Validity and Reliability of the Indonesian Version of SF-36 Quality of Life Questionnaire on Patients with Permanent Pacemakers

Simon Salim[1], Muhammad Yamin[1], Idrus Alwi[1], Siti Setiati[2]

[1] Department of Internal Medicine, Faculty of Medicine Universitas Indonesia - Cipto Mangunkusumo Hospital, Jakarta, Indonesia.
[2] Clinical Epidemiology and Evidence-based Medicine Unit, Cipto Mangunkusumo Hospital, Jakarta, Indonesia.

Corresponding Author:
Simon Salim, MD. Division of Cardiology, Department of Internal Medicine, Faculty of Medicine Universitas Indonesia-Cipto Mangunkusumo Hospital. Jl. Diponegoro 71, Jakarta 10430, Indonesia. email: simonsalim@gmail.com.

ABSTRAK

Tujuan: membentuk dan memvalidasi kuesioner SF-36 versi bahasa Indonesia. *Metode:* penelitian ini merupakan penelitian potong lintang yang terdiri atas 2 tahap: 1) adaptasi budaya dan bahasa; dan 2) uji validitas dan reliabilitas. Kami mengevaluasi 32 orang pada tahap awal dan 20 orang pada tahap akhir dari bulan September 2014 hingga Agustus 2015. Subjek merupakan pasien dengan pacu jantung permanen. Kami mengikuti petunjuk adaptasi lintas budaya untuk menghasilkan kuesioner SF-36 versi bahasa Indonesia. Kuesioner terjemahan akhir akan diperiksa validitasnya terhadap Tes Jalan 6 Menit dan pemeriksaan NT pro-BNP. *Hasil:* kuesioner SF-36 bahasa Indonesia memiliki korelasi positif antara tes jalan 6 menit dengan domain PF (Physical Functional) (r=0,363; p=0,001), dan memiliki korelasi negatif antara NT pro-BNP dengan domain GH (General Health) (r=-0,269; p=0,020) dan MH (Mental Health) (r=-0,271; p=0,019). Konsistensi internal kuesioner SF-36 bahasa Indonesia, yang diukur dengan Cronbach's alpha dinilai baik dengan nilai >0,70. Uji repeatability antara hari 1 dan hari 8 dinilai baik dengan korelasi positif kuat (r=0,626; p=0,003) dan tidak ada perbedaan bermakna pada level item, domain, dan keseluruhan kuesioner. *Kesimpulan:* kuesioner SF-36 bahasa Indonesia dapat digunakan sebagai kuesioner umum untuk menilai kualitas hidup pasien dengan pacu jantung permanen.

Kata kunci: kualitas hidup, kuesioner, SF-36, pacu jantung, Indonesia.

ABSTRACT

Aim: to construct and validate Indonesian version of SF-36. *Methods:* this is a cross-sectional study, which consist of 2 stages process: 1) language and cultural adaption; and 2) validity and reliability evaluation. We evaluated 32 pacemaker patients during language and cultural adaptation stage and 20 pacemaker patients during validity and reliability evaluation stages from September 2014 to August 2015. We followed cross-cultural adaptation guideline to produce Indonesian version of the questionnaire. The final translated questionnaire was checked by assessing the correlation of SF-36 and 6-minutes walking test (6MWT) and NT pro-BNP result. *Results:* Indonesian version of SF-36 showed positive correlation between 6MWT result and physical functioning (PF) (r=0.363; p=0.001) and negative correlation between NT pro-BNP score with general health (GH) (r=-0.269; p=0.020) and mental health (MH) (r=-0.271; p=0.019). The internal consistency of Indonesian version of SF-36 questionnaire, which measured by Cronbach's alpha, was good with value of >0.70. Repeatability between day 1 and day 8 was good, with strong positive correlation (r=0.626; p=0.003). *Conclusion:* the Indonesian version of SF-36 could be used as a general questionnaire to assess quality of life in patients with permanent pacemaker.

Keywords: quality of life, questionnaire, SF-36, pacemaker, Indonesia.

INTRODUCTION

As cardiovascular diseases (CVD) morbidity and mortality increased, the need of health intervention is increasing as well. Permanent pacemaker, as one of innovation in CVD management, has been widely used for arrhythmia therapy. More than 500,000 procedures have been performed per year worldwide.[1] Many studies showed that the permanent pacemaker has improved survival in patients with CVD, especially in the case of complete AV block.[2] According to World Health Organization (WHO), health is a state of complete physical mental and social well-being and not merely the absence of disease or infirmity.[3] Hence, physicians need to consider quality of life in recommending a therapy.

