

Title

Training in hospitals: what do GP Specialist Trainees think of workplace-based assessments?

Authors

Abigail Sabey, Michael Harris.

Abstract**Background**

Workplace-based assessment (WPBA) was introduced in 2007 as a new approach to monitoring competence of GP Specialist Trainees (GP STs). It includes a raft of assessments carried out in the workplace to assess what a trainee actually does in clinical practice. The assessment tools used are adapted from other contexts of doctors' training but little is known about how they function in day-to-day practice within GP training or how valid and useful they are found to be by trainees.

Aim

To establish how the new system of WPBA is working in day-to-day practice for GP STs in hospital posts.

Design

A mixed methods design including quantitative and qualitative phases of data collection.

Setting

Two training locations with Severn Deanery.

Methods

A questionnaire was completed by 52 GP STs (67% response rate) currently in hospital posts. Twenty-two took part in focus groups and semi-structured interviews to explore key findings from the questionnaire in greater depth.

Results

There is value in the face-to-face contact between trainees and senior doctors. However, quality and depth of feedback are not consistent and there is evidence of poor use of the tools, reducing the value of the assessments. The system is further undermined by a clear perception of bias and lack of honesty in judgements which limit the scope for assessment to lead to

learning. Overall, these weaknesses may impair the validity and usefulness of the system and its potential to improve the performance of doctors.

Conclusions

GP trainees in this study have a low opinion of how WPBA assessments function in the hospital setting. Changes are needed to optimise the potential of WPBA to improve the performance of doctors in training and to increase its credibility.

Keywords:

Professional competence; graduate medical education; general practice.

What is already known in this area

- Workplace-based assessment is a new approach to monitoring competence that has been widely implemented in specialist training, including with General Practice Specialist Trainees (GP STs).
- Knowledge is lacking about trainees' views of the assessments in day-to-day practice in GP training.

What this work adds

- Trainees find face-to-face contact and verbal discussions useful
- However, they place a low value on rating scale scores and they perceive a lack of honesty in assessments, as well as bias and a 'box ticking' attitude that undermines the credibility of WPBA.
- Change is needed to ensure trust in the assessment system, with emphasis on qualitative elements.

Suggestions for future research

- Further investigation into the views of assessors undertaking WPBA in hospitals will help to shape training for supervisors.
- Ways to optimise the formative use of WPBA need to be explored.

Introduction

General practice in the United Kingdom has seen the introduction of a competency-based training curriculum, spanning the three years of GP training in hospital and general practice posts,¹ and a new three-part end-point assessment process culminating in membership of the Royal College of General Practitioners, MRCGP. The three elements of assessment comprise:

an Applied Knowledge Test to assess a trainee's knowledge base; a Clinical Skills Assessment to assess a trainee's ability to integrate and apply the information and skills necessary for everyday situations in general practice; and thirdly, a system of Workplace-Based Assessment (WPBA).

WPBA is a system of performance assessment for doctors that is argued to distinguish between the 'can do' (competence) and the 'does do' (performance) of medical professionals.² GP Specialist Trainees (GP STs) are required to collate evidence of competence in core areas of the GP curriculum via assessments undertaken in day-to-day practice during hospital and GP placements. Standardised tools widely used in other contexts of doctors' training including the Foundation Programme^{3,4,5,6,7} are used to capture evidence. These include: Case-based Discussion (CbD), Mini-Clinical Evaluation Exercise (mini-CEX), Multi-source Feedback (MSF), and Direct Observation of Procedural Skills (DOPS). Specified numbers of each assessment are completed during each placement, with a Clinical Supervisor's Report obtained at the end.

The reliability and validity of the tools within doctors' training has been evaluated^{8,9,10,11,12,13} and they are considered fit for the purpose of WPBA² with some evidence supporting their use in speciality training in the UK including the medical specialties^{14,15} and the use of certain tools with GP registrars.¹⁶ However, there is no knowledge about how the tools actually function in the reality of daily practice for GP STs, or how valid and useful the assessments are perceived to be by trainees, especially during hospital posts. Whereas GP-based trainers are formally trained for the assessment role, in the hospital setting trainees have multiple assessors who may not be. Although change is expected, with standards for trainers in secondary care recently being set,¹⁷ anecdotal reports have suggested that assessments of GP STs in hospital posts are particularly problematic. Recent evidence of challenges in assessment during hospital-based Foundation training¹⁸ gave further impetus to focus this enquiry on the hospital setting.

