Quality of Life Research (2019) 28:2851–2857
https://doi.org/10.1007/s11136-019-02225-6

BRIEF COMMUNICATION

Evaluation of the responsiveness of Short Form-12 Health Survey version 2 (SF-12v2) in Chinese patients with hypertension in primary care

Eric Yuk Fai Wan[1] · Esther Yee Tak Yu[1] · Weng Yee Chin[1] · Edmond Pui Hang Choi[2] · Tingting Wu[1] · Cindy Lo Kuen Lam[1]

Accepted: 1 June 2019 / Published online: 5 June 2019

Abstract

Purpose There is limited evidence on the responsiveness of the Short Form-12 Health Survey version 2 (SF-12v2) in hypertensive patients. This study aimed to evaluate both the responsiveness of the SF-12 measures in Chinese hypertensive patients.
Methods A prospective longitudinal study was conducted on hypertensive patients managed in public primary care clinics between 2012 and 2013. A total of 583 and 431 patients were surveyed and completed SF-12v2 at baseline and at 12-month follow-up interviews, respectively. Using global rating of change scale as an external anchor, the responsiveness was assessed by linear mixed effect models, multiple linear regression models, and receiver operating characteristic (ROC) curve analysis.
Results SF-12v2 managed to detect negative changes among hypertensive patients in worsened general health group but failed to identify changes among hypertensive patients in improved general health group. Meanwhile, some domains of SF-12v2 detected a significant difference in difference between patients of worsened and stable/improved group and between patients of stable and improved group, but none of the domains and the summary scales reached the recommended standard of 0.7 in any comparisons in ROC analysis.
Conclusions The SF-12v2 was responsive to worsening of HRQOL but not to improvements in HRQOL among hypertensive patients. The overall responsiveness of SF-12v2 in hypertensive patients is unsatisfactory. Further studies are needed to identify HRQOL measures with good internal and external responsiveness for hypertensive patients.

Keywords Health-related quality of life (HRQOL) · Primary care · Responsiveness · Short Form-12 Health Survey version 2 (SF-12v2) · Hypertension (HT)

Abbreviations

BP	Bodily pain
DM	Diabetes mellitus
ES	Effect size
GH	General health
GRS	Global rating of change scale
HA	Hospital authority
HRQOL	Health-related quality of life
HT	Hypertension
ICPC-2	International classification of primary care-2
MCS	Mental component summary
MH	Mental health
PCS	Physical
PF	Physical functioning
RE	Role emotional
RP	Role physical
SD	Standard deviation
SF	Social functioning
SF-12	12-item Short Form Health Survey
SF-12 v2	SF-12 Health Survey Version 2
VT	Vitality

Eric Yuk Fai Wan and Esther Yee Tak Yu are the Joint first authors.

✉ Eric Yuk Fai Wan
yfwan@hku.hk

✉ Esther Yee Tak Yu
ytyu@hku.hk

[1] Department of Family Medicine and Primary Care, The University of Hong Kong, 3/F Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong, Hong Kong

[2] School of Nursing, The University of Hong Kong, Hong Kong, Hong Kong

Introduction

Hypertension (HT) is a worldwide epidemic that affects approximate 1.13 billion people globally [1]. Several indicators, not only traditional clinical parameters, such as blood pressure and lipid profile, but also the health-related quality of life (HRQOL) were established to monitor the chronic disease burden. HRQOL is a multi-dimensional concept that measures the impact of health status on one's quality of life, in terms of physical, mental, emotional, and social functioning [2]. A population-based study from Spain suggested that HT patients displayed poorer HRQOL in physical functioning and general health compared to non-HT patients [3]. Similarly, a Swedish study also found that HT is significantly associated with lower HRQOL scores [4]. HT patients with relatively higher HRQOL scores usually have a better commitment to self-management and adherence to antihypertensive medications [5], and thus are expected to have lower mortality and morbidity compared with those with relatively poorer HRQOL scores [6]. Therefore, HRQOL is an important patient-centered outcome measure. Improving HRQOL should be one of the main goals for HT interventions.

