Analysis of referral behaviour: responses to simulated case histories may not reflect real clinical behaviour

D C MORRELL

M O ROLAND

SUMMARY. In an attempt to develop a measure of the referral behaviour of general practitioners, 21 simulated case histories were constructed which presented between three and six stages at which a decision was required whether or not to refer a 'patient' to hospital. Twenty general practitioners completed the case histories and their responses were compared with their referral rates to outpatient departments. No significant correlations were found between the doctors' responses to the vignettes and their actual referral rates, and the repeatability of the vignettes was disappointing. The fact that the case histories appeared realistic to the doctors who completed them was not reflected in objective measures of their validity. Simulated case histories should not be used as a method of measuring doctors' behaviour without establishing their validity and reliability.

Introduction

ONE of the problems in carrying out research into the ways in which doctors make clinical decisions is that it is often difficult to compare doctors' performance in real life situations. One way round this problem is to construct simulated case histories on paper. If the responses of doctors to these 'paper patients' reflect their behaviour in clinical practice, then simulated case histories or vignettes can be a valuable research tool.

Performance on simulated case histories has been claimed to reflect real clinical behaviour in a number of situations. Kirwan, for example, found that assessment of disease activity in rheumatoid arthritis in simulated case histories closely reflected judgements made in clinical practice.[1,2] Indeed, he found that the responses to simulated cases appeared to correlate better with doctors' actual behaviour than doctors' descriptions of their own management policies.[3]

In a study of simulated case histories of patients with respiratory disease, Howie found overall agreement between the examinations which general practitioners said they would carry out on the simulations, and those which they carried out in clinical practice.[4] Good agreement on management and prescribing decisions was claimed by Chaput de Saintonage in a study comparing paper simulations and real patients with otitis media.[5] However, the actual levels of agreement quoted in a subsequent paper[6] were quite low, and some participants expressed doubts about the validity of the simulations.

D C Morrell, FRCGP, Wolfson professor of general practice, United Medical and Dental Schools of Guy's and St Thomas's, London. M O Roland, DM, director of studies in general practice, University of Cambridge School of Clinical Medicine.
Submitted: 17 March 1989; accepted: 9 November 1989.

In other situations, doubts have been expressed about the use of simulated case histories as a valid reflection of clinical practice. Page and Fielding found that community pharmacists responded differently to paper simulations than to actors who portrayed the same simulations incognito.[7] Goran and colleagues found that doctors collected more clinical information about a single simulated case history relating to urinary infection, and were more likely to order investigations than they did when faced with similar patients in real life.[8] Norman and Feightner found that the responses of medical students were different depending on whether the simulation was on paper or presented by an actor,[9] and Newble and colleagues suggested that the poor performance on paper simulations of junior doctors compared with medical students was evidence against the validity of the simulations.[10]

In relation to referral behaviour, Kuyvenhoven and colleagues[11-13] found a correlation between general practitioners' performance on three simulated patients and referral rates. However, these papers are difficult to interpret because the behaviour which correlated with referral behaviour was a global index of tendency to 'risk causing unnecessary harm' to the simulated patients, and the method by which this was assessed appears subjective and is not clearly described.

It is technically difficult to collect reliable data on the referral rates of a representative sample of general practitioners, and it would therefore be useful if it were possible to construct a tool which could be used as a valid indirect measure of referral behaviour. This paper reports an attempt to validate such a tool. We developed simulated case histories with several stages at which a decision was required whether or not to refer the patient to hospital. These were designed to measure the likelihood that a general practitioner would refer a patient to hospital in a given clinical situation. We looked for an association between responses to these simulated case histories and actual referral behaviour in a group of general practitioners with known rates of referral to hospital.

Method

Vignettes

A series of simulated case histories (vignettes) were constructed in the following manner. Each vignette consisted of a series of clinical descriptions, ranging from 15 to 50 words. The layout of the vignettes was such that the responder could only see one part of the case history at a time. After reading the first clinical description, the doctors were asked to say whether or not they would refer the patient to hospital. If a doctor marked 'yes', he or she completed no further parts of that vignette, and moved on to the next one. If he decided not to refer the patient at that point, he tore off a strip of coloured paper to reveal the next brief clinical description. He was then asked again whether or not he would refer the patient to hospital. The total number of steps in the vignettes ranged from three to six. The vignettes covered a wide range of medical problems. In some vignettes the 'patient' developed progressively more serious symptoms, signs or investigation results as the case history unfolded. In

others, the doctor completing the exercise was exposed to increasing pressure from the 'patient' for referral.

