Development and Validation of a Perioperative Satisfaction Questionnaire

Pascal Auquier, M.D., Ph.D., Nicolas Pernoud, M.D.,† Nicolas Bruder, M.D.,‡ Marie-Claude Simeoni, M.D., Ph.D.,§
Jean-Pierre Auffray, M.D.,‖ Christian Colavolpe, M.D.,# Georges François, M.D.,** François Gouin, M.D.,††
Jean-Claude Manelli, M.D.,‡‡ Claude Martin, M.D.,§§ Christophe Sapin, M.Sc.,‖‖ Jean-Louis Blache, M.D.##*

Background: Satisfaction is considered a valuable measure of outcome of healthcare processes. Only a few anesthesia-related validated questionnaires are reported. Because their scope is restricted to specific clinical contexts, their use remains limited. The objective of the current study was to develop and validate a self-reported questionnaire, Evaluation du Vécu de l'Anesthésie Générale (EVAN-G), assessing the satisfaction of the perioperative period surrounding general anesthesia.

Methods: Development of the EVAN-G questionnaire comprised a phase of item generation and a phase of psychometric validation. The patient sample was generated to be proportionally matched to the population of patients undergoing general anesthesia in France. The structure of the questionnaire was identified studying interitem, item–dimension, and interdimension correlations and factor analyses. Data were concurrently gathered to assess external validity. The discriminant validity was determined by comparison of scores across well known patient groups. Reliability was assessed by computation of Cronbach α coefficients and by test–retest.

Results: Eight hundred seventy-four patients were recruited in eight anesthesia departments. The EVAN-G includes 26 items; six specific scores and one global index score are available. Correlations between EVAN-G scores and other concurrent measures supported convergent validity. The EVAN-G correlated poorly with age, American Society of Anesthesiologists physical status, total anesthesia time, and number of previous anesthesias. Significantly higher satisfaction was reported by patients older than 65 yr, belonging to the laryngeal mask group. Reliability and reproducibility were shown.

Conclusion: The EVAN-G adds important information oriented toward patients' perceptions. The authors' approach provides a novel, valid, and reliable tool that may be used in anesthesia practice.

HUMAN satisfaction is a very complex concept, involving many components such as physical, emotional, mental, social, and cultural factors. Multiple definitions of this concept and theories of satisfaction have emerged mainly from the field of behavioral sciences and consumerism.[1–3] Satisfaction is now considered a valuable measure of outcome of healthcare processes.[4] In several countries, it has become a major step in health institution accreditation processes.[2] Moreover, satisfaction may have an important influence on various aspects of patients behavior, such as global consumption of healthcare resources, compliance with treatments, or steadiness of relationship with practitioners.[5,6] However, the complexity of the concept, as well as the multidimensional nature of satisfaction, has often impaired the development of reliable evaluation tools.[7] In anesthesia, the strong emotional context, the potential effect of drugs on cognition, and the short time interval of the anesthesia process make it even more difficult to assess satisfaction.[8] Only a few anesthesia-related validated questionnaires have been reported.[9] However, because their scope is either restricted to specific clinical contexts[7] or focused on specific anesthetic managements,[10] their use remains limited. Most of these tools have been developed relying on the views of experts to determine essential patient concerns.[10–15] Furthermore, they are not always psychometrically sound.[16] Such weaknesses may lead to unreliable results and may partly explain the high satisfaction ratings usually reported.

Satisfaction should be measured as a multidimensional concept and should reflect the exclusive concerns of the patients. The definition of satisfaction, relying on the theory of expectations,[3] satisfies this demand because it focuses the discrepancies between patients' perceptions and their expectations. At last, this tool should meet the needs of clinicians and should therefore be usable for any type of surgery and in routine clinical practice.

The main objective of this study was to develop and validate, following international guidelines,[17–19] a self-reported questionnaire, Evaluation du Vécu de l'Anesthésie Générale (EVAN-G), assessing the satisfaction of the perioperative period surrounding general anesthesia, that would fulfill those requirements.

Materials and Methods

Patients were recruited between January 2000 and February 2001, in eight anesthesia departments (including four university hospitals) in southeastern France. The study was piloted by a steering committee including nine anesthetists, two surgeons, one psychiatrist, and three epidemiologists from different centers. Approval was obtained from the local ethics committee. The cri-

* Professor of Public Health, § Associate Professor of Public Health, ‖‖ Statistician, Department of Public Health, ‡ Professor of Anesthesiology, ** Emeritus Professor of Anesthesiology, †† Professor of Anesthesiology, Department of Anesthesia, Timone University Hospital. † Research Fellow, ## Chief of Department, Department of Anesthesia, Paoli-Calmettes Institute, Marseilles, France. ‖ Professor of Anesthesiology, Department of Anesthesia, Sainte Marguerite University Hospital, Marseilles, France. # Staff Anesthesiologist, ‡‡ Professor of Anesthesiology, Department of Anesthesia, Conception University Hospital, Marseilles, France. §§ Professor of Anesthesiology, Department of Anesthesia, Nord University Hospital, Marseilles, France.

