Development of the Functional Recovery Index for Ambulatory Surgery and Anesthesia

Jean Wong, M.D., F.R.C.P.C., Doris Tong, M.D., F.R.C.P.C.,* Yoshani De Silva, B.Sc.,† Amir Abrishami, M.D.,‡ Frances Chung, M.D., F.R.C.P.C.§*

Background: It is increasingly important to evaluate patients' recovery after ambulatory surgery. The authors developed the Functional Recovery Index (FRI) to assess postdischarge functional recovery for ambulatory surgical patients.

Methods: The scale development involved four phases: item generation, item selection, reliability, and validity testing. A draft questionnaire was tested and revised. Items were selected through testing endorsement frequency, factor analysis, and testing internal consistency. The interrater reliability was calculated. Construct validity was tested by multiple hypotheses on convergent validity, extreme groups, and discriminant validity. Responsiveness was assessed by measuring the FRI postoperatively and comparing minor *versus* more extensive surgery. The rate of response and the time for completion of the questionnaire were recorded.

Results: The final questionnaire had 14 items grouped under 3 factors. Each item was scored from 0 to 10, with 0 = no difficulty and 10 = extreme difficulty with the activity. The 3 factors were summated for a total score. Internal consistency for the 3 factors (pain and social activity, lower limb activity, and general physical activity) was as follows: Cronbach α = 0.90, 0.89, and 0.86, respectively. Interrater reliability was 0.99. Convergent validity for FRI *versus* verbal rating scale pain score was 0.76. Discriminant validity testing showed that the type of surgery was significant and that intermediate (β = 0.138) and major surgery (β = 0.337) were associated with higher FRI scores than minor surgery. The time to complete the questionnaires ranged between 4 min 10 s and 4 min 35 s.

Conclusions: The FRI had excellent reliability, good validity, responsiveness, and acceptability, indicating that this questionnaire will be a good instrument for assessing functional recovery of ambulatory surgical patients.

AS more complex procedures on higher risk patients are performed as ambulatory surgery, it is increasingly important to evaluate patients' recovery after their hospital discharge. Data on postdischarge recovery are crucial not only as indicators for quality of care, but also as outcome measurements for the evaluation of new surgical and anesthetic techniques being developed for ambulatory surgery.

In-hospital morbidity, such as unanticipated admission, delayed discharge,[1–3] and postdischarge morbidity, such as readmission and symptom severity, have been used as adverse outcomes for ambulatory anesthesia.[1] However, the advances in both surgical and anesthetic techniques, particularly in ambulatory surgery, have made mortality and major morbidity rare events; therefore, the patient's quality of life, *i.e.*, the ability to resume normal activities after discharge home, should be considered one of the principal endpoints after ambulatory surgery and anesthesia. Because functional recovery of various aspects of patients' daily living is a subjective assessment, it is imperative that the assessment be made from the patients' perspective. Therefore, as with the measurement of any other subjective outcome, an instrument that has undergone vigorous methodologic development and testing for reliability and validity is needed. The existing instruments to evaluate postoperative recovery after ambulatory surgery do not meet all of these criteria or were not designed specifically for ambulatory surgical patients.[4] The use of a standardized instrument across clinical trials for ambulatory surgery and anesthesia would allow better comparisons between trials.

The aim of this study was to develop a new instrument, the Functional Recovery Index (FRI), to assess postdischarge functional recovery for ambulatory surgical patients. The reliability and validity of this instrument was determined using conventional psychometric methodologies.

Materials and Methods

The scale development involved four phases: item generation, item selection, reliability, and validity testing.

Patients were recruited from two sites, The Toronto Western Hospital and The Princess Margaret Hospital, Toronto, Ontario, Canada. The total number of patients recruited for developing the FRI was 688.

Patients were eligible for inclusion if they were older than 16 yr, undergoing any ambulatory surgical procedure with same-day discharge or 23-h stay, had the ability to participate in a postoperative phone interview, and were able to read English.

Patients were excluded if they had a history of alcohol and drug abuse, active mental dysfunction or cognitive deficiency, or any serious perioperative complication necessitating admission.