A total of 23 vignettes were originally constructed. In pilot studies it became apparent that the responses of doctors to two of the vignettes were almost identical (for example a child who fairly clearly had meningitis on the first clinical description), and these vignettes were omitted from the main study. In two further vignettes, small alterations were made to the text after the pilot studies in order to increase the complexity of decision required (for example a man of 60 years old with severe central chest pain was aged by 10 years to make him 70 years old). The results reported in this paper therefore relate to 21 vignettes, which contained a mean of around four steps at which a decision about referral was required. A typical vignette is shown in Figure 1.

The vignettes were scored by adding up the number of steps at which a doctor decided not to refer a 'patient', and subtracting this from the total number of steps in those vignettes. This gave a scale in which doctors who tended to refer 'patients' early received high scores and vice versa.

In addition to giving each doctor a score on all 21 vignettes, three subscales were defined. These were: (1) the uncertainty score, which was the score on those seven vignettes which were characterized by substantial diagnostic uncertainty; (2) the life-threatening illness score, which was the score on those 11 vignettes where there was a clear possibility of life-threatening illness; and (3) the gatekeeper role score, which was the score on those four vignettes where the doctor was under some pressure from the 'patient' to refer illness which was unlikely to be serious, but might be associated with discomfort or disability (one vignette covered two areas). These subscales were defined by the authors of the vignettes who independently made judgements about which of the 21 vignettes were characterized particularly by diagnostic uncertainty, by the possibility of life-threatening illness, and by patient pressure for 'something to be done'. The authors were in almost complete agreement about which vignettes should contribute to each of the subscales, and they were defined accordingly.

Mrs M is 53. Her last normal period was six months ago. One week ago she had what appeared to be a period. It was preceded by some swelling and tenderness in her breasts. The bleeding was not preceded by sexual intercourse. Pelvic examination is normal and a cervical smear is taken which is subsequently reported as normal.

DO YOU REFER?

No Yes

Three months later she has a further episode of vaginal bleeding, lasting five days, which follows sexual intercourse. Vaginal examination again reveals no abnormality

DO YOU REFER?

No Yes

Six months later she has another episode of vaginal bleeding 'like a period'. Again vaginal examination is unrewarding

DO YOU REFER?

No Yes

One year later Mrs M has had no further trouble apart from the fact that she is experiencing hot flushes and some vaginal dryness on intercourse

Figure 1. *Example of vignette to test referral behaviour (the doctor is only able to see one part of the vignette at a time).*

Participants

Twenty doctors completed the 21 vignettes. Thirteen doctors were principals in four Cambridge group practices, five were principals in one south London group practice and two were trainees in the London practice. Nineteen of the 20 doctors completed the vignettes approximately one year later in order to measure the repeatability of the simulations.

Referral rates

Two different methods were used to measure the actual referral rates of the 20 participating doctors. Nine of the Cambridge doctors worked in a practice where a separate file was kept of all typed referral letters. The number of new outpatient referral letters was counted over 12 months. The great majority of referral letters from these doctors are typed, but no account could be made of handwritten referral letters. In the remaining practices, details of numbers of outpatient referrals were collected prospectively over periods ranging from six to 12 months. Referral rates were calculated by dividing the number of referrals made by a doctor by the number of surgery consultations carried out by that doctor during the specified time period. Home visits were not included in the denominator as figures on visits were not available for two of the practices.

In analyses involving correlation coefficients, a logarithmic transformation of referral rates was carried out as the distribution of referral rates was positively skewed.

Results

Without exception, the general practitioners found the vignettes enjoyable to complete, and many commented that they seemed realistic. In a few cases, general practitioners commented that they had difficulty in completing a vignette because the course of action taken by the 'doctor' in the vignette was not what they personally would have chosen to do in that clinical situation. There was a spread of scores on the vignettes which indicated at least a two-fold variation in the doctors' behaviour as judged by overall scores and scores on the three subscales (Table 1). However, there was a much larger, four-fold variation in the referral rates of the same doctors (Figure 2).

