How consultants, hospitals, trusts and deaneries affect pre-registration house officer posts: a multilevel model

I C McManus,[1] B Winder[1] & Elizabeth Paice[2]

Objective To examine variation in reported quality and working conditions of pre-registration house officer (PRHO) posts. To use multilevel modelling to determine how much of the variation was due to the effects of consultant firms, hospitals, trusts and deaneries, as well as variation at the level of the individual doctor.

Design Questionnaire survey of national sample of PRHOs at the end of the pre-registration year. Multilevel modelling was carried out using MLwiN.

Setting PRHOs working in UK approved medical and surgical posts in NHS hospitals from August 1996 to January 1999.

Participants One thousand, four-hundred and thirty-five PRHOs who had previously been part of an extended cohort study of medical student selection and training. They reported on a total of 4926 posts, on 2721 identifiable consultant firms, in 336 hospitals, in 264 trusts and in 17 deaneries.

Main outcome measures PRHOs' perceptions of the overall quality of all PRHO posts they had worked in, and working conditions in the current post.

Results Twenty percent of PRHO posts were described as excellent and 34% as very good, through to 6% reported as not very good, poor or bad. The overall rating of the post showed highly significant variation at the level of hospital/trust, and consultant firm. Other aspects of posts, such as working conditions showed variation at the level of deanery, trust, hospital and firms, with different patterns for the different measures. Assessments of stress in PRHOs (GHQ, burnout, response to uncertainty), and a desire to leave medicine, showed no variation at the level of deanery, trust, hospital or firm.

Conclusion Some aspects of a PRHO post, in particular overall rated quality, but also many aspects of working conditions, show substantial variation at the level of hospital, trust and consultant firm, suggesting that differences reflect local variation in working practices and treatment of PRHOs, with the possibility of change and improvement. The lack of any variation at the level of deanery, trust, hospital or firm in stress of PRHOs suggests that these responses are idiosyncratic, individual responses by doctors themselves, rather than a general characteristic of posts, hospitals, trusts or deaneries.

Keywords Hospital staff, medical/*standards; *job satisfaction; workload; stress, psychological; cohort studies; Great Britain; questionnaires.

Introduction

Pre-registration house officer (PRHO) jobs have been controversial since their introduction in the 1950s,[1] the quality of training being extremely varied, the work load excessive and the stated educational aims often being absent.[2] In the 1990s a concerted effort was made at improvement, with the GMC issuing guidelines for posts;[3] at the same time universities and postgraduate deans, who were jointly responsible for the posts, introduced many reforms.[4]

A fundamental question in evaluating PRHO posts concerns the interpretation of reasons for variation. If a job is rated as poor, working conditions said to be inadequate, work loads described as excessive, or PRHOs report they are stressed or burned out, what is the locus of that problem? Does it reflect the individual doctor, the consultant 'firm' for which they

[1]Academic Centre for Medical Education, Centre for Health Informatics and Multi-professional Education (CHIME), Royal Free and University College Medical School, London, UK
[2]North Thames Department of Postgraduate Medical and Dental Education, London, UK

Correspondence: I C McManus, Professor of Psychology and Medical Education, Academic Centre for Medical Education, Centre for Health Informatics and Multi-professional Education (CHIME), Royal Free and University College Medical School, The Archway Campus, Highgate Hill, London N19 3UA, UK. Tel.: 020 7288 3363; Fax: 020 7288 3322; E-mail: i.mcmanus@ucl.ac.uk

are working, the hospital, the trust of which the hospital is a part, or the postgraduate deanery? It should be noted that although we are asking this question specifically in the context of PRHO posts, the problem is an extremely general one, and is found throughout complex hierarchical organisations, such as the NHS, medical schools, universities, etc., and the approach described here is applicable to all of them.

Data of this type are multilevel,[5] and can be examined statistically by means of multilevel modelling (MLM). MLM was developed by educational statisticians, to differentiate the effects of classes, teachers, schools, and education authorities upon children. Health data are likewise hierarchical,[6] and in the case of PRHO posts, variance can in principle be separated at the level of individual doctor, firm, hospital, trust and deanery, which are 'nested' inside each other. In essence, the method asks whether reports of two PRHOs working for the same firm of consultants are more similar than doctors at the same hospital working for different firms; and likewise, whether reports from two firms in a hospital are more similar than those from two firms at different hospitals within a trust, whether reports from two hospitals in a trust are more similar than those from different trusts within the same deanery, and whether reports from two trusts in the same deanery are more similar than those from trusts in different deaneries.

