Validation of the EQ-5D in a general population sample in urban China

**Hong-Mei Wang · Donald L. Patrick ·
Todd C. Edwards · Anne M. Skalicky ·
Hai-Yan Zeng · Wen-Wen Gu**

Abstract

Purpose To evaluate the reliability and validity of the EQ-5D in a general population sample in urban China.

Methods Thousand and eight hundred respondents in 18 communities of Hangzhou, China were recruited by multi-stage stratified random sampling. Respondents self-administered a questionnaire including the EQ-5D, the SF-36, and demographic questions. Test–retest reliability at 2-week intervals was evaluated using Kappa coefficient, the intraclass correlation coefficient. The standard error of measurement (SEM) was used to indicate the absolute measurement error. Construct validity was established using convergent, discriminant, and known groups analyses.

Results Complete data for all EQ-5D dimensions were available for 1,747 respondents (97%). Kappa values were from 0.35 to 1.0. The ICCs of test–retest reliability were 0.53 for the EQ-5D index score and 0.87 for the EQ VAS score. The SEM values were 0.13 (9.22% range) and 4.20 (4.20% range) for the EQ-5D index and EQ VAS scores, respectively. The Pearson's correlation coefficients between the EQ-5D and the SF-36 were stronger between comparable dimensions than those between less comparable dimensions, demonstrating convergent and discriminant evidence of construct validity. The Chinese EQ-5D distinguished well between known groups: respondents who reported poor general health and chronic diseases had

worse HRQoL than those without. Older people, females, people widowed or divorced, and those with a lower socioeconomic status reported poorer HRQoL. Respondents reporting no problems on any EQ-5D dimension had better scores on the SF-36 summary scores than those reporting problems.

Conclusions The Chinese version of the EQ-5D demonstrated acceptable construct validity and fair to moderate levels of test–retest reliability in an urban general population in China.

Keywords EQ-5D · General population · HRQoL · China

Abbreviations
HRQoL   Health-related quality of life
VAS   Visual analogue scale

Introduction

The EQ-5D is an established health-related quality of life (HRQoL) instrument, used frequently in both clinical trials and health services research [1]. The validity of the Chinese EQ-5D has been assessed in mainland China [2–4] and elsewhere [5–7]; its reliability, however, is not well reported. We evaluated the reliability and validity of the EQ-5D in a sample of the general population in urban China.

Methods

Sample and study design

The survey using a multi-stage stratified random sampling approach was conducted in Hangzhou, China in 2008. Nine

H.-M. Wang (✉) · H.-Y. Zeng · W.-W. Gu
Institute of Social Medicine and Family Medicine, School of Medicine, Zhejiang University, 866 Yuhang Tang Rd., 310058 Hangzhou, People's Republic of China
e-mail: rosa@zju.edu.cn; hongmei@u.washington.edu

D. L. Patrick · T. C. Edwards · A. M. Skalicky
Department of Health Services, University of Washington, 1208 43rd Street, Box 359455, Seattle, WA 98195, USA

"Jiedao" (sub-district neighborhood) were randomly selected from Xiacheng district (central), Gongshu district (sub-central), and Yuhang district (suburb), three for each. Two communities from each "Jiedao" and 70 households from each community were randomly selected. The total sample size was 1,800, with 200 in each "Jiedao". All residents 14 years old and above, living in a sampled household for at least 6 months were eligible to participate until the quota for each "Jiedao" was met. Participants self-administered a questionnaire containing the Chinese EQ-5D and SF-36. Trained interviewers administered questions regarding existence of chronic diseases. Sixty respondents were randomly sampled among respondents on the first survey day that would be willing to self-administer the EQ-5D and SF-36 in a 2-week period. Written consent was obtained from all respondents for this study approved by Zhejiang University School of Medicine Ethics Committee.

The EQ-5D comprises five health dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/ depression, each with three response categories: no, some or extreme problems) and a 0–100 points visual analogue scale (EQ VAS) [8]. Scores for the five dimensions can be converted into a utility index score by applying the scores from preference weights elicited from the UK general population [9]. The SF-36 is a validated [10, 11] 36-item instrument yielding eight scales and two summary scores. Higher scores indicate better health status.

