François Lehmann, MD, CCFP Denise Fontaine, RN
André Bourque, MD, CCFP Luc Côté, MSW

Measurement of Patient Satisfaction: The Smith-Falvo Patient-Doctor Interaction Scale

SUMMARY

Patient satisfaction has been considered important for many years. This article provides a review of various methods that have been developed to measure patient satisfaction and describes the use of the Smith-Falvo scale in determining patient satisfaction with the medical services provided by residents in the Verdun Family Practice Program. In view of the limited range of scores provided by the use of this scale, the authors recommend that further research be done to develop a method of assessment of patient satisfaction that will take into account the duration of patient-physician interviews. (*Can Fam Physician* 1988; 34:2641–2645.)

RÉSUMÉ

Depuis de nombreuses années, on accorde une importance marquée à la satisfaction du patient. Cet article passe en revue les différentes méthodes mises au point afin de mesurer la satisfaction du patient et décrit l'échelle de Smith-Falvo utilisée à l'Unité de médecine familiale de Verdun pour déterminer la satisfaction des patients vus par les résidents. À la lumière des résultats limités fournis par cette échelle, les auteurs recommandent de poursuivre les recherches afin de mettre au point une méthode d'évaluation qui tiendrait compte de la durée de l'entrevue patient-médecin.

Key words: patient satisfaction, methods of assessment, residency training

Dr. Lehmann is an Assistant Professor of Family Medicine at the University of Montreal. Ms. Fontaine is a member of the teaching team at the Verdun Family Practice Clinic. Dr. Bourque is an Associate Professor of Family Medicine at the University of Montreal. Mr. Côté is a family therapist at the Verdun Family Medicine clinic and a Lecturer in the Department of Family Medicine of the University of Montreal. Requests for reprints to: Dr. François Lehmann, Family Medicine Clinic, Verdun General Hospital, 4000 Lasalle Blvd., Verdun, Québec. H4G 2A3

PATIENT SATISFACTION is a concept that has been addressed as important for many years. As early as 1968, Korsch[1] conducted interviews with the parents of pediatric patients in order to measure their satisfaction. A patient who is satisfied is thought to be more likely to comply with treatment and/or experience a more favourable outcome. From an administrator's viewpoint, patient satisfaction might well influence the use of the different components of the health-care system and aid in better planning. From an educational viewpoint, patient satisfaction is one of the desired outcomes of a training program, especially in the discipline of family medicine, where the quality of the doctor-patient relationship is so highly valued.

We at the Verdun Family Medicine Clinic needed a reliable and valid measure of patient satisfaction to evaluate our residents' interviews and to compare our findings with our own methods of evaluating patient-doctor interaction. In 1978, Ware[2] reviewed more than one hundred articles on patient satisfaction that had been published within the preceding 35 years. He points out that underlying the use of satisfaction data is the assumption that "satisfaction" questionnaires measure patient satisfaction reliably and validly. Yet he found that only 11 of 81 empirical studies reported reliability estimates for patient satisfaction measures, and those that did so report suggest poor

reliability for single-item measures. Ware also notes that the validity of satisfaction scores as dependent variables in relation to specific characteristics of health-care providers are strictly limited.

Since the publication of this 1978 review, others have tried to produce better measurement devices. In 1984, Feletti[3] constructed a scale, the items of which were chosen by the researchers. No explanation was given for the choice of those particular items, which are descriptive of physician conduct without reference to the importance of this conduct to the patient. In 1979, Biehn[4] published a scale, the items of which, once again, were chosen by the researchers and not by patients. Moreover no evaluation of reliability or validity was included. Comstock[5] interviewed patients and asked about their preferences before choosing the items for her satisfaction scale, but there is no documentation of reliability and validity measures. The Di Matteo scale[6] represents a clear improvement over the earlier ones, for although the individual items are chosen by researchers and not by patients, concurrent validity is shown, and every patient was asked whether he/she wished to return to see the same doctor. Since responses were anonymous and patients were assured that their physician would not see their answers, their planning to return is probably good evidence of satisfaction. Unfortunately, this scale's value is considerably decreased because reliability could not be demonstrated.

