Systematic Review of Questionnaires Measuring Patient Satisfaction in Ambulatory Anesthesia

Pratamaporn Chanthong, M.D.,* Amir Abrishami, M.D.,* Jean Wong, F.R.C.P.C.,† Francisco Herrera, M.D.,* Frances Chung, F.R.C.P.C.‡

Background: Patient satisfaction has become an important component of quality improvement in ambulatory anesthesia services. However, it is difficult to measure due to its subjective and complex psychological construct. Psychometric methodology has been successfully used to evaluate this outcome. The authors conducted a systematic review to evaluate questionnaires to measure patient satisfaction with ambulatory anesthesia.

Methods: A systematic literature search of The Cochrane Library, MEDLINE, EMBASE, CINAHL, HAPI, PsycINFO, and Dissertation Abstracts was performed to identify studies on questionnaires evaluating patient satisfaction after ambulatory anesthesia. The authors included the articles that used multiple-item questionnaires, and the questionnaires were assessed with the strategy of psychometric questionnaire construction, validity, reliability, and acceptability.

Results: The authors scanned 131 articles yielded by our search strategy. Eleven articles were included in the study. Two questionnaires, IOWA Satisfaction with Anesthesia Scale and Evaluation du Vecu de I’Anesthesie Generale, fulfilled the criteria, but the latter was not developed specifically for ambulatory anesthesia, whereas Iowa Satisfaction with Anesthesia Scale was designed only for monitored anesthesia patients.

Conclusions: In a large number of trials, patient satisfaction has been evaluated using overall satisfaction or nonvalidated questionnaires. Only a few studies have developed questionnaires with rigorous psychometric methods to measure patient satisfaction with anesthesia care. At this time, there is still no valid or reliable questionnaire for measuring patient satisfaction in ambulatory anesthesia. Further study should be conducted to develop standardized instruments to measure this outcome.

IN 2007, ambulatory surgery comprised more than 60% of all surgery in the United States.§ Although traditional outcomes of the quality of anesthesia care that have been measured are patient morbidity and mortality, these outcomes are rare because of improvements in the surgical techniques and use of new anesthetic agents. Researchers in all medical specialties are increasingly studying nontraditional, patient-centered outcomes such as patient satisfaction and quality of life1 to assess quality of healthcare.

Patient satisfaction is an important indicator of outcome of health care and evaluation of the quality of services in anesthesiology.2 The American Society of Anesthesiologists Committee on Ambulatory Surgical Care and the Task Force on Office-based Anesthesia includes patient satisfaction as one of their outcome indicators. However, it is difficult to assess this outcome because satisfaction is a multi-dimensional concept with determinants that are not yet clearly defined.3

Although the role of patient satisfaction in anesthetic care has been increasingly investigated, many studies use only simple overall questions to assess satisfaction, leading to high score results. The reliability of single-item global satisfaction ratings is poor and inadequate to address the complexity of satisfaction.4,5 The lack of standardized, reliable, and valid questionnaires to assess patient satisfaction in anesthesia has been emphasized in many reviews.4,6,7 It is important to use a reliable and valid instrument to evaluate the outcome that researchers intend to assess. Psychometric methodology has been successfully used to create valid and reliable questionnaires to measure complex structures such as satisfaction with nursing care.8

We conducted a systematic review of questionnaires used for measuring patient satisfaction after ambulatory anesthesia to evaluate the psychometric properties of the questionnaires and advise on the selection of the most appropriate instrument for research and clinical use.

Materials and Methods

A systematic literature search of MEDLINE (from 1950 to March 1, 2008), EMBASE (from 1980 to March 2008), CINAHL (from 1982 to March 1, 2008), HAPI (from 1985 to March 1, 2008), PsycINFO (from 1967 to March 1, 2008), The Cochrane Library, and Dissertation Abstracts (from 1985 to March 08 2008) were reviewed to identify outcome measurements with patient satisfaction after ambulatory anesthesia. The following terms were combined for the search: ambulatory surgical procedures, ambulatory surgery/anesthesia, day surgery, same day, outpatient surgery, patient satisfaction, consumer satisfaction, questionnaires, surveys and instruments. The search was restricted to English publications only. A citation search of each identified questionnaire was performed to identify how frequently the questionnaire was used by other researchers.

