

Commentary: How robust are rankings? The implications of confidence intervals

Colin Sanderson, Martin McKee

Even if data were immaculate and risk adjustment perfect, performance indicators based on the numbers of “outcome events” commonly found in NHS institutions would still be vulnerable to the play of chance. Provision of confidence intervals draws attention to this. The authors of this paper have derived confidence intervals for performance rankings and show that league tables suffer from similar problems. The implication is that in 1993-4 the success rate of the in vitro fertilisation clinic at Bourne Hall, with a substantial 1315 treatment cycles, could actually be anywhere between 15th and 36th out of the 52 clinics examined. Newham General, with only 68 cycles, is ranked near the middle, but its place is consistent with a “true” rank of anywhere between 3rd from the bottom and 3rd from the top.

The technical interest of this paper lies in the method used to calculate these 95% confidence intervals for ranks. This was done by a process known as the Monte Carlo technique—the use of *sampling experiments* based on random numbers. This technique was originally developed by mathematicians interested in “random walks,” legendarily characterised as how far will the drunk be from the lamppost after a given number of irregular zigzags. It was taken up by physicists, and by operational researchers investigating complex queuing systems. Now its use by statisticians, as a way of deriving confidence intervals when they are not available from theory, is on the increase.

How was the technique used here? The starting point is that the success rate observed for clinic X is consistent with a range of “underlying” values, some more plausible than others. The relative plausibility of each value is characterised by a distribution. A random number is then used to “sample” from this distribution, and the resulting value is clinic X’s “simulated” underlying success rate. Repeating this for all the other clinics provides a *set* of simulated success rates and hence a simulated ranking for each clinic. The process is repeated with a second set of random numbers, generating a new set of ranks, and repeated again a large number of times. A distribution of plausible “underlying” ranks for each clinic is gradually built up, from which confidence intervals can be derived.

Why use league tables at all? The main advantage is that they are easy to read. One can see at a glance who is at the top and who is at the bottom. But if the information is both high impact and misleading, poor decisions are made and the source loses credibility. If tables are to be published it may well be better to order the entries on some other basis than indicated performance—geography or case mix perhaps. Each row should include the institutional indicator or the rate in question, the rate adjusted for case mix if the methods are available, and prominent confidence intervals. The inevitable public interest in league position could be dealt with by including ranks for a number of recent years, to give a rough but ready indication of their instability.

Health Services Research Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, London WC1E 7HT

Colin Sanderson, senior lecturer in health services research

Martin McKee, professor

Correspondence to: Dr Sanderson c.sanderson@lshtm.ac.uk

Underperforming doctors: a postal survey of the Northern Deanery

George Taylor

Abstract

Objectives: To discover the perceived size of pool of doctors considered to be underperforming in general practice in the Northern Deanery and to discover whether these perceptions are based on formal assessments.

Design: Postal questionnaire.

Setting: Area covered by the Northern Deanery.

Subjects: Seven health authority directors of primary care, seven secretaries of local medical committees, and 14 chief officers of community health councils.

Results: The response rate was 100% for directors of primary care and secretaries of local medical committees and, after one reminder, 92% for chief officers of community health councils. Numbers of doctors perceived to be underperforming ranged from none to over 15 in different health authority

areas. Main areas for concern were communication skills, clinical skills, and management skills. Patients’ representatives were concerned about lack of power of patients and health authorities and doctors’ lack of accountability. Health authorities were concerned about lack of power, identification of underperforming doctors, and doctors’ professional loyalty. Local medical committees were concerned about the problem of identifying underperformance. A number of methods were used for identification, and there was no common method applied.

Conclusions: The number of doctors thought to be underperforming was small. Work still needs to be done on developing tools that can be used in everyday practice to enable doctors to confirm for themselves, their colleagues, and their patients that they are providing an adequate level of care.

See Editor’s choice

Postgraduate Institute for Medicine and Dentistry, University of Newcastle, Newcastle upon Tyne NE2 4AB
George Taylor, deputy director general practice
g.b.taylor@ncl.ac.uk

BMJ 1998;316:1705-8



Study questionnaire is available on the internet

Introduction

The General Medical Council's performance procedures were introduced in the summer of 1997.¹ These new procedures give the GMC, for the first time, the power to discipline doctors whose performance is found to be seriously deficient. If doctors are found to be underperforming, the GMC now has the power to suspend the doctors' registration and make the lifting of this suspension conditional on a period of retraining.² There are, however, a number of uncertainties around these new procedures, not least the scale of the problem and the type of retraining that will be required.

In an effort to quantify the size of the problem in general practice in the Northern Deanery and identify the areas in which training may be required in the future, I carried out a postal survey of interested parties.