Data analysis

Reliability and validity of the EQ-5D were assessed according to established guidelines [12]. To evaluate reliability, it was assumed that the health status between two measurements was stable. The percentages of agreement and kappa coefficients for the five dimensions were calculated. Kappa values below 0.2 indicate a slight agreement, 0.21–0.4 fair, 0.41–0.6 moderate, 0.61–0.8 substantial, and 0.81–1.0 an almost perfect agreement [13]. Test–retest reliability of the EQ-5D index and EQ VAS scores was determined using the intraclass correlation coefficient (ICC; two-way mixed-effect model/absolute agreement definition $ICC_{2,1}$) [14]. An ICC greater than 0.70 is considered appropriate for group comparison [15]. The standard error of measurement ($SEM_{agreement}$) was used to assess variability, i.e., the absolute measurement error [16, 17]. It was also expressed as a percent of the measurement range (SEM%) likely to be encountered in actual research [12].

To evaluate construct validity, we first assessed convergent and discriminant evidence by examining relationships with the SF-36 using Pearson's correlation. It was expected that comparable dimensions, e.g., EQ-5D pain/ discomfort and SF-36 bodily pain, would correlate better, compared with less comparable dimensions, such as EQ-5D mobility and SF-36 mental health. Pearson's correlation coefficients of 0.50 or above were regarded as strong, 0.30–0.49 as moderate, and lower than 0.30 as weak [18].

Second, construct validity was assessed by comparing the EQ-5D index (both the UK [9] and Japanese preference weights [19] were used) and EQ VAS scores for subgroups of respondents with differing self-reported overall health and number of chronic diseases using ANOVA. It was also expected that older people, females, people widowed or divorced, and those with a lower socioeconomic status would report poorer health [2, 3, 20–22]. The relationships between the EQ-5D and the demographic variables were examined using ANOVA, t-test, or chi-square test. Finally, Mean SF-36 summary scores for respondents reporting no problems for any EQ-5D dimension were compared with those for respondents reporting problems using t-test, and higher SF-36 scores was expected in the first case [23].

Results

Among the 1,800 respondents from 1,260 selected households, complete data for all EQ-5D dimensions were available for 1,747 respondents (97%) and analyzed for the present study. The estimated response rate was 71.4% (two eligible individuals in each household on average [11]). The mean age was 47.5 years (SD 17.5, range 14–99), with 51.6% being women. Compared with Hangzhou urban area demographic statistics for year 2008 [24], our sample had similar sex ratio, older age, and higher educational attainment (Table 1).

EQ-5D response

The majority of respondents reported no problems (ceiling effects), ranging from 78.0% for the pain/discomfort dimension to 96.7% for the self-care dimension (Table 2). The mean EQ-5D index score was 0.92 (SD 0.17, range − 0.59 to 1), and the mean EQ VAS score was 84.44 (SD 13.0, range 8.50–100).

Test–retest reliability

In the retest samples, 48 of 60 respondents returned the retest questionnaire, and the data from 31 respondents whose scores of the first question of the SF-36 (self-reported overall health) was the same at two measurements were analyzed. The median interval of test–retest measurement was 13 days (interquartile range: 12–15 days). Kappa values for EQ-5D items regarding mobility, self-care, usual activities, pain/discomfort, and anxiety/depression between

Table 1 Characteristics of the study sample ($n = 1,747$)

	Sample		Urban population[b]
	n	%	%
Age group (years)			
14–44	743	42.9	56.3
45–64	680	39.3	30.9
65+	309	17.8	12.8
Gender			
Male	837	48.4	50.4
Female	894	51.6	49.6
Marital status			
Single	265	15.4	
Married/co-habiting	1,361	79.0	
Widowed	74	4.3	
Divorced/separated	23	1.3	
Education level[a]			
Low level	322	18.8	
Middle level	855	49.8	
High level	539	31.4	16.6[c]
Annual household income			
($1US = 6.8Yuan)			
<60,000 Yuan	1,261	80.6	66,572 Yuan[d]
≥60,000 Yuan	303	19.4	
Employment status			
Employed or self-employed	796	48.0	53.3
Other	863	52.0	46.7
Self-reported overall health			
Excellent	218	12.5	
Very good	463	26.6	
Good	495	28.4	
Fair	524	30.1	
Poor	40	2.3	
Number of chronic diseases			
0	711	41.1	
1	666	38.5	
2+	354	20.4	

