J Anesth (2015) 29:463–466
DOI 10.1007/s00540-014-1945-0

SHORT COMMUNICATION

Feasibility, reliability, and validity of the Japanese version of the Postoperative Quality of Recovery Scale: a first pilot study

Yusuke Naito · Yuu Tanaka · Noriyuki Sasaoka · Toshio Iwata · Yuko Fujimoto · Nozomi Okamoto · Satoki Inoue · Masahiko Kawaguchi

Received: 17 June 2014 / Accepted: 25 October 2014 / Published online: 11 November 2014

Abstract The Postoperative Quality of Recovery Scale (PQRS) is a multi-domain tracking scale to assess recovery after surgery. The PQRS is used in seven countries and five languages; however, the Japanese version of the PQRS (PQRSj) has not been established. We therefore translated the PQRS into Japanese, and examined the feasibility, reliability and validity of the PQRSj. The time taken to complete the test was measured to assess feasibility. Cronbach's alpha was calculated to assess reliability. The Mini Mental State Examination (MMSE) and the Japanese version of the Quality of Recovery Scale 40 (QoR-40 J) were performed for comparison with the PQRSj (validity). Fifty-one patients were enrolled in the study. The mean completion time was 3.9 min for baseline (feasibility). Cronbach's alpha was between 0.40 and 0.94 in each domain (reliability). A relationship was shown between cognitive domain and MMSE at baseline ($r = 0.65$, $P < 0.01$); however, no relationship was found between the other domains and the MMSE and QoR-40 J. Ceiling effects were observed in 78 % of the questions. These results indicate that the PQRSj can be used to assess recovery after surgery, although it may be better to revise some of the questions to improve the validity of the PQRSj.

Y. Naito (✉) · Y. Tanaka · S. Inoue · M. Kawaguchi
Department of Anesthesiology, Nara Medical University, 840 Shijo-cho Kashihara, Nara, Japan
e-mail: schneider.yusuke@gmail.com

N. Sasaoka · T. Iwata · Y. Fujimoto
Department of Anesthesiology, Nara Prefectural Mimuro Hospital, Nara, Japan

N. Okamoto
Department of Community Health and Epidemiology, Nara Medical University, Kashihara, Nara, Japan

Keywords Recovery · Scale · Translation · Reliability · Validity

Postoperative recovery is rapidly recognized as an important factor after surgery [1–10]. There are several scales to assess recovery but most of them are written in English [3]. When translating English scales to Japanese, it is important that the questions are translated well linguistically, and also adapted culturally to maintain validity [4, 5]. The Postoperative Quality of Recovery Scale (PQRS) is a multi-domain tracking scale, which is used in five languages and seven countries [1]. It consists of six domains with each domain comprising a series of questions. The PQRS is designed to be used multiple times after surgery to follow recovery (Appendix). Baseline assessment is mandatory because recovery is defined as a score more than or equal to baseline.

Since the assessment of patient recovery after surgery is crucial, it is important to use a reliable and valid scale. Therefore, we translated the PQRS into Japanese and assessed the feasibility, reliability, and validity of the Japanese version of the PQRS (PQRSj).

After the study was approved by the Institutional Review Board of Nara Prefectural Mimuro Hospital, Japan, it was performed in two stages. In the first stage, the scale was translated into Japanese and culturally adapted. In the second stage, the PQRSj was tested among patients undergoing elective surgery under general anesthesia. Since this is a first pilot study, patients were enrolled in one hospital. The study was registered in the UMIN Clinical Trial Registry before recruitment of the first subject (registration number R000012577).

The translation and cross-cultural adaptation were performed following guidelines by Beaton et al. [11]. It is

composed of four stages—translation, synthesis, back translation, and expert committee review. Two anesthesiologists performed the initial translation into Japanese individually (translation). The translations were combined into one common translation based on the original scale (synthesis). The common translation was then translated back into English by two English translation experts who worked independently and did not have any background on the medical concepts or any knowledge of the PQRS (back translation). An expert committee then reviewed the two back translations and combined them into one common English version. The English back-translated version was compared with the original version and checked for consistency (expert committee review). The final Japanese version together with the back-translated version was finally approved and accepted for use.