In the last 2 decades, health-related quality of life measurement has a strong impact in health care and medical intervention evaluation. Quality of life measurement is also essential to predict effectiveness and successfulness of therapy.[4,5] There are several studies that assessed the successfulness of pacemaker implantation, not only based on the survival of the patients but also the improvement of their quality of life.

Quality of life involves several dimensions: physical function, emotional state, social interaction, and somatic sensation.[5] The instrument to measure quality of life must not only be valid, appropriate, reliable, and responsive, but also capable of being interpreted.[6] There are two types of instruments that can be used to measure quality of life, generic and disease specific instruments. Using combination of both instruments may be more appropriate to evaluate quality of life, especially in patients with pacemaker. SF-36 has been well-known as a generic instrument which become a standard for measuring quality of life. It measures eight domains from physical and mental aspects: General Health (GH), Physical Functioning (PF), Role-Physical (RP), Bodily Pain (BP), Vitality (VT), Social Functioning (SF), Role-Emotional (RE), and Mental Health (MH). SF-36 questionnaire has been translated into several languages.[7] Up to our knowledge, there is no satisfactory Indonesian version of SF-36. A group of researchers had translated SF-36 into Indonesian version for cancer study and suggested most questions to be reconstructed.[8] The objective of this study was to translate and evaluate validity and reliability of the Indonesian version of SF-36 for assessing quality of life, especially in Indonesian language speaking pacemaker patients.

METHODS

This is a cross-sectional study, performed in an outpatient cardiology clinic of the Cipto Mangunkusumo Hospital, within a 1-year period between September 2014 and August 2015. This study had been approved by the Ethical Committee on Health Research, Faculty of Medicine Universitas Indonesia - Cipto Mangunkusumo Hospital on September 14th, 2014 with a reference number of 629/UN2.F1/ETIK/2014.

The study enrolled patients aged over 18 years with permanent pacemakers, which have been implanted for at least 3 months. We excluded patients with congestive heart failure NYHA III-IV, cognitive impairment, physical disability, not optimal echo window, and other comorbidities,such as pericardial effusion and pleural effusion. This study was divided into 2 steps, language and cultural adaptation and validity and reliability testing. The minimum of 30 patients were needed for the language and cultural adaptation process suggested by Beaton et al.[9]. For validity and reliability testing, the minimum samples size were determined by calculating from confidence level 95% with power 80% and r=0.6. Samples were collected consecutively. We finally involved 32 patients for language and cultural adaptation process and 20 patients for validity and reliability testing.

Step 1: Language and Cultural Adaptation

The procedures for translation into the Indonesian language were modified from Guillemin and Beaton.[9,10] This step consists of 6 stages: initial translation, translation synthesis, back translation, committee review, pretesting, and submission and appraisal of all written reports to the committee (**Table 1**). Two Indonesian translators performed the initial translation process and they synthesized one

11

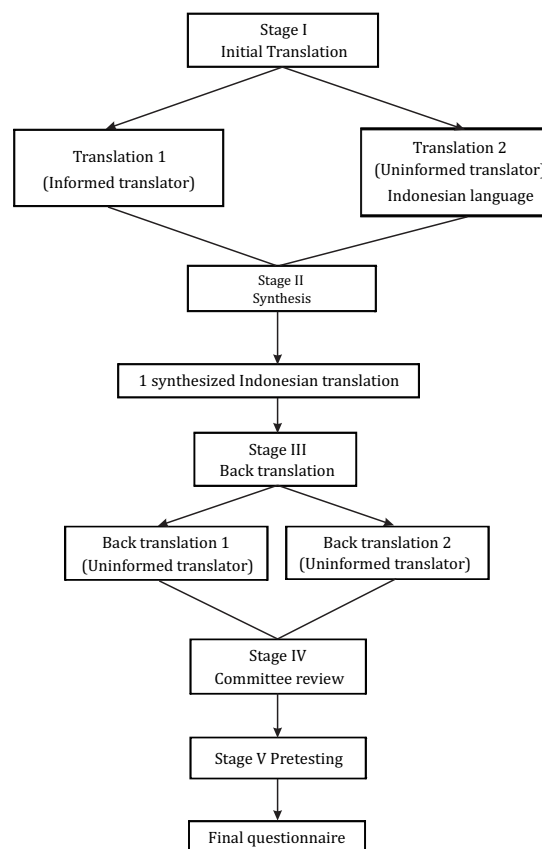**Table 1.** Stages of the language and cultural adaptation[9,10]