This study aimed to establish how the new system of WPBA is working in day-to-day practice for a cohort of GP STs in hospital posts, seeking their views of the process and experience of assessments, their perceptions of assessors' understanding and skills, and their suggestions for improvement.

Methods

A mixed methods design, using quantitative and qualitative approaches, was undertaken in two phases. The design was informed by a participatory approach typical of action research, to maximise the co-operation and support of those whose views were being sought and produce a useful understanding.¹⁹

Participants

From two centres in the Severn Deanery, all GP trainees in their hospital posts (78 trainees) were targeted for Phase 1. In Phase 2, the same trainees were given the opportunity to take part in focus groups. One centre is in a large city with a medical school, the other in a smaller city with a district general hospital.

Phase 1 data collection

As no suitable questionnaire existed, a new tool was designed to gain information from the trainees in hospital posts about undertaking assessments and their overall views of WPBA. A small group of trainees was consulted about the selection of topics for the questionnaire to increase face validity. This confirmed the relevance of themes arising in previous research relating to assessment within the Foundation Programme.¹⁸ The final tool (Appendix 1) was reviewed by a panel of GP, education and research colleagues experienced in questionnaire design to increase content validity.²⁰

The questionnaire was distributed at GP education sessions between January and March 2009. Trainees not in attendance were sent a questionnaire to ensure that all had the opportunity to respond, helping to reduce bias. Further

responses were prompted by GP educators and/or email, with three reminders being sent. Fifty-two responses were received, giving a response rate of 67%. Twenty-six came from each centre; 38 (73%) were female, reflecting exactly the gender balance of the cohort as a whole, and 29 (56%) were in ST2.

Phase 2 data collection

Further information was sent to the 31 trainees who expressed an interest in Phase 2. During April and May 2009, twenty participants took part in three focus groups (15 females, 5 males). Recruitment to this phase was challenging, attributed to the busy time of year when many trainees were moving rotation, and on two initial occasions only one participant attended for the focus group (1 female, 1 male). In each case a semi-structured interview was completed as it was deemed valuable to include as much data as possible. One of these participants subsequently helped to recruit further participants for focus groups. Discussion in groups and interviews was based around key issues highlighted in the findings of Phase 1 (see Box 1) and this topic guide did not require refinement during data collection. All discussions were facilitated by the lead researcher (AS) and lasted 45-75 minutes. Discussions were recorded and transcribed verbatim, with all transcripts anonymised.

Box 1: Main topics addressed in focus groups/interviews

- What makes good written feedback?
- How are rating scales being used? How does this affect the value of WPBA?
- How could the assessment tools be made more relevant to general practice?
- How could the process of WPBA be improved?

Data analysis

Phase 1

The data from questionnaires were entered into SPSS 15 by AS. The analysis process and output files were reviewed by a research colleague for integrity and completeness.

Phase 2

A thematic content analysis of the data was undertaken, facilitated by QSR NVivo 2.0. Following the work of Barbour (2007),²¹ an initial stage of preliminary coding by AS generated 12 top-level codes, each having two to four sub-codes. Two transcripts were coded independently within the research team. A high degree of consensus was seen in the interpretation of transcripts and the coding frame was confirmed without changes. NVivo was used to extract and compile coded data for the final stage of interpretation led by AS, resulting in four main themes. The research team independently reviewed the final analysis with no changes made and a participant was consulted on the interpretation of the Phase 2 data, strongly endorsing the findings, making only two minor changes to wording. Further validation of the findings was achieved through an early presentation at a national conference. Trainers and colleagues working in WPBA were present as well as some trainees, with verbal feedback offering clear endorsement of the findings.