Written informed consent was obtained from each patient before enrollment in the study. Criteria for the study included patients undergoing elective surgery under general anesthesia, able to speak and write Japanese, American Society of Anesthesiologist Performance Status 1–3, and aged >20 years. Exclusion criteria included patients planned to transfer to the intensive care unit, not extubated in the operating room, requiring spinal anesthesia, and patients undergoing neurosurgery that could possibly reduce their cognitive function. The study period extended from August 2012 to July 2013.

Baseline measurements of the PQRSj were performed within 2 days before surgery. The PQRSj assessment was then carried out at 15 and 40 min after the end of the surgery (T_{15m} and T_{40m}, respectively) in the operating room or in the recovery unit. At these time points, we assessed the physical domain and one question (time, date, or place) in the cognitive domain. Full assessments were performed at 1 day and 3 days after surgery (T_{1d} and T_{3d}, respectively). Assessment was not performed at 3 months after surgery (T_{3m}) because almost all the patients were discharged and refused to be assessed over the telephone.

We collected completion times and incompletion rates to assess feasibility. Reliability was assessed in two ways. First, it was assessed by calculating Cronbach's alpha. Second, the recovery rate was compared with that of original scale to assess reliability. In order to assess validity, the Mini Mental State Examination (MMSE), which is a scale for the diagnosis of cognitive decline and the Japanese version of the Quality of Recovery Scale 40 (QoR-40 J), which is a postoperative recovery score, were assessed at baseline and T_{3d} and the results of the PQRSj were compared with the MMSE and QoR-40 J results.

Statistical analyses were performed with Statflex software (version 6.0 for Windows; Artech Co., Tokyo, Japan). Paired *t* test and Spearman rank correlation tests were used where appropriate. A ceiling effect or floor effect was limited to 15 % [4, 12]. A *P*-value <0.05 was considered statistically significant.

Table 1 Baseline patient characteristics and operative variables

	Range	Mean (SD)
Age (years)	42–88	68.7 (10.7)
Weight (kg)	34–96	60.9 (11.4)
Height (cm)	145–177	159.5 (8.2)
Duration of anesthetic (min)	42–553	195.7 (139.4)
	Frequency	Percent
Gender (male)	27	64.3
American Society of Anesthesiologist status		
1	4	9.5
2	37	88.1
3	1	2.4
Surgery type		
General	12	28.6
Ear, nose, and throat	5	11.9
Urologic	6	14.3
Opthalmology	2	4.8
Thoracic	1	2.4
Orthopedic	15	35.7
Plastic	1	2.4

Demographic and operative variables for the 42 patient who were able to participate in baseline assessment

Fifty-one patients were enrolled in the study. Baseline and demographic details of the cohort are shown in Table 1. Incompletion rates were 17.7 %, 9.6 %, and 23.7 % for baseline, T_{1d}, and T_{3d}. The mean completion times were 3.9 s (SD 1.0), 5.5 s (SD 1.5), and 5.5s (1.5) for baseline, T_{1d}, and T_{3d}. The main reason for incompletion was lack of privacy (47.1 %) followed by pain or nausea (33.5 %).

Reliability which is concerned with the ability of a test to measure consistency was measured using Cronbach's alpha which is a standard analysis to assess reliability in the field of psychometric tests. Cronbach alpha >0.6 is a reliable value. The Cronbach alpha results in this study were 0.40, 0.46, 0.93, 0.94, 0.81, and 0.65 for physical, nociceptive, activity of daily living, emotional, overall perspective and cognitive domains. Full recovery for all domains was observed in 3.4 % of patients in T_{1d} and 11.1 % in T_{3d}. Recovery rates for each domain are shown in Fig. 1.