The partitioning of variance is not merely an academic exercise. If variance is principally at a particular level, for example the hospital, then that is the level at which change can probably be implemented most effectively, since existing differences must have causes which can be understood. The absence of variation at a particular level most probably means that local policy differences have little impact on the outcome, so merely changing policies at that level is unlikely to effect change. Alternatively, the lack of variation at a level might be construed as regulation from a higher level (just as there is little variation in the basic structure of the PRHO year since that is controlled by statute). If therefore variation is, for instance, entirely at the level of the individual doctor, then it might be unrealistic to expect individual deaneries, trusts, hospitals or firms to be able to intervene to cause change since little pre-existing variance is present at those levels.

In this paper we describe a large-scale national survey of a cohort of PRHOs. Their reports of working conditions and the quality of their posts are analysed by means of multilevel modelling.

Methods

The study began as a prospective cohort study of medical student selection, looking at 6901 applicants to UK medical schools in Autumn 1990 for admission in October 1991.[7,8] Three thousand, three hundred and thirty-three applicants entered medical school, 2961 in 1991 and a further 372 in 1992. Two thousand, four hundred and fifty-six of the applicants graduated in 1996 and 1997. One thousand, two hundred and thirty began their first PRHO post in August 1996, 140 in February 1997, 1017 in August 1997 and 69 in February 1998. PRHOs were sent a questionnaire approximately 8 weeks before the end of their second PRHO post, in May 1997, November 1997, May 1998 or November 1998. Non-respondents were sent 3 reminder letters. Other aspects of the PRHO survey have been reported elsewhere.[9,10]

The questionnaire was kept brief in order to ensure a good response rate, and comprised a single sheet of folded A3 paper. Applicants were asked to describe the consultant firm, the hospital, and the speciality, and to rate the overall quality of the job for each of their separate PRHO posts. They were asked to describe the working conditions (hours, on call rota, sleep on call, numbers of patients clerked, and numbers of doctors on the firm) for just their current post. Jobs in obstetrics were coded as surgical, and jobs in general practice were omitted from the analysis. The questionnaire also included a 12-item General Health Questionnaire (GHQ),[11] an abbreviated Maslach Burnout Inventory[12] with three items on each of the three subscales, and four questions on each of two subscales on uncertainty in medicine.[13] PRHOs were also asked whether they were considering a career outside of medicine, about the quality of careers advice during their PRHO post, and about the experience of harassment. Additional questions not reported here asked about personality, stress, coping, and career preferences.

Conventional statistical analysis used SPSS for Windows 9·0, and multilevel modelling used MLwiN 1·02.[14] Significance levels in multiple regression were set at 0·01 to correct in part for alpha inflation.

Results

Questionnaires were sent to 2456 PRHOs, of whom 1435 (58·4%) sent usable replies. Comparison of respondents and non-respondents on baseline measures suggested, as in previous studies,[8] that there was little bias. The 1435 respondents reported on a total

Table 1 The ratings made by the PRHOs of each individual medical and surgical job that they carried out. The first two columns show the ratings separately for medical and surgical posts, and the third column for all posts. The fourth and fifth columns show the ratings made by PRHOs at the best and worst hospitals, and the sixth and seventh columns show the ratings made by PRHOs working for the best and worst consultant firms

	Medical	Surgical	Total	Best-rated hospital	Worst-rated hospital	Best-rated firm	Worst-rated firm
Excellent	553 (23·0%)	401 (17·3%)	954 (20·2%)	13 (45%)	1 (7%)	5 (56%)	–
Very good	917 (38·1%)	709 (30·6%)	1626 (34·4%)	11 (38%)	3 (20%)	3 (33%)	–
Good	610 (25·3%)	646 (27·9%)	1256 (26·6%)	4 (14%)	5 (33%)	1 (11%)	1 (25%)
Adequate	228 (9·5%)	352 (15·2%)	580 (12·2%)	1 (3%)	–	–	–
Not very good	50 (2·1%)	124 (5·4%)	174 (3·7%)	–	4 (25%)	–	1 (25%)
Poor	30 (1·2%)	52 (2·2%)	82 (1·7%)	–	1 (7%)	–	1 (25%)
Bad	19 (0·8%)	33 (1·4%)	52 (1·1%)	–	1 (7%)	–	1 (25%)
Total	2407	2317	4724	29	15	9	4

of 4926 posts, which were held for a mean of 3·4 months (SD 1·4), on 2721 identifiable consultant firms, in 336 hospitals, in 264 trusts and in 17 deaneries.

Overall rating of posts

Two jobs in general practice were omitted from the present analysis. Table 1 shows PRHOs' overall ratings of each post in which they had worked. 6·5% (308/4724) of jobs were rated as not very good, poor or bad; medical jobs were rated more highly than surgical jobs.

