Jacqueline J Hill, Anthea Asprey, Suzanne H Richards and John L Campbell

Multisource feedback questionnaires in appraisal and for revalidation:

a qualitative study in UK general practice

Abstract

Background
UK revalidation plans for doctors include obtaining multisource feedback from patient and colleague questionnaires as part of the supporting information for appraisal and revalidation.

Aim
To investigate GPs' and appraisers' views of using multisource feedback data in appraisal, and of the emerging links between multisource feedback, appraisal, and revalidation.

Design and setting
A qualitative study in UK general practice.

Method
In total, 12 GPs who had recently completed the General Medical Council multisource feedback questionnaires and 12 appraisers undertook a semi-structured, telephone interview. A thematic analysis was performed.

Results
Participants supported multisource feedback for formative development, although most expressed concerns about some elements of its methodology (for example, 'self' selection of colleagues, or whether patients and colleagues can provide objective feedback). Some participants reported difficulties in understanding benchmark data and some were upset by their scores. Most accepted the links between appraisal and revalidation, and that multisource feedback could make a positive contribution. However, tensions between the formative processes of appraisal and the summative function of revalidation were identified.

Conclusion
Participants valued multisource feedback as part of formative assessment and saw a role for it in appraisal. However, concerns about some elements of multisource feedback methodology may undermine its credibility as a tool for identifying poor performance. Proposals linking multisource feedback, appraisal, and revalidation may limit the use of multisource feedback and appraisal for learning and development by some doctors. Careful consideration is required with respect to promoting the accuracy and credibility of such feedback processes so that their use for learning and development, and for revalidation, is maximised.

Keywords
appraisal; general practice; multisource feedback; performance assessment; qualitative research.

INTRODUCTION

The UK General Medical Council (GMC) has outlined the revalidation process.[1] All practising doctors will be regularly required to provide evidence demonstrating they are up to date and fit to practise in order to retain their licence. Evidence will be gathered in 5-year cycles via strengthened appraisal[1,2] and clinical governance systems.[3] Multisource feedback from colleague and patient questionnaires will be required as supporting information.[2,4]

Multisource feedback questionnaires have been put forward as feasible, reliable, and valid tools for the assessment of doctors in a variety of settings.[5–7] However, a systematic review summarising the effect of workplace-based assessment on doctors' education and performance concluded that there is conflicting evidence regarding whether doctors' change their practice following multisource feedback.[8] Moreover, among doctors who receive negative feedback, some fail to accept it and may question its accuracy and credibility.[9–13] The context in which multisource feedback is undertaken may influence how it is received.[14] Research in non-medical settings has indicated that participants' responses to feedback may be influenced by a variety of factors, including:

- the degree to which an organisation is perceived to be supportive;[15]
- perceptions of fairness;[16,17]
- opportunities for the person being assessed to 'voice' their own views;[18] and
- the degree to which multisource feedback is specific and timely.[19,20]

Within appraisal, multisource feedback is viewed as a formative, rather than summative, assessment.[2,4] If multisource feedback is to guide personal development within the context of revalidation, a greater understanding of those elements that promote its use and acceptance is required. This qualitative study sought to address this question by exploring GPs' and appraisers' views and experiences of using the GMC patient and colleague multisource feedback questionnaires,[21] as well as their views regarding plans to embed multisource feedback within appraisal and revalidation processes.

METHOD

Evidence regarding the utility, reliability, and validity of the GMC patient and colleague questionnaires can be found elsewhere.[22,23]

Doctors nominated up to 20 colleagues to complete a questionnaire. A patient questionnaire was also distributed to 45 consecutive consulting patients. Responders used five-point scales to assess aspects of performance and could record narrative comments. Upon completion, doctors were sent a confidential, personalised report, including the quartile in which their mean score for each item was located, compared with benchmark data.

JJ Hill, MSc, research fellow; **A Asprey**, MA, associate research fellow; **SH Richards**, PhD, senior lecturer; **JL Campbell**, MD, professor of general practice and primary care, Peninsula College of Medicine and Dentistry, Exeter.

Address for correspondence
John L Campbell, Primary Care Research Group, Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth, Smeall Building, St Luke's Campus, Exeter EX1 2LU.