The abstracts of all articles were reviewed by three authors. (PC, AA, and JW) Those studies measuring pa-
tient satisfaction with ambulatory surgery or that included ambulatory surgical patients were selected. Pediatric ambulatory studies were excluded. Full texts of the selected articles were reviewed by two authors (PC and JW). Patient satisfaction is a complex construct composed of multidimensional concepts; as a result, we only included articles that used multi-item questionnaires (i.e., more than two dimensions or questions) to measure patient satisfaction with anesthesia care. We excluded studies that used only a single or two questions or only evaluated the preoperative assessment or postoperative pain as measures of patient satisfaction.

The identified questionnaires were assessed for psychometric questionnaire construction and psychometric evaluation. The components of questionnaire construction that each instrument was evaluated for were:

**Item and Dimension Generation**

The item generation phase is undertaken to develop a pool of items that should include all important elements of patient satisfaction by reviewing the existing questionnaires, literature, opinions from anesthesia providers and focus groups of patients. Items that emerge directly from input from patients represent what patients truly value, and opinions from providers can ensure that significant elements of care have not been missed. Once generated, these items are then grouped into dimensions of care to develop an initial questionnaire.

**Pretest and Pilot Testing with Statistical Analysis to Revise the Final Questionnaire**

The process of pretest and pilot testing is for revision of the questionnaire into the final validated version by using the response from the pretest group, considering variables and statistical analysis from the pilot test. Items with ambiguous meanings can be eliminated to maximize the reliability and validity of the questionnaire. This process will result in a shorter, validated final version of the questionnaire.

**Retest the Final Version of the Questionnaire**

The final questionnaire should be tested in a new group of patients to determine if scores continue to exhibit reliability and validity.

**Validity**

Content and face validity refers to whether the items of the instrument cover all relevant and important contents of patient satisfaction in anesthetic care. Content validity is usually judged by the panel after literature reviews and focus groups interviews. Face validity is a subjective assessment by investigators if their items appear to measure outcomes that they claim to measure.

Criterion validity refers to the correlation of the measure with another criterion measure that is accepted as gold standard.

Construct validity is a process of hypothesis testing that includes convergent and discriminate validity. Convergent validity involves correlating the instrument with other indexes that measure similar aspects of the same construct. If the instrument is measuring what it is intended to measure, we can link the scores to other attributes by a hypothesis. If the prediction is confirmed, both the hypothesis and the instrument are validated. Because there is no standard for measuring patient satisfaction, researchers usually compare their questionnaire with other validated instruments or other related questions for this correlation. Discriminant validity requires that the construct should not show correlation with dissimilar variables.

**Reliability: Internal Consistency, Test-retest Reliability**

Reliability of the instruments is the consistency of instruments to produce the same results when applied to the same subjects at different times. The tests for reliability include:

Internal Consistency. Internal consistency involves testing for homogeneity, i.e., the correlations between items in the scale and correlations between the items and the total score. Cronbach’s $\alpha$ should be reported, the value should be 0.7–0.9, as a value above 0.9 may indicate that the questionnaire is too narrow in the scope.

Test-retest Reliability. Observations on the patient on two occasions separated by some interval of time. The minimum value of correlation coefficient should be 0.7.

Intra-rater and Inter-rater Agreements. Intrarater agreement is the agreement between observations made by the same rater on two different occasions. Interrater agreement is the degree of agreement between different observers.

Feasibility/Acceptability

The acceptability of a questionnaire has been frequently cited in the standard list of patient-based outcome measurements that can be evaluated by the response rate and the time to complete a questionnaire.

**Results**

The search strategy identified 2,962 articles on MEDLINE, 1,602 articles on CINAHL, 992 articles on EMBASE, 438 articles on HAPI, 275 articles from PsycINFO, and 545 articles from Dissertation Abstracts. A total of 6,757 abstracts were collected; those articles not related to patient satisfaction and ambulatory surgical patients were excluded. We reviewed 379 abstracts about patient satisfaction and ambulatory surgery. There were 208 articles that only evaluated surgical care, 13 articles only...
reporting satisfaction with preoperative assessment, and 27 articles only reporting satisfaction with postoperative pain management that were excluded. Therefore, the full texts of 131 articles were reviewed by the authors. We included articles that used multi-item questionnaires (i.e., more than two questions) to assess patient satisfaction in ambulatory anesthesia or studies that included ambulatory anesthesia (fig. 1). There were 120 articles that were excluded because the questionnaires contained only one or two questions about overall satisfaction or the studies were developed only for inpatients. Therefore, we found 11 articles that used multi-item questionnaires to measure patient satisfaction for anesthetic care after ambulatory surgery (table 1).