Multi-level modelling

The primary interest of the study was in the PRHOs' ratings of the various jobs they had carried out. A five-level model was fitted, individual ratings being at the lowest level, nested inside consultant firms, which were nested inside hospitals which were nested inside trusts which were nested inside deaneries. Since, from discussion with PRHOs, it was expected that surgical jobs would be less popular than medical jobs, the difference was modelled by a fixed effect at the lowest level, and it was highly significant. Exploratory modelling showed no evidence of differences between medical and surgical jobs at any of the higher levels. Table 2 shows the estimated components in the final model. The numbers in the table show the estimated effect in the model present at that level, along with the standard error of the effect. Variances are only left in the model if they are significant at the 5% level. Because variances can only be positive or zero, it is conventional in MLM to remove variance components from the model which

are not statistically significant and replace them with zero values.[*]

In fitting the model it was found that the effect of deanery was not significant and it was therefore dropped from the model. The effects of trust and hospital were somewhat confounded due to many trusts only having a single hospital; however, the joint effect of hospital and trust was highly significant but it was not possible to distinguish the individual components in a useful way.[**] The effect of consultant firm was highly significant.

It is clear from the MLM that there are significant differences between hospitals/trusts which account for about 24% of the variance above the level of the individual response. Figure 1 shows the estimated mean rating at each of 302 hospitals. There are highly significant differences between the best and the worst rated hospitals. The highest rated hospital was in Yorkshire, where 13/29 (45%) had described the posts as excellent, and none as not very good, poor or bad. In contrast at the worst rated hospital, in Scotland, only 1 of 15 PRHOs (7%) described their post as excellent, whereas 6 (40%) described the posts as not very good, poor or bad. The remaining 76% of variance above the

---

[*]MLM estimates two different types of effect. Fixed effects are similar to the effects estimated in multiple regression, and can be regarded as means or slopes, and can take negative or positive values. Random or variance effects are variances and hence can only take positive or zero values. In the present study the fixed effect of medicine versus surgery jobs is primarily a nuisance effect, and hence the description of the analysis is primarily in terms of the variance effects.

[**]As a result the standard errors of the variances due to hospital and trust are relatively large. However removal of either results in the other being significant.

Table 2 Partitioning of variance for each measure, by multilevel modelling, into random (or variance) components due to deanery, trust, hospital, firm and residual (measurement error and doctor−level variance), and a fixed effect due to differences between medical and surgical posts. Figures indicate the estimated effect and the standard error of the effect. Note that although fixed effects can be negative or positive, random effects being variances can only be positive

		Surgery vs. Medicine	Deanery	Trust	Hospital	Firm	Residual†
Working conditions	Rating of job	·335 (·040)	0**	·019 (·027)*	·042 (·029)*	·186 (·034)	1·262 (·039)
	Average contracted working week (hrs)	·131 (·755)	0	0	7·22 (3·79)	0	134·59 (3·79)
	Typical working day length when not on call	·547 (·049)	·041 (·021)	0	·118 (·025)	0	·424 (·050)
	Time of starting work	−·712 (·021)	0	0	·029 (·005)	·137 (·015)	·089 (·005)
	Time of finishing work	−·169 (·042)	·024 (·013)	0	·059 (·016)	0	·393 (·020)
	On call rota (1 in N)	0	0	·106 (·053)	·108 (·053)	0	·418 (·022)
	Hours sleep on call	3·709 (·195)	·167 (·089)	0	·560 (·111)	·673 (·177)	1·293 (·161)
	Emergency admissions per week	−2·33 (·452)	3·17 (1·65)	6·11 (1·77)	0	19·94 (3·37)	21·08 (2·82)
	Routine admissions per week	6·33 (·35)	2·25 (·98)	0	0	5·44 (2·51)	21·72 (2·56)
	Day care patients per week	·871 (·337)	0	0	0	7·05 (2·29)	18·65 (2·28)
	Outpatients clerked per week	·770 (·175)	0	0	0	3·05 (·544)	3·795 (·491)
	Consultants on firm	·146 (·074)	·180 (·079)	·213 (·053)	0	·320 (·108)	·847 (·103)
	Registrars on firm	·206 (·057)	·033 (·021)	·121 (·031)	0	·200 (·064)	·501 (·061)
	SHOs on firm	−·151 (·061)	·148 (·059)	0	·047 (·027)	·296 (·071)	·521 (·064)
	PRHOs on firm	·277 (·063)	·202 (·082)	0	·190 (·041)	0	·842 (·081)
Stress	GHQ (0−1−2−3 scoring)	0	0	0	0	0	32·5 (1·45)
	Emotional exhaustion	0	0	0	0	0	5·79 (·28)
	Depersonalisation	0	0	0	0	0	7·32 (·32)
	Personal accomplishment	0	0	0	·43 (·18)	0	6·05 (·31)
	Response to uncertainty: stress	0	0	0	0	0	8·26 (·37)
	Response to uncertainty: reluctance to disclose	0	0	0	0	0	5·35 (·24)
Career	Quality of careers advice	0	0	·086 (·048)	0	0	1·873 (·092)
	Consideration of a career outside medicine	0	0	0	0	0	·998 (·45)

*Hospital and trust are jointly significant (p<·05) but cannot be totally resolved.