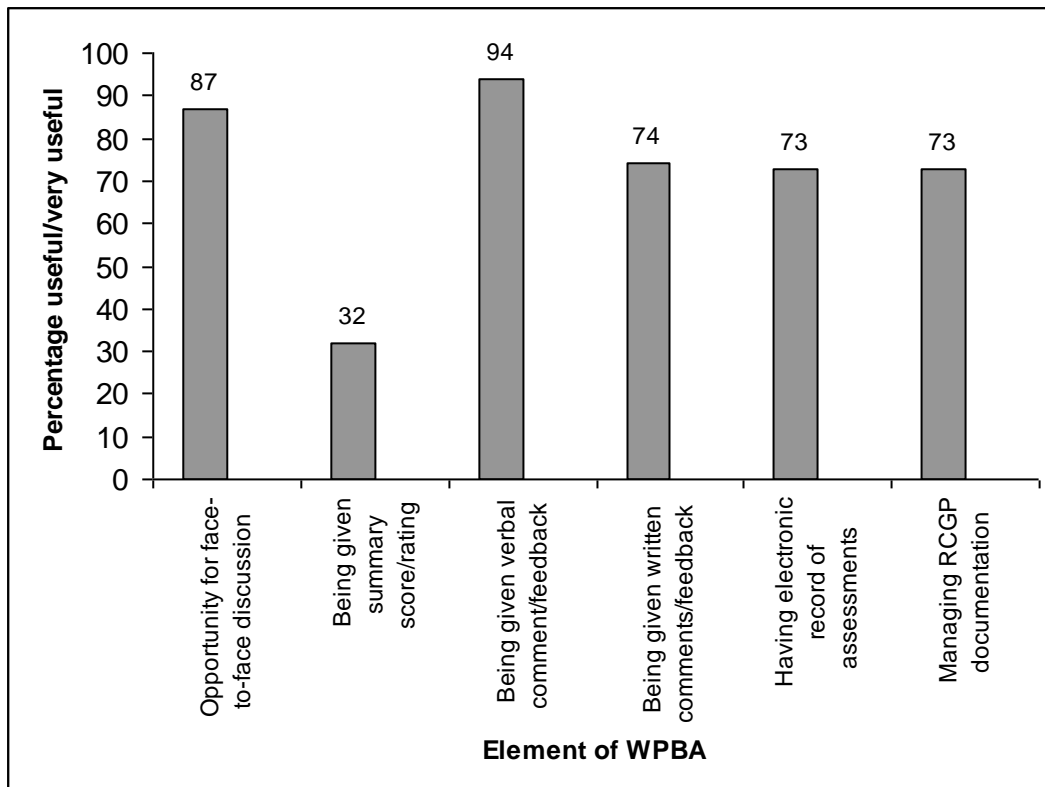
Results

Phase 1

Phase 1 yielded data on four main issues.

1. What different elements of WPBA do trainees find useful?

Figure 1. Trainees' views of usefulness of different elements of WPBA (%)

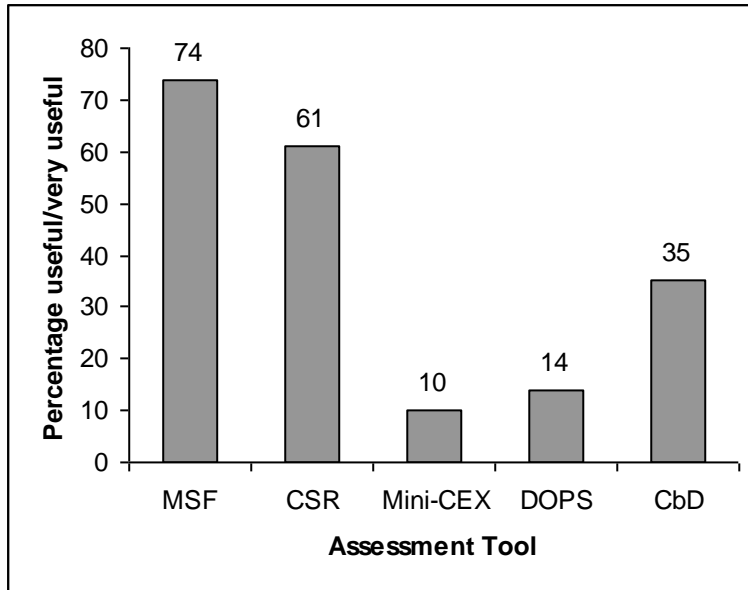


The face-to-face contact and discussion generated by assessment were highly valued by trainees. 87% of GP STs rated the opportunity for face-to-face discussion as very useful/useful and 94% rated verbal feedback the same way (Figure 1). 74% agreed that being given written comments/feedback was useful/very useful, with the electronic record of the assessments being similarly valued. In contrast, the summary scores/ratings generated by assessments were only thought to be useful by 32% of trainees, perhaps related to only 37% of trainees reporting that 'some or most assessors' used the full range of the scales (e.g. 'poor' to 'excellent').