There was no significant association between the doctor's referral rates and either overall vignette scores or scores on any of the three subscales. The Pearson correlation between log transformed referral rate and overall vignette score was only 0.17 (95% confidence interval − 0.33 to + 0.67) and correlations

Table 1. Range of overall scores on the vignettes and subscale scores for 20 general practitioners.

	Overall score	Uncertainty score	Life-threatening score	Gatekeeper role score
Mean score	37.9	13.7	25.1	7.0
Range	24–48	9–21	18–35	2–11
Maximum possible range	0–107	0–36	0–58	0–21
Pearson correlation coefficients between vignette scores and log referral rate	0.17	0.17	0.19	0.07

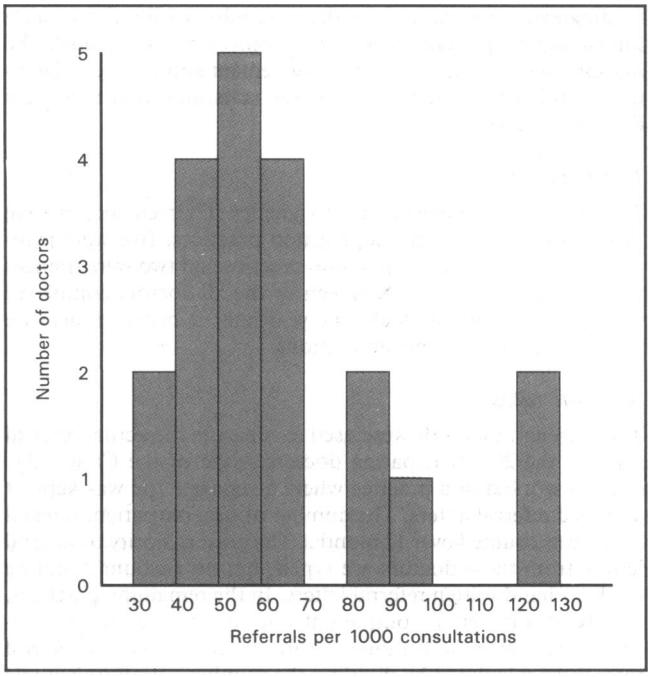

Figure 2. *Range of referral rates.*

for the subscales ranged from 0.07 to 0.19 (Table 1). Thus, even though the sample size in this study was relatively small, the probability of vignette scores explaining a substantial proportion of the variance in referral rates was very small.

The repeatability of the vignettes was disappointing. The mean difference in scores between the two sets of vignettes completed 12 months apart was only 1.6. However, the standard deviation of the difference was 5.1 and individual general practitioners' scores varied by up to 12 points between the two sets of vignettes. The correlation coefficient between the two sets of scores was only 0.69, indicating that less than half the variance of one set of scores could be explained by scores on the other set.

Discussion

Research on a number of aspects of general practice would be greatly simplified if the behaviour of general practitioners could be measured in simulated clinical situations rather than in real life. Referral behaviour is only one example of this. In designing the vignettes, we reasoned that if they were sufficiently realistic, doctors' responses on paper might reflect the actual decisions they made in clinical practice. Therefore, their responses to a series of decisions about hospital referral would be related to the actual rate at which they referred patients to hospital. However, the results of this study lend no support to the idea of using paper simulations for this purpose. There appeared to be virtually no correlation between what the doctors said they would do and what they actually did.

There are a number of possible explanations for this. First, it could be that the construction of the vignettes was faulty. Had more detailed information been given, for example by using photographs or videotape, it could be that the doctors' responses to the vignettes would more closely have represented what they do in real life. In the study by Kuyvenhoven and colleagues,[11] each simulated case history contained around 1000 questions and answers which the doctors could access, a much greater number than in the present study. Alternatively, it could be that the fault was defining the information available to the doctor. It is known (in simulations at any rate) that doctors vary greatly in the amount of clinical information which they use in making decisions.[14,15] A format which allowed the doctors to select information on which to base their referral decisions might have been more closely related to real life. The fact that the doctors in this study said that the vignettes felt true to life when completing them was clearly not a guarantee of their validity.

A second possibility is that the doctors chosen for this study were too similar in their referral behaviour. Had it been possible to include doctors with referral rates which differed by a factor of more than four, differences in behaviour on the vignettes might have been apparent. The referral rates of the doctors in this study were relatively low compared with those quoted in national studies,[16] which would also suggest that they were atypical. A further problem is that the case mix of problems in the vignettes might not have adequately reflected the case mix in the doctors' own practices.

