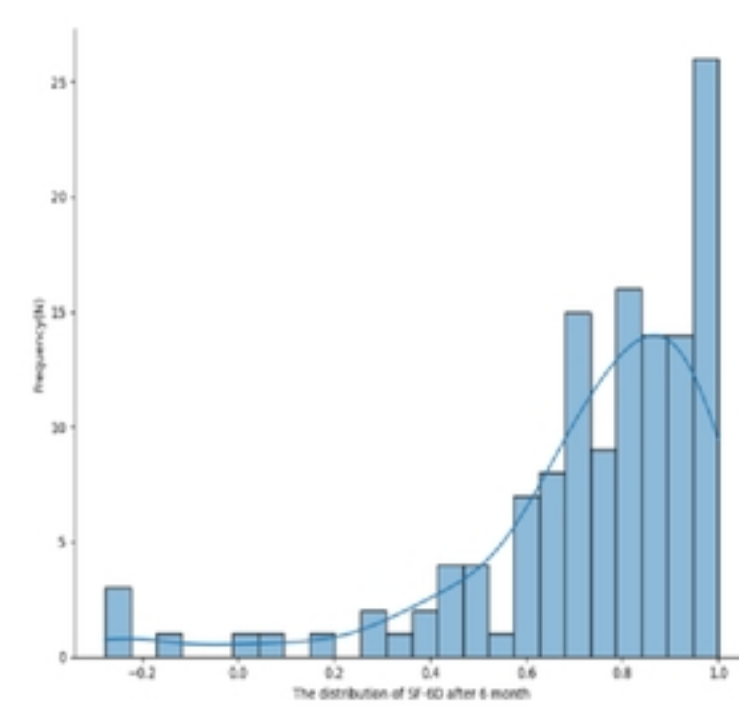
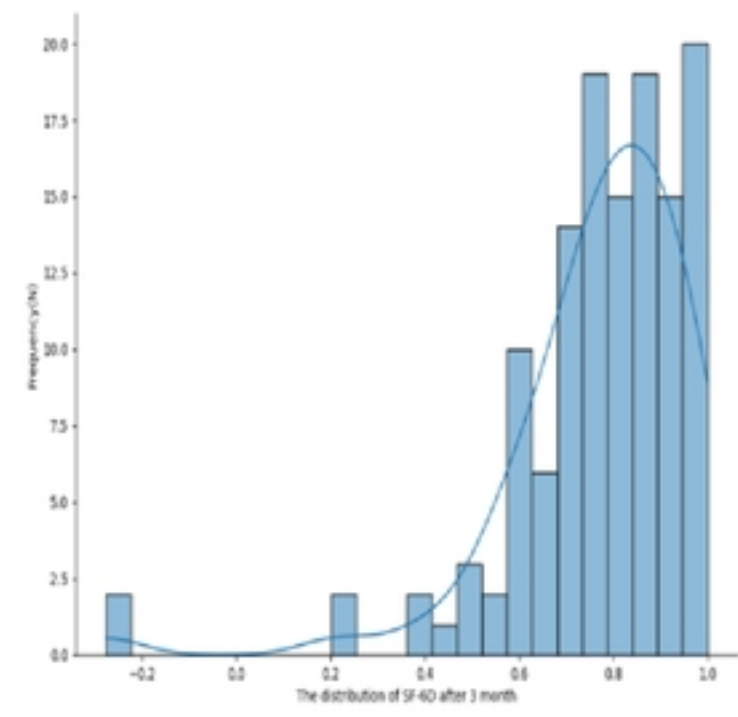
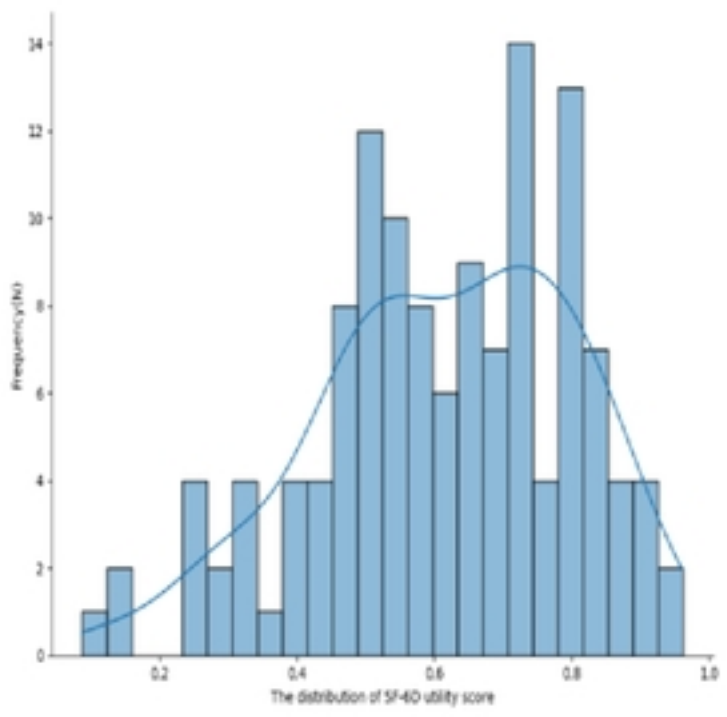
516 **S1 Table.** The ICCs of EQ-5D-5L and SF-6Dv2 index scores



