PRACTICE OBSERVED

*Practice Research***Simulated patients in general practice: a different look at the consultation**

J J E RETHANS, C P A VAN BOVEN

Abstract

To develop a better empirical basis for developing quality assessment in general practice three simulated patients made appointments with 48 general practitioners during actual surgery hours and collected facts about their performance. The simulated patients were indistinguishable from real patients and presented a standardised story of a symptomatic urinary tract infection. Two months later the same general practitioners received a written simulation about a patient who had the same urinary tract infection and were asked how they would handle this in real practice. Both results were scored against an existing consensus standard. The overall score for both methods did not show any substantial differences. A more differentiated analysis, however, showed that general practitioners performed significantly better with simulated patients. It also showed that general practitioners answering the written simulation performed significantly more unnecessary and superfluous actions.

The results of this study show that the use of simulated patients seems to show the efficient performance of general practitioners in practice.

Introduction

In general practice as well as in hospitals efforts are being made to develop "ideal" standards and equivalent, logical, branching flow charts for clinical performance. These developments aim at improving the performance of individual tasks, to promote a more uniform performance by doctors, and to provide standards for the quality of care in solving clinical problems.^{1,2} The use of such standards has important consequences—for example, a positive or negative evaluation of an individual doctor's performance and a reduced cost of health care for the community.

Most methods for developing standards are based on consensus models, in which a group of general practitioners discusses the complaint or medical problem in question and decides on a standard procedure for dealing with it. Studies of the use of such standards, however, show that doctors seem to do less well than such standards deem to be desirable.^{3,4} General practitioners may not be acting as efficiently as the designers of the standard would like, but the standard may also be too academic in that it does not allow for the circumstances actually occurring in general practice. It is therefore desirable to establish a standard of performance for general practitioners based on the conditions of real life.

Methods of assessing actual performance may be either indirect, by sending a questionnaire to general practitioners to answer questions about problems with managing patients, or direct, by actually observing doctors while they are consulting. Comparisons have shown a discrepancy between doctors' clinical performance as assessed by indirect methods and what they actually do when they are observed in their practice.³ Observed doctors omit actions that were expected from them from the results of the indirect methods. These studies, however, compared different groups of doctors dealing with different medical problems.

Though there is a substantial experience with the use of simulated patients for educational purposes,^{5,8} and the simulated patient method has been described as the best method to assess the

Department of General Practice, University of Limburg, Postbus 616,
6200 MD Maastricht, The Netherlands

J J E RETHANS, MD, general practitioner

Department of Medical Microbiology, University of Limburg, Maastricht,
The Netherlands

C P A VAN BOVEN, PHD, professor

Correspondence to: Dr Rethans.

management of patients by doctors,⁹ few studies have used simulated patients to research problems in actual practice.⁹⁻¹² This study aimed at comparing the performances of a group of general practitioners when dealing with a particular medical problem by an indirect method of assessment, using written details about a simulated patient, and a direct method, using simulated patients in practice. We hypothesised that the general practitioners would do less in a consultation with the simulated patient than they would have indicated in their answers to the written problem.

The study therefore had the following purposes: to investigate the possibility of using trained simulated patients to collect data about the reality of the consultation; to measure the difference between the performances of general practitioners when dealing with the simulated patient (what the doctor does) and the actions that the same general practitioners said that they would take when dealing with a written problem about a similar patient (what the doctor says he would do); and to measure the agreement between an already existing standard based on a consensus model and the actual performance of general practitioners as shown by data gathered using simulated patients.

Subjects and methods

The medical problem in this study concerned a 30 year old woman who complained about painful micturition for one and a half days. This problem was constructed as a standardised role for three simulated patients, who were paid to participate in this project and had no medical background. The problem was also written down in the form of a "simulation of initial medical problem solving."

A During your surgery Mrs Cox appears. You don't know her: she is with her husband and two children on holiday. To you she looks about 30 years old. She complains of painful micturition for the past day and a half.

What would you like to ask her about her history?