**Estimates of zero indicate that the term was not significant and has been forced to zero.

†Residual consists of components specific to the PRHO and confounded measurement error.

Figure 1 Plot of rating of PRHO posts at 302 hospitals, taking into account differences between medical and surgical posts, and differences between firms within hospitals. Points are plotted ±1·4 standard errors, so that hospitals whose error bars do not overlap are significantly different at the 5% level 15. Note: since Excellent is scored as 1 and Bad as 7, high scores indicate *worse-rated* hospitals.
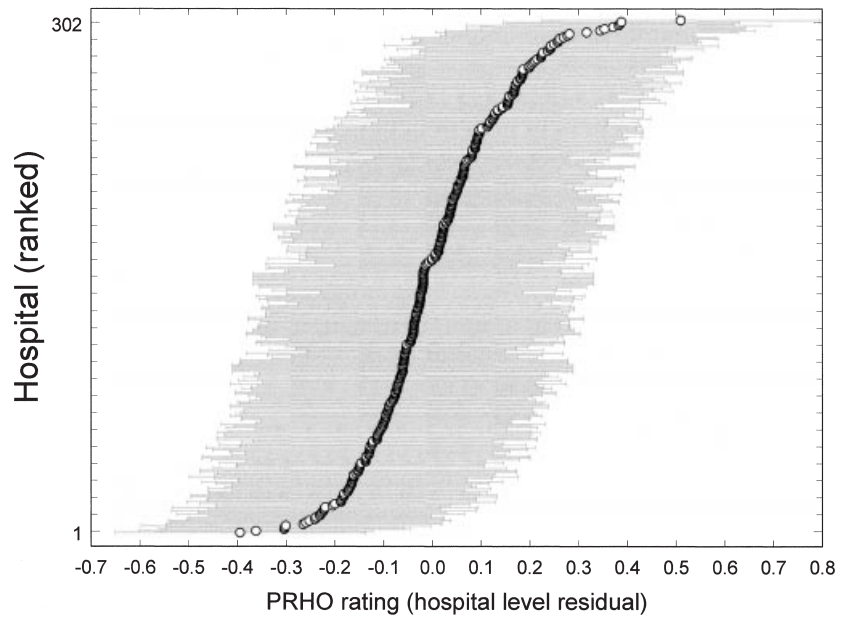
Figure 2 Plot of mean rating of PRHOs for 2555 consultant firms, taking into account differences between medical and surgical posts, and differences between hospitals. Points are plotted ±1·4 standard errors, with lines and points to left and right indicating upper and lower confidence intervals. Firms whose intervals do not overlap are significantly different at the 5% level.[15] Note: since Excellent is scored as 1 and Bad as 7, high scores indicate *worse-rated* consultant firms.
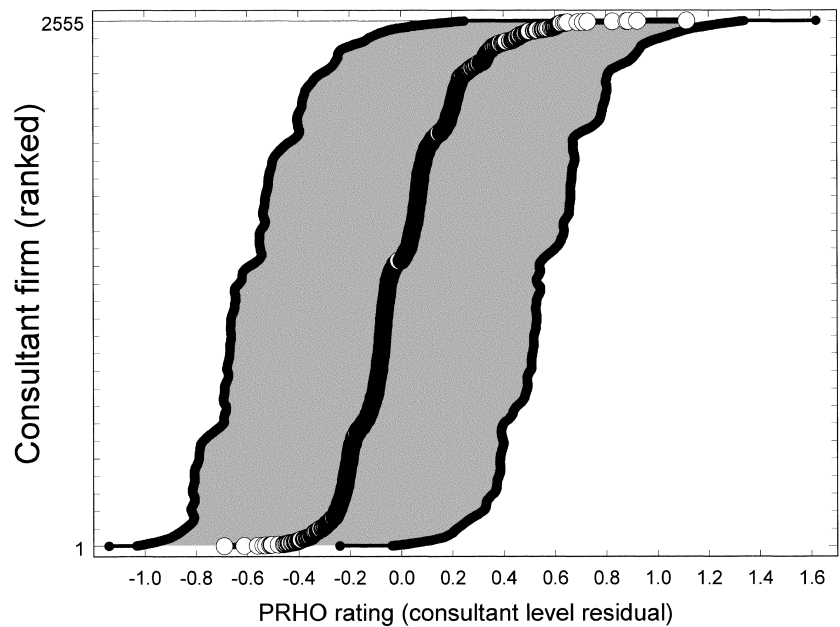
level of individuals was due to consultant firm. Figure 2 shows the effect for each of 2555 firms. For the highest rated firm, 5 of 9 PRHOs had described it as excellent, compared with the worst firm which, 3 of 4 PRHOs had described as not very good, poor or bad (Table 1). As is conventional in MLM,[15] 1·4 standard error bounds are indicated for estimates, and those which do not overlap are individually statistically significant with $P < 0.05$. It is perhaps worth emphasising that although there is no doubt that there are

significant differences between hospitals and firms, the majority are not significantly different from the rest.