E-mail: john.campbell@pms.ac.uk

Submitted: 10 November 2011; **Editor's response:** 28 November 2011; **final acceptance:** 24 January 2012.

This is the full-length article (published online 30 Apr 2012) of an abridged version published in print. Cite this article as: **Br J Gen Pract 2012; DOI: 10.3399/bjgp12X641429.**

How this fits in

GPs will be required to collect multisource feedback as part of the supporting information for appraisal and for revalidation. Research has indicated that participants' responses to multisource feedback may be influenced by a variety of factors (for example, perceptions of fairness, credibility of process) and the perceived impact of linking appraisal (including multisource feedback) with revalidation processes is unknown. Doctors and appraisers value multisource feedback when it is undertaken within the context of formative assessment and for appraisal, although concerns remain regarding aspects of the General Medical Council multisource feedback methodology, which may undermine its credibility and could apply to other multisource feedback tools. Proposals linking formative appraisal to summative revalidation may limit the potential of both multisource feedback and appraisal for learning and development. Some doctors and appraisers may become more guarded in relation to what is openly discussed and formally recorded in appraisal.

An independent survey organisation, CFEP UK, managed doctor recruitment and data collection. Doctors were informed that their participation in multisource feedback, or any subsequent research, was voluntary and that anonymised survey data would be shared with the academic team.

Sampling and recruitment

GPs were sampled from two primary care organisations operating established appraisal systems. Doctors who had completed multisource feedback between April and December 2009, and who had completed their appraisal within 3 months of this date, were eligible, as were their appraisers. Participants were recruited independently — that is, a doctor could participate and their appraiser decline participation and vice versa. This study aimed to recruit 10–15 doctors and 10–15 appraisers. Analysis was undertaken in parallel with interviews; sampling continued until no new themes emerged from the data.

Participants were not instructed on how to use the results of their multisource feedback in the context of their appraisal, however, as part of recruitment, the medical director and/or chief executive wrote to all doctors emphasising that multisource feedback might usefully contribute to their appraisal. To ensure anonymity, CFEP UK and the primary care organisations identified participants who were eligible and CFEP UK sent them a study information pack inviting participation. Those wishing to take part returned a reply slip to the research team. A researcher then contacted them to answer any questions prior to interview. As a consequence of this recruitment process, no descriptive data (for example, age and sex) are available to describe the sample of eligible appraisers and appraisees who were approached to take part.

Interview process and schedule

Semi-structured telephone interviews were conducted by an experienced qualitative interviewer. This allowed participants to describe their views in confidence and to share perceptions that might be difficult to discuss as part of a group or with someone they knew.[24] The semi-structured approach enabled the interviewer to clarify the meaning of the participant's response and to elicit more detail on themes arising during the interview.[25]

Box 1. Summary of key interview topics

Doctors were asked for their views on the General Medical Council (GMC) multisource feedback process, including:

- Were they able to complete all elements of the survey required?
- How easy was it to carry out the GMC multisource feedback process in their practice setting?
- Reactions to their personalised feedback report.

Doctors and appraisers were asked for their views and experience of using the GMC multisource feedback data as part of appraisal, including:

- The extent to which they had used patient and colleague feedback, and the self-assessment exercise
- What (if anything) stopped them from using the GMC multisource feedback in appraisal?
- Ideas for improving the use of multisource feedback as part of appraisal, including 'one piece of advice' for the appraisee and 'one piece of advice' for an appraiser about how to use multisource feedback in appraisal

Participants who had used the GMC multisource feedback data in appraisal were also asked:

- How easy was it to discuss their or the doctor's GMC multisource feedback results in appraisal?
- How easy was it to discuss unexpected or negative feedback?
- Whether they planned to use the discussion about multisource feedback in appraisal over the next weeks and months — if so, how?
- What (if any) future action was agreed for the doctor?
- Challenges and/or benefits of multisource feedback (as part of appraisal) for revalidation purposes

Different schedules were designed for doctors and appraisers (Box 1). Questions were developed through consultation with an academic with expertise in GP appraisal systems. Interviews were tape-recorded with the participant's permission. Before commencing the interview, doctors were told that they need not disclose their multisource feedback results and that any multisource feedback data obtained for the study would be anonymised. Appraisers were informed that the performance of individual doctors would not be discussed.