Stage	Participants	Product
Initial translation	2 translators (informed and uniformed translator) Into target language	Translation 1 (T1) Translation 2 (T2)
Translation synthesis	2 initial translators and observer	1 synthesis translation (T12)
Back translation	2 translators with English as the first language and unfamiliar to the questionnaire	Back translation 1 (BT1) Back translation 2 (BT2)
Committee review	Research team (methodology expert, health expert, and translators)	Pre-final questionnaire
Pretesting	32 patients	Written report
Additional stage	Submission of documentation to the committee	

Figure 1. Language and cultural adaptation process

Indonesian version. In the back translation stage, the synthesis questionnaire was then translated back to English by two English speaking mother tongue translators. The committee, which consists of methodological expert, clinical expert, and translators, reviewed the original at collaboration with the European questionnaire and each translation together with corresponding written reports. After being reviewed by the committee, the questionnaire was edited to be the pre-final version (**Figure 1**).

The pre-final translated version was then distributed to 32 pacemaker patients. The researcher documented any difficulties that the patients experienced during the time the patients completed the questionnaire. The documentations were reviewed and used to modify the questionnaire into the final version.

Step 2: Validity and Reliability Testing

The validity of the final version of Indonesian language was assessed through the participation of 20 patients by comparing SF-36 questionnaire with functional class and 6-minutes walking test (6MWT). Internal consistency was assessed by item-to-item correlation, item-to-domain, item-to-total correlation, and Cronbach's a coefficient. It was considered to be acceptable when item-to-

item correlation was >0.2–0.4, item-to-domain correlation and item-to-total correlation were >0.2, and Cronbach's alpha was 0.7 or greater. Repeatability was assessed by test-retest method by asking patient to fill the same questionnaire under the same condition in the next 7 days. If the correlation between two separate administration of the questionnaire is >0.8, the questionnaire is then considered has good test-retest reliability.

RESULTS

Step 1: Language and Cultural Adaptation

A total of 32 patients participated in the language and cultural adaptation process with mean age of 65.8 (SD 15.096); in that, 53.1% were females and 28% were patients with bachelor degree or higher. Most patients (62.5%) have pacemaker implantation for high degree or total AV block and majority of them (81.3%) have dual chamber pacemaker. This step involved 2 native Indonesia translators, which one of them has medical background, and 2

native non-medical English translators.

In the Indonesian version of SF-36, most of the questions were translated without any significant changes (**Appendix 1**: http://www.actamedindones.org/index.php/ijim/article/view/275/pdf). Some questions were clarified by adding some explanations. For example: the word "flight of stairs" in question number 6 and 7 was translated into '*rangkaian tangga*' (lit. set of stairs), which normally consists of 10 to 12 stairs. But to facilitate patients understanding, the translators decided to add the word '*lantai*' (lit. floor). "Several flight of stairs" was explained by '*satu lantai atau lebih*' (lit. one floor or more), while "one flight of stairs" was explained by '*setengah lantai*' (lit. half floor). Another example, the question number 9, "walking more than a mile", was explained by adding 1.6 kilometers since Indonesian people were not familiar with "mile". The translators also added explanation for the word "blocks" by adding 100 meter (lit. one hundred meters) for one block and >100 meter (lit. more than one hundred meters) for several blocks. For number 21, the word "pain" has been translated into two different words by both Indonesian translators. In Indonesian language, the word "pain" can be translated into *nyeri* or *sakit*. Even though Indonesian people more familiarized with the word *sakit* but it also has another meaning, which is "sickness". Both translators finally decided to use the word *nyeri* to avoid misunderstanding.

The translators also made some changes for options number 23–31. They used the words *sepanjang waktu* (lit. all of the time), *sering* (lit. often), *cukup sering* (lit. quite often), *kadang-kadang* (lit. sometime), *jarang* (lit. seldom), and *tidak pernah* (lit. never) to express the words "All of the time", "Most of the time", "A good bit of the time", " Some of the time", "A little of the time", and "None of the time" in the original questionnaire. For number 36, the word "excellent" was translated into *sangat baik* (lit. very good).

Step 2: Validity and Reliability Testing

A total of 20 patients, with the mean age 62.35 (SD 16.69), were involved in validity and reliability testing. Patients characteristics can be seen in **Table 2**.