Address correspondence to Dr. Auquier: Service de Santé Publique. Faculté de Médecine, 27, bd Jean Moulin, 13385 Marseille, France. Address electronic mail to: pascal.auquier@medecine.univ-mrs.fr. Individual article reprints may be purchased through the Journal Web site, www.anesthesiology.org.

teria for patient inclusion were consent to participate in the study, elective surgery (except obstetric) or endoscopic procedure requiring general anesthesia (excluding monitored anesthesia care, regional anesthesia), age older than 18 yr, and ability to understand and read French and to fulfill a self-administered questionnaire within the 48 h after general anesthesia. To maximize the return rate, all questionnaires were administered and collected before the patients left the hospital. An independent investigator made sure the patients could complete the questionnaire without any external help or influence. Development of the EVAN-G questionnaire comprised two steps: a phase of item generation and reduction and a phase of psychometric validation.

Step 1. Questionnaire Development: Item Generation and Reduction

Face-to-face semistructured interviews were performed by a trained interviewer within the 48 h after surgery. The interviews comprised two parts. First, patients were asked to report about their satisfaction in a nondirective way. The second part of the interview was based on a guideline issued from the literature review.[19,20] This guideline listed the main domains related to satisfaction as reported in literature and did not consist of a set of directive questions. The interviews addressed the impact of the perioperative process with general anesthesia on patients' satisfaction referring to the theory of expectation, defining satisfaction as the discrepancy between expectations and current life experience.[2,3] Therefore, the views of the patients were gathered regarding both their current experience of health care in the perioperative period and their corresponding expectations. All interviews were recorded and transcribed. Content analysis was performed by three members of the steering committee who were skilled in textual analysis, complemented by a computerized textual analysis (Alceste® software; IMAGE, Toulouse, France). It aimed at identifying recurrent themes that were then used to generate individual questions within the questionnaire.[21] Patient interviews were also used to determine the wording in question stems and the range of response options. Interviews were conducted until no new ideas emerged in the content analysis performed in real time, up to a total of 24 patients.

Seventy-five questions were selected from these interviews (item generation). These items were answered using a five-point Likert scale, where $1 =$ much less than expected, $2 =$ less than expected, $3 =$ as expected, $4 =$ more than expected, and $5 =$ much more than expected.

The acceptability of this 75-item self-administered questionnaire was pretested in a sample of 29 consecutive patients who had not participated in the interviews. Items that were ambiguous, misunderstood, or rarely answered were suppressed, leading to a preliminary questionnaire comprising 66 remaining items.

This preliminary 66-item version of the questionnaire was administered to a new sample of 171 randomly selected patients undergoing surgery. Eleven questions were deleted with regard to redundancy (interitem correlation), low response rates (missing data over 20%), or skewness of the distribution of the answers (floor–ceiling effect).

Concurrently, this pilot study ensured content validity and guaranteed that the questionnaire was a true reflection of the patients' experience. Patients were asked to comment on any aspects of questionnaire (content, wording, response choices) that they thought were irrelevant or prone to improvement.

Step 2. Questionnaire Validation

Population and Data Collection. The primary objectives of the validation phase were to check that the questionnaire actually (validity) and accurately (reliability) measured the concept it has been designed for, *i.e.*, patient satisfaction. Assessment of the validity of an instrument is meant to evaluate the systematic error of measure (drift), whereas assessment of the reliability is meant to evaluate the random error of measure (scatter). A secondary objective of the validation phase was to further reduce the number of items to make the instrument shorter and usable in routine clinical practice. Patients included in the first stage could not be included in this validation phase.

The patient sample regarding surgical procedures was generated to be proportionally matched to the yearly population of patients undergoing general anesthesia in France,[22] excluding monitored anesthesia care and regional anesthesia. A total of 977 patients who met the inclusion criteria were included, over 15 consecutive days in eight anesthesia departments in southeastern France. Sociodemographic and other clinical data were collected by a senior anesthetist.

Within 4–48 h after surgery, the patients were given the self-administered 55-item questionnaire, along with two other questionnaires detailed in the next section.

Statistical Analysis: Psychometric Validation. The multidimensional structure of the questionnaire was identified studying interitem, item–dimension, and interdimension correlations (Pearson r) and principal component factor analyses with Varimax rotation.[18,19] More specifically, item internal consistency was assessed by correlating each item with its scale, which was corrected for overlap. A correlation of 0.4 is recommended as the standard for supporting item internal consistency. Item discriminant validity was assessed by determining the extent to which items correlate more highly with the dimensions they are hypothesized to represent than with the other ones.[18,19] The construct of each dimension was assessed using Rasch Rating Scale model analy-

ses.[23,24] This model allows one to ensure the unidimensionality of each dimension. The more meaningful and psychometrically sound solution was kept to produce the final version of the EVAN-G questionnaire, consisting of 26 items assessing six dimensions: attention, privacy, information, pain, discomfort, and waiting times (appendix: EVAN-G questionnaire).