Steps of Questionnaire Construction

Item Generation Process. This is the process of creating the items of the questions and grouping of items into dimensions of care. Four of eleven studies described the process of item generation in their studies (table 2). Two questionnaires, Fung\textsuperscript{13} and Evaluation du Vecu de l’Anesthesie Generale (EVAN-G),\textsuperscript{14} were developed with direct interviews with patients. Fung studied ambulatory patients, whereas EVAN-G included both inpatients and ambulatory patients receiving general anesthesia. Iowa SATisfaction with Anesthesia Scale (ISAS)\textsuperscript{15} was developed from advice from anesthesia-providers, experts in satisfaction questionnaire development, and literature searches. All providers interviewed patients as part of routine postoperative care. This instrument was tested and recommended for monitored anesthesia care patients only. The questionnaire by Hadjistavropoulos et al.\textsuperscript{16} was developed by modifying the existing questionnaire, Wascana Client-Centered Care Survey (WCCS). The Wascana Client-Centered Care survey was constructed from the items obtained after discussion with patients and health providers.\textsuperscript{17} The items were then grouped into dimensions of care that emerged from the process of item generation in the studies from Hadjistavropoulos, Fung, ISAS, and EVAN-G. Other studies used multiple questions in variable dimensions to evaluate satisfaction such as physical discomfort, information, and pain, but they did not group those questions into dimensions of care. Seven studies did not describe details of how the items or dimensions used in their studies were generated.\textsuperscript{18–24} Two of the authors did not reply to our request for more information, and we could not contact the remaining five authors because the address or contact information was not available or up to date.

Pretest, Pilot Testing, and Revision. EVAN-G was generated with 75 questions from the interview. Pretest and pilot testing was then done reducing items to the final version, a 26-item questionnaire by validity and reliability testing. ISAS was constructed with 18 questions from the literature and advice from colleagues. The questionnaire was pretested with 61 patients, the results were analyzed, and the items were reduced to 11 questions. Fung et al. selected 36 items from the telephone interviews with the patients and tested 10 subjects for comprehensiveness of the questionnaire. Three of the four studies that mentioned the item-generation process pretested their questionnaires.\textsuperscript{15–17} All seven studies that did not describe the item-generation process also did not conduct the pretest or pilot test of their study questionnaires.

Retest the Final Version. After the pretest phase, the final questionnaire should be retested in other groups of patients and be evaluated with psychometric tests. The final version of each questionnaire is described in table 1. All questionnaires were used to evaluate patient satisfaction with anesthetic care in their studies. Five articles tested their questionnaires with validity and/or reliability, whereas six studies did not test their questionnaires (table 2).

Validity. Content validity was assessed in EVAN-G and ISAS by asking patients if there were other aspects of care that were not mentioned in the questionnaires. Construct validity was tested in EVAN-G by the correlation of pain and discomfort items of the questionnaire with the validated questionnaire, McGill Pain Questionnaire (MPQ). ISAS was tested for construct validity by comparing the overall scores with the scores predicted by anesthesia providers and with the statement "I was satisfied with my anesthetic care."

Reliability. Internal consistency was assessed with Cronbach’s $\alpha$ in four instruments. The values were found to be more than 0.7 in three instruments.\textsuperscript{14–16,18}
Test-retest reliability was analyzed in EVAN-G, ISAS, and the Pestey study, and the values were more than 0.7 in all these three articles.

**Acceptability.** EVAN-G required 9 ± 7 min to complete and ISAS required 4.6 ± 2.1 min to complete the questionnaires. Response rates of the papers included in our review exceeded 50%, except the report by Hadjistavropoulos et al., which had a response rate of only 26%.

A citation search showed that five studies used ISAS as an instrument measuring patient satisfaction, whereas none of the other instruments were cited by other studies.
Thus these questions may not be truly representative of patient satisfaction in other areas of medicine.4 For example, this method has been used to develop questionnaires measuring health status, quality of life, and patient satisfaction.8 Our systematic review included eleven studies that used multiple-item questionnaires (more than two questions or dimensions) to assess patient satisfaction in ambulatory anesthesia. Of the identified instruments, only two instruments, EVAN-G and ISAS, met the established criteria necessary for good psychometric questionnaire development. However, both instruments have limitations; EVAN-G was developed for both inpatients and ambulatory surgical patients under general anesthesia, whereas ISAS was designed only for monitored anesthesia care patients.