The Short Form-12 Health Survey (SF-12) is a generic and well-established HRQOL measure that has been commonly used in both general populations and various patient cohorts [7]. Ideally, the psychometric properties (e.g., validity, reliability, and responsiveness) of SF-12 should be carefully assessed before its application in either trials or observational studies [8]. It is especially true for responsiveness, an ability of a study instrument to detect the change over time, as HRQOL measurements with poor responsiveness may induce greater type II error (false negative) in clinical trials and cause inaccurate estimation of HRQOL in longitudinal observational studies. Research on the responsiveness of SF-12 in patients with diabetes mellitus (DM) [9], prostate cancer [10], etc., has already been studied; however, there is no study assessing the responsiveness of SF-12 in hypertensive patients. Hence, this study evaluated the responsiveness of SF-12 in Chinese patients with HT.

Methods

Study design

A prospective longitudinal study was performed using convenience sampling to prospectively recruit patients diagnosed with HT. HT patients who were treated in public primary care clinics under the Hospital Authority (HA) from May 2013 to March 2014 were identified through the HA database by the International Classification of Primary Care-2 (ICPC-2) code of "K86" or "K87." Patients were excluded if they had other comorbidities, including DM, cardiovascular disease, and chronic renal disease. Patients who were under 18 years old, non-Cantonese speakers, had cognitive impairment, or refused to give consent to participate were also excluded. Patients were approached by trained research assistants, who explained the nature of the study. Baseline assessments of consented patients were completed within 2 weeks and were done by phone if necessary. The baseline assessment included collecting patients' gender, age, marital status and smoking status, and completing Chinese (Hong Kong) SF-12 Health Survey Version 2 (SF-12v2, a validated HRQOL measure for general Hong Kong Chinese adults [7, 11]). At 12 months post-enrollment, both SF-12v2 and the Global Rating of Change Scale (GRS) were administered. GRS, serving as an external criterion, was used to assess minimal clinically significant difference and responsiveness of SF-12v2 [12–18].

Sample size calculation

Assuming a small effect size of 0.3 for the difference in SF12v2 between the stable, improved, and worsened groups, a minimum of 59 patients per group were required to detect this between-group difference with 95% power at 5% level of significance.

Data analysis

Responsiveness reflects the ability of an instrument to detect changes in health status [19]. Thus, a single self-reported health change anchor, GRS, was adopted at 12 months post-recruitment. Patients were asked to rate, with respect to HT, their current health status against their health status at baseline on a scale from -3 (worsened) to 3 (improved). Patients were divided into three subsets, including 1) "stable" (rating of 0); 2) "improved" (rating of 1–3); and 3) "worsened" (rating of -3 to -1) [19].

For the assessment of responsiveness within group, paired *t* test was used to investigate the mean changes during the 12-month follow-up period in "worsened," "stable," and "improved" groups. Cohen's d effect size (ES) were used separately to test differences in the HRQOL scores between the baseline and 12-month follow-up assessments for each group [20]. The value of ES was interpreted as trivial for <0.2, small for ≥ 0.2, and <0.5, moderate for ≥ 0.5 and <0.8, or large for ≥ 0.8 [21, 22]. Responsiveness was supported if these changes were interpreted as small or above. 95% bootstrap bias corrected and accelerated confidence

intervals for ES were obtained using the bootstrapping estimation method with 2000 replications [23].

In the assessment of external responsiveness, it was hypothesized that (i) global health status would be a factor associated with the change in the HRQOL score; (ii) by comparing the "improved versus worsened group," the change in HRQOL score would be the largest; and (iii) that of the "improved versus stable group" would have the smallest change in the HRQOL scores. To test these hypotheses, the unadjusted differences in difference in HRQOL between baseline and 12 months of groups were evaluated by independent *t* test first. Then, the adjusted difference in difference in HRQOL between baseline and 12 months of groups was investigated by the multivariable linear regression adjusted for patients' baseline characteristics (i.e., gender, age, marital and smoking status, and clinical characteristics). The models satisfied the model assumptions of normality of residuals, homoscedasticity, and multi-collinearity. Responsiveness was also examined by the multiple receiver operating characteristic (ROC) curve analysis [23], where an area under the curve (AUC) ≥ 0.7 was considered adequate [24].

All statistical analyses were conducted using Stata Version 13.0 (StataCorp LP, College Station, Tex). All significance tests were two-tailed and P value < 0.05 was considered statistically significant.