To assess external validity of the EVAN-G, other external data were concurrently gathered from patients and clinicians. Relations were investigated between specific potential dimensions of EVAN-G (*e.g.*, pain, anxiety) and other instruments, such as validated questionnaires (McGill Pain Questionnaire [MGPQ],[25] Spielberger State-Trait Anxiety Inventory [STAI][26]) or specific visual analog scales (VAS) assessing miscellaneous domains: anxiety, fear, pain, discomfort, trust in medical staff, ability to ask for information, quality of information delivered, kindness of medical staff, paid attention, and overall satisfaction with care. The underlying assumption was that dimension scores of the EVAN-G were higher correlated with scores of similar dimensions from the other concurrent instruments than with dissimilar ones (convergent validity).

The discriminant validity of EVAN-G was determined by dimension mean scores of dimensions across patient groups that were known to differ in their sociodemographic (age, sex) or clinical features (American Society of Anesthesiologists [ASA] physical status, number of previous anesthesias, anesthesia duration) using analysis of variance, the Student *t* test, the Mann–Whitney U test, or the Pearson correlation.

To check content validity, patients were asked, in an open-ended question at the end of the questionnaire, to point out the different important domains of their life that were not mentioned on the EVAN-G.

Reliability was assessed by computation of Cronbach α coefficients (internal consistency) for each dimension score, and by test–retest using an intraclass correlation coefficient conducted on a subsample of 36 patients evaluated twice with a 15-day interval. To ensure data quality by excluding those answers likely to be unreliable, the validation analysis was performed on records with less than 25% of missing data on the questionnaire to be validated, as recommended in international guidelines. This procedure reinforces the generalizability of the final validated questionnaire.[18,19]

Results

To avoid confusion with intermediate development versions of the questionnaire, only the results of the final psychometric validation phase are reported.

Sample Characteristics

Of the 977 recruited patients, 103 (10.5%) answered less than 75% of the questions and were excluded from the analysis. The 874 remaining patients included in the validation analyses underwent a wide range of surgical procedures: 102 gynecologic (11.7%), 83 orthopedic (9.5%), 80 digestive (9.2%), 71 ear, nose, and throat (8.1%), 71 vascular (8.1%), 65 esthetic (7.4%), 50 endocrine (5.7%), 44 maxillofacial (5.0%), 41 ophthalmologic (4.7%), 30 spine (3.4%), 29 urologic (3.3%), 14 intracranial (1.6%), 12 thoracic (1.4%), 160 endoscopic (18.3%), and 22 other (2.5%). The mean patient age was 51 ± 17 yr, 397 of the patients (45.4%) were male, 678 (77.5%) received a premedication, 118 (13.5%) were ambulatory patients, and the mean number of previous surgeries was 3 ± 2. ASA physical status scores, rated by physicians, were as follows: I, 422 (48.5%); II, 371 (42.6%); III, 74 (8.5%); IV/V, 3 (0.4%). The mean MGPQ score was 6.3 ± 8.6 (range, 0–57), and the STAI mean global score was 51.3 ± 2.9. The mean duration of anesthesia was 110 ± 86 min (median, 90; range, 10–720).

Scoring

The scores for negatively worded items were reversed so that higher scores indicated a higher level of satisfaction. For each individual, the score of each dimension was obtained by computing the mean of the item scores of the dimension. If fewer than one half of the items were missing, the mean of the nonmissing items was substituted for the missing items. All dimension scores were linearly transformed to a 0–100 scale, with 100 indicating the best possible level of satisfaction and 0 indicating the worst. The global satisfaction score was computed as the mean of the dimension scores.

Validity

Principal component analyses after Varimax rotation isolated 26 questions, grouped into six dimensions. The six-factor structure accounted for 64% of the total variance (table 1). Each dimension was named according to its constitutive items: attention (5 items), privacy (4 items), information (5 items), pain (5 items), discomfort (5 items), and waiting (2 items).

As expected, correlations between items and their corresponding dimension (item internal consistency) ranged from 0.55 to 0.92, whereas correlations between items and the other dimensions (item discriminant validity) ranged from 0.02 to 0.53 (table 2). Correlations between dimension scores were low to moderate, ranging from 0.14 to 0.57 ($P < 0.001$). The global mean satisfaction score was 75 ± 14. The worst mean dimension score was found for information (64 ± 22), and the best was found for discomfort (84 ± 19) (table 2).

Overall, the levels of correlation between EVAN-G scores and other concurrent measures confirmed the convergent validity of the instrument (table 3). High levels of correlation were found between the MGPQ and pain ($r = -0.52$, $P < 0.01$) and between the MGPQ and discomfort ($r = -0.36$, $P < 0.01$). Each EVAN-G dimen-

Table 1. Principal Component Analysis (Varimax Rotation) of the 26-Item EVAN-G Questionnaire

Item No.	Attention	Information	Privacy	Pain	Discomfort	Waiting
21	**0.84**					
6	**0.80**					
10	**0.80**					
15	**0.77**					
22	**0.73**		0.34			
3		**0.81**				
1		**0.80**				
2		**0.80**				
5	0.31	**0.71**				
4	0.34	**0.68**				
23			**0.81**			
9			**0.78**			
7		0.32	**0.70**			
24			**0.65**			
20			**0.46**	0.41		
19				**0.87**		
13				**0.73**		
18				**0.66**		
14			0.37	**0.43**		
12					**0.77**	
8					**0.71**	
11					**0.66**	
16				0.37	**0.57**	
17				0.36	**0.47**	
26						**0.88**
25						**0.87**
% Variance explained (total = 64%)	29.9	12.4	6.9	6.5	4.6	4.1