(A)

(B)

(C)



(D)

(E)

(F)

Figure 1

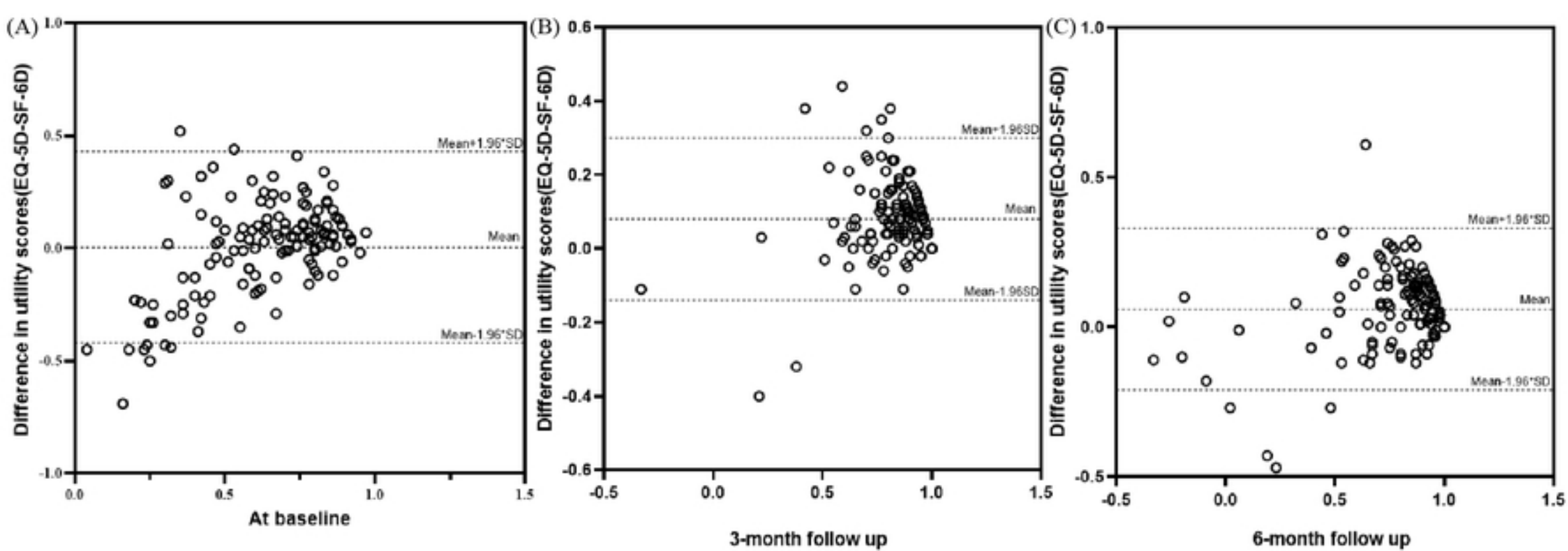


Figure 2