B During your surgery hours Mrs Smith appears. She is 30 years old, and you have not seen her for a long time. She lives with her husband and two children in their own house in a new part of the town. She tells you that she has been well for the past few years. She now complains of painful micturition for the past day and a half. You question her and suspect that she has an uncomplicated urinary tract infection.

- 1 If you usually perform a physical examination what exactly would you do?
- 2 If you usually perform laboratory tests what exactly would you do?

After your completed history and examination you conclude that she has indeed got an uncomplicated urinary tract infection.

- 3 What would be your further management (in terms of what you would prescribe and for how long, what you would tell her, and how you would arrange a follow up)?

Written simulation of initial medical problem solving used in study.

The written form, meant to be used to assess medical competence, was a simplification of a modified essay question (E de Graaf, unpublished observations). The figure shows the format used in this study. The written format made it necessary to mention the suspected urinary tract infection before the doctor continued with the questions on management (during a real consultation the doctor receives answers from the patient). For this reason part A and part B were presented separately to the doctors (while reading part A they were unable to see part B).

The facts collected about history, results of physical and laboratory examination, instructions given to the patient, treatment, and follow up were scored according to an existing standard on urinary tract infection developed at the University of Nijmegen.¹³ The standard is divided into obligatory actions (considered to be necessary) and intermediate actions (not essential but not harmful either). All other actions are considered to be superfluous (see tables I and II for the items listed in the standard used).

The simulated patients were trained in their role as a patient and in

reporting reliable and valid facts about the consultation. To check the reliability among the simulated patients we assessed three test consultations. The patients completed a checklist of items (based on the standard) immediately after such a consultation, and the same consultation was also scored by three independent doctors. For each consultation the scores obtained by the three independent doctors were considered to be the gold standard for that particular consultation. Subsequently, the score of the simulated patient was compared with this gold standard to assess their reliability, which was 1.0 ($\kappa=1.0$), 0.89 ($\kappa=0.78$), and 0.89 ($\kappa=0.78$), respectively. To assess the consistency of individual patients each patient was retested after six weeks with the same consultation (recorded on videotape). Their consistencies were 0.96 ($\kappa=0.92$), 0.89 ($\kappa=0.78$), and 0.85 ($\kappa=0.70$), respectively. During the study the simulated patients were themselves responsible for making appointments with the doctors. They presented themselves as Mrs Cox (figure).

Most Dutch general practice units have their own small laboratory, where medical secretaries perform simple tests. Simulated patients are obviously unable to undergo such tests, so we asked the secretaries of the participating doctors what kind of laboratory investigations their doctor would order them to do to diagnose a urinary tract infection. The reliability of the answers received was assessed by asking the secretaries the same question one week later. The agreement between their first and second answers was 99.25%.

TABLE I—Number (%) of general practitioners ($n=25$) performing items listed in the standard procedure for urinary tract infection during consultation with simulated patient and in answers to written simulation

Item No	Obligatory(O), intermediate (F), or superfluous (S) action	Performed	
		During consultation	In written answer
History:			
Chief complaint	1 O	25 (100)	25 (100)
Frequency	2 O	24 (96)	18 (72)*
Onset of complaint	3 O	13 (52)	11 (44)
Ask about general wellbeing	4 O	14 (56)	11 (44)
Ask if this is first time	5 F	22 (88)	22 (88)
Ask for history of urogenital tract problems	6 F	4 (16)	10 (40)
Vaginal discharge	7 F	5 (20)	11 (44)
Superfluous actions	8 S	5 (20)	12 (48)
Physical examination:			
Measure blood pressure	9 F	1 (4)	1 (4)
Feel flank for pain	10 F	3 (12)	9 (36)
Superfluous actions	11 S	5 (20)	13 (52)*
Instructions to patient:			
Explain diagnosis	12 O	23 (92)	6 (24)**
Explain prognosis	13 O	12 (48)	2 (8)*
Advise voiding after intercourse	14 F	1 (4)	2 (8)
Advise drinking more fluids	15 F	11 (44)	9 (36)
Discuss micturition habits	16 F	2 (8)	5 (20)
Ask for fresh midstream specimen	17 O	1 (4)	
Superfluous actions	18 S	3 (12)	2 (8)
Treatment:			
Antibiotics	19 O	25 (100)	23 (92)
Superfluous actions	20 S		
Follow up:			
Standard	21 O	20 (8)	19 (76)
After treatment	22 O	18 (72)	16 (64)
If complaint persists despite drugs return earlier	23 O	1 (4)	3 (12)
Superfluous actions	24 S	1 (4)	

* $p<0.05$, ** $p<0.0005$.