For clarity, factor loadings above 0.3 are not reported in the table. For each column, bold numbers are the factor loadings of the items participating in the computation of the corresponding dimension.

sion score correlated higher with its domain-related VAS. VAS score assessing overall satisfaction with care correlated with all of the EVAN-G dimensions ($r = 0.16 – 0.36$, $P < 0.01$) and the global score ($r = 0.40$, $P < 0.01$). The EVAN-G score was not correlated with the STAI score, except for attention. All these results are reported in table 3

Overall, EVAN-G scores correlated poorly with age, ASA physical status, total anesthesia time, and number of previous anesthesias. The only significant correlations were reported between (1) age and pain ($r = 0.18$, $P < 0.01$), discomfort ($r = 0.20$, $P < 0.01$), waiting ($r = 0.18$,

$P < 0.01$), and global index ($r = 0.18$, $P < 0.01$); (2) total anesthesia time and pain ($r = -0.29$, $P < 0.01$), discomfort ($r = -0.13$, $P < 0.01$) and global index ($r = -0.13$, $P < 0.01$); and (3) ASA physical status and pain ($r = 0.10$, $P < 0.01$) and waiting ($r = 0.13$, $P < 0.01$).

Findings regarding the comparisons of EVAN-G scores among clinical groups met the assumptions expressed *a priori* by the steering committee relying on clinical experience and literature. Significantly higher global satisfaction was reported by patients older than 65 yr, belonging to the laryngeal mask group and undergoing minor surgery. Information dimension score was signif-

Table 2. Item-internal Consistency, Item-discriminant Validity, Percentage of Missing Values, Cronbach α, and Intraclass Correlation Coefficient of the EVAN-G Dimension Scores and Global Index

Dimension	Mean ± SD	% MV	IIC	IDV	α	ICC†
Attention	76 ± 23	2.1	0.84–0.90*	0.13–0.53*	0.91	0.81
Privacy	70 ± 21	0.9	0.75–0.85*	0.02–0.47*	0.83	0.74
Information	64 ± 22	2.6	0.78–0.84*	0.04–0.48*	0.87	0.78
Pain	73 ± 20	0.9	0.63–0.77*	0.07–0.48*	0.74	0.80
Discomfort	84 ± 19	0.6	0.55–0.76*	0.06–0.37*	0.73	0.75
Waiting	83 ± 26	2.7	0.88–0.92*	0.02–0.26*	0.81	0.72
Global index	75 ± 14	9.8	‡	‡	0.90	0.77

Data are presented as mean ± SD. EVAN-G scores range from 0 (lowest satisfaction) to 100 (highest satisfaction). Item-discriminant validity (IDV): correlation between items scores of a given dimension with the other dimension scores. Numbers are lowest–highest Pearson correlation coefficient. Item-internal consistency (IIC): correlation between items scores and their dimension score (corrected for overlap). Numbers are lowest–highest Pearson correlation coefficient.

* $P < 0.01$.   † On a subsample of 36 patients.   ‡ Not computable.

ICC = intraclass correlation coefficient; % MV = percentage of missing values.

Table 3. Correlations between EVAN-G Scores and McGill Pain Questionnaire, State-Trait Anxiety Inventory, and Visual Analog Scales

Dimension	MGPQ	STAI	Anxiety	Pain	Confidence in Staff	Ability to Ask Questions	Kindness of Staff	Fear	Discomfort	Attention	Quality of Explanation	Global Satisfaction
Attention	−0.07	0.11†	−0.13†	−0.07	0.36†	0.34†	0.29†	−0.08	−0.15†	0.34†	0.39†	0.36†
Privacy	−0.02	0.04	−0.04	−0.07	0.19†	0.14†	0.14†	0.00	−0.06	0.20†	0.17†	0.18†
Information	−0.10*	0.06	−0.11†	−0.08*	0.22†	0.30†	0.17†	−0.05	−0.11†	0.18†	0.31†	0.23†
Pain	−0.52†	0.09*	−0.25†	−0.68†	0.19†	0.13†	0.15†	−0.23†	−0.41†	0.09†	0.17†	0.16†
Discomfort	−0.36†	0.06	−0.26†	−0.37†	0.25†	0.24†	0.22†	−0.26†	−0.36†	0.16†	0.21†	0.23†
Waiting	−0.18†	0.05	−0.10*	−0.14†	0.23†	0.30†	0.28†	−0.07	−0.16†	0.21†	0.29†	0.31†
Global index	−0.33†	0.11*	−0.23†	−0.35†	0.39†	0.40†	0.33†	−0.18†	−0.32†	0.30†	0.41†	0.40†

Numbers are Pearson correlation coefficients.

* $P < 0.05$.   † $P < 0.01$.