Table 2. Patients' characteristics in validity and reliability testing

Variables	Value (n=20)
Age (year), mean (SD)	62.35 (16.69)
Gender (male), %	55.00
Educational Background, %	
- Elementary School	25.00
- Junior High School	30.00
- Senior High School	10.00
- Diploma	5.00
- Bachelor Degree	25.00
- Master Degree	5.00
Indication for Pacemaker, %	
- AV Block (High Degree or Total)	75.00
- Sick Sinus Syndrome	10.00
- Symptomatic Bradycardia	15.00
Type of Permanent Pacemaker, %	
- Single Chamber	0.00
- Double Chamber	100.0

Data were retrieved between March 2015 and April 2015. SF-36 questionnaire consists of 36 items question, which grouped into 8 health domains. Because each item has different scale, coding was used to equalize the weight of each item. Item number 1, 2, 20, 22, 34, 36, which has 5 scale, were converted into 100, 75, 50, 25, 0; item number 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, which have 3 scale, were converted into 0, 50, 100; item number 13, 14, 15, 16, 17, 18, 19, which have 2 scale, were converted into 0 and 100; item number 21, 23, 26, 27, 30, which have 6 scale, were converted into 100, 80, 60, 40, 20, 0; item number 24, 25, 28, 29, 31 were converted into 0, 20, 40, 60, 80, 100; and item number 32, 33, 35 were converted into 0, 25, 50, 75, 100.

Validity of the questionnaire was measured by correlation of questionnaire and other supportive measures, 6MWT and NT pro BNP. Correlation was considered very strong if $r = 0.80–1.00$; strong if $r = 0.60–0.79$, moderate if $r = 0.40–0.59$; weak if $r = 0.20–0.39$; and very weak if $r = 0.00–0.19$. Most of all domains in SF-36 Indonesian version were not normally distributed, except GH and VT domain. Meanwhile, the correlation between 6MWT and NT pro BNP was tested using Kendall's tau method.

The result showed a very significant weak positive correlation between PF domain and

Table 3. Correlation between SF-36 domains and 6MWT and NT pro-BNP

	GH	PF	RP	RE	SF	BP	VT	MH	Total SF-36
6mwt	0.173	0.363**	0.028	-0.002	0.039	0.123	0.092	0.077	0.207
NT pro-BNP	-0.269*	-0.179	-0.131	-0.231	-0.16	-0.199	-0.14	-0.271*	-0.261*

*) $p < 0.05$ (2-tailed); **) $p < 0.01$ (2-tailed)

6MWT (r=0.363; p=0.001). On the other hand, there were 2 domains (GH and MH), which showed a significant weak negative correlation with NT pro BNP. Indonesian version of SF-36, as a questionnaire, only showed weak negative correlation with NT pro BNP and did not have any significant correlation with 6MWT (**Table 3**).

Meanwhile, reliability of the questionnaire was measured by internal consistency and repeatability test. Internal consistency of the questionnaire was tested by Cronbach α and inter-item and inter-domain correlation. If α coefficient is 0.70 or higher, it is considered acceptable. This result showed that Cronbach α was higher than 0.7 for all domains (range: 0.751–0.922) except for Social Functioning (α=0.614) and Vitality (α=0.434) domains (**Table 4**). Cronbach α for total SF-36 was 0.789.

Table 4. Internal consistency reliability

Domains	Cronbach α
GH	0.751
PF	0.869
RP	0.878
RE	0.842
SF	0.614
BP	0.922
VT	0.434
MH	0.824
Total SF-36	0.789

Inter item correlation were calculated in each domain. All items in domain RP, RE, SF, and BP were significantly correlated. Meanwhile in other domains, such as GH, PF, VT, and MH, some items were not significantly correlated to each other.