Validity is the extent to which a test measures what it claims to measure and it is vital for a test to be valid in order for the results to be accurately applied and interpreted. There are several types of validity and it is not determined by a single statistic. Construct validity includes the degree of correlation between an instrument and other scales that assess similar concepts. The PQRS was compared with the QoR-40 J for domains other than cognitive

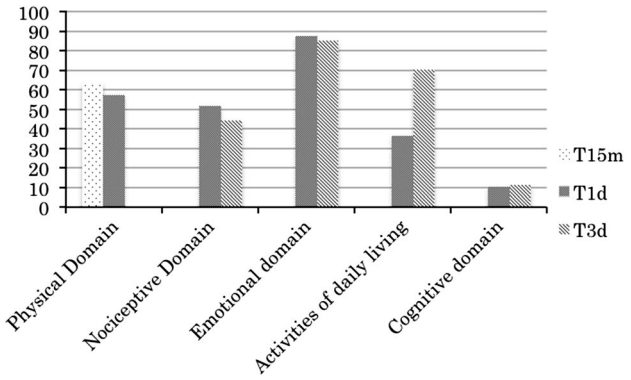

Fig. 1 Recovery rate for each domain. The percentage of the 42 patients showing recovery in each domain. Recovery is defined as return to baseline values. Note that only the physical domain is fully assessed at T_{15m}. The physical domain is not assessed at T_{3d}

domain, and with MMSE for cognitive domain. The correlation coefficient was calculated using Spearman rank correlation test. A relationship was found between the cognitive domain of the PQRSj and MMSE at baseline measurement ($r = 0.65$, $P < 0.01$); however, no relationship was found between the other domains of the PQRSj and the MMSE and QoR-40 J. The ceiling effect is the upper limit of a questionnaire. It is the top score a patient can score on a test regardless of the patient's ability. When a patient reaches the ceiling of a test, it means that the questions on the test were unable to measure true ability. The ceiling effects for all the questions on the questionnaire were calculated as—baseline, 14 out of 18 at T_{15m}, 6 out of 10 at T_{40m}, 7 out of 10 at T_{1d}, 11 out of 22 at T_{3d}, and 4 out of 13 questions had a ceiling effect of $\geq$15 %.

Our results show that feasibility was fair but a high incompletion rate should be addressed. Reliability was confirmed as excellent, and validity can be improved by modifying questions.

The completion time was short and similar to that of the original scale; most of the patients were able to complete the test within 5 min. A total of 43.1 % patients, however, refused to start the test, mainly due to lack of privacy. On the other hand, if we had provided more privacy for the patient, the incompletion rate may have been higher because some of the patients were not allowed to walk freely in the hospital, especially on POD1. A similar score was also found for the cognitive domain in patients on POD1 and POD3. Therefore, it might be feasible to test cognitive domain only in POD3 in future research.

Reliability using Cronbach's alpha was excellent. Although the physical and nociceptive domains had a value <0.6, each question in these domains represents a different concept (e.g., blood pressure is not related to SpO_2).

Content validity was measured by comparison with other scales. The cognitive domain showed a positive relationship with MMSE at baseline. No relationship was seen between the QoR-40 J and the PQRSj; however, this may be due to recall bias. Since the QoR-40 J was retrieved on POD5, it might have contained all the complaints from PODs 0–5. Another reason may be the 'learning effect' as patients were asked to answer the same questions multiple times. Physical factors such as requirement of airway reached the ceiling by almost 100 %. In Japanese hospitals, which usually have no post-anesthetic care unit, it is common to transfer the patient to the general ward directly after surgery. Therefore, it is uncommon to discharge patients who require an airway device. These factors may not be necessary in our situation.

There are some limitations to this study. First, it was conducted in one facility. Our hospital has approximately 300 beds and is classed as being intermediate scale in Japan. The results in the study may not reflect all hospitals in Japan. Second, the anesthesiologist who was responsible for direct patient care performed most of the study. This might have increased the score, for example in patient satisfaction in the overall perspective domain.