MGPQ = McGill Pain Questionnaire; STAI = State-Trait Anxiety Inventory.

icantly higher in the premedicated group. The pain dimension score was significantly higher in the following groups of patients: ASA physical status greater than II, minor surgery, outpatient, unpremedicated, and laryngeal mask. Discomfort dimension scores were significantly higher in patients older than 65 yr, males, patients with ASA physical status greater than II, patients undergoing minor surgery, outpatients, and patients in unpremedicated and laryngeal mask groups. Waiting dimension scores were significantly higher in patients older than 65 yr, patients with ASA physical status greater than II, and patients in unpremedicated and laryngeal mask groups (table 4).

The steering committee confirmed that the content of each dimension was meaningful and that the six-factor structure dealt with the major domains reported in the patients' interviews and open comments, reflecting a relevant content validity.

Reliability

The internal consistency reliability of EVAN-G was high: The Cronbach α value ranged from 0.73 to 0.91 (table 2). The stability of the EVAN-G score evaluated by test–retest correlation conducted on 36 patients without any intercurrent health events was satisfactory: The intraclass correlation coefficient ranged from 0.72 to 0.81.

Acceptability

The average time of completion of the EVAN-G was 9 ± 7 min, fully compatible with clinical practice. Missing data per dimension were low, ranging from 0.6 to 2.7% (table 2).

Table 4. Comparisons of EVAN-G Scores according to Age, Sex, and Clinical Features

	n	Attention	Privacy	Information	Pain	Discomfort	Waiting	Global Index
Age*						‖	‖	‖
≤ 65	616	75 ± 23	70 ± 21	63 ± 22	73 ± 21	82 ± 19	82 ± 27	74 ± 14
> 65	172	79 ± 22	70 ± 21	67 ± 24	76 ± 19	88 ± 14	88 ± 22	78 ± 12
Sex*						‖		
Male	359	75 ± 22	70 ± 21	64 ± 20	75 ± 19	87 ± 15	84 ± 24	76 ± 12
Female	429	76 ± 22	70 ± 21	64 ± 23	71 ± 22	81 ± 21	82 ± 27	74 ± 15
ASA*					‖	§	‖	
I	385	75 ± 22	71 ± 21	65 ± 21	71 ± 21	82 ± 18	80 ± 27	74 ± 14
II–IV	400	76 ± 23	69 ± 21	63 ± 23	75 ± 20	85 ± 18	87 ± 24	76 ± 14
Heaviness SP†					‖	‖		§
Minor	269	75 ± 23	70 ± 21	63 ± 21	79 ± 17	87 ± 16	82 ± 26	76 ± 13
Average	304	75 ± 22	68 ± 21	63 ± 22	70 ± 21	81 ± 19	82 ± 28	73 ± 14
Major	62	78 ± 22	69 ± 20	64 ± 19	63 ± 24	80 ± 22	83 ± 26	73 ± 15
Patient status*					‖	‖		
Inpatient	678	76 ± 22	70 ± 21	64 ± 22	71 ± 21	83 ± 18	83 ± 26	75 ± 14
Outpatient	104	73 ± 24	71 ± 21	64 ± 21	84 ± 15	88 ± 18	83 ± 26	77 ± 13
Premedication‡				§	‖	‖	§	
No	97	73 ± 22	66 ± 20	60 ± 21	79 ± 18	89 ± 14	87 ± 24	76 ± 11
Yes	621	76 ± 22	70 ± 21	64 ± 22	71 ± 21	82 ± 19	82 ± 27	74 ± 14
Airway*					‖	‖	‖	§
Intubation	104	76 ± 23	70 ± 21	64 ± 22	69 ± 21	81 ± 19	81 ± 28	74 ± 15
LM	535	73 ± 23	71 ± 21	65 ± 23	77 ± 17	88 ± 15	89 ± 20	77 ± 13

Values are presented as mean ± SD. EVAN-G scores range from 0 (lowest satisfaction) to 100 (highest satisfaction).

* Independent samples *t* test.   † Heaviness of surgical procedure: one-way analysis of variance.   ‡ Mann–Whitney U test.   § $P < 0.05$.   ‖ $P < 0.01$.

ASA = American Society of Anesthesiologists (physical status); LM = laryngeal mask.

Discussion

The objective of this study was to report the stages of development and validation of a self-administered satisfaction questionnaire of the perioperative period surrounding general anesthesia based on issues pertinent to patients.

The first issue to be discussed refers to the psychometric validity of the questionnaire. None of the satisfaction instruments commonly used in anesthesia studies concurrently include development from the exclusive point of view of patients, perioperative period as time frame, and data demonstrating high psychometric validity.

The perspectives to which the questionnaire refers are an important issue because of discrepancies between patients' and physicians' points of view.[27] Therefore, physicians should favor standardized questionnaires that are not only self-completed by the patients but also ask them questions about their own concerns. The EVAN-G was developed with 24 individual interviews, conducted by a nonclinician who was not part of the medical staff, including a wide range of surgical procedures, followed by cognitive debriefing with patients at different steps of the development. These qualitative data were collected. Content validity was ensured by developing items on the basis of in-depth interviews with patients rather than relying on literature or expert opinions. The content of the questionnaire encompasses experiences of great importance to patients. Any items that were criticized by patients for being inappropriate or ambiguous were removed. The domains covered by the EVAN-G express the two-fold originality of its development process, conveying patients' views and focusing on the perioperative period. The EVAN-G describes themes commonly reported when evaluating patients' satisfaction or perceptions related to surgery, such as pain and discomfort. In that, there is overlap with the Quality of Recovery questionnaire,[28] for example. Nevertheless, the EVAN-G explores complementary aspects specific to satisfaction, such as attention and privacy. These dimensions of anesthesia are seldom evaluated in studies on patients' satisfaction relying on expert-based questionnaires, but patients believed these aspects of anesthesia care were important.