TABLE II—Number (%) of medical secretaries ($n=25$) performing items listed for laboratory examination in standard procedure for urinary tract infection, and number (%) of supervising general practitioners ($n=25$) indicating in answers to written simulation that these items are performed

Item No	Obligatory(O), intermediate (F), or superfluous (S) action	Performed	
		By secretary	In written answer
Clean midstream urine specimen	1 O	1 (4)	2 (8)
Microscopical urine examination	2 O	21 (84)	20 (80)
Urine concentration of:			
Glucose	3 O	21 (84)	13 (52)
Nitrite	4 F	9 (36)	10 (40)
Albumin	5 F	23 (92)	14 (56)*
Dipslide or urine culture	6 F	1 (4)	
Superfluous actions	7 S	1 (4)	

* $p<0.05$.

TABLE III—Number and range of actions listed in standard procedure for urinary tract infections performed by general practitioners ($n=25$) during consultation with simulated patient and in answers to written simulation (excluding laboratory examination). Each item of standard procedure has value of one point

Actions	Mean (SD) No performed	Range	% Of standard
Total:			
Simulated patient	9.88 (3.44)	4-19	
Written simulation	10.04 (3.37)	4-18	
Obligatory:			
Simulated patient	7.04 (1.54)**	4-9	64
Written simulation	5.24 (1.80)	2-8	47
Intermediate:			
Simulated patient	1.96 (1.24)*	0-6	24
Written simulation	2.76 (1.36)	0-6	34
Superfluous:			
Simulated patient	0.88 (1.74)*	0-8	
Written simulation	2.04 (2.16)	0-8	

* $p<0.05$, ** $p<0.005$, Student's t test for paired design.

All 378 general practitioners working in the same county as our university received information about this study four and a half months before the actual consultations with the simulated patients took place. They were told about the possibility of a visit by a simulated patient, who would collect facts about the consultation, during the next 12 months. The content of the case was specifically not mentioned. If a doctor did not want to participate in the study he or she had to withdraw actively; otherwise we considered him or her to be a possible participant.

Because one of us (JJER) was teaching at a school for medical secretaries during this study we made the criterion for including a general practitioner in the study the fact that he or she employed a secretary still connected with the school. This opened up the possibility of collecting information directly from the secretaries about the laboratory part of the consultation without bothering their supervising general practitioners and without disclosing the actual design of the study.

The participating doctors were visited by the simulated patients over a period of four and a half weeks. Two months after the visits all doctors were sent the written simulation and were requested to return it to us. The written problems that were returned were scored by a doctor according to the urinary tract infection standard. A select sample of the scores was also scored again independently by another doctor. Their coefficient of agreement was 0.94 ($\kappa=0.89$).

The two tailed McNemar test for paired data was used to compare the proportion of general practitioners performing an action during a consultation with a simulated patient and the proportion of general practitioners who explicitly said that they would carry out the same action in reply to the written simulation (significant at $p<0.05$). The two tailed Student's t test for paired data was used to compare the total number of actions scored by the general practitioners with the two methods (significant at $p<0.05$).

Results

Of the 378 doctors receiving notification of the study, 30 actively withdrew, among them three of the original 69 doctors who fitted the inclusion criterion. Of the remaining 66 doctors, 14 were excluded because of having connections with the department of general practice of our university or because their practice was based too far away. Of the remaining 52 doctors chosen for visits, four could not be visited: one simulated patient knew the doctor that she was supposed to visit, one patient was unable to get an appointment with her doctor, and two doctors could not be visited within the time allowed for this study. Thus at the end of the study 48 general practitioners had been visited. Of the 378 doctors receiving written simulation, 219 (57%) returned it, including 27 (56%) of the 48 doctors who were visited. Two of these 27 wrote spontaneously to say that they had detected the simulated patient; these two were excluded from the statistical analysis.