Analysis of interview data

Interviews were transcribed verbatim and anonymised. A thematic analysis[26,27] was undertaken using NVivo (version 2.0, QSR International). Data were analysed to identify salient, overarching themes. Codes were developed iteratively, such that interview scripts were continually revisited in light of subsequent scripts to ensure that the codes were comprehensively applied — the 'constant comparison technique'[28] was applied to the data analysis[29] — and that contrasting data could shape the coding. Two researchers independently coded a sample of transcripts to ensure consistency.

Interview findings were combined into a summary that was sent to all participants with a structured feedback form inviting comments on the veracity of the interpretation of the study findings. Participants were not asked to provide their name on the form.

Analysis of participating doctors' multisource feedback data

Patient and colleague data for the doctors contributing to the interview study, as well as data from GPs contributing to the wider study in the two sampling areas were obtained and anonymised. A description of their age, sex, and summary performance scores[23] is provided to contextualise this research.

Table 1. Age, sex, and summary performance scores of 12 GPs contributing to this study and 116 GPs completing GMC multisource feedback processes in the same sites

Characteristic	N[a]	GP participants	N[a]	Other GPs
Age group, years, *n* (%)[b]	12		116	
30–39		0 (0)		21 (18)
40–49		7 (58)		65 (56)
50–59		5 (42)		26 (22)
≥60		0 (0)		4 (3)
Sex, *n* (%)[b]	12		113	
Female		4 (33)		52 (46)
Mean summary performance, score (SD)[c]				
Patient questionnaire	421	4.85 (0.33)	4277	4.82 (0.38)
Colleague questionnaire	201	4.66 (0.34)	1755	4.61 (0.40)

[a]Refers to the number of responders who were eligible for each analysis. For age group and sex, n *is the number of doctors with valid age and sex data; sex details were missing or spoilt for three (3%) 'Other GPs'. For mean patient questionnaire summary performance score,* n *is the number of patients with a score and for mean colleague questionnaire summary performance score,* n *is the number of colleagues. [b]Age group and sex details were self-reported. [c]Summary performance scores were obtained from doctors', patient, and colleague data. GMC = General Medical Council. SD = standard deviation.*

RESULTS

Sampling

Of 33 GPs and 25 appraisers who were identified as eligible, 12 doctors (36%) and 12 appraisers (48%) participated; 16 participants were male. Participating appraisees worked with an average of five other GPs in their practice (range 2–9). Nine doctors had discussed their multisource feedback data in appraisal and all the appraisers had discussed multisource feedback data with one or more appraisees. The age, sex, and summary performance scores for GP participants, and the wider sample of GPs who undertook the GMC multisource feedback process within the two localities, are shown in Table 1.

Data collection

Interviews were completed between November 2009 and February 2010; each interview lasted 30–40 minutes. All participants gave permission for the interview to be recorded and transcribed.

Twelve participants completed a feedback form commenting on this study's findings. Most were satisfied that the summary accurately represented their views and experiences that were expressed at interview. Seven responders provided additional comments on the summary (data not provided) which were jointly considered by the academic lead and two researchers on this study. After checking transcripts, it was agreed that these comments represented diversity within this study's overarching themes. No amendments were made to the summary and the authors' interpretation of the interview data was unchanged.

Two overarching themes were identified:

- the utility of multisource feedback in appraisal; and
- linking revalidation, multisource feedback, and appraisal systems.

Utility of multisource feedback in appraisal

This theme reflected on the benefits (or otherwise) of using multisource feedback within appraisal. It comprised five sub-themes:

- benefits of incorporating multisource feedback in appraisal;

- validity of the colleague questionnaire;
- validity of the patient questionnaire;
- difficulty understanding and accepting 'low' benchmark scores;
- desire for qualitative feedback.