Informed consent was obtained, and approval for the study was obtained from the University Health Network Research Ethics Board, Toronto, Ontario, Canada.

Scale Development

Items were generated from three sources: literature search, patients, and content expert interviews.

596

The literature was searched to find validated instruments for assessing postdischarge functional recovery after ambulatory anesthesia. MEDLINE 1966–2007, HealthSTAR 1966–2007, and EMBASE 1966–2007 were searched with the following text words: *recovery, functional, function, outcome, measure, measuring, measurement, health status, symptom distress, quality of life.* Established instruments were reviewed to provide relevant dimensions and items. Measurement-related articles in anesthesia[5-13] were checked for relevant citations.

Patients were interviewed using open-ended questions (appendix 1). Patients were encouraged to volunteer any concepts they considered important to postoperative recovery and any expressions they would use to describe postoperative recovery. The principle for the interviews was "sampling to redundancy,"[14] *i.e.,* the interviewing process was terminated when no new items were generated from at least three consecutive interviews. Any items suggested by at least one patient were included in the draft scale. Forty patients were interviewed for this stage.

The draft scale, consisting of 34 items divided among six dimensions and formatted using the basic structure of the Medical Outcomes Study 36-Item Short-Form Health Survey,[15] was forwarded to 11 content experts in Canada and the United States (appendix 2).

These experts were committee members of the Society of Ambulatory Anesthesia or well-known experts in ambulatory anesthesia. The experts were encouraged to review the scale to suggest additional items and to modify existing items.

Items retained after content expert review were scaled from 0 (no difficulty) to 10 (extreme difficulty) and were pretested on a minimum of 50 patients on day 3 after surgery for feasibility (time, ease of administration) and ambiguity of items (jargon, double-barreled questions, negative/value-laden wordings, length of items). A continuous scale (0–10) was chosen because a categorical scale can result in a loss of information and reliability.[16]

Items were selected through the following steps: testing endorsement frequency, factor analysis, and testing internal consistency.

The preliminary FRI scale was pretested to determine the proportion of patients choosing each score on the scale for each item (endorsement frequency). A histogram of the responses was constructed, and items with an endorsement rate of less than 0.2 or greater than 0.8 were eliminated.

Retained items were subjected to factor analysis to determine the underlying factor/dimensions. Items with communalities of 0.6 or greater were retained. Factor analysis was performed with orthogonal rotation, assuming that the factors were mutually exclusive and selected with a Catell Scree plot.[16]

The critical value for retention of items was 0.39 based on $5.152/\sqrt{N-2}$ where N indicates sample size.[16] Items not loading on any factor (*i.e.,* factor loading < 0.39), items loading on more than one factor, and factors with less than three items per factor were eliminated. If many items loaded on one factor, redundant items were eliminated. However, a minimum of three items per factor had to be retained.

After factor analysis, the internal consistency of each item within a factor was determined. A Cronbach α was calculated, and items with an α less than 0.7 or greater than 0.9 were eliminated.

Scale Evaluation

Each patient was interviewed by telephone by each of two research assistants, 30 min apart from each other to determine the interrater reliability of the FRI. These patients were interviewed on day 3 after surgery by telephone. The intraclass correlation coefficient was determined.

Face and content validity were determined through patient and content expert interviews during the scale development phase.

To determine construct validity, we tested five hypotheses:

1. That the FRI would correlate with the verbal rating scale for pain (0 = no pain, 10 = severe pain), because pain constitutes one aspect of postoperative recovery (convergent validity).
2. That the FRI would correlate positively with the hours of restricted activity, duration in recovery room, and duration of hospital stay.
3. That age, American Society of Anesthesiologists physical status, duration of surgery, intermediate surgery, *e.g.,* knee arthroscopy, and major surgery, *i.e.,* microdiscectomy and mastectomy and axillary dissections, were significant predictors of FRI (discriminant validity).
4. That function would improve over time and the FRI would decrease in response to increasing function (lower score indicating an improvement in functional recovery [responsiveness]).
5. That the FRI change score would improve most dramatically after more extensive surgery when compared with minor surgery. As such, the change in the score of patients having intermediate surgery, *e.g.,* knee arthroscopy, and microdiscectomy (major surgery) should be greater than those having dilatation and curettage and eye procedures (minor surgery).