Internal consistency reliability of the six domains and reproducibility have been shown to be high. Construct validity was explored by the comparison to well-established measures (ASA, STAI, MGPQ, and others). Moreover, because there was no accepted criterion that can serve as a definitive standard of satisfaction, following the lead of previous investigators, we chose to compare the EVAN-G to the VAS score representing different domains of satisfaction.[29] To limit the risk of bias due to improper patient selection, a large sample was taken from eight hospitals, and the recruitment mimicked the repartition of the surgical population in France.[22] This reinforced the reliability and validity of the EVAN-G as a measure of patient perioperative satisfaction.

The acceptability of the EVAN-G was good. The rate of missing data of the final 26-item version, less than 3% for all the dimension scores, and the rate of spontaneous refusal were low, ranging from 1 to 2%, along with the various steps of development and depending on the surgical procedure. The average completion time of the 26-item EVAN-G is short enough (approximately 9 min) to allow its use in routine practice, unlike other instruments, comprising 40 items, which are too lengthy.[28,30]

The second issue concerns the relevance of developing a specific questionnaire assessing satisfaction in the perioperative period and the clinical interest of the EVAN-G. The previous results did not aim at explaining the links between clinical and sociodemographic status and satisfaction; clinical and sociodemographic data were only required to test hypotheses to reinforce the validity of the questionnaire. However, our findings are consistent with those reported in literature, especially regarding sex and age, described as high predictors of patients' satisfaction level.[31–33]

The perioperative period, considered the time frame of interest, lasting from the first visit with the anesthetist before surgery up to the 48th hour after surgery, is remarkable. This issue has often been discussed in previous published works.[32] So, it becomes possible to take into account the perioperative period, making it possible to examine the factors related to anesthesia care that influence satisfaction, in addition to the more usual factors reflecting surgical procedures. Moreover, the perioperative period is a unique process and a global experience from a patient's point of view, and EVAN-G items cover both the successive steps in the care process and the different caregivers involved in the process.

Restriction of the period of questionnaire administration to the first 48 h after surgery is also found in other studies exploring anesthesia care management.[34] Conversely, the assessment of satisfaction at a later point could mainly reflect perceptions related to surgery. This probably led to an underestimation of the reliability of the questionnaire using the test–retest procedure in 36 patients at a 15-day interval.

Increasingly, clinical researchers and healthcare providers have focused on measuring patients' perceptions of the outcomes of care, namely patient-reported outcomes. In anesthesiology, efficacy and safety drive anesthesia providers' determination of the quality improvement. Anesthesia-related mortality has been dramatically reduced during the past 20 yr,[35] but still, little consideration is given to integrating patients' satisfaction in the decision-making process. Nevertheless, the evaluation of services by patients should be an integral part of continuous quality improvement in anesthesia. Therefore, it is important to identify the factors for patient dissatisfaction. An important question is, what exactly does the

questionnaire evaluate? Because there is no accepted standard of patient satisfaction, the global value of the EVAN-G in a population is not an absolute value reflecting a given level of satisfaction. Moreover, one can imagine that the simple fact of having passed the perioperative period without any major adverse event would overestimate patient satisfaction. This is probably true and may explain very high levels of satisfaction when the question is whether a patient is "satisfied," "somewhat dissatisfied," or "dissatisfied."[36] But a 26-item questionnaire based on patients' expectations is obviously less contaminated by the happy feeling of a good surgical outcome. It could also be argued that one severe discomfort (*e.g.*, postoperative pain or frequent nausea or vomiting) might lead to a low satisfaction level even if the other expectations were met, but the multidimensional construction of the scale, with low correlation values between dimensions, should allow analysis of the causes of dissatisfaction.

Furthermore, the EVAN-G may be useful to assess the comparative effects of alternative anesthesia procedures. From this perspective, one obvious limitation of the EVAN-G is that it cannot forecast for regional anesthesia or monitored anesthesia care. Further work should explore patients' satisfaction in these specific contexts. One alternative would be to develop modules specific to these special anesthetic contexts in addition to a core questionnaire. Currently, research with patients undergoing regional anesthesia has already been conducted, confirming the relevance of this modular approach.

Last, for anesthetists, one major interesting quality of the EVAN-G is the wide range of surgical procedures studied. In most satisfaction studies, validated questionnaires focus on a single type of surgery, meeting the specific needs of surgeons but restricting the usefulness of the questionnaire with regard to anesthetists' decision making. For example, the pain component is very important for major surgical procedures but is less relevant to procedures associated with a low level of pain. Therefore, the same global score could reflect various perceptions of the anesthesia process.