Benefits of multisource feedback in appraisal. All appraisers were supportive of using some form of multisource feedback to enable GPs to reflect on their practice, and all the doctors viewed multisource feedback as providing a useful opportunity to learn and develop. Some had used GMC multisource feedback data to change their practice:

'... the bit where I came into the lower quartile ... was the record keeping. So one of the things I've done since is that we have voice-activated software where we can dictate into our notes and I actually make much more comprehensive notes ... I'm quite up-front and honest and I said, "This is the bit, and this is what I thought about it, and this is what I've done as a result." To which he [the appraiser] *said, "Excellent, that's what appraisals are about; that's what feedback is about".'* (GP 107)

Although participants were supportive of using multisource feedback to guide learning and development, most expressed some concerns about elements of the GMC multisource feedback methodology, which might limit its use in appraisal. These will be dealt with in the following sections.

Validity of the GMC colleague questionnaire. Participants were concerned about the extent to which colleagues' questionnaire responses accurately reflected a doctor's performance; some colleagues might be unfamiliar with the doctor's clinical practice, making it impossible to provide accurate and informed feedback:

'I found the hardest part of it [GMC multisource feedback] [was] *nominating 20 individuals ... many of them would be unable to comment meaningfully on my clinical abilities.'* (GP 113)

Some interviewees expressed concerns that doctors identify colleagues rather than this being done on an independent basis. Interviewees felt that, intentionally or unintentionally, this may influence the feedback given:

'It's basically very flawed, partly because you're going to select people [colleagues] *that are going to give the answers that you want. Maybe not deliberately but because they're the people that you know. So it's not going to give a very true picture, necessarily, of your performance.'* (GP 108)

A few interviewees were concerned also that some colleagues may not complete the questionnaire honestly for fear that their negative (albeit truthful) feedback would have a detrimental impact on their colleague and/or undermine working relations:

'I was chatting about this ... at another surgery, and they said, "Oh yeah. I filled in 'excellent' for all my colleagues even though I don't think they are ... because I don't want them to feel dispirited because I've got to work with them".' (GP 109)

Validity of the GMC patient questionnaire. The extent to which patient responses are a valid reflection of a doctor's performance was raised. Some interviewees commented that unfavourable feedback might be received simply because of the patient group with whom doctors work:

'If you were in an area where there's an immigrant population with language problems, people who aren't literate, you're going to score very badly in this because the people who are coming in won't be the sort of people who are trying to be nice about you, don't care what they're doing, or they won't understand, and you'll be judged badly, not because your practice is poor but because your practice population is different.' (GP 108)

One participant believed that patient feedback might be a reaction to the 'interpersonal style' of the doctor rather than their skills, so a doctor who is well liked by patients may receive very positive feedback.

Difficulty understanding and accepting 'low' benchmark scores. Some doctors and appraisers found it difficult to understand the GMC benchmark data, stating it was unclear if a low benchmark score represented real cause for concern or if, as absolute differences in scores between the upper and lower quartiles were small, those differences had marginal importance:

'So quite what do those things [absolute scores and the quartile in which the score is located] *mean? Does that mean that there was a real issue? Did somebody spot some subtlety ... that it would have been useful for me to know about?* (GP 106)

Some participants reported that they, or a doctor whom they had appraised who ranked in the two quartiles below the sample mean (three or four) in one or more areas of their performance, had been very distressed:

'Of course, what happens is you get the Summary of Evaluation Results thing and you see "Benchmark Performance Band" and you see a "3" and you think, "They think I'm terrible!" In fact, I stopped reading it. I felt annoyed. I felt peeved. I actually put it away, I just thought, "I can't read this now".' (GP 109)

Desire for qualitative feedback. Some interviewees reported that the qualitative feedback was useful to help doctors understand the reasons why they had been scored as they had:

'The comments are probably as useful or more useful than the scores, really, because they do give you a little bit more flesh on the bones of the scores to look at what's behind them.' (GP 108)

Narrative comments, particularly those provided by patients, were seen as a rare source of direct encouragement:

'The free-text feedback was quite helpful. I think when you're working away at the coalface, you actually get very little praise for what you do in, often, quite difficult circumstances. And to actually see in writing that people are grateful for what you've done and grateful for the way you treat them, is actually something that I hadn't actually had in 20 years of being a doctor, so it's actually quite a boost really.' (GP 110)

Some interviewees reported that comments provided a clear indication of where someone could improve performance:

'That's the whole point of these issues, that you actually look at something that you thought was not an area of concern and to allow comment and change your practice ... I've shared it with the team and we're looking at how we might be more aware of that and what we might change".' (GP 103)

Linking revalidation, multisource feedback, and appraisal processes

Interviewees reflected on both the link between appraisal and revalidation, and the role of multisource feedback in appraisal and for revalidation. Participants' views were mainly regarding multisource feedback in general, that is, multisource feedback data collection using any appropriate survey instrument. A few participants expressed views that specifically focused on the GMC multisource feedback methods.