[a] *Education* low level included no school and elementary school; middle level included middle school and high school or vocational training; high level referred to college or above

[b] Estimates for the urban population are based on Hangzhou statistical yearbook 2008

[c] Percentage of talented personnel among population of 14 years and older

[d] Annual household income for urban population in Hangzhou, China

measurements were 1.00, 0.65, 0.87, 0.35, and 0.63, respectively. The ICCs of test–retest reliability were 0.53 for the EQ-5D index score and 0.87 for the EQ VAS score, respectively. The SEM values (SEM%) were 0.13 (9.22%) and 4.20 (4.20%) for the EQ-5D index and EQ VAS scores, respectively, (Table 3).

Table 2 Distribution of responses to EQ-5D dimensions

EQ-5D dimension	Response (%)		
	No problems	Some problems	Extreme problems
Mobility	1,619 (92.7)	119 (6.8)	9 (0.5)
Self-care	1,690 (96.7)	47 (2.7)	10 (0.6)
Usual activities	1,628 (93.2)	108 (6.2)	11 (0.6)
Pain/discomfort	1,363 (78.0)	367 (21.0)	17 (1.0)
Anxiety/ depression	1,532 (87.7)	204 (11.7)	11 (0.6)

Validity

The Pearson's correlation coefficients between the EQ-5D and the SF-36 were stronger between comparable dimensions (e.g., −0.59 between EQ-5D pain/discomfort and SF-36 BP and −0.44 between EQ-5D mobility and SF-36 PF) than those between less comparable dimensions (e.g., −0.26, −0.20 between EQ-5D mobility, self-care, and SF-36 MH, respectively) with a few exceptions, demonstrating convergent and discriminant evidence of construct validity. The EQ-5D index and EQ VAS scores had moderate or strong correlations with all SF-36 scores (all $P < 0.001$, Table 4).

Respondents who reported poor general health and chronic diseases had significantly lower EQ-5D index and EQ VAS scores (Table 5). The discrepancy between the UK and the Japanese versions of EQ-5D index scores was much smaller for better health states, but larger for worse health status. Older people, females, people widowed or divorced, and those with a lower socioeconomic status reported poorer HRQoL as expected (Table 6). Respondents reporting no problems on any EQ-5D dimension had better scores on the SF-36 summary scores, respectively, than those reporting problems (all $P < 0.001$, data not shown).

Discussion

The study assessed the reliability and validity of the Chinese EQ-5D in a large urban population in China. Compared with most EQ-5D studies in general population [21, 25, 26], our sample covered 14-to 18-year-old adolescents. The Chinese EQ-5D youth version is not available; it is therefore suitable to apply the EQ-5D in adolescents primarily to allow for follow-up and comparisons over a wide range of ages.

Construct validity of the EQ-5D was demonstrated using convergent, discriminant, and known groups analyses. The EQ-5D showed fair to moderate levels of test–retest

Table 3 Test–retest of reliability of the EQ-5D ($n = 31$)

Dimension	Agreement between test and retest	Kappa	95% CI
Mobility	31 (100.0%)	1.00	1.00–1.00
Self-care	30 (96.7%)	0.65	0.02–1.00
Usual activities	30 (96.8%)	0.87	0.62–1.00
Pain/discomfort	23 (74.2%)	0.35	0.04–0.66
Anxiety/depression	29 (93.6%)	0.63	0.16–1.00

	Mean difference (standard deviation) between test and retest	ICC	95% CI	SEM$_{agreement}$
EQ-5D index	0.02 (0.18)	0.53	0.22–0.74	0.13
EQ VAS	−1.19 (5.91)	0.87	0.76–0.94	4.20

ICC intraclass correlation coefficient; *CI* confidence interval; *SEM$_{agreement}$* standard error of measurement including systematic differences