The EVAN-G is not intended to replace conventional outcome measures such as mortality, adverse events, or other information dealing with recovery. The EVAN-G adds important information, oriented toward patients' views and perceptions, to the information traditionally collected in anesthesia. Further work is needed to test its strengths and weaknesses in different cultural contexts. However, reliability and validity of the EVAN-G are evidenced by our results. Our approach provides a novel, valid, and valuable tool that may be used in routine anesthesia practice.

References

1. Wu CL, Naqibuddin M, Fleisher LA: Measurement of patient satisfaction as an outcome of regional anesthesia and analgesia: A systematic review. Reg Anesth Pain Med 2001; 26:196–208

2. Pascoe G: Patient satisfaction in primary health care: A literature review and analysis. Eval Prog Planning 1983; 6:185–210

3. Calnan M: Towards a conceptual framework of lay evaluation of health care. Soc Sci Med 1988; 27:927–33

4. Schug SA: Patient satisfaction-politically correct fashion of the nineties or a valuable measure of outcome? Reg Anesth Pain Med 2001; 26:193–5

5. Carr-Hill RA: The measurement of patient satisfaction. J Public Health Med 1992; 14:236–49

6. Williams B: Patient satisfaction: A valid concept? Soc Sci Med 1994; 38:509–16

7. Le May S, Hardy JF, Harel F, Taillefer MC, Dupuis G: Patients' perceptions of cardiac anesthesia services: A pilot study. Can J Anaesth 2001; 48:1127–42

8. Whitty PM, Shaw IH, Goodwin DR: Patient satisfaction with general anaesthesia: Too difficult to measure? Anaesthesia 1996; 51:327–32

9. Heidegger T, Husemann Y, Nuebling M, Morf D, Sieber T, Huth A, Germann R, Innerhofer P, Faserl A, Schubert C, Geibinger C, Fluckiger K, Coi T, Kreienbuhl G: Patient satisfaction with anaesthesia care: Development of a psychometric questionnaire and benchmarking among six hospitals in Switzerland and Austria. Br J Anaesth 2002; 89:863–72

10. 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

11. Auquier P, Blache JL, Colavolpe C, Eon B, Auffray JP, Pernoud N, Bruder N, Gentile S, Francois G: A scale of perioperative satisfaction for anesthesia: I. Construction and validation. Ann Fr Anesth Reanim 1999; 18:848–57

12. Pernoud N, Colavolpe JC, Auquier P, Eon B, Auffray JP, Francois G, Blache JL: A scale of perioperative satisfaction for anesthesia: II. Preliminary results. Ann Fr Anesth Reanim 1999; 18:858–65

13. Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM: Patient satisfaction after anaesthesia and surgery: Results of a prospective survey of 10,811 patients. Br J Anaesth 2000; 84:6–10

14. Shevde K, Panagopoulos G: A survey of 800 patients' knowledge, attitudes, and concerns regarding anesthesia. Anesth Analg 1991; 73:190–8

15. Tong D, Chung F, Wong D: Predictive factors in global and anesthesia satisfaction in ambulatory surgical patients. ANESTHESIOLOGY 1997; 87:856–64

16. Le May S, Hardy JF, Taillefer MC, Dupuis G: Patient satisfaction with anesthesia services. Can J Anaesth 2001; 48:153–61

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

18. Nunnally JC, Bernstein ICH: Statistical foundations, Psychometric Theory, 3rd edition. Edited by Nunnally JC, Bernstein IH. New York, Mc Graw-Hill, 1994, pp 31–208

19. Juniper EF, Guyatt GH, Jaeschke R: How to develop and validate a new health related quality of life instruments, Quality of Life and Pharmacoeconomics in Clinical Trials, 2nd edition. Edited by Spilker B. Philadelphia, Lippincott-Raven, 1996, pp 49–56

20. Fung D, Cohen M: What do outpatients value most in their anesthesia care? Can J Anaesth 2001; 48:12–9

21. Kaid LL, Wadsworth AJ: Content analysis, Measurement of Communication Behavior. Edited by Emmert P, Baker LL. New York, Longman, 1989, pp 97–217

22. Auroy Y, Clergue F, Laxenaire MC, Lienhart A, Pequignot F, Jougla E: Anesthetics in surgery. Ann Fr Anesth Reanim 1998; 17:1324–41

23. Rasch G: A structural model for items of a test, Probabilistic Models for Some Intelligence and Attainment Tests. Edited by Rash G. Chicago, Mesa Press, 1993, pp 62–75

24. Rasch G: An item analysis which takes individual differences into account. Br J Math Stat Psychol 1966; 19:49–57

25. Melzack R: The short-form McGill Pain Questionnaire. Pain 1987; 30:191–7

26. Spielberger CD, Sydeman SJ: State Trait Anxiety Inventory and State-Trait Anger Expression Inventory, The Use of Psychological Tests for Treatment Planning and Outcome Assessment. Edited by Maruish ME. Hillsdale, LEA, 1994, pp 292–21

27. Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM: Who should measure quality of life, the doctor or the patient? Br J Cancer 1988; 57:109–12