Since SF-36 consists of 8 domains, correlation between each domain was also calculated. Domain score was obtained from total questionnaire in day 1 and day 8. The result showed strong positive correlation between domain RE and BP (0.614; p<0.01). Some domains have positive moderate correlation (GH and SF 0.422; p<0.01; PF and RP 0.489; p<0.01; PF and SF 0.463; p<0.01; RP and RE 0.574, p<0.01; RP and SF 0.414; P<0.01; SF and BP 0.416; p<0.01; SF and MH 0.445; p<0.01) and the others have weak positive correlation (GH and PF 0.282; p<0.05; GH and VT 0.311; p<0.01; GH and MH 0.308; p<0.05; PF and RE 0.376, p<0.01; PF and BP 0.394, p<0.01; PF and VT 0.354, p<0.01; RP and BP 0.335, p<0.05; RE and SF 0.29; p<0.05; SF and VT 0.269; p<0.05; BP and MH 0.302; p<0.05, VT and MH 0.360; p<0.01). However, all domains were very significantly correlated with total questionnaire (GH and total SF-36 0.410; p<0.01; PF and total SF 36 0.694; p<0.01; RP and total SF 36 0.582; p<0.01; RE and total SF 36 0.549; p<0.01; SF and total SF 36 0.552; p<0.01; BP and total SF 36 0.446; p<0.01; VT and total SF 36 0.426; p<0.01, MH and total SF -36 0.422; p<0.01).

Repeatability

Repeatability test was conducted by asking the patients to fill in the questionnaire in 2 different times with the same condition. The patients were asked to visit the hospital 7 days after the first meeting. We analyzed the correlation, by using Kendall's test, and the difference, by using Wilcoxon, of the items, domains, and total questionnaire. From the repeatability test, it was found that 20 items were correlated significantly between day 1 and day 8 (item 1 (0.517, p<0.05), item 2 (0.604, p<0.01), item 3 (0.821, p<0.01), item 4 (0.41, p<0.05), item 5 (0.514, p<0.05), item 10 (0.668, p<0.01), item 11 (0.498, p<0.05), item 12 (1.00, p<0.01), item 15 (0.685, p<0.01), item 18 (0.579, p<0.05), item 19 (0.577, p<0.05), item 21 (0.431, p<0.05), item 24 (0.454, p<0.05), item 25 (0.702, p<0.01), item 27 (0.535, p<0.01), item 28 (0.889, p<0.01), item 30 (0.649, p<0.01), item 32 (0.625,

p<0.01), item 34 (0.570, p<0.01), and item 36 (0.476, p<0.05)).

In domain level, all domains were correlated significantly between day 1 and day 8, except RP domain (GH 0.398 (p<0.05), PF 0.572 (p<0.01), RE 0.478 (p<0.05), SF 0.479 (p<0.05), BP 0.434 (p<0.05), VT 0.376 (p<0.05), and MH 0.616 (p<0.05)). The total questionnaire was correlated significantly between day 1 and 8 with a strong positive correlation (r=0.626; p=0.003). Meanwhile, there was no significant difference of the repeatability test in 20 patients in item level, domain level, and total questionnaire.

DISCUSSION

In the translation and back-translation process, there was no big discrepancy among the 2 Indonesian versions and the 2 back-translation versions performed in this study. In the cultural adaptation process, we did not make any significant changes since Indonesian people are already familiar with some activities mentioned in the questionnaire. We only added several explanations, such as distances, because Indonesian people might have a different perception in those regards.

The validity evaluation of SF-36 Indonesian version showed this questionnaire is valid to be used as a quality of life assessment tool in patients with permanent pacemaker. In validation process, this questionnaire has been through content validity, face validity, criterion validity, and construct validity. The content validity was assessed by receiving the input from patients as well as the expert about the content of the questionnaire. Both of patients and the experts agreed that the content of the questionnaire are logic, comprehensive, and has been covered the characteristics to be measured. The face validity process was conducted by the observation of the researcher that the questions used in this questionnaire were relevant, reasonable, unambiguous, and clear. For the criterion validity, we evaluate the correlation with other criteria, which has been considered as gold standard, 6 minutes walking test and NT pro BNP. We found that SF-36 Indonesian version has no significant correlation with 6MWT but has an inverse or negative correlation with NT pro BNP. This

finding was different with other studies, which showed a moderate-strong correlation between 6MWT and Physical Function domain.[11] In construct validity, we and the expert assessed that this questionnaire has good convergent validity and discriminant validity because some domains were correlated to assess the same thing and some other were different to assess the different concept.

In the reliability evaluation, the internal consistency of this questionnaire was good. There were 2 domains, SF and VT, which showed the Cronbach α score were less than 0.7. Low reliability of SF and VT domains had been observed in other studies. A study of Moroccan-Arabic language groups in Netherland reported reliabilities of 0.54 for VT and 0.63 for SF.[12] The same result was also found in Thailand, which reported reliabilities of 0.68 for VT and 0.55 for SF.[13] Translating the concept of social functioning is difficult since there are some cultural differences in Indonesia. Social functioning has also been rated as difficult in translations of SF-36 items in a cross-cultural comparison of 10 countries.[14] Meanwhile, items in vitality domain are intended to measure physical and mental energy and fatigue. The low reliability of VT might be caused by different perception of patient's own condition.