To assess acceptability, the rate of response and the time for completion of the questionnaire were recorded.

The final score was adjusted for items that were not answered or not applicable for the patient.

Statistical Analysis

Nonparametric or parametric tests were used where appropriate. Data were analyzed with SAS 9.1 for Windows (SAS Institute, Cary, NC) or SPSS 16.0 for Windows

(SPSS Inc., Chicago, IL). A *P* value less than 0.05 was considered significant.

A minimum of 50 patients were required to test endorsement.[14] Factor analysis requires a sample size of 5–10 subjects per item on the scale. Because the draft scale contained 34 items, approximately 180 patients were required. To determine internal consistency for the correlation of each item with each other within the factor/dimension, Cronbach α was calculated. Because grouping of items under dimension is essentially another test for internal consistency, the scores from the same 180 patients were used. The intraclass correlation coefficient was calculated using the following formula: intraclass correlation coefficient = σ^2 patients/(σ^2 patients + σ^2 observers + σ^2 error), where σ^2 patients = variance attributed to differences among patients; σ^2 observers = variance attributed to differences among observers; and σ^2 error = variance attributed to random error that is inversely related to reliability.[17] Assuming the intraclass correlation coefficient to be at least 0.7 with a confidence interval ± 0.1 and α error of 0.05, the sample size requirement for interrater reliability testing was 130 patients.[18] Assuming a Pearson/Spearman correlation of 0.5–0.6 with a confidence interval of ± 0.1 and α error of 0.05, the sample size requirement for testing convergent validity was 189–247 patients.[17] For discriminant validity testing for multivariable analysis, the sample size requires 10 subjects per item analyzed; because there were 14 items, 140 patients were required. To calculate responsiveness, the Friedman test (the nonparametric equivalent of a one-way analysis of variance for comparing repeated measures) was used to analyze the changes in the scores during the postoperative period (postoperative days 1, 3, 5, and 7) as compared with the baseline, *i.e.*, the preoperative value. Responsiveness was evaluated in 100 patients. The change in FRI score on postoperative day 7 from baseline (preoperative) was compared between patients undergoing minor surgery (dilatation and curettage or eye surgery) and those having more extensive surgery using the Mann–Whitney *U* test, and a *P* value less than 0.05 was considered statistically significant.

Results

Scale Development

Thirty-four preliminary scale items were compiled from a review of the literature and interviews with patients. A further 21 were compiled after content expert input, for a total of 55 items comprising the draft FRI. These 55 items were grouped under six factors/dimensions, including basic activities of daily living, intermediate activities of daily living, role limitation, social function, mental function, and symptom distress. The 55-item draft questionnaire was pretested with 76 patients, and 9 of the 55 items were eliminated because they had an endorsement frequency of greater than 0.8.

Because the modified questionnaire now had 46 items,

Table 1. Cohort A (n = 324)

Age, yr	46 (13–85)
ASA physical status, I/II/III*	136/153/16
Sex, M/F	132/192
Duration of surgery, min	45 (15–165)
Duration of PACU stay, min	45 (10–240)
Duration of hospital stay, min	145 (40–1,820)
Duration of restricted activity, h	36 (0–72)
Type of surgery	
Dilatation and curettage	28 (8.6%)
Hysterectomy	15 (4.6%)
Laparoscopy	26 (8.0%)
Laparoscopic cholecystectomy	5 (1.5%)
Inguinal hernia repair	9 (3%)
Breast surgery	7 (2.2%)
Mastectomy	13 (4%)
Arthroscopy	91 (28.1%)
Microdiscectomy	16 (5%)
Plastic surgery	37 (11.4%)
Urologic surgery	10 (3.1%)
Eye surgery	53 (16.4%)
Other	14 (4.3%)

Values are median (range), ratio, or number (percentage).