Thirdly, it could be that the data in this study were collected during periods which did not adequately reflect the doctors' real referral behaviour. However, there is said in general to be consistency of doctors' referral behaviour over time,[17] and the periods of data collection in this study were sufficiently long for the confidence limits of individual doctors' referral rates to be narrow in relation to the observed range of referral rates. There are therefore a number of possible explanations for the lack of validity of the vignettes in this study. However, equally disturbing was the poor repeatability of the vignette scores.

The studies quoted in this paper which have claimed to produce valid simulated cases have, on the whole been dealing with well circumscribed clinical situations, for example which examinations to carry out on patients with respiratory symptoms,[4] whether to prescribe for otitis media,[5] and an index of severity in rheumatoid arthritis.[1] The process of making a decision whether or not to refer a patient to hospital is much more complex than these, and is likely to depend on factors such as the availability of services, the clinical state of the patient, the clarity of the diagnosis and pressure from the patient. It could be that our attempts to include all these factors in short vignettes was too simplistic.

The fact remains that, whatever the explanation, the vignettes used in this study were neither valid nor reliable, despite considerable care in constructing and piloting them. The doctors found the clinical situations with which they were presented realistic, but this should certainly not be taken as an indication of validity in future studies of this type as we found no evidence that responses to simulated case histories reflected real clinical behaviour. Our results suggest that investigators wishing to make use of simulated case histories should not do so without careful measurement of their validity and reliability.

References

1. Kirwan JR, Chaput de Saintonge DM, Joyce CRB, Currey HLF. Clinical judgement in rheumatoid arthritis. 1. Rheumatologists' opinions and the development of 'paper patients'. *Ann Rheum Dis* 1983; **42:** 644-647.
2. Kirwan JR, Bellamy N, Condon H, *et al.* Judging current disease activity in rheumatoid arthritis — an international comparison. *J Rheumatol* 1983; **10:** 901-905.
3. Kirwan JR, Chaput de Saintonge DM, Joyce CRB, *et al.* Inability of rheumatologists to describe their policy for assessing rheumatoid arthritis. *Ann Rheum Dis* 1986; **45:** 156-161.
4. Howie JGR. Further observations on diagnosis and management of general practice respiratory illness using simulated patient consultations. *Br Med J* 1974; **2:** 540-543.
5. Chaput de Saintonge DM, Hathaway NR. Antibiotic use in otitis media: patient simulations as an aid to audit. *Br Med J* 1981; **283:** 883-885.
6. Chaput de Saintonge DM, Hattersley LA. Antibiotics for otitis media — can we help doctors agree? *Fam Pract* 1985; **2:** 205-212.

7. Page GG, Fielding DW. Performance on PMPs and performance in practice: are they related? *J Med Educ* 1980; **55:** 529-537.
8. Goran MJ, Williamson JW, Gonnella JSA. The validity of patient management problems. *J Med Educ* 1973; **48:** 171-177.
9. Norman GR, Feightner JW. A comparison of behaviour on simulated patients and patient management problems. *Med Educ* 1981; **15:** 26-32.
10. Newble DI, Hoare J, Baxter A. Patient management problems: issues of validity. *Med Educ* 1982; **16:** 137-142.
11. Kuyvenhoven MM, Jacobs HM, Touw-Otten FWMM, Van Es JC. Written simulation of doctor-patient encounters. 1. Research instrument for registration of the performance of general practitioners. *Fam Pract* 1984; **1:** 14-19.
12. Jacobs HM, Kuyvenhoven MM, Touw-Otten FWMM, Van Es JC. Written simulation of doctor-patient encounters. 2. Assessment of the performance of general practitioners. *Fam Pract* 1984; **1:** 20-24.
13. Kuyvenhoven MM, Jacobs HM, Touw-Otten FWMM, Van Es JC. Written simulation of doctor-patient encounters. 3. Comparison of the performance in the simulation with prescription and referral data in reality. *Fam Pract* 1984; **1:** 25-29.
14. Taylor TR, Aitchison J, McGirr EM. Doctors as decision makers: a computer assisted study of diagnosis as a cognitive skill. *Br Med J* 1971; **3:** 35-40.
15. Jones RH. Data collection in decision making: a study in general practice. *Med Educ* 1987; **21:** 99-104.
16. Office of Population Censuses and Surveys, Department of Health and Social Security and Royal College of General Practitioners. *Morbidity statistics in general practice. Third national study 1981-82.* London: HMSO, 1986.
17. Crombie DL, Fleming DM. General practitioner referrals to hospital: the financial implications of variability. *Health trends* 1988; **20:** 53-56.