Table 4 Correlations between the EQ-5D and SF-36

SF-36 Scales/summaries	EQ-5D dimensions					EQ-5D index	EQ VAS
	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression		
PF	−0.44	−0.36	−0.40	−0.32	−0.26	0.42	0.42
RP	−0.39	−0.28	−0.38	−0.39	−0.33	0.44	0.42
BP	−0.47	−0.36	−0.44	**−0.59**	−0.42	**0.62**	**0.53**
GH	−0.35	−0.24	−0.34	−0.39	−0.36	0.43	**0.60**
VT	−0.33	−0.23	−0.32	−0.39	−0.38	0.43	**0.59**
SF	−0.35	−0.32	−0.35	−0.26	−0.34	0.40	0.34
RE	−0.29	−0.26	−0.31	−0.30	−0.42	0.41	0.33
MH	−0.26	−0.20	−0.22	−0.24	−0.37	0.32	0.43
PCS	−0.41	−0.33	−0.41	−0.39	−0.34	0.47	0.35
MCS	−0.30	−0.22	−0.28	−0.34	−0.36	0.39	**0.55**

Strong correlations (≥ 0.50) indicated in bold. Pearson's correlation, $P < 0.001$

PF physical functioning, *RP* role physical, *BP* bodily pain, *GH* general health, *VT* vitality, *SF* social functioning, *RE* role emotional, *MH* mental health, *PCS* physical component summary, *MCS* mental component summary

Table 5 Comparison of EQ-5D index and EQ VAS scores for subgroups of respondents with differing health status

Health condition	N	EQ-5D index[a]	EQ-5D index[b]	EQ VAS	Assumption
Self-reported overall health					
Excellent	218	0.99	0.99	94.92	
Very good	463	0.98	0.97	90.76[d]	Excellent > Very good
Good	495	0.94	0.93	84.53	Very good > Good[d]
Fair	524	0.87	0.86	76.95	Good > Fair[d]
Poor	40	0.39	0.51	51.26	Fair > Poor[d]
Number of chronic diseases					
0	711	0.95	0.95	87.88	
1	666	0.92	0.91	82.94	0 > 1[d]
2+	354	0.89	0.88	80.24	1 > 2+[c]

[a] Using UK TTO value set

[b] Using Japanese plain main effects model; ANOVA

[c] $P < 0.05$

[d] $P < 0.001$

Table 6 Responses to the EQ-5D by sociodemographic variables ($n = 1,747$)

Variables	N	EQ-5D index		EQ VAS		Percentages of reporting problems (some or extreme)				
		Mean	SD	Mean	SD	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression
Age group(age)		d		d		d	d	d	d	d
14–44	743	0.96	0.11	89.38	10.02	1.9	0.9	2.0	10.1	9.7
45–64	680	0.92	0.17	83.15	12.46	5.7	2.5	6.2	25.7	11.9
65+	309	0.84	0.23	75.58	14.92	23.6	10.7	19.4	42.4	19.7
Gender		b		c		c			c	
Male	837	0.93	0.16	85.33	12.64	5.6	3.0	6.2	18.9	11.6
Female	894	0.91	0.18	83.64	13.24	8.9	3.5	7.4	24.7	13.1
Marital status		d		d		d	d	d	d	b
Single	265	0.96	0.15	89.72	11.49	3.0	2.3	4.9	9.1	8.7
Married/co-habiting	1,361	0.92	0.17	84.14	12.58	6.6	2.8	6.2	23.2	12.5
Widowed	74	0.81	0.23	70.98	14.38	35.1	14.9	25.7	48.6	21.6
Divorced/separated	23	0.93	0.13	82.52	13.43	8.7	4.3	4.3	17.4	17.4
Education level[a]		d		d		d	d	d	d	
Low level	322	0.87	0.22	79.97	14.61	15.2	8.4	13.7	35.1	13.4
Middle level	855	0.92	0.17	84.47	13.20	6.8	2.6	6.4	22.6	13.3
High level	539	0.95	0.12	87.01	10.73	3.0	1.1	3.0	13.5	9.5
Annual household income		d		d		c	b	c	c	
<60,000 Yuan	1,261	0.91	0.18	83.84	13.70	8.9	3.8	8.0	24.4	13.3
≥60,000 Yuan	303	0.94	0.12	86.69	11.26	3.3	1.3	3.6	16.2	11.6
Employment status		d		d		d	d	d	d	d
Employed or self-employed	796	0.96	0.11	87.53	10.63	1.8	0.9	2.4	12.8	9.0
Other	863	0.88	0.21	81.34	14.35	12.6	5.3	11.0	31.3	16.1