* Nineteen patients had missing American Society of Anesthesiologists (ASA) physical status values.

PACU = postanesthesia care unit.

with a sample size of 5 subjects per item, a minimum of 230 patients were to be interviewed for factor analysis.

Three hundred twenty-four patients (cohort A) were interviewed for factor analysis. The demographic data and types of surgery for cohort A are shown in table 1. The surgery was performed with general anesthesia in 56% of the patients, 15% of the patients received regional anesthesia, and 29% of the patients had monitored anesthesia care.

Diagnostic checks removing items with communalities less than 0.6 removed 12 items, leaving 34 items. Principal component analysis was used to extract factors. The factor analysis with 4 factors was performed and was found to be satisfactory because a total of 77% of the variance of the data were explained by the 4 factors. Factor loading matrix was determined between the 34 items and 4 factors. Varimax rotation was performed to optimize the loading distribution. One factorially complex item was eliminated. Seventeen items loaded well on factor 1; however, of these, 7 items with the lowest values were eliminated to shorten the questionnaire to 26 items. Initially, 4 factors were chosen; however, because factor 4 (which included 4 items) had an α of only 0.43, it was eliminated. The factor analysis with 3 factors showed that a total of 74% of the variance of the data were explained by the 3 factors. The factors identified did not confirm the original dimensions assigned. Factor 1 suggested pain and social activity, factor 2 suggested lower limb activity, and factor 3 suggested general physical activity.

Internal consistency testing of the revised scale in cohort A patients led to elimination of another 8 items, leaving 14 items grouped under 3 factors (table 2). Each item was scored from 0–10 with 0 = no difficulty to 10 = extreme

Table 2. Factors and Interpretation of Scores

Factor	Number of Items	Number of Levels	Meaning of Scores	
			Low	High
1: Pain and social activity	7	0–10	Performs normal social activities without interference due to pain or physical or emotional problems	Frequent interference with normal social activities due to pain or physical or emotional problems
2: Lower limb activities	4	0–10	Performs lower limb activities without difficulty	Performs lower limb activities with limitations
3: General physical activity	3	0–10	Performs general physical activities without difficulty	Performs general physical activities with limitations

difficulty with the activity. The 3 factors were summated for a total score, and the final score was adjusted for items that were not answered or not applicable for the patient. A lower score indicates better recovery, whereas a higher score indicates greater difficulty with recovery; the maximum score would be 140.

Internal consistency for the 3 factors was as follows: Cronbach α = 0.90, 0.89, and 0.86, respectively. The final questionnaire had 14 items (appendix 3).

Scale Evaluation

Interrater reliability as assessed in 264 patients (cohort B; table 3) by intraclass correlation was found to be 0.99, indicating high reliability.

Validity testing was performed on cohort B. The convergent validity for FRI *versus* verbal rating scale was 0.76 (Spearman correlation coefficient), and that for hours of restricted activity was 0.72. The FRI score did not correlate well with postanesthesia care unit duration or hospital stay (0.39 and 0.39, respectively).

Table 3. Cohort B (n = 264)

Age, yr	44 (13–85)
ASA physical status, I/II/III*	187/55/6
Sex, M/F	95/169
Duration of surgery, min	45 (10–165)
Duration of PACU stay, min	45 (10–240)
Duration of hospital stay, min	147 (40–1,820)
Duration of restricted activity, h	24 (0–72)
Type of surgery	
Dilatation and curettage	26 (9.8%)
Hysterectomy	16 (6.1%)
Laparoscopy	14 (5.3%)
Arthroscopy	37 (14.0%)
Inguinal hernia repair	13 (4.9%)
Breast surgery	5 (1.9%)
Mastectomy	21 (8.0%)
Plastic surgery	33 (12.5%)
Urologic surgery	19 (7.2%)
Eye surgery	35 (13.3%)
Laparoscopic cholecystectomy	16 (6.1%)
Other	28 (10.6%)

Values are median (range), ratio, or number (percentage).

* Some cases had missing American Society of Anesthesiologists (ASA) physical status values.

PACU = postanesthesia care unit.

