Written Simulation of Patient–Doctor Encounters. 3. Comparison of the Performance in the Simulation with Prescription and Referral Data in Reality

M M KUYVENHOVEN, H M JACOBS, F W M M TOUW-OTTEN AND J C VAN ES

Kuyvenhoven M M (Department of General Practice, University of Utrecht, Mariahoek 6, 3511 LD Utrecht, The Netherlands), Jacobs H M, Touw-Otten F W M M and van Es J C. Written simulation of patient-doctor encounters. 3. Comparison of the performance in the simulation with prescription and referral data in reality. *Family Practice* 1984, 1:

The performance of 19 general practitioners in a written simulation of patient-doctor encounters[1] was compared with real data on prescription and referral from 'sick-fund' patients. Two issues were studied: the amounts of certain drugs prescribed (analgesics and antirheumatics; hypnotics/sedatives and tranquillizers; neuroleptics and antidepressants) and the number of referrals to specialists. In the simulation, experts assessed the 'risk of unnecessary harm to the patient' induced by the therapeutic procedures of these general practitioners, as described previously.[2] A close correlation was found between the ratings of the risk of unnecessary harm in the simulation and the prescription and referral data in reality for general practitioners who were consistently generous or frugal in prescribing the above drugs and referring the patients to specialists. The tentative conclusion is that the written simulation with its rating procedure discriminates reliably between general practitioners who induce less risk of unnecessary harm from those who are likely to cause harm to the patient.

Two previous articles have described a research instrument for registration and comparison of the performance of general practitioners confronted with vague complaints.[1,2] This consisted of a written simulation of patient-doctor encounters and an assessment by experts of two aspects of performance: the attention paid to somatic aspects or causes of the complaint, and the risk of inducing unnecessary harm to the patient.

In a test with 19 general practitioners the content validity of the simulation and the reliability of the rating procedures were found to be high.[1,2] Both are important parameters in the assessment of the validity of the assumptions of the research technique.[3] A third parameter is the comparison of the performance in the simulation with that in reality. Having no valid counterpart this comparison had to be made indirectly by resorting to 'sick-fund' prescription and referral data. One might expect there to be a rational basis for not only the prescription of drugs[4] but also the referral of patients to specialists. In the view of Mapes and Williams[5] 'the profession responds more widely to reports on the good effects of a drug than it does to its adverse effects' and this may also apply to the results obtained by medical specialists. Both referral to a medical specialist and the prescription of a drug can induce untoward results: for example, side effects of the drug, lesions caused by invasive clinical examinations, anxiety in patients and sick role fulfilment. The advantages to the patient do not always outweigh the disadvantages—it is often 'a matter of weighing chances'[6]

Department of General Practice, University of Utrecht, Mariahoek 6, 3511 DL Utrecht, The Netherlands.

—and general practitioners differ in their assessment of the benefits and risks.

This article considers whether the practitioners confronted with vague complaints in the simulation induced as much (or as little) risk of unnecessary harm as they did in their normal prescription and referral behaviour. To attempt to answer this question, real prescription and referral data were selected which reflected the extent to which general practitioners induced a risk of unnecessary harm to patients; these data were obtained from 'sick-fund' patients (all citizens and their dependants whose annual income is below about £10 500). They give a good example of the prescription and referral behaviour of Dutch general practitioners because some 70% of their patients are sick-fund patients.

A risk of unnecessary harm to the patient is induced by a general practitioner who prescribes drugs where the indication is either arbitrary or undetermined and which may induce side effects or adverse effects in chronic use. This risk often applies to analgesics, antirheumatics, hypnotics, sedatives, tranquillizers, neuroleptics and antidepressant drugs.

Referral to a specialist can also induce a risk of unnecessary harm. Many complaints—such as gastric and abdominal pain, dyspnoea, arthralgia —have no readily definable pathological origin and a first examination by a specialist may reveal no definite causes. In this situation specialists often continue the investigation with more invasive methods, which may entail both inconvenience and risk of harm to the patient.