Results

A total of 583 and 431 patients completed the baseline test and 12-month follow-up interviews, respectively (Table 1). Baseline characteristics between the follow-up group and the attrition group did not differ significantly.

Table 2 displays the mean change and responsiveness statistics for each group. In the worsened group, scores of the two summary scores and all domains except the domain of VT decreased significantly (deterioration in HRQOL), with the values of ES ≥ 0.2. In the stable group, the scores of RP and RE domains decreased significantly (deterioration in HRQOL), with all the responsiveness statistics < 0.2. In the improved group, the score of the GH domain increased significantly (improvement in HRQOL), with the responsiveness statistics > 0.2. There was no difference in the results between the unadjusted and adjusted analysis (table not shown).

Table 3 shows the unadjusted and adjusted differences in the difference in HRQOL between baseline and 12 months of the (1) worsened and stable group, (2) worsened and improved group, and (3) stable and improved group. Generally, the summary scores of the SF-12v2 managed to detect the significant difference in mean change between worsened and stable/improved group, but failed to identify the changes between the stable and improved group. Table 3 also shows the result of the AUC of ROC (1) between stable and worsened groups, (2) between improved and worsened groups, and (3) between improved and stable groups. No domain and summary scale of the SF-12v2 detected the difference in mean change in three comparisons with an AUC standard of at least 0.7.

Discussion

SF-12v2 validated the negative change in general health among HT patients. In comparison with other domains and component summaries, the RP and BP domains showed the highest responsiveness to the negative change in the worsened group. It was probably because these two domains have a stronger correlation with GRS when compared with other domains and two component summaries in this subgroup. Negative changes were also found in RP and RE domains in the stable group with the ES value < 0.2. However, the detected negative changes were too small and probably not clinically meaningful, as these changes may due to "noises."

SF-12v2 illustrated poor responsiveness in capturing positive changes in this study. Only the GH domain detected the positive change in improved group with a statistically significant coefficient and the ES value > 0.2. It is reasonable because the correlation between GH and GRS was stronger than that between GRS and other domains as well as summary scores. Moreover, it was found that the remaining seven domains and summary scales failed to identify positive change over 12 months in improved group. It may be because that these domains and the two summary scores were high at baseline. In fact, the baseline SF-12 scores of improved group were similar to or even higher than the Hong Kong population norm of the SF-12 v2 [11]. Therefore, there was little room for improvements during a 12-month period for improved group. Also, some of the high baseline SF-12 scores may influence the reliability of the study results.

SF-12v2 also had unsatisfactory external responsiveness in patients with HT. This is a common deficiency of generic HRQOL measures that are unable to capture the domains specific to HT patients. Indeed, the poor external responsiveness of the SF-12v2 was also declared in the previous literature that evaluated responsiveness of SF-12v2 in patients with DM and prostate cancer [9, 10].

Limitation

GRS is a generic HRQOL measure, so that it may not be sensitive to the change of hypertension-specific conditions. Moreover, some confounders, such as duration of HT and medications, were unavailable.