Discriminant validity testing using multiple regression of the square root transformed scores of the FRI score *versus* age, American Society of Anesthesiologists physical status, sex, type of anesthetic, and types and duration of surgery demonstrated that the type of surgery was the only significant variable ($P < 0.0001$). The types of surgery were classified into three groups: minor (dilatation and curettage and eye surgery), major (mastectomy and axillary dissections and microdiscectomy), and the rest were considered intermediate surgery. The coefficients for intermediate surgery ($\beta = 0.138$) and for major surgery ($\beta = 0.337$) indicated that intermediate and major surgery was associated with higher FRI scores than minor surgery.

Responsiveness

One hundred patients (cohort C; table 4) were evaluated for responsiveness of the FRI. The Friedman test showed that there was a significant difference between preoperative and postoperative day 1, 3, and 7 scores (all $P < 0.001$), but not postoperative day 5 (P = not significant), because the scores returned to baseline (fig. 1). *Post hoc* analysis also showed that there is a statistically significant difference in the FRI score among the different postoperative days (all $P < 0.01$).

Table 4. Cohort C (n = 100)

Age, yr	43 (18–78)
Sex, M/F	50/50
ASA physical status, I/II/III	53/40/70
Duration of surgery, min	65 (45–125)
Duration of PACU stay, min	25 (15–70)
Duration of hospital stay, min	113 (80–265)
Type of surgery	
Arthroscopy	9 (9%)
Dilatation and curettage	7 (7%)
Eye surgery	16 (16%)
Inguinal hernia repair	4 (4%)
Laparoscopic cholecystectomy	12 (12%)
Upper limb surgery	18 (18%)
Lower limb surgery	6 (6%)
Microdiscectomy	5 (5%)
Urologic surgery	5 (5%)
Other	18 (18%)

Values are median (range), ratio, or number (percentage).

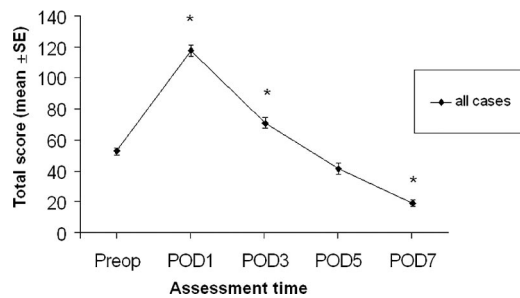ASA = American Society of Anesthesiologists; PACU = postanesthesia care unit.

Fig. 1. Functional Recovery Index scores during the postoperative period. * Significant difference (*P* < 0.001) from baseline (*i.e.*, preoperative [Preop]). POD = postoperative day.

There were only 5 patients having major surgery (microdiscectomy), *versus* 23 patients in the minor group and 72 patients in the intermediate group. Because the unequal numbers in each group may lead to misleading results, we compared knee arthroscopy (9 patients), representative of intermediate surgery, with minor surgery (dilatation and curettage and eye surgery). There was a significant difference between the scores among the knee arthroscopy and minor surgical procedures at each point of time (*P* < 0.01; fig. 2) for preoperative and postoperative days 1, 3, and 5 but not for postoperative day 7 assessments (*P* = 0.08). The change in FRI scores on postoperative day 7 compared with baseline was greater for knee arthroscopy *versus* minor surgery, *i.e.*, median (range) 32.6 (−7.0 to 66.0) *versus* 15.3 (−26 to 68) (*P* = 0.027).

Acceptability

Follow-up was complete for 92% of the cohort A patients and 87% of the cohort B patients tested for validity and reliability. Of the 100 patients involved in the evaluation of responsiveness, all patients completed at least one of the four follow-up telephone questionnaires. The time to complete the questionnaires was similar at the various times of administration (days 1, 3, 5, and 7) and on average ranged between 4 min 10 s and 4 min 35 s.

Discussion

We have developed a 14-item FRI to evaluate postoperative functional recovery after ambulatory anesthesia and surgery. The FRI demonstrated excellent reliability and good convergent validity for verbal rating scale pain scores and hours of restricted activity. Discriminate validity testing revealed that the FRI score could discriminate between major and intermediate surgery when compared with minor surgery. The questionnaire demonstrated good responsiveness to detect the changes in functional recovery during the postoperative period and between different types of ambulatory surgical procedures.