Inter-item correlation in each domain showed weak to strong correlations. In RP, RE, SF, and BP domains, the correlation between items were significant. On the other hand, some of the items in GH, PF, VT, and MH domains were not significantly correlated. Based on study Health Care Financing Review, items in Vitality and Mental Health domain have strong correlation with each other.[15] This discrepancy might be caused by different perception of the patients in assessing their condition or embarrassment of answering the question. Inter-domain correlation showed weak to strong correlation and it was in accordance to other study. The repeatability of SF-36 Indonesian version was good because there was a good correlation between questionnaire obtained in day 1 and day 8. Even though there were some differences in item 2, 3, 23 and 26, these differences were not meaningful if we compared to the domain and total SF-36. Since

SF-36 questionnaire was designed to evaluate the health condition in 1 last month, the difference measurement in 1-week away would not give a meaningful difference.

CONCLUSION

The Indonesian version of SF-36 has good validity and reliability and could be used as a general questionnaire to assess quality of life in permanent pacemaker patients. From this study we assumed that Indonesian version of SF-36 questionnaire could be applied to patients who have chronic disease or potentially having disability. The assessment of quality of life should be taken into consideration, especially in giving medical treatment to the patients. However, this study has several weaknesses. We did not evaluate the correlation between the receptiveness of the patients with their education level and ethnical background. We also excluded the patients with disability, especially those who could not conducted 6MWT. The validity and reliability of this questionnaire were good. However, we cannot conclude directly that this Indonesian version of SF-36 can be used to monitor improvements, as it is not in the scope of our study. We also urge the need to validate our questionnaire in other patient groups, as our sample was very homogenous.

ACKNOWLEDGMENTS

REFERENCES

1. Moller M, Arnsbo P, Asklund M, et al. Quality assessment of pacemaker implantations in Denmark. Europace. 2002;4(2):107-12.
2. Vardas PE, Auricchio A, Blanc JJ, et al. Guidelines for cardiac pacing and cardiac resynchronization therapy: The task force for cardiac pacing and cardiac resynchronization therapy of the European Society of Cardiology. Developed in collaboration with the European Heart Rhythm Association. Eur Heart J. 2007;28(18):2256-95.
3. World Health Organization. Constitution of The World Health Organization. Basic Documents, Supplement 2006.
4. Viens AM. Quality of life as a paradigm for health in a developed society. Internet J Law, Healthcare Ethics. 1999;1(1).
5. Linde C. How to evaluate quality of life in pacemaker patients: problems and pitfalls. PACE. 1996;19:391-7.
6. Chen T-h, Li L, Kochen MM. A systematic review: How to choose appropriate health-related quality of life (HRQOL) measures in routine general practice? J Zhejiang Univ SCI. 2005;6B(9):936-40.
7. Ware JE, Jr., Gandek B. Overview of the SF-36 health survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol. 1998;51(11):903-12.
8. Perwitasari DA. Development the validation of Indonesian version of SF-36 questionnaire in cancer disease. Indonesian J Pharm. 2012;23(4):248-53.
9. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25(24):3186-91.
10. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993;46(12):1417-32.
11. Alison JA, Kenny P, King MT, et al. Repeatability of the six-minute walk test and relation to physical function in survivors of a critical illness. Phys Ther. 2012;92(12):1556-63.
12. Hoopman R, Terwee CB, Deville W, Knol DL, Aaronson NK. Evaluation of the psychometric properties of the SF-36 health survey for use among Turkish and Moroccan ethnic minority populations in the Netherlands. Qual Life Res. 2009;18(6):753-64.
13. Lim LL, Seubsman SA, Sleigh A. Thai SF-36 health survey: tests of data quality, scaling assumptions, reliability and validity in healthy men and women. Health Qual Life Outcomes. 2008;6:52.
14. Wagner AK, Gandek B, Aaronson NK, et al. Cross-cultural comparisons of the content of SF-36 translations across 10 countries: results from the IQOLA Project. International quality of life assessment. J Clin Epid. 1998;51(11):925-32.
15. Gandek B, Sinclair SJ, Kosinski M, Ware JE, Jr. Psychometric evaluation of the SF-36 health survey in medicare managed care. Health Care Fin Rev. 2004;25(4):5-25.