The likelihood of a link being established between appraisal and revalidation was widely accepted:

'I think it is a good way to do revalidation rather than have a test. So I think, I can only speak for our areas where our appraisal has really been quite — we've been in the forefront of appraisal and we have quite a good system in this particular area and people have done quite well. So, it does seem to work in our area. I don't know whether it will work in every area ... but I think it's a good thing".' (Appraiser 209)

Participants reported that multisource feedback could make a positive contribution to revalidation:

'I think it is good, politically, that the nation knows that it is being asked what it thinks of doctors, and that it knows that also our colleagues are asked to tell us what they think of us.' (Appraiser 208)

However, some expressed concerns that surveys may not be able to identify doctors who are performing poorly or those deliberately concealing dangerous performance:

'I'm just a bit afraid that we're all going to spend half our lives filling in forms about each other and there isn't very robust evidence that they're going to be useful in catching out people with problems. Lots of people will have said this but, you know, Harold Shipman could probably find 20 people to say nice things about him.' (GP 108)

One participant suggested that if multisource feedback is used for revalidation, it may deter responders from providing critical — albeit honest — feedback. Many interviewees perceived tensions between the formative role of appraisal (including multisource feedback) and the summative function of revalidation. If appraisal data are used as evidence for revalidation, some felt this would inhibit doctors from openly exploring difficulties or professional limitations:

'In terms of being an appraiser, I'm absolutely certain that, in a sense, the

formative value of appraisal is going to go down as a result of revalidation. People are going to be extremely careful what significant events they log and stuff like that.' (Appraiser 212)

One appraiser was already advising appraisees to exercise caution about what to write in appraisal documentation:

'I'm advising people not to put anything in [appraisal documentation] *that they might find a bit exposing, because the questions are a bit intrusive. Because, at the moment, as far as I can see, if there is a concern about the doctor they can subpoena the appraisal documentation, and I don't want to put anybody in a position where something that was written down ... could be used in evidence against them.'* (Appraiser 209)

However, another participant reflected that doctors should be prepared to receive honest, constructive feedback to improve their clinical practice:

'This is about people looking at their clinical practice, evaluating themselves, imposing change themselves, rather than it being a very judgmental process. Now, yes, somebody's going to look at it and ensure that your practice is safe and you keep up to date, look at what you're doing. I personally think that's only fair and reasonable. We're put in a huge position of trust and respect by patients and I think that's something that we don't deserve. We earn it by demonstrating things like this.' (GP 107)

DISCUSSION

Summary

GPs and their appraisers broadly supported the formative use of multisource feedback within appraisal, although most were concerned about some elements of the GMC multisource feedback methodology. Issues included the self-selection of colleagues, and patients' and colleagues' ability to provide objective feedback. Some reported difficulties in interpreting benchmark data and doctors were sometimes upset by their scores. Most participants accepted emerging links between appraisal systems in general practice and revalidation, although many identified tensions between the formative role of appraisal and the summative function of revalidation.

Strengths and limitations

By interviewing GPs and appraisers shortly before the introduction of revalidation, participants' views in this important area of changing regulation were elicited. The study purposefully recruited in sites with well-developed appraisal systems to enable participants to reflect on using multisource feedback in appraisal and revalidation, rather than focusing on problems with their appraisal process. Although recent international research has explored the relationship between multisource feedback, appraisal, and changing practice,[11,13,30,31] this exploratory study provides new insight into how the emergent plans for the revalidation of UK doctors might influence the relationship between multisource feedback and appraisal.