Job characteristics

The working conditions vary in PRHO posts, in particular in the on call rota, the working hours, the number of patients admitted and the number of other doctors on the firm. It is possible that these are related to the perceived quality of a PRHO post. Figure 3
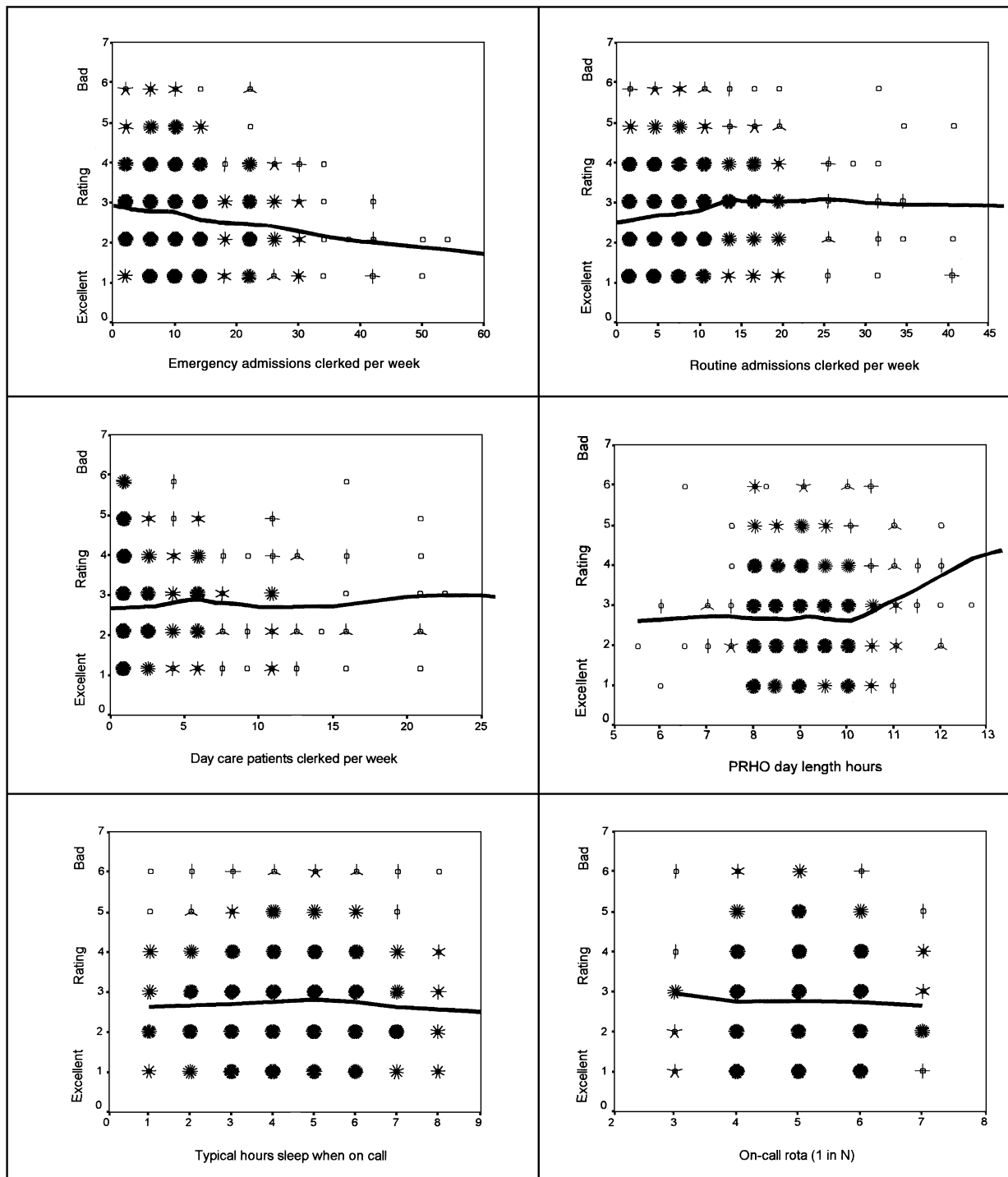
Figure 3 Rated PRHO job quality in relation to summary measures of working conditions. Each petal on the sunflowers represents one PRHO. Fitted lines are Lowess curves.