In 1978, Wolf[7] developed a scale of a higher quality as reflected by it's use by other researchers, among them Henbest.[8] Fifty patients critically assessed the appropriateness of each item in the scale. Internal consistency was demonstrated, as was reliability, by means of Cronbach's coefficient alpha. But, as Wolf himself points out, "further research is needed to assess [the instrument's] clinical validity". Moreover — and this is a factor that we should have looked at more attentively — both the median and the means were above .8, though Wolf states that the results are less skewed than those obtained by means of other scales. Finally, since the scale includes 46 items, its length might discourage its use.

Fully realizing that there is no perfect satisfaction scale, we were nevertheless impressed with the one constructed by Falvo in 1983.[9] She and Smith developed the scale after interviewing patients from a family practice, generating 1540 descriptions of patient-preferred and not-preferred physician behaviours. In a second interview, other family-practice patients rated categories of behaviour by preference, and from the results a satisfaction scale was constructed.[10] Reliability was assessed by the test-retest method, and internal reliability was measured by Cronbach's alpha. Concurrent validity was assessed by correlating the scale's scores with the patient's reported intention to return to the physician for further health care. Convergent validity was assessed by correlating the scores with those of Wolf's scale, described in the previous paragraph.

Because of its strengths, we chose to use the Smith-Falvo scale.

Method

The Verdun Family Practice Program lasts two years and accepts 24 residents at any one time. Of the 14 residents participating in the study, eight were in their first year of training, and six were in their second year. Their average age was 25 years. Ten of the participating trainees were women, and they conducted 19 of the 28 interviews. All were selected because their particular rotations gave them the time required to participate in the study when we were ready to proceed. None of them knew the purpose of the research project.

Our program makes use of community-based family practices. Thirty-seven patients from two community-based practices and one from our own hospital-based unit were approached and asked to participate in our study. Seven of these patients were coming for a regular follow-up visit and 21 others for a walk-in consultation. There was no patient selection, except for age: we ensured representation of all age groups by including four patients in each of the following age brackets: 0–3 years, 4–11 years, 12–17 years, 18–35 years, 38–84 years, 85–74 years, and over 75 years of age. A member of our team approched each patient in the appropriate age groups as soon as he or she had registered with the medical secretary. Ten patients refused to participate. They were all in the 12–17-year age bracket. For the 28 patients who accepted, the average number of years of schooling was 9.5.

The patients had a single encounter with a resident whom they had never seen before. They knew that they would not be seen by this doctor again in the foreseeable future.

The interview, but not the physical exam, was filmed. Filming is a procedure we regularly use in our teaching unit. The residents were supervised in the usual way, by direct discussion with the supervisor after taking a history and performing physical exam.

Immediately following the doctor-patient encounter, the patient (or his/her parent for those under 15 years of age) completed the Smith-Falvo questionnaire in private. A nurse, always the same one, was available to help any patient who needed assistance in interpreting the questionnaire (Figure 1).

Twenty-four of the interviews were conducted in the French language. The questionnaire was translated by six bilingual instructors (two physicians, three nurses and a family therapist), and this translation was submitted for checking to three other bilingual physicians. Their comments were considered and appropriate corrections were made. No countertranslation was done.

The questionnaire was scored according to the same method used by Smith-Falvo, giving each answer a rating of one to five (Figure 1).

Results

All residents obtained a score above 75% on the Smith-Falvo scale; 13 of the 28 interviews were scored above 85%. Assuming a normal distribution, the 99% confidence interval of the mean is $.905 > M > .834$.

A T-test pairing of the answers for all the interviews indicated that questions 7 and 11 on the Smith-Falvo scale scored significantly lower than the other answers. Re-weighting these two items by giving them more relative importance did not decrease the mean of the distribution because for the lowest-scored interviews these two questions had not been completed. Moreover, after re-weighting these two items, item-total correlations (Pearson) for four of the 19 items were unsatisfactory.

Figure 1

Dr. _____

Smith-Falvo Patient-Doctor Interaction Scale

It is important to our resident physicians to know what you, their patients, feel about your interaction with them. Only with your help can the physicians be aware of what areas they should try to improve and in what areas they are especially good. Please help us give them this feedback by filling out the following questionnaire. Your physician will not see this questionnaire and will not be aware what you, as an individual, said about him/her, but only what patients as a group said. Complete confidentiality will be maintained.