Table 1 Baseline characteristics of subjects

Factor	Overall ($N=583$)	Follow-up ($N=431$)	Attrition ($N=152$)	P value
Socio-demographic (n, %)				
Gender				0.592
Female	346 (59%)	253 (59%)	93 (61%)	
Male	237 (41%)	178 (41%)	59 (39%)	
Age (mean (SD)) (years)	63 (9.10)	63 (9)	62 (10)	0.418
Marital status				0.197
Not married	157 (27%)	110 (26%)	47 (31%)	
Married	426 (73%)	321 (74%)	105 (69%)	
Smoking status				0.930
Non-smoker	536 (92%)	396 (92%)	140 (92%)	
Smoker	47 (8%)	35 (8%)	12 (8%)	
Lifestyle (n, %)				
No. of fruits and vegetables eaten per day				0.090
<5	480 (82%)	348 (81%)	132 (87%)	
≥ 5	103 (18%)	83 (19%)	20 (13%)	
Exercise				0.255
None	123 (21%)	84 (19%)	39 (26%)	
<120 min per week	27% (158)	121 (28%)	37 (24%)	
≥ 120 min per week	52% (302)	226 (52%)	76 (50%)	
BP monitoring				0.110
No	146 (25%)	100 (23%)	46 (30%)	
On irregular basis	233 (40%)	171 (40%)	62 (41%)	
On regular basis	204 (35%)	160 (37%)	44 (29%)	
Clinical characteristics (mean (SD))				
SBP (mmHg)	132 (12)	132 (13)	131 (11)	0.268
DBP (mmHg)	78 (10)	78 (10)	77 (9)	0.721
LDL-C (mmol/L)	3.2 (0.8)	3.2 (0.8)	3.1 (0.8)	0.530
TC/HDL-C ratio	4.0 (1.1)	4.0 (1.2)	3.9 (1.1)	0.663
BMI (kg/m^2)	26 (4.)	26 (4)	26 (4)	0.627
Triglyceride (mmol/L)	1.4 (0.8)	1.5 (0.8)	1.4 (0.6)	0.256
Treatment modalities (n, %)				
Use of lipid-lowering agents				0.520
No	489 (84%)	359 (83%)	130 (86%)	
Yes	94 (16%)	72 (17%)	22 (14%)	
SF-12v2 scores (mean (SD))				
Physical functioning	83 ± 27 (583)	84 ± 25 (431)	81 ± 30 (152)	0.233
Role physical	83 ± 24 (583)	84 ± 23 (431)	81 ± 27 (152)	0.191
Bodily pain	78 ± 25 (583)	79 ± 23 (431)	75 ± 29 (152)	0.096
General health	36 ± 22 (583)	36 ± 22 (431)	35 ± 22 (152)	0.441
Vitality	59 ± 32 (583)	59 ± 31 (431)	57 ± 34 (152)	0.371
Social functioning	87 ± 25 (583)	88 ± 24 (431)	84 ± 28 (152)	0.168
Role emotional	86 ± 23 (583)	87 ± 21 (431)	85 ± 26 (152)	0.308
Mental health	76 ± 22 (583)	76 ± 22 (431)	74 ± 24 (152)	0.261
PCS	47 ± 9 (583)	47 ± 9 (431)	46 ± 10 (152)	0.169
MCS	54 ± 11 (583)	54 ± 11 (431)	53 ± 13 (152)	0.327

HT hypertension, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *LDL-C* low density lipoprotein–cholesterol, *TC* total cholesterol, *HDL-C* high density lipoprotein–cholesterol, *BMI* body mass index, *PCS* physical component Summary, *MCS* mental component summary