28. Myles PS, Hunt JO, Nightingale CE, Flechter 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

29. Myles P, Weitkamp B, Jones K, Melick J, Hensen S: Validity and reliability of a post operative quality of recovery score: The QoR-40. Br J Anaesth 2000; 84:11–5

30. Oakes CL, Ellington KJ, Oakes KJ, Olson RL, Neill KM, Vacchiano CA: Assessment of postanesthesia short-term quality of life: A pilot study. AANA J 2002; 70:267–73

31. Huppe M, Beckhoff M, Klotz KF, Heinzinger M, Prussmann M, Gerlach K, Ocker H, Schmucker P: Reliability and validity of the Anaesthesiological Questionnaire for electively operated patients. Anaesthesist 2003; 52:311–320

32. Thierbach A, Fichtner K, Kugler A, Dick W: Patient's emotional satisfaction with anaesthetic care. Anesthesiol Intensivmed Notfallmed Schmerzther 2003; 38:85–9

33. Myles PS, Reeves MD, Anderson H, Weeks AM: Measurement of quality of recovery in 5672 patients after anaesthesia and surgery. Anaesth Intensive Care 2000; 28:276–80

34. Eberhart LH, Greiner S, Geldner G, Wulf H: Patient evaluation of postoperative recovery: An evaluation of the QoR scores in 577 patients. Anaesthesist 2002; 51:463–6

35. Lienhart A, Auroy Y, Péquignot F, Benhamou D, Jougla E: Premières leçons de l'enquête «mortalité» Sfar-Inserm. Conférences d'actualisation, Sfar ed. Paris, Elsevier, 2003, pp 203–18

36. Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM. Patient satisfaction after anaesthesia and surgery: Results of a prospective survey of 10,811 patients. Br J Anaesth 2000; 84:6–10

Appendix: List of the 26 EVAN-G Items

No.	French Items	Item General Meaning
	Avant l'opération, lors des visites avec l'anesthésiste:	During the preoperative visits with the anesthetist:
1	J'ai reçu de l'information sur ce qui allait se passer	I received information about what was going to happen
2	J'ai pu poser les questions que je voulais	I was able to ask questions I wanted
3	Je me suis senti rassuré, détendu, mis en confiance	I felt reassured, relaxed, confident
	Avant l'opération, lors des visites avec le chirurgien,	During the preoperative visits with the surgeon:
4	J'ai reçu de l'information sur ce qui allait se passer	I received information about what was going to happen
5	Je me suis senti rassuré, détendu, mis en confiance	I felt reassured, relaxed, confident
6	Le chirurgien a été attentionné	The surgeon was attentive
	Des premières consultations jusqu'à la descente au bloc	From the first visits until I entered the operating room:
7	Mon intimité a été respectée	My privacy was respected
	A l'arrivée au bloc:	When entering the operating room:
8	J'ai été gêné dans mon confort: froid, chaud, mal installé sur la table. . .	I felt uncomfortable: cold, warm, badly postured on the operating table. . .
9	Mon intimité a été respectée	My privacy was respected
10	Les médecins ont été attentionnés	The medical staff was attentive
	En salle de réveil:	In the recovery room:
11	J'ai ressenti des sensations désagréables comme: soif, faim, nausées, vomissements, maux de tête	I had unpleasant feelings such as: thirst, hunger, nausea, headache. . .
12	J'ai été gêné dans mon confort: froid, chaud, mal installé dans le lit. . .	I felt uncomfortable: cold, warm, badly postured on the bed. . .
13	J'ai eu mal	I had pain
14	J'ai eu un réveil agréable	I had a pleasant awakening
15	Les médecins et le personnel soignant ont été attentionnés	The medical and nursing staff were attentive
	Depuis mon retour dans le service:	Since I came back to my bedroom:
16	J'ai ressenti des sensations désagréables comme: soif, faim, nausées, vomissements, maux de tête	I had unpleasant feelings such as: thirst, hunger, nausea, headache
17	J'ai été gêné dans mon confort: froid, chaud, mal installé sur le lit. . .	I felt uncomfortable: cold, warm, badly postured on the bed. . .
18	J'ai eu des difficultés à effectuer les gestes de la vie quotidienne: communiquer, manger, se laver, aller aux toilettes. . .	I had difficulty achieving everyday tasks: communicating, eating, washing myself, going to the toilet
19	J'ai eu mal	I had pain
20	J'ai été soulagé de ma douleur	I was relieved of my pain
21	Les médecins ont été attentionnés	The medical staff was attentive
22	Le personnel soignant a été attentionné	The nursing staff was attentive
23	Mon intimité a été respectée	My privacy was respected
	Globalement, depuis la première consultation jusqu'à ce jour:	Overall, since my first visit until today:
24	J'ai pu voir mon entourage (famille, amis)	I was able to meet my family or friends
	Les délais d'attente m'ont paru anormalement élevés:	Waiting times in the hospital seemed too long:
25	Pour avoir un rendez-vous avec l'anesthésiste ou le chirurgien	To obtain an appointment with the anesthetist or surgeon
26	Lors des consultations avant l'opération	During the preoperative visits