shows job quality (of the current post only) in relation to summaries of these measures. It is clear that the rating of a post shows little relation to working conditions. Stepwise multiple regression of job ratings on the 14 measures of working conditions shown in Table 2 taking medical/surgical differences into account found only one that was significant at $P < 0.01$; jobs were rated more highly if there were more
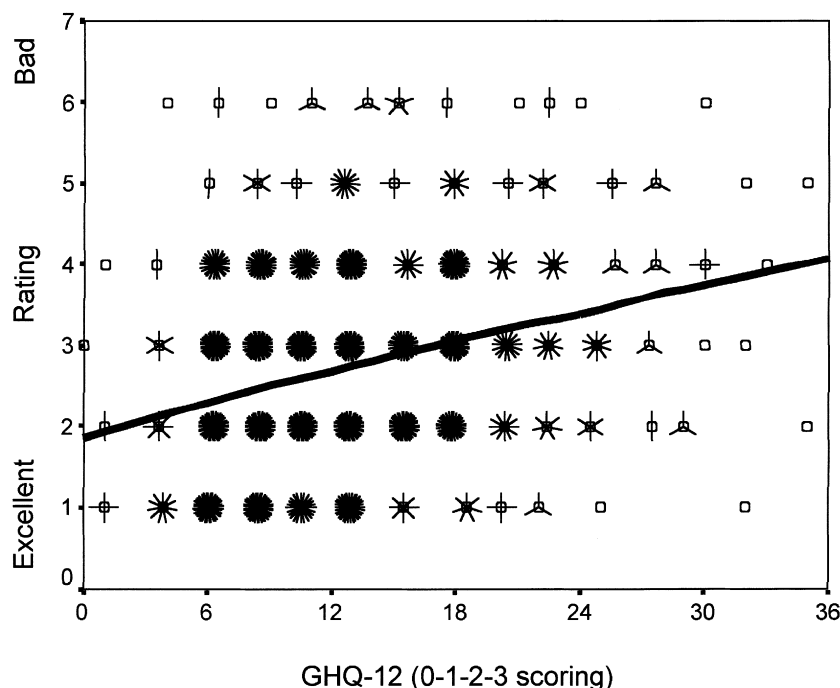
Figure 4 GHQ-12 score in relation to rating of current PRHO post. Each petal on the sunflowers represents one PRHO. The fitted line is a Lowess curve. A score of 1 indicates a rating of excellent and a score of 7 a rating of bad. A high score on the GHQ indicates that the PRHO is more stressed.

emergency admissions per week (beta = 0·131, $z = 4·66, P < 0·001$).

Multilevel modelling shows that the working conditions of PRHOs are affected by the consultant firm, by hospital, by trust, and also by deanery (Table 2).

Measures of stress

The GHQ-12 is conventionally scored in several different ways. For epidemiological purposes the four responses on each GHQ question can be scored as 0-0-1-1 (i.e. the two least stressed responses are scored as zero and the two most stressed responses are scored as one). Psychiatric 'caseness' is then calculated for epidemiological purposes using a 3/4 threshold (i.e. scores of 4 or more are cases, scores of 3 or less are not). Overall 31·7% (422/1330) of the PRHOs met the GHQ criterion for psychiatric caseness. The 0-0-1-1 scoring of the GHQ results in a very skewed distribution in which most individuals score zero. For a more powerful and detailed statistical analysis it is common to score the GHQ on a 0-1-2-3 basis, giving a score of zero for the least stressed response and 3 for the most stressed response. The distribution is then far more Gaussian with very few individuals scoring zero. Figure 4 shows there is a strong relationship between current job rating and GHQ score (0-1-2-3 scoring), high scores on the GHQ indicating greater stress levels. Multiple regression of current job rating on the six stress-related measures shown in Table 2 found that

only GHQ was a significant predictor ($z = 3·381$, $P = 0·001$). Multilevel modelling of the stress-related measures showed, with a single exception, that none of the measures had variance at the level of firm, hospital, trust or deanery, the variance all being at the level of the individual doctor. The sole exception was that personal accomplishment (the 'positive' subscale of the burnout inventory) also had variance at the hospital level.

Careers advice and a career outside medicine

PRHOs were asked if they had considered a career outside medicine: 26·2% had not considered it at all, 44·1% had thought about it occasionally, 22·0% had thought about it quite frequently, 3·8% were considering it quite seriously, 3·6% were considering it very seriously, and 0·4% had a definite intention not to practice in medicine. There was a small but very significant association between the rating of the current PRHO post and an intention not to practice medicine (Kendall's tau = 0·146, $P < 0·001$). Multilevel modelling found that none of the variance in intention not to practice in medicine was at the level of firm, hospital, trust or deanery (Table 2). PRHOs were also asked about the quality of the careers advice they had received during their PRHO year: 16·9% described it as very good, 12·4% as good, 19·0% as fair, 33·7% as poor, 14·1% as very poor, and 3·9% had had none. Better careers advice was correlated with a better rating of the current PRHO post (Kendall's tau = 0·127,

$P < 0.001$), and a lower likelihood of a career outside medicine (Kendall's tau = 0·092, $P < 0.001$). Multi-level modelling of careers advice found no evidence of variance at the level of firm, hospital or deanery, but there was significant variance at the level of the trust (Table 2).