Methods

In early 1988 I conducted a postal survey of the three groups perceived to be interested in general practice in the Northern Deanery: NHS commissioners, represented by the seven directors of primary care at the relevant health authorities; general practitioners, represented by the seven secretaries of local medical committees; and patients, represented by the 14 chief officers of community health councils. The response rate was excellent, with 100% responses from the directors of primary care and the secretaries of local medical committees and 11/14 responses from the secretaries of community health councils to the first request for information. The response from the community health councils rose to 13/14 after one reminder.

I used standard development methods for the questionnaire, including piloting to ensure clarity, question structure, and time to complete.³ The first series of questions related to whether responders had referred or were planning to refer general practitioners to the GMC under the procedures for seriously deficient performance. The respondents were also

Table 1 Respondents' areas of concern in relation to doctors referred, or being considered for referral, to GMC for deficient performance. Values are number of responses

Areas of concern	Respondent		
	Health authority	Local medical committee	Community health council
Clinical skills	3	1	6
Communication skills	4	0	6
Records	4	1	2
Concerns about prescribing	1	0	0
Concerns about referral	1	0	0
Management	3	1	4
Other	0	0	1

("poor doctor")

asked to identify the areas in which these doctors were perceived to be underperforming. The development of this part of the questionnaire was based on views obtained during unstructured interviews with a sample population of general practitioners and health authority directors of primary care. The facility to add further areas of concern was allowed, as was further free comment.

Respondents were then asked to gauge how many doctors in their area needed help with their performance but not at such a level as to require referral; they were asked to indicate the size of the population of underperforming doctors by circling a range. A further question asked them to identify areas that they perceived to be problematic in this population of doctors; they were able to identify more than one area.

Respondents were asked what methods, either formal or informal, were used to identify underperforming doctors. In the final question respondents were asked to identify up to three deficiencies in the current systems relating to the identification and management of poorly performing doctors, and these were analysed with standard qualitative analysis techniques.⁴

Results

Referral of doctors to the GMC

The seven health authorities had already referred five doctors under the new GMC procedures and were also considering referral in five further cases. The community health councils had referred six doctors and were considering the case of five others. The local medical committees had not referred any doctors but were considering the cases of two. It is impossible to say from the data if there was overlap of referral between organisations, but the numbers involved, while small, are not insignificant. The total number of general practitioners in the Northern Deanery is 1633.⁵

Table 1 shows the respondents' areas of concern relating to these underperforming doctors. Respondents were allowed to identify more than one area of concern. Clinical skills and communication skills were common areas of concern, and practice management was also often felt to be an area of deficiency. Health authorities, but less so community health councils, identified poor record keeping as an area of concern.

Doctors needing help with their performance

Figure 1 shows the respondents' estimates of the number of doctors who were in need of help with their performance but not at such a level as to require

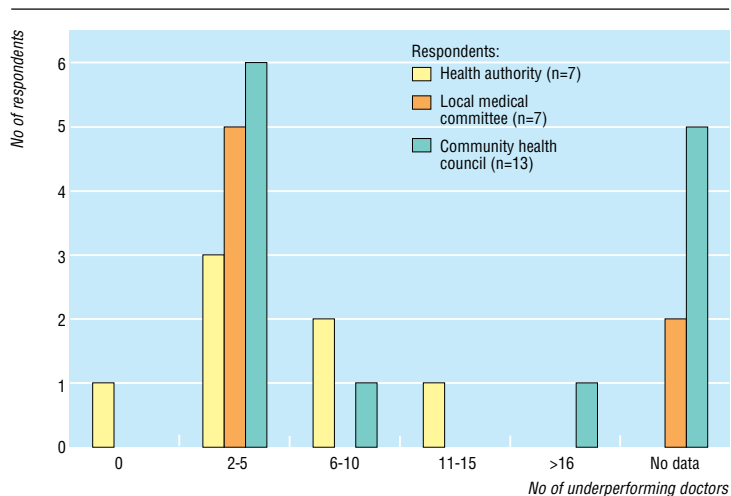


Fig 1 Respondents' estimates of the number of local general practitioners in need of help for underperformance

referral, and figure 2 shows the respondents' areas of concern about these underperforming doctors. The areas of concern identified were similar to those listed for the doctors who had been referred to the GMC or considered for referral. The health authorities and local medical committees were also concerned about the doctors' prescribing, and to a lesser degree health authorities about referrals. Concerns about management (which was not defined in the questionnaire) were common, especially among the patients' representatives.

Methods used to identify underperforming doctors

Health authorities listed many different formal mechanisms for identifying underperforming doctors, including targets, practice inspections, input from the pharmaceutical advisor, and a number of quality systems such as King's Fund organisational audit,⁶ the Royal College of General Practitioners fellowship by assessment,⁷ and a local health authority practice accreditation scheme. Two health authorities also used the complaints mechanisms. Of the 13 community health councils that responded, six used the complaints mechanisms to identify underperforming doctors, and this was the only formal mechanism used by the councils. The local medical committees had no formal mechanisms other than one being involved with the local practice accreditation scheme.