In the present study, we translated the PQRS into Japanese. We found that the PQRSj has a fair level of feasibility and reliability was excellent, suggesting that the PQRSj can be used for recovery after surgery. However, it may be better to revise some of the questions to improve validity of the PQRSj.

Appendix

See Table 2

Table 2 Six domains and the questions in each domain of the PQRS

Domains	Questions	Assessed time points
Physical domain[a]	Blood pressure, heart rate, ventilation rate, temperature, SpO_2, airway, agitation, consciousness, response	Baseline, T_{15m}, T_{40m}, T_{1d}
Nociceptive domain	Pain, nausea	Baseline, T_{15m}, T_{40m}, T_{1d}, T_{3d}, T_{3m}
Emotional domain	Depressed, anxious	Baseline, T_{15m}, T_{40m}, T_{1d}, T_{3d}, T_{3m}
ADL[b] domain	Ability to stand, walk, eat, dress	Baseline, T_{40m}, T_{1d}, T_{3d}, T_{3m}
Overall perspective	Ability to work, clarity of thought, activities of daily living, satisfaction with anesthetic care	T_{1d}, T_{3d}, T_{3m}
Cognitive domain	Name, date, place, digits forward, digits backward, Word list, word generation	Baseline, T_{15m}, T_{40m}, T_{1d}, T_{3d}, T_{3m}

[a] Consciousness and response are not assessed at T^{b}_{1d}

ADL Activities of daily living

Time zero (T_0) is defined as the time point after which anesthesia is no longer required. Assessment is performed at 15 and 40 min after T_0 as T_{15m} and T_{40m}. These time points are mainly designed to assess recovery at the point of discharge from the operating room or recovery unit. Acute-phase recovery is assessed at time points on postoperative day (POD) 1 as T_{1d} and POD3 as T_{3d}. Long-term recovery is assessed at 3 months (T_{3m})

References

1. Royse CF, Newman S, Chung F, Stygall J, McKay RE, Boldt J, Servin FS, Hurtado I, Hannallah R, Yu B, Wilkinson DJ. Development and feasibility of a scale to assess postoperative recovery: the post-operative quality recovery scale. Anesthesiology. 2010;113:892–905.
2. Tanaka Y, Wakita T, Fukuhara S, Nishiwada M, Inoue S, Kawaguchi M, Furuya H. Validation of the Japanese version of the quality of recovery score QoR-40. J Anesth. 2011;4:509–15.
3. Herrera FJ, Wong J, Chung F. A systematic review of postoperative recovery outcomes measurements after ambulatory surgery. Anesth Analg. 2007;105:63–9.
4. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.
5. Garratt AM, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. Br Med J. 2002;324:1417–9.
6. Myles PS, Hunt JO, Nightingale CE, Fletcher H, Beh T, Tanil D, Nagy A, Rubinstein A, Ponsford JL. Development and psychometric testing of a quality of recovery score after general anesthesia and surgery in adults. Anesth Analg. 1999;88:83–90.
7. Talamini MA, Stanfield CL, Chang DC, Wu AW. The Surgical Recovery Index. Surg Endosc. 2004;18:596–600.
8. Hogue SL, Reese PR, Colopy M, Fleisher LA, Tuman KJ, Twersky RS, Warner DS, Jamerson B. Assessing a tool to measure patient functional ability after outpatient surgery. Anesth Analg. 2000;91:97–106.
9. Swan BA, Maislin G. Traber KB Symptom distress and functional status changes during the first seven days after ambulatory surgery. Anesth Analg. 1998;86:739–45.
10. Oakes CL, Ellington KJ, Oakes KJ, Olson RL, Neill KM, Vacchiano CA. Assessment of post anesthesia short-term quality of life: a pilot study. AANA J. 2002;70:267–73.
11. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25:3186–91.
12. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995;4:293–307.