There are some limitations. This study recruited a small, self-selected sample of GPs and appraisers working in two areas with established appraisal systems. Participants' views may not apply to specialities other than that of family medicine, or to other UK primary care organisations — particularly those areas with less-developed systems of appraisal in place. Doctors who feel more confident about their professional performance may be more likely undertake multisource feedback;[22] likewise, participants holding polarised views about multisource feedback, appraisal, and/or revalidation may have been more likely to volunteer for this study.

Detailed plans regarding the mechanisms for implementing the revalidation process were published subsequent to this research.[1,2] Although these data provide important insight into the likely impact of revalidation, the participants not only reflected on a process they had not actually experienced, but also did so in the absence of a detailed understanding of how it might be implemented.

Comparison with existing literature

Some participants raised concerns about the validity of the multisource feedback methodology: some thought doctors might receive less-positive patient ratings because of the population with which they work, some questioned whether colleagues who were unfamiliar with a doctor's clinical practice were able to provide objective ratings, and others felt that 'self' selection of colleagues might introduce bias. Tensions were identified between maintaining good working relations with the doctor being assessed, the willingness to provide critical feedback, and its impact on the doctor.

In a parallel paper, using multivariate regression models, it was identified that less-favourable patient scores were predicted by doctors with lower proportions

of patients of white ethnicity who provided feedback.[32] For selected items from the colleague questionnaire, managers, administrative staff, and non-medical health professionals had more favourable views of doctors' performance than medical peers;[23] other research examining the impact of responder characteristics on ratings achieved by individual doctors has found similar results.[33–37]

Research working with volunteer samples of doctors suggests that an assessor's ability to provide accurate ratings may depend on their ability to observe relevant behaviours.[38] There is also some evidence that peers tend towards leniency to minimise bad feeling or to avoid damaging colleagues' careers.[15,39] Early work indicated that the self-selection of colleagues was not an issue of major importance,[40] although this has recently been contested.[41,42]

Some participants reported difficulties interpreting benchmark data and, in particular, whether a 'low' benchmark score represents a real cause for concern. Patient and colleague ratings collected using the GMC questionnaires[22,23] and multisource feedback tools of a similar intent[5,6,43–46] consistently report highly skewed, positive performance ratings. In this context, low scores identified by norm-referencing approaches may not necessarily reflect unacceptable performance.[23] It is unknown if this observation will apply when benchmarks are derived from non-volunteer samples.

Some participants reported that they or another doctor had been upset by their feedback, a finding that is consistent with previous research.[12,47] Sargeant *et al* found that among physicians who receive negative feedback, some eventually accept it; others do not, but express concern with the assessment process and/or describe barriers to change.[12]

Implications for practice

Doctors and appraisers value multisource feedback when it is undertaken within the context of formative assessment and when it is intended to contribute towards appraisal. However, concerns remain regarding aspects of the GMC multisource feedback methodology, which might undermine its credibility and also apply to other multisource feedback tools of similar intent. If multisource feedback, as part of learning and for revalidation, is to be maximised, guidance on its conduct and on how to interpret feedback is vital to support its acceptance and use.

Proposals linking formative appraisal to summative revalidation may limit both multisource feedback and appraisal for learning and development. Some doctors and appraisers may become more guarded regarding what is openly discussed and formally recorded in appraisal. One participant expressed concerns regarding the legal status of multisource feedback data — it may be wise for the UK to learn from an established multisource feedback process in Alberta, Canada, where, although doctors are mandated to undertake the feedback, primary legislation states such data cannot be subpoenaed in a court of law.[48] This approach may alleviate some UK doctors' concerns in this regard and enhance the formative potential of the multisource feedback process.

Funding

The study was funded by the UK General Medical Council as an unrestricted award (Reference: RR100218).