METHOD

Prescription and Referral Data

The following data from the sick-funds was analysed:

a. The amount of drug prescribed per sick-fund patient for the three categories of drugs: analgesics and antirheumatics; hypnotics/sedatives and tranquillizers; neuroleptics and antidepressants.

b. The number of referrals per 1000 sick-fund patients to specialists in: internal medicine, cardiology, gastroenterology and metabolic disorders, lung diseases, rheumatology.

Collectively, these data may be postulated as giving a good indication of the frequency with which general practitioners make therapeutic decisions which entail a risk of unnecessary harm. It can be assumed, therefore, that general practitioners who show restraint in prescribing and referring induce less risk of unnecessary harm to the patient than those who prescribe generously and refer frequently. It is impossible, however, to determine from these data whether the risk in prescribing analgesics and antirheumatics exceeds that in prescribing, say, neuroleptics and antidepressants. Nor is it possible to determine whether the risk in prescribing drugs of all three categories together exceeds that in referring patients to medical specialists. In other words, the weight to be assigned to the separate data as measurements of the 'risk of unnecessary harm' remains unknown.

Analysis of Prescription and Referral Data

Prescription and referral data were taken from the same four municipalities in which the 19 general practitioners who formed the test group in the simulation practised. One municipality was eliminated (see later) but the data from virtually all general practitioners in the remaining three municipalities was used (number of practitioners, $n_1 = 59$, $n_2 = 24$, $n_3 = 12$).

The analysis was in three stages:

1. To establish whether the selected prescription and referral data were stable over a period of time the correlations between the data for the years 1977 and 1978 was calculated. The strength of the correlations was expressed using Pearson's product-moment correlations;[7] when these correlations are high, the prescription and referral data are stable in time.
2. The correlation between the prescription and the referral data was examined in order to determine how consistent the general practitioners were in prescribing the three categories of drugs and in referring the patients to medical specialists. The correlation was described with the aid of Spearman's rank correlation coefficients.[7] When the rank correlations between, say, the prescribing of analgesics and antirheumatics and that of hypnotics/sedatives and tranquillizers is high, then the doctors involved are consistent in prescribing drugs of these two categories.

The strength of the correlation between the collective prescription data was calculated with the aid of Kendall's coefficient of concordance W.[7] This expresses the correlation between distinct rank orderings. A high coefficient W

indicates that the doctors involved are consistent in the amounts of drugs they prescribe in the three categories. The same applies to the collective prescription and referral data.

3. The general practitioners in each municipality were classified into three groups:

> Consistent: practitioners who seldom prescribe the drugs in the three categories or refer patients to medical specialists, and those who frequently prescribe drugs or refer patients. The former are consistently 'frugal' while the latter are consistently 'generous'. These doctors (by definition) show a strong positive correlation between the prescription and referral data.
>
> Alternating: practitioners who are generous in prescribing drugs of the three categories but frugal in referring patients to medical specialists, or vice versa.
>
> Inconsistent: practitioners who are generous with drugs of one category, modest with drugs of a second, and frugal with drugs of a third category. The correlation between the prescription and the referral data for these doctors cannot be interpreted unequivocally.

Comparison between Performance in Reality and in the Simulation

Comparison of performance in the simulation with that in reality was confined to the general practitioners defined as 'consistent' by the above criteria. Consistently frugal general practitioners induce less risk of unnecessary harm than do consistently generous general practitioners. The extent to which 'alternating' and 'inconsistent' practitioners induce a risk of unnecessary harm cannot be determined, because it is impossible to determine the relationship between performance in the simulation and prescription/referral data in reality.

In the simulation the rating of induction of a risk of unnecessary harm is the sum of the expert ratings of the 'risk of unnecessary harm' with three of the patients.[2] For the comparison, the places of these general practitioners in the quartile distributions in their municipality (with regard to prescription and referral data) was considered. The comparison determines whether there is a high correlation between prescription/referral data and the ratings of the risk of unnecessary harm in the simulation and also whether this correlation is as strong as the (by definition) strong correlation between the prescription and referral data. Since a correlation between different measurements of variables was considered here, Kendall's coefficient of concordance W was used again to express the strength of the correlation.