Table I shows, for each item of the urinary tract infection standard, the proportion of general practitioners who actually performed the action as reported by the simulated patient compared with the proportion indicating by the written method that they would perform the action. Table II compares separately the comments made by the doctors and their secretaries about laboratory tests. We do not present the results of all 48 doctors who were visited because 21 did not return the written simulation, and we cannot therefore make any comparison. The proportion of doctors per item of these 48 general practitioners, however, does not deviate substantially from that of the 25 doctors reported on here.

Table III shows the mean number of actions performed as reported by the simulated patients compared with the number performed as assessed by the written method. It shows also the number performed as a percentage of the number recommended by the standard.

Discussion

This study shows that simulated patients can produce reliable and valid data about the performance of doctors in actual practice. Only two of the 48 general practitioners who were visited reported afterwards that they detected the simulated patient. We have no reason to suspect that others also did, though we admit that we cannot be sure. Some doctors, on receiving the written simulation, may have recognised the visit of the simulated patient because they were able to recall the features of the case. To prevent the doctors from recalling the consultation with the simulated patient we introduced Mrs Smith in the written method. It could be argued that the section on management in the written simulation no longer applies to the visit of the simulated patient Mrs Cox. This is true, but there is no strict medical reason for the management of Mrs Smith and Mrs Cox to differ in the items recorded here.

Table I shows clearly that in taking the patient's history, giving the patient instructions, prescribing treatment, and following up the doctors performed more obligatory actions during the visit by the simulated patient than indicated by their corresponding intentions in the written simulation (except for item 23). The total score for both methods, however (table III), shows that more acts were intended to be performed than were performed in real life, but the difference was minimal and not significant. A more differentiated analysis, however, shows that for the obligatory acts doctors do more in reality than they say that they intend to do as measured by the written simulation, a finding that is at variance with our initial hypothesis. For the intermediate and superfluous actions, however, the findings were just the opposite.

These results are surprising. Most surprising is the finding that there seems to be a clear distinction between obligatory actions on the one hand and intermediate and superfluous actions on the other. We think that perhaps during the consultation with the simulated patient, under the pressure of time, the doctors deal efficiently with a routine problem. During the written simulation, however, they tend to want to show their knowledge, exhibiting an examination technique and forgetting about normal routine procedures. The actions and intentions for laboratory examinations show fewer differences, but this is the only category in which doctors are not being compared with themselves. Such results give reason to believe that using simulated patients shows a pattern of efficient performance (for the most essential issues) by general practitioners.

We are well aware that content specificity could affect the results.¹⁴ More research is therefore necessary to investigate whether the simulated patient method also shows this special pattern of efficient performance when more doctors are dealing with other medical problems. Practical problems, such as devising different ways to enter practices and deciding on which medical problems can be used in a simulation, still have to be resolved.

This study shows that the use of simulated patients in practice results in more (and probably more relevant) information about the performance of general practitioners than other existing methods. This is shown most clearly by the category "instructions to patients." The answers obtained by the indirect method would lead to the conclusion that doctors forget to tell their patients about the diagnosis, while the direct method shows that doctors are giving excellent instructions.

During consultation with a simulated patient the participating doctors performed only 64% and 24% of the obligatory and intermediate actions, respectively (table III). We have no reason to suspect the participating doctors of having low standards of care. We therefore think that these results indicate that the consensus standard does not take into account sufficiently the reality of practice for general practitioners. The standard should have a better empirical basis. The simulated patient method could be the method of providing this basis. Standards should almost certainly be

evolved in a practical setting even if developed initially in an abstract manner by discussion groups. In any case, the measurement of performance against standards should be carried out practically and not by an abstract questionnaire.

The results of this study encourage the use of the simulated patient method as a basis for establishing the actual performance patterns of general practitioners. The full documentation of these patients in daily practice still seems to be a great challenge.