Ethical approval

The multisource feedback survey and this sub-study were considered by the Devon and Torbay NHS Research Ethics Committee but judged not to require a formal ethics submission.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors gratefully acknowledge the cooperation and support of the doctors and appraisers who contributed to this study, along with their patients, colleagues, and supporting administrative staff, who contributed to the main survey. We acknowledge Matthew Taylor and Louise Coleman (both from CFEP UK) in supporting the delivery of the project and we are grateful for the work of Antoinette Buisman and Emily Fletcher (both from the Peninsula College of Medicine and Dentistry) in coding a sample of the interview transcripts. We thank Professor Malcolm Lewis, who commented on the development of the project protocol and, in addition to him, are grateful to Nick Lyons and Susie Caesar for their help in finalising supporting materials and promoting the study.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: **http://www.rcgp.org.uk/bjgp-discuss**

REFERENCES

1. General Medical Council. *The Good Medical Practice Framework for appraisal and revalidation*. London: GMC, 2011.
2. General Medical Council. *Supporting information for appraisal and revalidation*. London: GMC, 2011.
3. General Medical Council, Health Foundation, Federation of State Medical Boards. *Revalidation: the way ahead*. Publication proceedings: International Revalidation Symposium: Contributing to the evidence base. London, 2–3 Dec 2010.
4. General Medical Council. *Annex 3- GMC principles, criteria and key indicators for colleague and patient questionnaires in revaldiation*. London: GMC, 2011.
5. Archer J, Norcini J, Davies H. Use of SPRAT for peer review of paediatricians in training. *BMJ* 2005; **330(7502):** 1251–1253.
6. Lockyer J, Fidler H. *Comparison of colleague and patient multisource feedback instruments designed for GPs in UK*. London: Royal College of General Practitioners, 2010.
7. Violato C, Lockyer J. Self and peer assessment of pediatricians, psychiatrists and medicine specialists: implications for self-directed learning. *Adv Health Sci Educ* 2006; **11(3):** 235–244.
8. Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. *BMJ* 2010; **341:** c5064.
9. Sargeant J, Mann K, Sinclair D, *et al*. Challenges in multisource feedback: intended and unintended outcomes. *Med Educ* 2007; **41(6):** 583–591.
10. Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness. *Med Educ* 2005; **39(5):** 497–504.
11. Sargeant J, Mann K, Sinclair D, *et al*. Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. *Adv Health Sci Educ* 2008; **13(3):** 275–288.
12. Sargeant J, Mann K, Van der Vleuten C, Metsemakers J. 'Directed' self-assessment: practice and feedback within a social context. *J Contin Educ Health Prof* 2008; **28(1):** 47–54.
13. Sargeant JM, Mann KV, van der Vleuten CP, Metsemakers JF. Reflection: a link between receiving and using assessment feedback. *Adv Health Sci Educ* 2009; **14(3):** 399–410.
14. Toegel G, Conger J. 360-degree assessment: time for reinvention. Academy of Management Learning & Education 2003; **2(3):** 297–311.
15. Wood L, Hassell A, Whitehouse A, *et al*. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. *Med Teach* 2006; **28(7):** e185–e191.
16. Kline T, Sulsky L. Measurement and assessment issues in performance appraisal. *Can Psychology* 2009; **50:** s161–s171.
17. Erdogan B, Kraimer M, Liden R. Procedural justice as a two-dimensional construct: an examination in the performance appraisal account. *J Appl Behav Sci* 2001; **37(2):** 205–222.
18. Elicker J, Levy P, Hall R. The role of leader-member exchange in the performance appraisal process. *J Manage* 2006; **32(4):** 531–551.
19. Stanton J. Reactions to employee performance monitoring: framework, review, and research directions. *Hum Perform* 2000; **13(1):** 85–113.
20. MacDonald H, Sulsky L. Rating formats and rater training redux: a context-specific approach for enhancing the effectiveness of performance management. *Can J Behav Sci* 2009; **41(4):** 227–240.
21. General Medical Council. *Good medical practice*. London: GMC, 2006.