[a] *Education* low level included no school and elementary school; middle level included middle school and high school or vocational training; high level referred to college or above

The significance of overall variable with ANOVA or *t*-test or chi-square test

[b] $P < 0.05$

[c] $P < 0.01$

[d] $P < 0.001$

reliability, with high percentage of respondents reporting same level of problems in the dimensions and satisfactory ICC for the EQ VAS score. However, the examination of reliability was compromised by high ceiling effects. Reliability coefficients not only reflect the degree of agreement between repeated measures, but also the degree to which a measurement instrument can differentiate among individuals [17, 27]. In a homogeneous population, the within-subject variance can easily overwhelm the between-subject variance, making for low reliability [28]. The SEM is relatively sample-independent and useful in the interpretation of HRQoL change [29, 30]. A higher SEM value for the EQ VAS score after 1 month was reported recently [31]. When applying Japanese preference weights [19], the ICC, SEM value, and SEM% for the EQ-5D index score were 0.64, 0.09, and 8.11%, respectively.

There are several studies where the EQ-5D has been used among the Chinese general population. Wang et al. [2] measured EQ-5D data among 2,994 respondents from one district of Beijing. Recently, Sun et al. [3] analyzed national EQ-5D data and provided norms for the Chinese general population. The reliability of EQ-5D was not measured in these two studies. Chang et al. [32] reported validation results in a representative sample of the 20–64 years Taiwanese population. Similar ICCs were reported (0.51 for the EQ-5D index score and 0.70 for the EQ VAS score), even though people more than 65 years old were not recruited.

This study had limitations. First, although there were a small number of non-respondents due to refusal or inaccessibility after three visits, no data were available for them and it is unclear whether characteristics of the non-respondents differed from the respondents. Second, although the estimated response rate was high, there might be selection or response bias [33]. Third, the retested sample size was small for the assessment of the reliability. Fourth, our sample was more representative of an older and educated general population.

We conclude that the Chinese EQ-5D demonstrated acceptable construct validity and fair to moderate levels of test–retest reliability in an urban general population in China.

Conflict of interest The authors declare no conflict of interests.