Acknowledgements

We would like to thank members of the Fenland research club and the doctors of the Lambeth Road group practice for their help in completing the vignettes.

Address for correspondence

Professor D C Morrell, United Medical and Dental Schools of Guy's and St Thomas's Hospitals, Department of General Practice, 80 Kennington Road, London SE11 6SP.

INFORMATION FOR AUTHORS AND READERS

Papers submitted for publication should not have been published before or be currently submitted to any other journal. They should be typed, on one side of the paper only, in double spacing and with generous margins. A4 is preferred paper size. The first page should contain the title only. To assist in sending out papers blind to referees, the name(s) of author(s), degrees, position, town of residence, address for correspondence and acknowledgements should be on a sheet separate from the main text.

Original articles should normally be no longer than 4000 words, arranged in the usual order of summary, introduction, method, results, discussion and references. Letters to the Editor should be brief — 400 words maximum — and should be typed in double spacing.

Illustrations of all kinds, including photographs, are welcomed. Graphs and other line drawings need not be submitted as finished artwork — rough drawings are sufficient, provided they are clear and adequately annotated.

Metric units, SI units and the 24-hour clock are preferred. Numerals up to 10 should be spelt, 10 and over as figures. Use the approved names of drugs, though proprietary names may follow in brackets. Avoid abbreviations.

References should be in the Vancouver style as used in the *Journal*. Their accuracy must be checked before submission. The title page, figures, tables, legends and references should all be on separate sheets of paper.

Three copies of each article should be submitted, with a small stamped addressed envelope (for acknowledgement), and the author should keep a copy. One copy will be returned if the paper is rejected.

All articles and letters are subject to editing.

Papers are refereed before a decision is made.

Correspondence and enquiries to the Editor
All correspondence to the Editor should be addressed to: The British Journal of General Practice, Royal College of General Practitioners, 12 Queen Street, Edinburgh EH2 1JE. Telephone (office hours; 24 hour answering service): 031-225 7629. Fax (24 hours): 031-220 6750.

Copyright
Copyright of all material in the *Journal* is vested in the *Journal* itself. However, authors may use minor parts (up to 15%) of their own work after publication without seeking written permission provided they acknowledge the original source. The *Journal* would, however, be grateful to receive notice of when and where such material has been reproduced. Authors may not reproduce substantial parts of their own material without written consent. However, requests to reproduce material are welcomed and consent is usually given. Individuals may photocopy articles for educational purposes without obtaining permission up to a maximum of 25 copies in total over any period of time. Permission should be sought from the Editor to reproduce an article for any other purpose.

Advertising enquiries
Display and classified advertising enquiries should be addressed to: Iain McGhie and Associates, 7a Portland Road, Hythe, Kent CT21 6EG. Telephone 0303 264803/262272. Fax: 0303 262269.

Circulation and subscriptions
The *British Journal of General Practice* is published monthly and is circulated to all Fellows, Members and Associates of the Royal College of General Practitioners, and to private subscribers. All subscribers receive *Policy statements* and *Reports from general practice* free of charge with the *Journal* when these are published. The annual subscription is £80 post free (£90 outside the UK, £100 by air mail). Non-members' subscription enquiries should be made to: Bailey Bros and Swinfen Ltd, Warner House, Folkestone, Kent CT19 6PH. Telephone: Folkestone (0303) 850501. Members' enquiries should continue to be made to: The Royal College of General Practitioners, 14 Princes Gate, Hyde Park, London SW7 1PU. Telephone: 071-581 3232.

Notice to readers
Opinions expressed in the *British Journal of General Practice* and the supplements should not be taken to represent the policy of the Royal College of General Practitioners unless this is specifically stated.

RCGP Connection
Correspondence concerning the news magazine, *RCGP Connection*, should be addressed to: RCGP Connection Editor, Royal College of General Practitioners, 14 Princes Gate, Hyde Park, London SW7 1PU. Telephone: 071-581 3232.