The final FRI questionnaire consisted of 14 items grouped under 3 factors. The FRI was designed to be administered by telephone interviews. Given the intended use of the instrument for phone interviews, it was decided
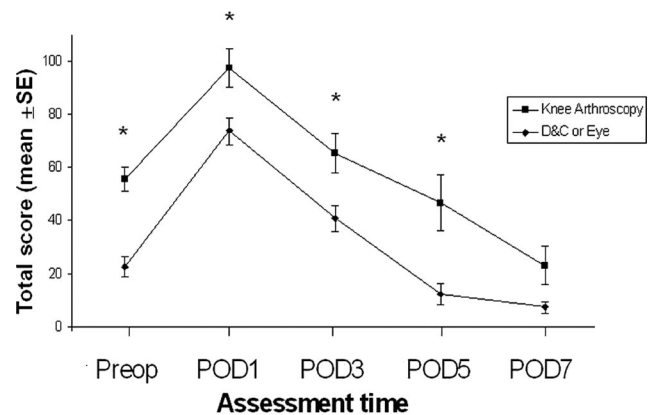

Fig. 2. Functional Recovery Index scores for knee arthroscopy and minor surgeries. * Significant difference (*P* < 0.01) in the Functional Recovery Index scores between knee arthroscopy and minor surgeries. D&C = dilatation and curettage; POD = postoperative day; Preop = preoperative.

that a familiar verbal rating format (0–10) for each item would be used. Nishisato and Torii[19] have performed simulation models on two variables of predefined correlations using different numbers of steps, and found that any scale with less than 5–7 steps is unreliable in assessing subjective outcomes. The questionnaire could be completed in less than 5 min, indicating good acceptability of the questionnaire as a practical tool to assess recovery after ambulatory surgery. The FRI was tested on many different types of surgery and various types of anesthetics, including general, regional, and monitored anesthesia care. The FRI items and scales were constructed for scoring using the Likert method of summated ratings. The use of a summary score should enhance feasibility and practical use of this instrument by clinicians and researchers. The final score is adjusted for items that were not answered or not applicable for the patient.

Health-related quality of life is increasingly used as an outcome in clinical trials and research on the quality of health care.[20] There is increasing evidence that measures of health-related quality of life are valid, reliable, and responsive to clinical changes.[21] Unfortunately, existing generic subjective health status measurements are not specifically designed for use after ambulatory surgery. For example, the Medical Outcomes Study 36-Item Short-Form Health Survey[14,15,22] and Sickness Impact Profile are widely accepted to be valid and reliable for measurement of disease severity in chronic medical and psychiatric conditions.[23,24] However, these instruments have been shown to have a ceiling effect on healthy subjects[24] and have never been tested or used in relation to recovery after ambulatory surgery and anesthesia.

In addition, verbal and visual analog scales have been widely used in the assessment of a significant number of postoperative outcomes, such as pain, emesis, and fatigue.[25] However, visual analog scales have not been validated for the assessment of functional recovery after ambulatory surgery and anesthesia.

An instrument that specifically measures functional recovery after ambulatory surgery is important to determine true endpoints for patient functionality so that the necessary changes in ambulatory anesthesia and surgery practices can be guided to achieve better patient outcomes.

The Quality of Recovery (QoR) 9 Score[4,5] is the most commonly cited instrument to assess postoperative recovery after ambulatory surgery and anesthesia. However, this instrument had only moderate validity and reliability (0.5–0.61).[5] Of the existing instruments to evaluate postoperative recovery after anesthesia, the QoR 40 was found to have the best psychometric development.[4] However, both the QoR 9 and the QoR 40 were not specifically developed for use in ambulatory surgical patients. The QoR 40 is too long to be used for phone interviews. Conversely, the FRI is a brief questionnaire (taking approximately 4 min to complete), and its reliability and convergent and discriminant validity are high.