Discussion

Although some 20% of PRHO posts are described as excellent, there are also some that are bad, 1% being explicitly described as such, with a further 2% poor and 4% not very good. The present study allows an explanation of the reasons for those differences. Despite PRHO posts varying quite substantially in working conditions, particularly in case load, case mix, normal working hours, on call commitment, and number of other doctors on the firm, almost none of these factors predict whether a job is rated as good or bad. The sole exception, that better jobs have a greater number of emergency admissions, dispels any simplistic notion that excessive workloads are the sole determinant of poor jobs. Our findings are clearly in agreement with the conclusion of others that working hours bear little relationship to stress levels,[16] but disagree in finding any direct relationship of stress to amount of sleep.[17]

Multi-level modelling of overall ratings of PRHO posts provides a clear picture that there is variance at the consultant level and at the hospital and/or trust level (hospital and trust not being clearly separable due to partial confounding). A poor rating is not therefore merely an idiosyncratic rating of a post by a single PRHO but instead reflects some consistent influence of consultants on PRHOs and of hospitals themselves on PRHO posts. Bad jobs are in part therefore due to both bad consultant firms and bad hospitals.

Of particular interest is that deaneries seem to contribute no variance to PRHO post ratings. That might be because deaneries have no impact at all on PRHO posts, although that seems unlikely in that characteristics such as working hours, number of admissions, hours of sleep on call, and numbers of doctors on firms, clearly show deanery related variance; in interpreting such a result it must however, be emphasised that deaneries are heavily confounded with geographical regions, which are probably correlated with other aspects of local NHS organisation and working conditions.

That trusts and hospitals do influence PRHO ratings is important, since it shows that organisational factors within the NHS can influence how satisfactory are doctors' assessments of their jobs – and that is over and above variance due to differences between consultant firms. Likewise as consultant firms clearly influence working conditions of PRHO posts, they also influence the perceived quality of jobs, albeit not through working conditions as such. That there is no consultant variance in measures such as contracted working week, on call rota, or length of working day (which consultants cannot be expected to influence) but that these measures do have variance at the hospital level, acts as a validation of the statistical method and the quality of the data.

The residual effect at the lowest level of the multi-level model must be interpreted with care, since it is a mixture of measurement error and systematic differences. It is possible to estimate how much of the variance is likely to be systematic if it is supposed that on repeated testing a PRHO had a 50% chance of using exactly the same rating point on the seven-point scale for the overall quality of the job, a 25% chance of scoring one point higher and a 25% chance of scoring one point lower. We have evidence from other questionnaire studies that such proportions are reasonable.†

Given the scoring scheme for the rating scale measuring the overall rating, then in such a case the variance due to measurement error would be 0·667. The systematic variance in the residual could then be calculated as 0·595 (i.e. 1·262–0·667), and the total systematic variance at all levels would be 0·842, of which about 7% would be at the level of the hospital, 22% at the level of the firm, and 71% at the level of the individual PRHO.

The PRHOs in our study show high levels of caseness on the GHQ, as in other studies of PRHOs,[18] although that may in part reflect the fact that our questionnaire was, in large part, explicitly about stress and methods of coping with it.[19] In interpreting these high GHQ scores, in addition to variation in other burnout and stress related measures, it is very important that no variance has been found at the level of consultant firm, hospital, trust or deanery. In other words two doctors on the same firm are no more likely to be similar in their stress levels than two doctors on different firms or at different hospitals. The implication is that differences in stress do not result primarily from organisational factors present in the administrative structure of hospitals, but instead reflect individual differences within the doctors themselves. That conclusion sets

---

†That means that if a job is rated 'good' on one occasion then on the next it would have a 50% chance of being rated 'good' again, a 1 in 4 chance of being rated 'very good', a 1 in 4 chance of being rated as 'adequate', and no chance of being rated as 'excellent', 'not very good', 'poor', or 'bad'. Such proportions seem reasonable if a questionnaire is being completed seriously.

important limitations on possible sources of action for influencing stress levels. The tiny exception to the rule is, however, intriguing. Personal accomplishment, the conceptual opposite of burnout, does show variation at the hospital level; hospitals can therefore help influence doctors to feel that they are achieving worthwhile results in their daily professional practice.