Sources of help for underperforming doctors

The questionnaire offered respondents a choice of organisations that might provide help to underperforming doctors, and table 2 shows their responses. They were able to identify more than one source of help. One community health council thought that the GMC itself should be a source of help to such doctors.

Table 2 Respondents' choices of organisations that might provide help to underperforming doctors

Source of help	Respondent		
	Health authority	Local medical committee	Community health council
Health authority	4	6	11
Local medical committee	5	8	10
Royal College of General Practitioners	4	4	8
University	4	3	5
General practitioner tutor	5	5	8
Others	Deanery (1), mentors (1)	0	Community health council (1), GMC (1)

Perceived deficiencies in present system

Figure 3 lists the deficiencies in the present system that respondents identified. The health authorities felt limited by their lack of power, had problems identifying underperforming doctors, and were limited by doctors' loyalty to colleagues. They also identified problems with resources, including time. The local medical committees saw the main problem as identification of underperforming doctors but also noted doctor loyalty as a problem. The community health councils, however, did not see identification as a problem but saw the major problems to be doctors' professional loyalty, their lack of accountability to the health service,

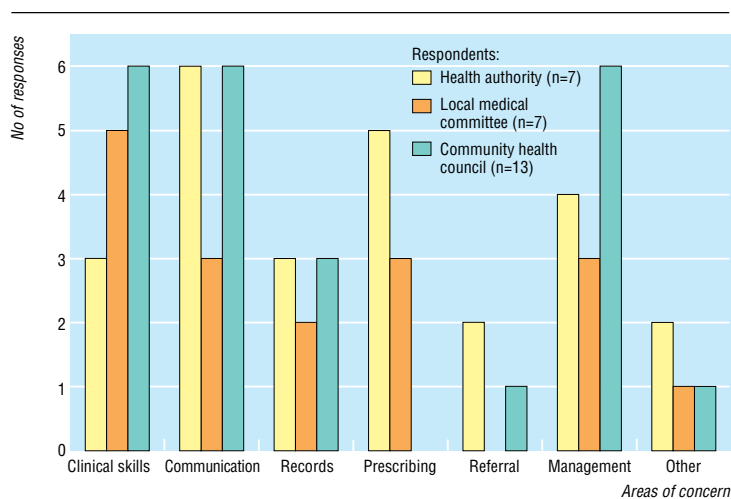


Fig 2 Respondents' perceptions of areas of concern in the population of underperforming doctors

and the lack of patients' and health authorities' power in these situations.

Discussion

The purpose of the performance procedures is to protect the public from doctors providing seriously deficient care, and the profession has a responsibility to help with this. It is also important to recognise that some doctors may need help to enable them to avoid being drawn into these procedures. It is unclear whether mechanisms exist at present to identify these doctors. Preventive help rather than punitive action should surely be the aim in the long term.

The results of this small survey seem to point to a small but significant number of doctors being referred under the performance procedures and a larger pool of doctors who may fall into this system if action is not taken to help them. The areas of concern cited were,

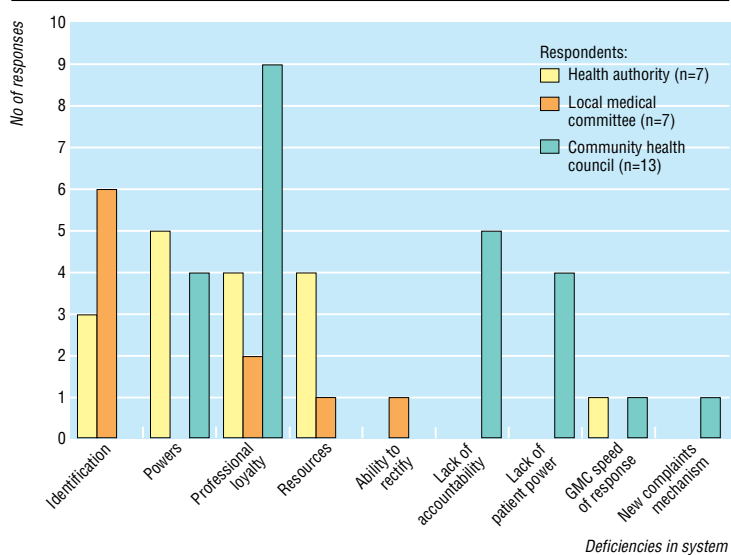


Fig 3 Respondents' choices of deficiencies in current system of dealing with underperforming doctors

however, only perceptions. There were few, if any, formal mechanisms being used for the routine assessment of performance of doctors in practice, and the tools used would seem to be crude measures of performance. It seems unlikely that poorly performing doctors would involve themselves in activities such as the King's Fund organisational audit or the Royal College of General Practitioners fellowship by assessment, both of which require a high level of commitment. Formal targets or practice inspections will provide, at best, crude measures of performance. Complaints will identify poor performance, but perhaps too late in a doctor's career.