We thank D L Crombie, C de Geus, J Hendrix, P Hobus, A Knottnerus, C Phaff, M Verwijnen, and W van Zutphen for their help during the study.

References

- Irvine D. Quality of care in general practice: our outstanding problem. *J R Coll Gen Pract* 1983;33:521-3.
- Wulff HR. *Probabilities and value judgements in clinical decision-making. Evaluation of efficacy of medical action.* Amsterdam: North-Holland Publishing, 1979.
- Tielens VCL, Mookink H, Mesker P. Het medisch diagnostisch handelen van de huisarts. *Huisarts en Wetenschap* 1985;28:44-8.
- Norman GR, Tugwell P, Feighner JW. A comparison of resident performance on real and simulated patients. *J Med Educ* 1982;57:708-15.
- Norman GR, Feighner JW. A comparison of behaviour on simulated patients and patient management problems. *Med Educ* 1981;15:26-32.
- Lamont CT, Hennen BKE. The use of simulated patients in a certification examination in family medicine. *J Med Educ* 1972;47:789-92.
- Barrows HS. *Simulated patients.* Springfield, Illinois: Charles C Thomas, 1971.
- Goran MJ, Williams JW, Gonella JS. The validity of patient management problems. *J Med Educ* 1973;48:171-7.
- Norman GR, Neufeld VR, Walsh A, Woodward CA, McConvey GA. Measuring physicians' performances by using simulated patients. *J Med Educ* 1985;60:925-34.
- Owen A, Winkler R. General practitioners and psychosocial problems: an evaluation using pseudopatients. *Med J Aust* 1974;2:393-8.
- Burri A, McCaughan K, Barrows HS. The feasibility of using the simulated patient as a means to evaluate clinical competence of practising physicians in a community. *Proc Annu Conf Res Med Educ* 1976;15:295.
- McClure CL, Gall EP, Meredith KE, Annel Gooden M, Boyer JT. Assessing clinical judgment with standardized patients. *J Fam Pract* 1985;20:457-64.
- Grol R, Mesker P, eds. *Huisarts en toetsing.* Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1984.
- Elstein AS, Shulman LS, Sprafka SA. *Medical problem solving: an analysis of clinical reasoning.* Cambridge, Massachusetts: Harvard University Press, 1978.

(Accepted 23 January 1987)

100 YEARS AGO

We all know that the love of quacks and quackery is deeply ingrained in the British mind, the British Islands being the happy hunting fields of the race. There are local newspapers in this kingdom that if cut off from the revenue they derive from quack advertisements, and the still more profitable source of emolument from quack "paragraphs," would find it impossible to exist. It appears, however, that Calcutta is not a whit behind us; that, on the contrary, quackery in its most homicidal form is rampant in that city. The subject, as we learn from the *Englishman*, one of the leading Calcutta newspapers, has been brought forward by Dr. Birch, in a very able manner, in a speech delivered at a meeting of the Calcutta Medical Society. The *Englishman*, commenting on Dr. Birch's speech, describes the system of medical practice in Calcutta "as one of permitted murder, and nothing else." In that city, the capital of British India, the seat of the Supreme Government, containing not only the largest number of educated Europeans of various nationalities, but also the home and centre of the various educational colleges which are acting on the minds of the higher native classes more and more every day, in that city, we say, there is not the shadow of any restriction on medical practice. As the article in the *Englishman*, to which we have referred, puts it, "Anyone can betake himself to the practice of prescribing medicines and attending patients in Calcutta; hundreds make a living in this way; men who have never received an hour's medical education or training in their lives. A signboard, a small room, a few jars, a lancet, or even an old penknife, this is all that is required to start with. The ignorance and superstition of the community do the rest." Dr. Birch, in his address, gives some terrible examples of the murderous use of drugs by ignorant and unlicensed practitioners. We give one or two of them in Dr. Birch's own words:

"Some time ago I was sent for by a medical man to see a patient, an old man, the head of a wealthy house, who lived in one of the most crowded thoroughfares in this city. I found him semi-collapsed, purging bloody serum copiously, vomiting bloody mucous, and suffering agonies. His abdomen was intensely tender, his mouth and fauces were white. We retired to consult. It was clear enough what the lesion was, and indications for treatment were almost as clear, but we lingered in the consultation-room, each not caring to express his doubts, till after a considerable time a paper was passed in to us headed, 'Further facts which have just come to our knowledge. Since the patient gave up the drink, he has been under the treatment of so-and-so, with gold and silver and other medicines which are reputed to be of a cooling nature.' We at once placed ourselves in communication with the members of the family, with the result that we were handed a box of large pills, which were of a scarlet colour and quite as heavy as lead; a paper containing a fawn-coloured powder and another packet containing a number of white powders. I sent the whole of them to my friend Dr. Warden, who lost no time in making an analysis of them. The red pills, that is the gold, were found to consist of sulphide of mercury—our old friend vermilion; the fawn-coloured powder was sulphide of tin; and the white powders, that is the silver, were simple carbonate of lime. The old man died the next day in great agony, and the self-styled 'medical practitioner' escaped, though I took some action to endeavour to bring him to justice. The culprit here was a native. Not very long ago I was consulted by an English lady, who lived in one of the best parts of Calcutta, for certain abdominal symptoms. I found she had previously been under the care of a well known physician, but, becoming impatient and hearing much of the skill of a certain

medical practitioner, she sent for him and submitted to his treatment, part of which consisted of an examination per vaginam. Tiring of him it came to my turn, and really it is hard to describe this poor lady's indignation and horror when she learnt from me that the individual who preceded me had never been for a single day a medical student! This I call one of the horrible cases. The culprit was *not* a native. Dr. Birch gives other instances equally horrible.

Professor D. B. Smith, formerly President of the Calcutta Medical College, now Professor of Military Medicine in the Army Medical School, relates a case occurring within his own experience, where a native quack gave strychnine under the impression that he was using santonin, with results which need not be described. The remedy—the partial remedy—that Dr. Birch proposes as at least a means of to some extent controlling this great evil, is the introduction of a Medical Registration Act for the town of Calcutta. It is with pleasure we observe that several of the native medical graduates of the Calcutta College supported Dr. Birch's proposition with much earnestness and ability. It is not pleasant to note that the speakers anticipated great opposition to the proposition, not only from those who have a vested interest in murder, but from the want of active support on the part of the authorities, who should be the most energetic in abating this terrible evil. We hope Dr. Birch will persevere in the face of opposition, come from what quarter it may, and that he will have the support of the medical profession in India, native and European, of the press, and of the more enlightened part of the native community, who in their own interest and that of their less fortunate fellow countrymen, should do their utmost to abate, if they cannot altogether put down, the murderous system described by Dr. Birch. (*British Medical Journal* 1887;ii:516.)

There are few consultants who have not felt the inconvenience of being without any ready and convenient statement of standard fees usually expected and accepted in performance of the ordinary duties of consulting practice. Young consultants, or those in smaller practice, are often timid in putting forward a scale which is not one of easy and general reference, and to which captious exceptions are sometimes taken by inexperienced and niggardly patients which they have no ready authority to answer. It is always inconvenient when any discussion of an individual or personal character arises on such points. Consultants in large practice lose much time, and have to sacrifice dignity in dealing individually and, as it were, *de novo* with each question of the kind. A consultant of the highest eminence addresses a letter to us this week which conveys what will be felt by all to be really useful detailed information on the subject, and puts out for consideration two alternative documents. He desires, and in that desire we feel sure that the profession generally will concur, that opinions should be gathered now in a form to settle an acknowledged basis which will meet with general acceptance. The source of this communication guarantees that it has already had consideration in the highest quarters. It is now put forward for comment and debate, and we commend it to the general attention of consultants. We shall be glad to receive and publish communications on the subject, whether *sub nomine* or confidential. We are inclined to think that the second tariff is that most likely to meet unanimous approval; but the whole matter is open to the freest discussion, and we express at present no opinion. (*British Medical Journal* 1887;iii:1057)