Discriminant validity testing using multiple regression analysis showed that only the type of surgery was a significant predictor of FRI. Age, sex, American Society of Anesthesiologists physical status, type of anesthetic, and duration of surgery were not predictors of FRI. A likely explanation for this finding may be because many of the elderly patients who participated in developing the FRI had minor surgery, *e.g.*, eye surgery. In contrast to the QoR 9 and QoR 40, we did not find that sex was a predictor of FRI. A possible explanation for this finding may be that Myles' questionnaires were developed primarily with inpatients having more extensive surgery.[5,12]

The development of shorter instruments that are easy to understand and administer is important for feasibility and applicability for both clinicians and researchers in ambulatory surgical settings.

Test–retest reliability, defined as the agreement on two occasions separated by some interval of time, was not performed because we expected patients' recovery to be rapid and it would not have changed between interviews on the same day. There is no accepted standard for assessment of postoperative functional recovery after ambulatory anesthesia; therefore, criterion/predictive validity could not be tested.

One of the limitations of the FRI is that the score does not adjust for patients who had prolonged recovery room stay (> 2 h), those who were fast-tracked to discharge, and those who were admitted for nonserious complications or readmitted to an acute care facility postoperatively. Unlike previous questionnaires that have included the item nausea and/or vomiting,[5,12] this was one of the items that was eliminated from the final FRI questionnaire after factor analysis.

In conclusion, the development of the Functional Recovery Index followed the steps recommended for rigorous psychometric questionnaire construction.[18,26] This instrument can be used as a postoperative outcome measure for the evaluation of new ambulatory anesthesia and surgical techniques in future randomized controlled trials. The impact of interventions and changes in practice should be assessed using outcomes of concern to patients. This validated and reliable instrument for postdischarge functional recovery after ambulatory surgery and anesthesia will provide a valuable tool for postoperative quality assessment, particularly with the increasing number and complexity of surgeries performed in the ambulatory setting.

References

1. Chung F, Un V, Su J: Postoperative symptoms 24 hours after ambulatory anaesthesia. Can J Anaesth 1996; 43:1121-7

2. Chung F: Recovery pattern and home-readiness after ambulatory surgery. Anesth Analg 1995; 80:896–902

3. Fortier J, Chung F: Unanticipated admission after ambulatory surgery: A prospective study. Can J Anaesth 1998; 45:612–9

4. Herrera FJ, Wong J, Chung F: A systematic review of postoperative recovery outcomes measurements after ambulatory surgery. Anesth Analg 2007; 105:63–9

5. 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

6. Dexter F, Aker J, Wright WA: Development of a measure of patient satisfaction with monitored anesthesia care: The Iowa Satisfaction with Anesthesia Scale. ANESTHESIOLOGY 1997; 87:865–73

7. Fung D, Cohen MM: Measuring patient satisfaction with anesthesia care: A review of current methodology. Anesth Analg 1998; 87:1089–98

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 K: 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. Kleinbeck SVM: Self-reported at-home postoperative recovery. Res Nurs Health 2000; 23:461–72

12. Myles PS, Weitkamp B, Jones K, Melick J, Hensen S: Validity and reliability of a postoperative quality of recovery score: The QoR-40. Br J Anesth 2000; 84:11–5

13. Talamini MA, Stanfield CL, Chang DC, Wu AW: The surgical recovery index. Surg Endosc 2004; 18:596–600

14. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD: The MOS 36-item Short-Form Health Survey (SF-36), III: Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994; 32:40–66

15. Ware JE Jr, Sherbourne CD: The MOS 36-item short-form health survey (SF-36), I: Conceptual framework and item selection. Med Care 1992; 30:473–83

16. Biostatistics: The Bare Essentials. Edited by Norman GR, Streiner DL. St. Louis, Mosby-Year Book, 1994, pp 129–48

17. Streiner DL: Sample-size formulae for parametric estimation. Percept Mot Skills 1994; 78:275–84

18. Streiner D, Norman G: Health Measurement Scales: A Practical Guide to Their Development and Use, 3rd edition. Oxford, Oxford University Press, 2003, pp 4–9, 19–22, 61–8, 119–23, 172–87