A variety of methods have been put forward to assess clinical competence for the purpose of re-certification.⁸ Southgate and Jolly, from the United Kingdom, recommend information sources such as practice logs, patterns of referral and prescribing, and direct observation, whereas Hays, from Australia, proposes practice audit, external audit, standardised patients, and direct observation by trained assessors.⁸ In our survey the health authorities were using some of the external audit data that they held, such as targets and prescribing. There does not seem, as yet, to be any mechanism to assess and use data relating to day to day practice. The NHS Executive have recently proposed a national framework for assessing performance.⁹ This does not, however, seem real or relevant to everyday practice. Is an admission rate for severe ear, nose, and throat infection a true indicator of management of acute care in primary care? It may well be that developments such as the Royal College of General Practitioners clinical practice evaluation programme (CPEP)¹⁰ will help in this area.

Our survey also seemed to identify deficiencies in the NHS systems for bringing about change in underperforming doctors. The medical profession clearly has a major responsibility in overcoming its natural reticence in "blowing the whistle" on seriously underperforming colleagues. It also needs to recognise its responsibility to help doctors whose performance is slipping to rectify their problems before these threaten their ability to practice to an acceptable level. The GMC and the medical profession must either reassure the public that present systems are able to protect them

Key messages

- To quantify the problem of underperforming general practitioners in the Northern Deanery, a postal survey was carried out among representatives of healthcare commissioners, doctors, and patients
- A small but not insignificant number of doctors were identified as providing a poor level of performance
- Main areas for concern were communication skills, clinical skills, and management skills.
- Various methods were used for identifying underperforming doctors, but there was no common method applied
- Perceived problems with the present system of dealing with underperforming doctors included identification, lack of power of patients and health authorities, and doctors' professional loyalty and lack of accountability

or else review the existing systems in the light of experience and make changes to increase the public's confidence.

Funding: None.

Conflict of interest: None.

- 1 *Medical (Professional Performance) Act 1995*. London: Stationery Office, 1997. (Commencement No 3.)
- 2 General Medical Council. *The new performance procedures*. London: GMC, 1997.
- 3 Fink A, Kosecoff J. *How to conduct surveys*. Newbury Park: Sage, 1985.
- 4 Britten N, Jones R, Murphy E, Stacy R. Qualitative research methods in general practice and primary care. *Fam Pract* 1995;12:104-14.
- 5 Department of Health, NHS Executive. *GMS statistics England and Wales*. Leeds: NHSE, 1997.
- 6 *King's Fund organisational audit primary care standards*. London: King's Fund, 1993.
- 7 Royal College of General Practitioners. *Fellowship by assessment*. London: RCGP, 1997. (Occasional Paper No 50.)
- 8 Newble D, Jolly B, Wakeford R, eds. *The certification and re-certification of doctors. Issues in the assessment of clinical competence*. Cambridge: Cambridge University Press, 1994:184, 192-4.
- 9 *The new NHS modern and dependable: a national framework for assessing performance*. Leeds: NHS Executive, 1998.
- 10 Royal College of General Practitioners. *RCGP clinical practice evaluation programme (CPEP)*. London: RCGP, 1997.

(Accepted 22 May 1998)

Fifty years ago

The new NHS: The plebiscite result

The present plebiscite, the results of which are analysed on another page, shows that the concessions made by Mr. Bevan have persuaded many medical men to change their attitude to the Act and their decision whether or not to enter the Service on July 5. In the plebiscite of February there was a 9 to 1 majority in an 84% poll against the Act. In this plebiscite approximately 2 out of every 3 doctors in a 77% poll disapprove of the Act, notwithstanding the concessions Mr. Bevan has made. The figures show that some of those who disapprove of the Act are nevertheless prepared to enter the Service on July 5. In answer to Question B—the most important question of the three—12,799 medical men in England and Wales and Scotland are in favour of accepting service, and 13,891 are not in favour. The group which

answered this question contains those directly affected by the Act—consultants and specialists, general practitioners, and whole-time voluntary hospital workers. There is an almost even division of opinions—48% willing to enter the Service and 52% unwilling. The number of general practitioners (principals and assistants) against accepting service is 9,588. Although there is an overall majority against accepting service, the majority does not include approximately 13,000 general practitioners. The majority the B.M.A. required if it was to continue to advise the profession not to enter the Service has, therefore, not been obtained.

(*Editorial*, 8 May 1948, p 885. See also editorial by Gordon Macpherson, 3 January 1998, p 6.)