Table 2 Comparison between HRQOL scores at baseline and 12-month follow-up

Measure/subscale[a]	Baseline (mean (SD))	At 12-month follow-up (mean (SD))	Paired difference (mean (SD))	*P* value	ES (95% CI)
Worsened GRS ($n=61$)					
SF-12v2 scores					
Physical functioning	70 ± 33	64 ± 31	−6.6 ± 25.0	0.045*	−0.20 (−0.42, 0.00)
Role physical	72 ± 27	52 ± 28	−20 ± 29	< 0.001*	−0.72 (−1.04, −0.45)
Bodily pain	69 ± 24	50 ± 29	−19 ± 29	< 0.001*	−0.72 (−1.04, −0.43)
General health	27 ± 17	23 ± 15	−4.1 ± 15.5	0.043*	−0.26 (−0.52, −0.02)
Vitality	50 ± 32	41 ± 30	−8.6 ± 37.3	0.077	−0.28 (−0.56, −0.03)
Social functioning	77 ± 30	58 ± 36	−20 ± 40	< 0.001*	−0.60 (−0.93, −0.28)
Role emotional	82 ± 22	67 ± 28	−15 ± 27	< 0.001*	−0.58 (−0.86, −0.30)
Mental health	72 ± 22	59 ± 24	−13 ± 24	< 0.001*	−0.58 (−0.84, −0.26)
PCS	42 ± 11	37 ± 11	−4.3 ± 9.7	0.001*	−0.38 (−0.63, −0.17)
MCS	53 ± 11	46 ± 11	−6.8 ± 10.9	< 0.001*	−0.61 (−0.90, −0.34)
Stable GRS ($n=245$)					
SF-12v2 scores					
Physical functioning	86 ± 22	85 ± 23	−0.6 ± 19.6	0.625	−0.03 (−0.14, 0.08)
Role physical	86 ± 21	83 ± 24	−3.6 ± 22.0	0.012*	−0.16 (−0.28, −0.05)
Bodily pain	80 ± 23	79 ± 23	−1.5 ± 25.4	0.347	−0.07 (−0.21, 0.07)
General health	37 ± 22	37 ± 22	0.3 ± 22.2	0.807	0.02 (−0.10, 0.14)
Vitality	60 ± 31	64 ± 28	3.8 ± 32.8	0.073	0.13 (0.02, 0.26)
Social functioning	90 ± 22	88 ± 24	−2.1 ± 25.9	0.196	−0.09 (−0.23, 0.05)
Role emotional	89 ± 19	87 ± 20	−2.6 ± 19.6	0.043*	−0.13 (−0.25, 0.00)
Mental health	79 ± 21	78 ± 19	−0.8 ± 20.0	0.524	−0.04 (−0.18, 0.08)
CS	48 ± 8	47 ± 9	−0.5 ± 7.5	0.344	−0.05 (−0.17, 0.06)
MCS	55 ± 10	55 ± 10	−0.3 ± 9.1	0.594	−0.03 (−0.15, 0.08)
Improved GRS ($n=125$)					
SF-12v2 scores					
Physical functioning	86 ± 24	89 ± 19	3.2 ± 22.2	0.110	0.14 (−0.04, 0.31)
Role physical	85 ± 22	83 ± 21	−2.3 ± 22.5	0.256	−0.11 (−0.29, 0.09)
Bodily pain	81 ± 22	82 ± 21	1 ± 19.7	0.571	0.05 (−0.11, 0.21)
General health	40 ± 25	47 ± 27	7 ± 23.6	0.001*	0.27 (−0.11, −0.43)
Vitality	62 ± 30	61 ± 30	−1.6 ± 34.2	0.601	−0.05 (−0.25, 0.15)
Social functioning	88 ± 24	87 ± 22	−0.6 ± 23.9	0.779	−0.03 (−0.21, 0.16)
Role emotional	85 ± 25	88 ± 20	2.2 ± 23.3	0.294	0.10 (−0.08, 0.28)
Mental health	74 ± 22	77 ± 21	3.2 ± 22.3	0.112	0.15 (−0.02, 0.34)
PCS	49 ± 8	49 ± 8	0.4 ± 7.6	0.538	0.05 (−0.13, 0.20)
MCS	53 ± 12	54 ± 10	0.8 ± 11.0	0.433	0.07 (−0.09, 0.26)

Mean scores and standard deviations are rounded to whole numbers to make them easier to read. Paired differences were rounded to one decimal point for better precision

PCS physical component summary, *MCS* mental component summary, *GRS* global rating scale, *SES* standardized effect size, *SRM* standardized response mean, *CI* confidence interval

*Significant difference at 0.05 level between baseline and 12-month follow-up by paired t-test

[a]Higher scores represent a higher level of functioning or a better HRQOL

Conclusion

This is the first study that examined the responsiveness of the SF-12v2 in HT patients. SF-12v2 demonstrated good responsiveness to probe negative changes among hypertensive patients with worsened general health, but showed an inadequate performance to capture changes in improved and stable general health group. SF-12v2 also showed poor external responsiveness among hypertensive patients.

Table 3 Unadjusted and adjusted difference in differences with different anchors at 12-month follow-up