19. Nishisato N, Torii Y: Effects of categorizing continuous normal distributions on the product-moment correlation. Jpn Psychol Res 1970; 13:45–9

20. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R: Quality of life measurement: Bibliographic study of patient assessed health outcome measures. BMJ 2002; 324:1417

21. Wilson IB, Cleary PD: Linking clinical variables with health-related quality of life. JAMA 1995; 273:59–65

22. McHorney CA, Ware JE Jr, Raczek AE: The MOS 36-Item Short-Form Health Survey (SF-36), II: Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31:247–63

23. Bergner M, Bobbitt RA, Carter WB, Gilson BS: The Sickness Impact Profile: Development and final revision of a health status measure. Med Care 1981; 19:787–805

24. Pollard WE, Bobbitt RA, Bergner M, Martin DP, Gilson BS: The Sickness Impact Profile: Reliability of a health status measure. Med Care 1976; 14:146-55

25. de Boer AG, van Lanschot JJ, Stalmeier PF: Is a single-item visual analogue

scale as valid, reliable, and responsive as multi-item scales in measuring quality of life? Qual Life Res 2004; 13:311–20

26. Liang MH: Longitudinal construct validity: Establishment of a clinical meaning in patient evaluation instruments. Med Care 2000; 38:S84–90

Appendix 1: Open-ended Questions

1. What is your experience in your postoperative recovery?
2. How would you describe your recovery after the surgery?
3. How could we make your recovery better for you the next time?
4. What would you say about the recovery after surgery to a family member who was about to have his or her first anesthetic?
5. What do you expect from your recovery?
6. Is the recovery up to your expectation? If not, what is the reason?
7. Before the operation, how are you prepared for the recovery? (What did the surgeons/anesthetists/nurses tell you about the recovery?)
8. What bothered you most when you recovered from the anesthesia?
9. What aspects of your activity/life are affected by the recovery? (Hint: at work, daily activity, social function, and mental capacity)

Appendix 2. Draft Scale of Functional Recovery Index

Basic ADLs	Intermediate ADLs	Role Limitation
Bathing	Moderate exercise	Return to work
Dressing	Vigorous exercise	Dependency
Walking indoors	Groceries	Cut down time
Lying down	Stairs	Accomplish less
Sitting	Bending	Extra effort
In/out of a chair	Housework	
Mobility*	Driving	
	Public transportation	

Social Function	Mental Function	Symptom Distress
Interfere	Mood*	Nausea and vomiting
Visit	Concentration	Dizzy
Community	Orientation	Sore throat
Caregiving	Drowsy*	Headache
	General*	Pain

* Negatively keyed scales.
ADL = activity of daily living.

Appendix 3. Functional Recovery Index

Factor 1: pain and social activity
Have you experienced difficulty with the following since your surgery:

No difficulty at all ——————— Extreme difficulty
0   1   2   3   4   5   6   7   8   9   10   NA

1   Returning to work
2   Visiting friends, relatives
3   Taking care of other family member
4   Driving a car
5   Moderate exercise, such as moving a table or pushing a vacuum cleaner
6   Lifting or carrying groceries
7   Since your surgery, has pain been a problem?

Factor 2: lower limb activity
Since your surgery, have you experienced difficulty in:

No difficulty at all ——————— Extreme difficulty
0   1   2   3   4   5   6   7   8   9   10   NA

1   Walking indoors or around the house
2   Moving in/out of a chair
3   Climbing stairs
4   Bending, kneeling, or stooping

Factor 3: general physical activity
Since your surgery, have you experienced difficulty in:

No difficulty at all ——————— Extreme difficulty
0   1   2   3   4   5   6   7   8   9   10   NA

1   Taking a bath/shower/cleaning yourself
2   Dressing
3   Lying down
Grand score = (total of all the scores) × 14/number of answered
    questions = _____

If patients do not normally perform such activities, *e.g.*, driving, choose not applicable (NA). If patients were instructed by surgeons not to perform any of the activities, choose NA.