RESULTS

Analysis of Prescription and Referral Data of the General Practitioners

1. The correlation coefficients of the selected prescription and referral data were found to be stable ($0.67 < r < 0.95$), therefore further analysis was confined to the data for one year, 1978.

2. The correlation between the prescription and referral data in the three municipalities was complex. There was a strong correlation between prescription of analgesics and antirheumatics on the one hand, and of hypnotics/sedatives and tranquillizers on the other. This is to say that many general practitioners were consistently either generous or frugal in prescribing drugs of these two categories. In two of the three municipalities this consistency extended to the prescription of neuroleptics and antidepressants and to referrals to medical specialists. In the third municipality the consistency in prescribing drugs of the two categories extended slightly to referrals to medical specialists but not to prescription of drugs of the third category.

3. The classification into 'consistent', 'alternating' and 'inconsistent' general practitioners showed that most general practitioners (75%) prescribed drugs consistently; there were a few (15%) practitioners who prescribed inconsistently.

In the two larger municipalities about half the general practitioners were both consistent in prescribing and in referring to medical specialists. In the smaller municipality this applied to virtually all general practitioners.

There were no indications of a correlation between generous or frugal prescribing of the three types of drugs and 'the number of sick-fund patients' or 'solo practice' or 'group practice'. The same applies to referrals to medical specialists. These findings agree with those of Taylor, who in a study of 14 general practitioners in 1974 likewise found no correlation between prescribing costs on the one hand, and practice characteristics on the other.[8]

Correlation between Performance in Reality and in the Simulation

The prescription and referral data from one

municipality proved to be useless, consequently three general practitioners were eliminated from the test group for the comparison between simulation and reality. The consistency of the remaining 16 general practitioners of the test group were classified; nine were 'consistent', six were 'alternating' and one was 'inconsistent'.

For the nine consistent doctors a strong correlation was found between the ratings of the risk of unnecessary harm and the prescription/referral data ($r=0.62$). The strength of the correlation roughly equalled that of the correlation between the prescription and referral indicators ($r=0.69$). The separate rank correlations (see Table 1) revealed that this correlation was based on positive correlations between the separate data and the ratings of the risk of unnecessary harm. This means that the practitioners who in reality were consistently frugal (or generous) in prescribing the harmful drugs and also frugal (or generous) in referring patients to medical specialists, indeed induced little (or much) risk of unnecessary harm in the simulation.

The two variables of performance—behaviour in the simulation and prescription/referral data in reality—indicate a common denominator; it is a plausible hypothesis that this common denominator is the induction of a risk of unnecessary harm. After all, the two variables were designed independently of each other to measure the risk of causing unnecessary harm.

It was useful to discriminate between 'consistent' and 'alternating' general practitioners, as demonstrated by the rank correlations between the ratings of the risk of harm and the separate prescription and referral data (Table 1). The consistent doctors showed a positive correlation between the risk of harm in the simulation and that in reality. For the alternating practitioners no such correlation was discerned.

DISCUSSION

The correlation between prescription and referral data in reality and the ratings of the risk of causing unnecessary harm in the simulation supplements the previous data on content validity[1] and on the reliability of the rating procedures.[2] It can be concluded with some caution that the research technique developed is a reliable method for discriminating between general practitioners who are more likely to cause harm to their patients and those who are less likely to induce unnecessary harm. The conclusion is presented cautiously because the data refers to only a small number of general practitioners: three of the 19 practitioners were eliminated from the test group because for them no reliable prescription and referral data were available. Moreover, the comparison between simulation and reality was confined to the nine 'consistent' general practitioners.

The classification into consistent, alternating and inconsistent general practitioners is a faithful reflection of the numerical proportions between these three groups. Selection of general practitioners on this the basis, however, would unfortunately pose privacy problems.

Further research will be needed to show whether the separate prescription and referral data are valid indicators of the extent to which doctors risk causing unnecessary harm to their patients. This can be established only in cooperation with medical specialists and pharmaceutical chemists.