The purpose of this review was to guide the selection of the most appropriate questionnaire for the evaluation of patient satisfaction after ambulatory anesthesia. The use of a patient-based outcome measure as one of the endpoints of a clinical study requires that the instrument has to fulfill the requirements for good psychometric questionnaire construction. Psychometric methodology has been shown to produce reliable and valid multiple-item instruments to evaluate the complex construct of patient satisfaction in other areas of medicine.9 For example, this method has been used to develop questionnaires measuring health status, quality of life, and patient satisfaction.8 Our systematic review included eleven studies that used multiple-item questionnaires (more than two questions or dimensions) to assess patient satisfaction in ambulatory anesthesia. Of the identified instruments, only two instruments, EVAN-G and ISAS, met the established criteria necessary for good psychometric questionnaire development. However, both instruments have limitations; EVAN-G was developed for both inpatients and ambulatory surgical patients under general anesthesia, whereas ISAS was designed only for monitored anesthesia care patients.

The next step of questionnaire development is pretesting or pilot testing the preliminary questionnaire. The goal of pilot testing is to refine and reduce the questions on the basis of analysis of reliability and validity of the questionnaire in the pilot phase. During this stage, problems with questions or wording can be clarified, the variability of answers can be considered, and presentation aspects of questionnaires and confounders such as social desirability biases or nonresponse bias should be checked.15–25 For example, EVAN-G started with 75 questions; after pretesting, ambiguous or misunderstood questions were deleted, leaving 66 items. Pilot, validity, and reliability testing led to elimination of more questions, leaving a 26-item questionnaire. The revised questionnaire version based on the pretest and pilot results could reduce exaggerated items and produce improvements in the measurement of outcomes of the study.26

After the pilot test, the final version of the questionnaire should be tested in other patients. The psychometric properties should be tested again. Validity and reliability are two qualities that health measurements should be tested for. Most validation studies begin with content validity; to assess content validity, patients are asked to review content of items in the questionnaire. The opened-end questions at the end of questionnaires can be useful for content validation,27 but content validity cannot be tested formally. The more formal statistical procedures are criterion validity and construct validity testing. Criterion validity considers whether scores or contents in the questionnaire correlate with the definitive standard measurement of the same outcome; however, there is no definitive standard for measuring patient satisfaction in any previous study.5 Therefore, we cannot use criterion validity to assess questionnaires measuring patient satisfaction. Construct validity can be used for subjective outcomes such as satisfaction and happiness where definitive standards do not exist. These
may be expressed as hypotheses indicating correlation between patient satisfaction, and other measurements such as pain dimension could be correlated with, for example, McGill Pain Questionnaire (MPGQ) in EVAN-G. Construct validity of ISAS was assessed by correlation with the statements “I would have the same anesthetic” and “I was satisfied with my anesthetic care.”

Some questionnaires were not tested for validity but were tested only for reliability. Reliability only assures the stability of the scales; it does not show exactly what the questionnaires measure. Questionnaires should be tested for both validity and reliability. The acceptability of a questionnaire by the patients is another important criterion to select the questionnaire. A high response rate indicates good acceptability; however, there are many factors that may affect the response rate from patients such as mode of administration and appearance of the questionnaire. Response rates of mailed back questionnaires are usually less than an interview mode. A response rate of 50% for mailing methods is considered adequate for analysis. The time to complete the questionnaire is another aspect of acceptability. Generally, the shorter the form and less time to complete the form, the greater is its acceptability. Questionnaires that can be finished within 5 minutes are considered to be superior. It is essential that instruments should be acceptable to patients to obtain high response rates and minimize bias from nonrespondents.

One theory of satisfaction is the differences between the expectation and the actual outcome. In our review, the most important factors for determining satisfaction were information and effective communication. Thus, we can improve patient satisfaction with adequate information and continuity of care for all phases of the perioperative period from preoperative to intraoperative and postoperative periods. Caljouw also found in his study that patient satisfaction is largely based on good information and staff-patient relationships. Education and information, to help patients after discharge were identified as factors for satisfaction in ambulatory surgical patients. From the included papers, we found that information, communication, and aspects of patient involvement had lower scores versus the other dimensions, which may indicate the areas in which improvement in ambulatory anesthesia care are needed.