	Stable versus worsened		Improved versus worsened		Improved versus stable		ROC AUC
	Unadjusted DD[a] (95% CI)	Adjusted DD[a,d] (95% CI)	Unadjusted DD[b] (95% CI)	Adjusted DD[b,d] (95% CI)	Unadjusted DD[c] (95% CI)	Adjusted DD[c,d] (95% CI))	
SF-12v2 scores							
Change in physical functioning	5.9* (0.1, 11.8)	5.5 (−0.6, 11.6)	9.8* (2.6, 16.9)	10.2* (2.6, 17.9)	3.8* (−0.6, 8.2)	4.0 (−0.7, 8.7)	0.57
Change in role physical	16.5* (9.9, 23.1)	15.4* (8.6, 22.1)	17.8* (10.2, 25.4)	19.4* (11.2, 27.5)	1.3 (−3.5, 6.1)	1.2 (−3.7, 6.1)	0.62
Change in bodily pain	17.3* (10.0, 24.7)	16.8* (9.1, 24.5)	19.9* (12.7, 27)	21.3* (13.8, 28.8)	2.5 (−2.6, 7.6)	3.1 (−2.3, 8.4)	0.62
Change in general health	4.4 (−1.5, 10.4)	2.7 (−3.5, 8.9)	11.1* (4.6, 17.7)	12.5* (5.7, 19.3)	6.7* (1.8, 11.6)	7.7* (2.7, 12.7)	0.57
Change in vitality	12.4* (2.9, 21.9)	12.2* (2.5, 21.9)	7 (−3.8, 17.9)	5.3 (−6.2, 16.7)	−5.4 (−12.6, 1.8)	−5.7 (−13.1, 1.7)	0.55
Change in social functioning	17.5* (9.3, 25.7)	16.4* (8.0, 24.8)	19.1* (9.8, 28.3)	19.4* (9.4, 29.3)	1.5 (−3.9, 7)	3.1 (−2.5, 8.7)	0.60
Change in role emotional	12.0* (6.0, 18.0)	11.1* (4.9, 17.3)	16.7* (9.2, 24.3)	16.0* (7.8, 24.1)	4.8* (0.2, 9.3)	4.8* (0.1, 9.5)	0.64
Change in mental health	12.3* (6.4, 18.2)	11.9* (5.7, 18.0)	16.3* (9.3, 23.4)	16.7* (9.0, 24.3)	4* (−0.5, 8.5)	4.1 (−0.6, 8.8)	0.62
Change in PCS	3.8* (1.6, 6.1)	3.5* (1.1, 5.8)	4.7* (2.1, 7.3)	5.3* (2.7, 8.0)	0.9 (−0.8, 2.5)	1.0 (−0.6, 2.7)	0.59
Change in MCS	6.5* (3.9, 9.2)	6.2* (3.5, 9.0)	7.6* (4.2, 11)	7.2* (3.6, 10.8)	1.1 (−1, 3.2)	1.2 (−1.0, 3.4)	0.63

PCS physical component summary, *MCS* mental component summary, *DD* difference in differences, *CI* confidence interval

*Significant difference at 0.05 level

[a]Difference between groups = stable−worsened

[b]Difference between groups = improved−worsened

[c]Difference between groups = improved−stable

[d]All adjusted difference in differences are adjusted by gender, age, marital status, smoking status, diet, exercise, BP monitoring, SBP, DBP, LDL-C, TC/HDL-C ratio, BMI, triglyceride, and use of lipid-lowering agent

Acknowledgements The authors hope to acknowledge the contributions of the multidisciplinary risk-stratification-based hypertension management programme team at the Hospital Authority head office, chief of service in primary care, programme coordinator in each cluster, and Statistics and Workforce Planning Department at the Hong Kong Hospital Authority.

Funding This study was funded by the Health Services Research Fund, Food and Health Bureau, HKSAR (Ref. Nos. EPC-HKU-2 and 13142471). No funding organization had any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation of the manuscript. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Compliance with ethical standards

Conflict of interest No known conflicts of interest relevant to this article were reported.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethics approval of this study was granted by all local Institutional Review Board in Hong Kong (HKE: HKEC-2016-006; HKW: UW15-259; KCC: KC/KE-16-0019/ER-2; KCC: KC/KE-16-0020/ER-2; KWC: KW/EX-16-0059(98-02); NTE: 2016.055; NTW: NTWC/CREC/16010).

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. World Health Orgnization (WHO). (2015) Raised Blood Pressure. Global Health Observatory (GHO) data. Retrieved from http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence/en/. Assessed 23 May, 2016.
2. Singh, K., Kondal, D., Shivashankar, R., et al. (2017). Health-related quality of life variations by sociodemographic factors and chronic conditions in three metropolitan cities of South Asia: The CARRS study. *British Medical Journal Open, 7,* e018424.