22. Campbell J, Richards S, Dickens A, *et al*. Assessing the professional performance of UK doctors: an evaluation of the utility of the General Medical Council patient and colleague questionnaires. *Qual Saf Health Care* 2008; **17(3):** 187–193.
23. Wright C, Richards SH, Hill JJ, *et al*. Multisource feedback in medical regulation: the example of the UK GMC Patient and Colleague Questionnaires. *Acad Med* 2012; in press.
24. Kitzinger J. Focus groups with users and providers of health care. In: Pope C, Mays N, (eds.). *Qualitative research in health care*. Malden, MA: Blackwell Publishers — BMJ books; 2006: 20–29.
25. Taylor M. Interviewing. In: Holloway I, (ed.). *Qualitative research in health care*. Maidenhead: Open University Press, 2011: 29–55.
26. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; **3(2):** 77–101.
27. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess R, (eds.). *Analyzing qualitative data*. London: Routledge, 1994: 173–194.
28. Strauss A, Corbin J. *Basics of qualitative research*. 2nd edn. Newbury Park, CA. US: Sage Publications, 1998.
29. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *BMJ* 2000; **320(7227):** 114–116.
30. Cohen S. Assessing the assessments: UK dermatology trainees' views of the workplace assessment tools. *Br J Dermatol* 2009; **161(1):** 34–39.
31. Overeem K, Lombarts MJMH, Arah OA, *et al*. Three methods of multi-source feedback compared: A plea for narrative comments and coworkers' perspectives. *Med Teach* 2010; **32(2):** 141–147.
32. Campbell J, Roberts M, Wright C, *et al*. Factors associated with variability in the assessment of UK doctors' professionalism: analysis of survey results. *BMJ* 2011; **343:** d6212.
33. Mead N, Roland M. Understanding why some ethnic minority patients evaluate medical care more negatively than white patients: a cross sectional analysis of a routine patient survey in English general practices. *BMJ* 2009; **339:** b3450.
34. Taira DA, Safran DG, Seto TB, *et al*. Asian-American patient ratings of physician primary care performance. *J Gen Intern Med* 1997; **12(4):** 237–242.
35. Lipner R, Blank L, Leas B, Fortna G. The value of patient and peer ratings in recertification. *Acad Med* 2002; **77(10 Suppl):** S64–S66.
36. Wenrich MD, Carline JD, Giles LM, Ramsey PG. Ratings of the performances of practicing internists by hospital-based registered nurses. *Acad Med* 1993; **68(9):** 680–687.
37. Sargeant JM, Mann KV, Ferrier SN, *et al*. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. *Acad Med* 2003; **78(10 Suppl):** S42–S44.
38. Mazor K, Clausen BE, Holman M, Margolis M. Evaluation of missing data in assessment of professional behaviours. *Acad Med* 2007; **82(Suppl 10):** S44–S47.
39. Peiperl MA. Getting 360 degrees feedback right. *Harv Bus Rev* 2001; **79(1):** 142–147, 177.
40. Ramsey P, Wenrich M, Carline J, *et al*. Use of peer ratings to evaluate physician performance. *JAMA* 1993; **269(13):** 1655–1660.
41. Archer J, McGraw M, Davies H. Republished paper: assuring validity of multisource feedback in a national programme. *Postgrad Med J* 2010; **86(1019):** 526–531.
42. Bullock AD, Hassell A, Markham WA,*et al*. How ratings vary by staff group in multi-source feedback assessment of junior doctors. *Med Educ* 2009; **43(6):** 516–520.
43. Hall W, Violato C, Lewkonia R, *et al*. Assessment of physician performance in Alberta: the physician achievement review. *Can Med Assoc J* 1999; **161(1):** 52–57.
44. Lockyer J, Clyman S. Multi source feedback (360 degree evaluation). In: Holmboe E, Hawkins R, (eds.). *Practical guide to the evaluation of clinical competence*. 1 edn. Philadelphia, PA: Mosby Elsevier; 2008: 75-85.
45. Mackillop LH, Crossley J, Vivekananda-Schmidt P, *et al*. A single generic multi-source feedback tool for revalidation of all UK career-grade doctors: does one size fit all? *Med Teach* 2011; **33(2):** e75–e83.
46. Lelliott P, Williams R, Mears A, *et al*. Questionnaires for 360-degree assessment of consultant psychiatrists: development and psychometric properties. *Br J Psychiatry* 2008; **193(2):** 156–160.
47. Hawkins RE, Katsufrakis PJ, Holtman MC, Clauser BE. Assessment of medical professionalism: who, what, when, where, how, and ... why? *Med Teach* 2009; **31(4):** 348–361.
48. Government of Alberta. *Alberta Evidence Act. Revised Statutes of Alberta 2000*. Chapter A–18. Edmonton, Alberta: Alberta Queen's Printer, 2010.