The comparison described in this article was

TABLE 1 *The rank correlations between the 'risk of unnecessary harm' ratings and the prescription and referral data for consistent and alternating general practitioners*

Factors causing a risk of unnecessary harm	Consistent and alternating general practitioners r values	Consistent general practitioners r values	Alternating general practitioners r values
Analgesics and antirheumatics	0.17	0.43	−0.07
Hypnotics/sedatives and tranquillizers	0.09	0.52	−0.28
Neuroleptics and antidepressants	−0.11	0.27	−0.55
Referrals to medical specialists per 1000 sick-fund patients	0.39	0.57	0.27

r values expressed in Spearman's rank correlation coefficients.

not aimed at determining an absolute similarity between the performance in the simulation and that in reality. We know from the studies of Goran et al.[9] and Norman and Feightner[10] that it would be unrealistic to expect such a similarity.

The correlation between prescription/referral data and the risk of unnecessary harm rated in the simulation implies that the relevant differences in real performance are also apparent in the simulation and this finding corroborates the construct validity of the research instrument we have developed.[11,12]

The conclusion from the results of the three validation steps is that the diagnostic and therapeutic procedures of the general practitioners in the simulation showed little distortion in comparison with reality, at least insofar as a general practitioner's performance was more directive than non-directive, and more complaint-oriented than person-oriented. It is precisely these directive and complaint-oriented facets of a general practitioner's performance which demonstrate how he examines and treats somatic aspects of a complaint and induces risks with his strategy.

ACKNOWLEDGEMENTS

We owe a debt of gratitude to Dr C. Spreeuwenberg, general practitioner and associate of the Department, for his valuable and inspiring help in the planning and writing of this article. We also thank M. M. van Nunen, general practitioner in Hoensbroek, for his contribution to the planning and execution of the analysis, and to Mrs E. M. A. van Hijfte, undergraduate in pharmaceutical chemistry, Dr I. Ph. L. Koperoerg, general practitioner (both former associates of the Department), and Prof Dr R. A. de Melker, professor of general practice at the Department, for their contribution to the analysis of the prescription and referral data.

REFERENCES

1 Kuyvenhoven M M, Jacobs H M, Touw-Otten F W M M, van Es J C. Written simulation of patient-doctor encounters. 1. Research instrument for registration of the performance of general practitioners. *Fam Practice* 1984; 1: (this issue).

2 Jacobs H M, Kuyvenhoven M M, Touw-Otten F W M M, van Es J C. Written simulation of doctor-patient encounters. 2. Assessment of the performance of general practitioners. *Fam Practice* 1984; 1: (this issue).

3 Groot A D de. Methodology. Foundations of inference and research in the behavioural sciences. The Hague: Mouton, 1969.

4 Editorial. *N Engl J Med* 1969; 280: 161-162.

5 Mapes R E, Williams W O. The changing pattern of general practitioner drug prescribing in the National Health Service in England from 1970 to 1975. *J R Coll Gen Pract* 1979; 29: 406-412.

6 Lubsen J. Het therapeutische experiment: effectief of niet, dat is de vraag. In: Querido A, en Roos J, eds. Controversen in de Geneeskunde. Utrecht: Wetenschappelijke uitgeveij Bunge, 1980.

7 Hays W L. Statistics for the social sciences. London: Holt, Rinehart, Winston, 1977.

8 Taylor R J. Prescribing costs and patterns of prescribing in general practice. *J R Coll Gen Pract* 1978; 28: 531-535.

9 Goran M J, Williamson J W, Gonella J S. The validity of patient management problems. *J Med Educ* 1973; 48: 171-177.

10 Norman G R, Feightner J W. A comparison of behaviour on simulated patients and patient management problems. *Med Educ* 1981; 15: 26-32.

11 Inbar M. Toward valid computer simulations of bureaucratized decisions. *Simulation and Games* 1976; 7: 243-260.

12 Gerritsma J G M, en Smal J A. Schriftelijke patiëntensimulaties voor onderwijs, evaluaties en onderzoek. *Medisch Contact* 1979; 34: 1217-1225.