3. Mena-Martin, J. F., Martin-Escudero, C. J., Simal-Blanco, L. F., et al. (2003). Health-related quality of life of subjects with known and unknown hypertension: results from the population-based Hortega study. *Journal of Hypertension, 21*(7), 1283–1289.
4. Bardage, C., & Isacson, D. G. L. (2001). Hypertension and health-related quality of life: An epidemiological study in Sweden: an epidemiological study in Sweden. *Journal of Clinical Epidemiology, 54*(2), 172–181.
5. Zyoud, S. H., Al-Jabi, S. W., Sweileh, W. M., et al. (2013). Health-related quality of life associated with treatment adherence in patients with hypertension: A cross-sectional study. *International Journal of Cardiology, 168*(3), 2981–2983.
6. National Centers for Disease Control and Prevention and Health Promotion and USA. (2016) HRQOL Concepts. Retrieved from https://www.cdc.gov/hrqol/concept.htm. Accessed 23 May, 2016.
7. Lam, C., Tse, E., & Gandek, B. (2005). Is the standard SF-12 Health Survey valid and equivalent for a Chinese population? *Quality of Life Research, 14*(2), 539–547.
8. Choi, E. P. H., Lam, C. L. K., & Chin, W. Y. (2014). The health-related quality of life of Chinese patients with lower urinary tract symptoms in primary care. *Quality of Life Research, 23*(10), 2723–2733.
9. Wan, E. Y. F., Choi, E. P. H., Yu, Y. E. T., et al. (2018). Evaluation of the internal and external responsiveness of Short Form-12 Health Survey version 2 (SF-12v2) in patients with type 2 diabetes mellitus. *Quality of Life Research, 27*(9), 2459–2469.
10. Choi, E., Wong, C., Wan, E., et al. (2016). The internal and external responsiveness of Functional Assessment of Cancer Therapy-Prostate (FACT-P) and Short Form-12 Health Survey version 2 (SF-12 v2) in patients with prostate cancer. *Quality of Life Research, 25*(9), 2379–2393.
11. Lam, C. L. K., Wong, C. K. H., Lam, E. T. P., Lo, Y. Y. C., & Huang, W. W. (2010). Population norm of Chinese (HK) SF-12 Health Survey_Version 2 of Chinese adults in Hong Kong. *Hong Kong Practitioner, 32,* 77–86.
12. Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: A critical review and recommendations. *Journal of Clinical Epidemiology, 53*(5), 459–468.
13. Kamper, S. J., Maher, C. G., & Mackay, G. (2009). Global rating of change scales: A review of strengths and weaknesses and considerations for design. *Journal of Manual & Manipulative Therapy, 17*(3), 163–170.
14. Wong, C. K. H., Mulhern, B., Wan, Y. F., & Lam, C. L. K. (2014). Responsiveness was similar between direct and mapped SF-6D in colorectal cancer patients who declined. *Journal of Clinical Epidemiology, 67*(2), 219–227.
15. Wan, E. Y. F., Choi, E. P. H., Yu, E. Y. T., Chin, W. Y., Fung, C. S. C., Chan, A. K. C., et al. (2018). Evaluation of the internal and external responsiveness of Short Form-12 Health Survey version 2 (SF-12v2) in patients with type 2 diabetes mellitus. *Quality of Life Research, 27,* 1–11.
16. Stavem, K., Frøland, S. S., & Hellum, K. B. (2005). Comparison of preference-based utilities of the 15D, EQ-5D and SF-6D in patients with HIV/AIDS. *Quality of Life Research, 14*(4), 971–980.
17. Marra, C. A., Rashidi, A. A., Guh, D., et al. (2005). Are indirect utility measures reliable and responsive in rheumatoid arthritis patients? *Quality of Life Research, 14*(5), 1333–1344.
18. Choi, E. P. H., Chin, W. Y., Lam, C. L. K., & Wan, E. Y. F. (2015). The responsiveness of the International Prostate Symptom Score, Incontinence Impact Questionnaire-7 and Depression, Anxiety and Stress Scale-21 in patients with lower urinary tract symptoms. *Journal of Advanced Nursing, 71*(8), 1857–1870.
19. Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology, 61*(2), 102–109.
20. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
21. Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases, 40*(2), 171–178.
22. Liang, M. H., Fossel, A. H., & Larson, M. G. (1990). Comparisons of five health status instruments for orthopedic evaluation. *Medical Care, 28*(7), 632–642.
23. Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association, 82*(397), 171–185.
24. Hand, D. J., & Till, R. J. (2001). A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning, 45*(2), 171–186.