Thinking about the visit you just had with your physician, please check the boxes that best describe whether you agree or disagree with the following statements:

	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree	Does Not Apply
1. The doctor went straight to my medical problem without first greeting me.	1	2	3	4	5	
2. The doctor greeted me pleasantly.	5	4	3	2	1	
3. The doctor seemed to pay attention as I described my condition.	5	4	3	2	1	
4. The doctor made me feel as if I could talk about any type of problem.	5	4	3	2	1	
5. The doctor asked questions that were too personal.	1	2	3	4	5	
6. The doctor handled me roughly during the examination.	1	2	3	4	5	
7. The doctor gave me an explanation of what was happening during the examination.	5	4	3	2	1	
8. The doctor explained the reason why the treatment was recommended for me.	5	4	3	2	1	
9. I felt the doctor diagnosed my condition without enough information.	1	2	3	4	5	
10. The doctor recommended a treatment that is unrealistic for me.	1	2	3	4	5	
11. The doctor considered my individual needs when treating my condition.	5	4	3	2	1	
12. The doctor seemed to rush.	1	2	3	4	5	
13. The doctor behaved in a professional and respectful manner towards me.	5	4	3	2	1	
14. The doctor seemed to brush off my questions.	1	2	3	4	5	
15. The doctor used words I did not understand.	1	2	3	4	5	
16. The doctor did not give me all the information I thought I should have been given.	1	2	3	4	5	
17. The doctor criticized me for not taking care of myself.	1	2	3	4	5	
18. I would recommend this doctor to a friend.	5	4	3	2	1	
19. I would return to this doctor for future health care.	5	4	3	2	1	

THANK YOU FOR TAKING TIME TO FILL OUT THIS QUESTIONNAIRE!

Discussion

Unusual distribution of scores

The Smith-Falvo scores obtained in our setting gave us a very high mean and showed little dispersion. We believe that these results fail to permit discrimination between good and "poor" interviews. To explain these results we considered the following hypotheses.

Poor translation of the Smith-Falvo scale

Since the scale was administered in French in 24 of the 28 interviews, the scale's validity could be affected if the translation were poor. However, since nine different bilingual health professionals critically evaluated the translation, major errors seem unlikely.

Residents' characteristics

Overall, the interview scores on the Smith-Falvo scale were outstanding. Are we dealing with an outstanding group of residents? The residents are young physicians-in-training. Are patients less critical of the young trainee in view of his/her lack of experience?

Although the residents were not told the specific purpose of the study, one usually expects that in the context of a taped interview, they exhibit their best behaviour. Are the scores on the interaction scale a reflection of patients' satisfaction with the residents' skills or with their effort to perform well?

The interviews were subsequently viewed and graded independently by four teachers of family medicine, as well as by the residents themselves. For this purpose a reliable scale of doctor-patient relationship was used.[11] With this scale there was good dispersion in performance, whether graded by the teachers or by the residents themselves. We are convinced that this dispersion more truly represents the real differences in performance.

Patient characteristics

The educational level of our patients is lower than that of the Smith-Falvo sample. There is one study which suggests that less highly educated patients are less critical of their doctors, but other studies do not report this finding.[2]

The type of encounter

Each patient was meeting the doctor for the first time. Are patients more tolerant during a first encounter, preferring to wait before making final judgements? Do these same patients hesitate to criticize because they are so rarely asked to do so?

The reputation of the community practice

All of the interviews except one were held in two Local Community Health Centres. Both Centres offer a variety of health services and enjoy an excellent reputation in their community. It is conceivable that patient satisfaction with the overall health services could have a "halo" effect on patient satisfaction with the doctor-patient encounter.