Dissatisfaction with PRHO posts has been claimed to influence doctors' decisions about leaving medicine as a career. That is supported in part by a correlation in our data between job ratings and thoughts of leaving medicine. However the variation in wanting to leave medicine shows no components at the level of firm, hospital, trust or deanery, suggesting that organisation at these levels is not influencing the decision, and implying that, perhaps like stress, it is a more personal, individual response at the level of the doctor alone.

Careers advice in our study showed much variation, and interestingly there is variance at the level of trusts, suggesting that some trusts are being more effective than others in making efforts to organise careers advice within their hospitals. In contrast, deaneries seem to be having little impact in an area where it might be expected that they could be very influential.

This paper has shown that the success or otherwise of a PRHO post can be assessed straightforwardly by means of a large scale survey, and that the variation can be partitioned into different levels, of firm, hospital, trust and deanery. If carried out systematically at a national level and repeated over a number of years, the data would be more powerful for partitioning effects, although it is clear that the present sample has adequate statistical power for demonstrating the nature of differences and the feasibility of the approach. The method is moreover applicable to many other areas in undergraduate medical education, postgraduate medicine, and NHS organisation. In a much earlier paper, one of the present authors referred to the then PRHO year as being 'chaos by consensus';[20] the present study shows how to dissect the chaos and the consensus and assess its level, from deanery down to individual firms. It also shows that the dissatisfaction which led us then to talk of chaos is, to some extent, still present.

Contributors

ICM, BCW and EP jointly developed the design of the study and contributed to the writing of the questionnaire. BCW and ICM were responsible for the day-to-day running of the study; ICM was primarily responsible for the statistical analysis. ICM, BCW and EP all contributed to the interpretation of the results and the writing of the manuscript.

References

1 Bynum WF. Pre-registration in historical perspective. *Med Educ* 1997;**31** (Suppl. 1):43–8..

2 Goldacre M, Stear S, Lambert T. The pre-registration year. The trainees' experience. *Med Educ* 1997;**31** (Suppl. 1):57–60.

3 General Medical Council. *The New Doctors. Recommendations on General Clinical Training.* London: General Medical Council; 1997.

4 Paice E. *Delivering the New Doctor.* Edinburgh: ASME; 1998.

5 Kreft I, Leeuw J. *Introducing Multilevel Modeling.* London: Sage; 1998.

6 Rice N, Leyland A. Multilevel models: applications to health data. *J Health Serv Res Policy* 1996;**1**:154–64.

7 McManus IC, Richards P, Winder BC, Sproston KA, Styles V. Medical school applicants from ethnic minorities. identifying if and when they are disadvantaged. *BMJ* 1995;**310**: 496–500.

8 McManus IC, Richards P, Winder BC. Intercalated degrees, learning styles, and career preferences: prospective longitudinal study of UK medical students. *BMJ* 1999;**319**:542–6.

9 Paice E, Rutter H, Wetherell M, Winder BC, McManus I. C. Stressful incidents, stress and coping strategies in the pre-registration house officer year. *Med Educ* 2002;**36**:56–65.

10 Paice E, Moss F, Heard S, Winder BC, McManus I. C. The relationships between pre-registration house officers and their consultants. *Med Educ* 2002;**36**:26–34.

11 Goldberg DP, Gater R, Sartorius N, Ustun TB, Piccinelli M, Gureje O et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psych Med* 1997;**27**:191–7.

12 Maslach C, Jackson SE. *Maslach Burnout Inventory.* Palo Alto, CA: Consulting Psychologists Press; 1986.

13 Gerrity MS, DeVellis RF, Earp JA. Physicians' reactions to uncertainty in patient care: a new measure and new insights. *Med Care* 1990;**28**:724–36.

14 Goldstein H, Rasbash J, Plewis I, Draper D, Browne W, Yang M, Woodhouse G, Healy M. *A User's Guide to Mlwin.* London: Institute of Education; 1998.

15 Goldstein H. *Multilevel Statistical Models*. London: Arnold; 1995.

16 Firth-Cozens J. Stress, psychological problems, and clinical performance. In: CA Vincent, M Ennis, RJ Audley, eds. *Medical Accidents*. Oxford: Oxford University Press; 1993: pp. 131–49.

17 Firth-Cozens J, Moss F. Hours, sleep, teamwork and stress. *BMJ* 1999;**317**:1335–6.

18 Kapur N, Borrill CS, Stride C. Psychological morbidity and job satisfaction in hospital consultants and junior house officers: multicentre, cross sectional survey. *BMJ* 1998;**317**:511–2.

19 McManus IC, Winder BC, Gordon D. Are UK doctors particularly stressed? *Lancet* 1999;**354**:1358–9.

20 McManus IC, Lockwood DNJ, Cruickshank JK. The pre-registration year: chaos by consensus. *Lancet* 1977;**i**: 413–7.