References

1. Rabin, R., & de Charro, F. (2001). EQ-5D: A measure of health status from the EuroQol Group. *Annals of Medicine, 33*(5), 337–343.
2. Wang, H., Kindig, D. A., & Mullahy, J. (2005). Variation in Chinese population health related quality of life: results from a EuroQol study in Beijing, China. *Quality of Life Research, 14*(1), 119–132.
3. Sun, S., Chen, J., Johannesson, M., Kind, P., Xu, L., Zhang, Y., et al. (2011). Population health status in China: EQ-5D results, by age, sex and socio-economic status, from the national health services survey 2008. *Quality of Life Research, 20*(3), 309–320.
4. Zhao, F. L., Yue, M., Yang, H., Wang, T., Wu, J. H., & Li, S. C. (2010). Validation and comparison of EuroQol and short form 6D in chronic prostatitis patients. *Value Health, 13*(5), 649–656.
5. Lubetkin, E. I., Jia, H., & Gold, M. R. (2004). Construct validity of the EQ-5D in low-income Chinese American primary care patients. *Quality of Life Research, 13*(8), 1459–1468.
6. Leung, B., Luo, N., So, L., & Quan, H. (2007). Comparing three measures of health status (perceived health with Likert-type scale, EQ-5D, and number of chronic conditions) in Chinese and white Canadians. *Medical Care, 45*(7), 610–617.
7. Luo, N., Chew, L. H., Fong, K. Y., Koh, D. R., Ng, S. C., Yoon, K. H., et al. (2003). Validity and reliability of the EQ-5D self-report questionnaire in Chinese-speaking patients with rheumatic diseases in Singapore. *Ann Acad Med Singap, 32*(5), 685–690.
8. Brooks, R. (1996). EuroQol: The current state of play. *Health Policy, 37*(1), 53–72.
9. Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35*(11), 1095–1108.
10. Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health survey manual and interpretation guide*. Boston, MA: The Health Institute, New England Medical Center.
11. Li, L., Wang, H. M., & Shen, Y. (2003). Chinese SF-36 health Survey: Translation, cultural adaptation, validation, and normalisation. *Journal of Epidemiology and Community Health, 57*(4), 259–263.
12. Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research, 11*(3), 193–205.
13. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.
14. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428.
15. Staquet, M. J., Hays, R. D., & Fayers, P. M. (1998). *Quality of life assessment in clinical trials: Methods and practice* (pp. 169–182). Oxford: Oxford University Press.
16. Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34–42.
17. de Vet, H. C., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology, 59*(10), 1033–1039.
18. Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.
19. Tsuchiya, A., Ikeda, S., Ikegami, N., Nishimura, S., Sakai, I., Fukuda, T., et al. (2002). Estimating an EQ-5D population value set: the case of Japan. *Health Economics, 11*(4), 341–353.
20. Johnson, J. A., & Coons, S. J. (1998). Comparison of the EQ-5D and SF-12 in an adult US sample. *Quality of Life Research, 7*(2), 155–166.
21. Kind, P., Dolan, P., Gudex, C., & Williams, A. (1998). Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ, 316*(7133), 736–741.
22. Kontodimopoulos, N., Pappa, E., Niakas, D., Yfantopoulos, J., Dimitrakaki, C., & Tountas, Y. (2008). Validity of the EuroQoL (EQ-5D) instrument in a Greek general population. *Value Health, 11*(7), 1162–1169.
23. Brazier, J., Jones, N., & Kind, P. (1993). Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Quality of Life Research, 2*(3), 169–180.
24. Hangzhou Bureau of Statistics. (2008). *Hangzhou statistical yearbook.* http://www.hzstats.gov.cn/web/. Accessed 15th September 2010.
25. Fryback, D. G., Dunham, N. C., Palta, M., Hanmer, J., Buechner, J., Cherepanov, D., et al. (2007). US norms for six generic health-related quality-of-life indexes from the National Health Measurement study. *Medical Care, 45*(12), 1162–1170.
26. König, H. H., Bernert, S., Angermeyer, M. C., Matschinger, H., Martinez, M., Vilagut, G., et al. (2009). Comparison of population health status in six european countries: Results of a representative survey using the EQ-5D questionnaire. *Medical Care, 47*(2), 255–261.
27. Kottner, J. (2009). Interrater reliability and the kappa statistic: a comment on Morris et al. (2008). *International Journal of Nursing Studies, 46*(1), 140–141.
28. Bartko, J. J. (1991). Measurement and reliability: statistical thinking considerations. *Schizophrenia Bulletin, 17*(3), 483–489.
29. Terwee, C. B., Roorda, L. D., Knol, D. L., De Boer, M. R., & De Vet, H. C. (2009). Linking measurement error to minimal important change of patient-reported outcomes. *Journal of Clinical Epidemiology, 62*(10), 1062–1067.
30. de Boer, M. R., de Vet, H. C., Terwee, C. B., Moll, A. C., Völker-Dieben, H. J., & van Rens, G. H. (2005). Changes to the subscales of two vision-related quality of life questionnaires are proposed. *Journal of Clinical Epidemiology, 58*(12), 1260–1268.
31. Mannion, A. F., Boneschi, M., Teli, M., Luca, A., Zaina, F., Negrini, S., et al. (2011). Reliability and validity of the cross-culturally adapted Italian version of the core outcome measures index. *Eur Spine J.* [Epub ahead of print].
32. Chang, T. J., Tarn, Y. H., Hsieh, C. L., Liou, W. S., Shaw, J. W., & Chiou, X. G. (2007). Taiwanese version of the EQ-5D: validation in a representative sample of the Taiwanese population. *Journal of the Formosan Medical Association, 106*(12), 1023–1031.
33. Sjöström, O., & Holst, D. (2002). Validity of a questionnaire survey: response patterns in different subgroups and the effect of social desirability. *Acta Odontologica Scandinavica, 60*(3), 136–140.