Deficiencies of the scale

Many of the questions in the Smith-Falvo scale address the issue of professional manners and patient respect in a rather basic sense (Questions 1–3, 5, 8, 13, 14, 17): "The doctor greeted me pleasantly" (Q.2); "The doctor handled me roughly during the examination" (Q.6); "The doctor criticized me for not taking care of myself" (Q.17). It is rather unusual for a resident to show flaws in these matters to the point of drawing criticism from a patient. Another series of questions addresses the issue of communication and information given to the patient. (Q. 7, 8, 15, 16): "The doctor gave me an explanation of what was happening during the examination." (Q.7) Public criticism directed to the medical profession on this point has been strong and repeated in the past years: again, although one hardly expects a supervised young physician to neglect this area grossly, serious deficiencies often appear. It may be that the questions as formulated in the scale are not sufficiently sophisticated to elicit a critical response by the patient. The patient may be influenced in this questionnaire to give a favourable assessment in an "all or nothing" manner. Furthermore the construction of the scale may tend to induce a halo effect and a proximity error in the direction of more favourable scoring.

No control for the length of the interview

Our interviews lasted from 20 to 80 minutes; no time limits were set for the residents. We have no solid proof, but an article by Kent Smith[12] and our own clinical experience tell us that patients overlook many shortcomings in a physician who spends a lot of time with them. Conversely, even if the history, physical exam, diagnosis and prescribed treatment are scientifically unassailable, the patient may be completely dissatisfied with the encounter if the physician spends too little time. We suggest that the variable duration of an interview may be the factor that *most* influences patient satisfaction.

After our study was completed we contacted one of the authors who conducted the initial study with the Smith-Falvo scale. She reported that by far the greater number of their own interviews were also highly rated and were not controlled for time.

Conclusion

In our study, the Smith-Falvo scale produced a dispersion of scores that severely limits discrimination between good and poor interviews. Even though "satisfaction" is a relative concept, and even if some discrimination may be possible, since nearly half the scores were above 85%, we must look for other variables that might have produced such high scores.

When measuring patient satisfaction, convergent and concurrent validity do not provide sufficient criteria by which to evaluate a scale. Dispersion of scores must permit discrimination between poor and excellent interviews. The Smith-Falvo scale, as administered by our group, does not permit this dispersion. Further re-

search should take into account the length of the doctor-patient encounter, a variable which could be more important than any other. ■

Acknowledgements

Thanks to:

- Cécile Pineault, who typed and retyped this paper so many times.
- The employees of the CLSC Centre Sud and the Pte St-Charles Community Clinic whose staff allowed us the use of their offices and their clientele.

References

1. Korsch BM, Gozzi EK, Francis V. Gaps in doctor-patient communication. I. Doctor-patient interaction and patient satisfaction. *Pediatrics* 1968; 42:855.

2. Ware JE, Davies-Avery A, Stewart AL, et al. Measurement and meaning of patient satisfaction. *Health Med Care Service Rev* 1978; 1:15.

3. Felleti GI, Corney SL. Evaluating patient's satisfaction with medical student's interviewing skills. *Med Educ* 1984; 18:15–20.

4. Biehn JT, Molineux JE. Patient evaluation of physician performance. *J Fam Pract* 1979; 83:585–9.

5. Comstock LM, Hooper EM, Goodwin JM, Goodwin JS. Physician behaviors that correlate with patient satisfaction. *J Med Educ* 1982; 57:105–12.

6. Di Matteo RM, Prince LM, Taranta A. Patients' perceptions of physician behavior: determinants of patient commitment to the therapeutic relationship.

7. Wolf MH, Putnam SM, James SA, Stiles WB. Medical interview satisfaction scale: development of a scale to measure patient perceptions of physician behavior. *J Behav Med* 1978; 1(4):391–401.

8. Henbest RJ. Study of the patient-centred approach in family practice. (thesis) London, Ont.: University of Western Ontario, 1985.

9. Falvo DR, Smith JK. Assessing residents' behavior science skills: patients' views of physician-patient interaction. *J Fam Pract* 1983; 17(3):479–83.

10. Smith JK, Salvo D, McKillip J, Pitz G. Measuring patient perceptions of the doctor-patient relationship: development of the P.D.I.S. *Evaluation and the health professions* 1984; 7(1):77–94.

11. Manuscript in preparation. The quality of doctor-patient interactions: a reliable and valid checklist. Submitted to *Can Fam Physician*.

12. Smith CK, Kent C, Polis E, Haldac RR. Characteristics of the initial medical interview associated with patient satisfaction and understanding. *J Fam Pract* 1981; 12(2):283–8.