One of the limitations of this review is that we excluded non-English language articles. We also excluded many studies that only evaluated the preoperative assessment, recovery, or postoperative pain alone as the only aspect of care to assess satisfaction in ambulatory anesthesia.

There are many studies that have developed questionnaires for measuring inpatient satisfaction after anesthetic care using psychometric construction. However, these papers only included inpatients, and some aspects of care for inpatients may not be applicable for ambulatory care. This has been shown in a study by Bost et al., where the Quality of Recovery from Anesthesia (QoR) instrument, developed primarily for inpatients, was not appropriate for his ambulatory patients. It is important for clinical investigators to select instruments that are suitable for the intended task.

EVAN-G and ISAS developed their questionnaires with the appropriate steps of psychometric questionnaire construction and tested questionnaires for reliability, validity, and acceptability. A citation search showed that ISAS was the only instrument from our 11 questionnaires that was used by other studies. In our opinion, EVAN-G was superior to ISAS for general anesthesia patients; it included more dimensions of care that were important for satisfaction, such as information and respect. However, neither EVAN-G nor ISAS were developed for all types of ambulatory anesthesia. EVAN-G was developed for general anesthesia patients, and ISAS was developed for monitored anesthesia care patients only. Regional anesthesia is another type of anesthesia that has been used with ambulatory surgery patients, but we did not find a validated questionnaire designed for evaluating patient satisfaction specifically for regional anesthesia. The lack of a valid, reliable instrument measuring patient satisfaction in regional anesthesia has been identified previously.

From our review, the dimensions of care that emerged from patient interviews were information provision, physical comfort/discomfort, emotional support (relaxed, reassured, and attentive), involvement of care and privacy. Information for ambulatory surgery patients emphasized pain and homecare management. These dimensions should be included in the questionnaire for assessing patient satisfaction after ambulatory anesthesia.

From our literature search, we found many articles that assessed patient satisfaction with anesthetic care but very few articles that used validated questionnaires to measure this outcome. There were 79 articles that used only one question or overall satisfaction to evaluate patient satisfaction for anesthesia care. However, a single global question cannot accurately measure the complexity of satisfaction. It is very important for researchers to use a multiple-item, valid, and reliable questionnaire to assess patient satisfaction, so we can truly compare this outcome across research studies.

In conclusion, further studies about patient satisfaction after ambulatory anesthesia should be conducted by following rigorous methodological psychometric questionnaire construction. The instrument should also be validated by other users, further confirming the validity and reliability of the instrument. We suggest further studies for developing questionnaires measuring satisfaction in ambulatory anesthesia. The process of validity and reliability testing may be lengthy, but it is worthwhile to create a standard questionnaire for evaluating this outcome for ambulatory anesthesia as ambulatory sur-
surgery comprises the highest proportion of surgery in North America.

The authors thank Marina F. Englesakis, B.A. (Hons), MLIS, Information Specialist, Libraries and Information Services, University Health Network, Toronto, Ontario, Canada, for her assistance to develop the search strategy and access to medical databases.

References

20. Carro JLP, Diaz-Guerra CS-C, Wong PL. Study of unilateral posthemorrho...
28. Fung D, Cohen M, Stewart S, Davies A: Can the Iowa Satisfaction with Anesthesia Scale be used to measure patient satisfaction with cataract care under topical local anesthesia and monitored sedation at a community hospital? Anesth analg 2005; 100:1637–43
31. Caljouw MA, van Beusekom M, Boer F. Patient’s satisfaction with peripera...
32. Rhodes L, Miles G, Pearson A: Patient subjective experience and satisfac...
35. Whitty PM, Shaw IH, Goodwin DR. Patient satisfaction with general anesthesia. Too difficult to measure? Anaesthesia 1996; 51:327–32
37. Bost JE, Williams BA, Bottega MT, Dang Q, Rubio DM. The 8-item Short-Form Health Survey and the physical comfort composite score of the quality of recovery 40-item scale: The most responsive assessments of pain, physical function, and mental function during the first 4 days after ambulatory knee surgery with regional anesthesia. Anesth analg 2007; 105:1693–708
39. Fung D, Cohen M, Stewart S, Davies A: Can the Iowa Satisfaction with Anesthesia Scale be used to measure patient satisfaction with cataract care under topical local anesthesia and monitored sedation at a community hospital? Anesth Analg 2005; 100:1637–43

Anesthesiology, V 110, No 5, May 2009

Copyright © by the American Society of Anesthesiologists. Unauthorized reproduction of this article is prohibited.