2. How useful do trainees find the WPBA assessment tools?

The trainees had mixed views about the usefulness of the four assessment tools and the end-of-placement Clinical Supervisor's Report (CSR) (Figure 2).

Figure 2. Percentage of trainees who rated each tool as 'useful' or 'very useful'

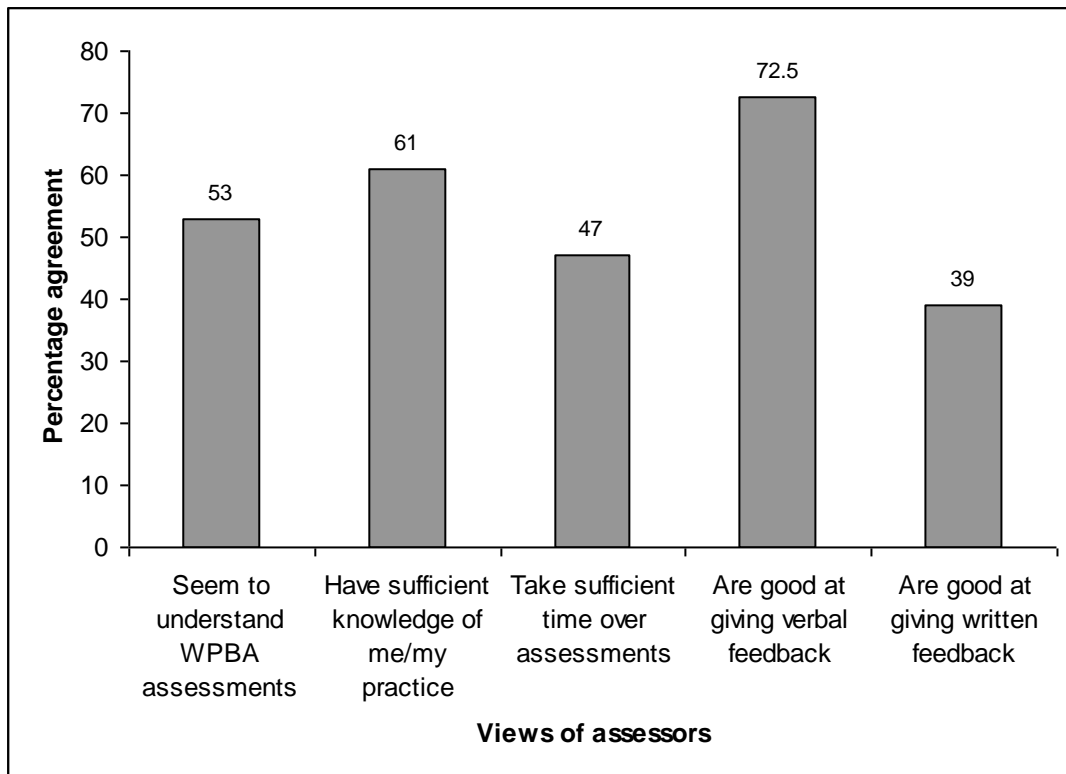


The majority of trainees found the MSF and CSR useful (74% and 61% respectively). However, the mini-CEX and DOPS assessments were only considered useful by a few (10% and 14%). Views about the CbD assessment were more mixed, with many (25%) expressing a neutral opinion. It is possible that the usefulness of CbD is related to how the tool is used. Only 2% of GP STs selected cases for CbD that had 'not gone well', though 55% reported that in day-to-day practice they learnt from such cases.

3. What do trainees think about their assessors?

Trainees experienced difficulties securing assessors for WPBA assessments, with 85% of the sample reporting difficulty in finding a suitable person at the time and/or getting their agreement to do the assessment. Trainees' judgements about their assessors' knowledge and skills are shown in Figure 3.

Figure 3. Trainees views of their assessors (% agreement)

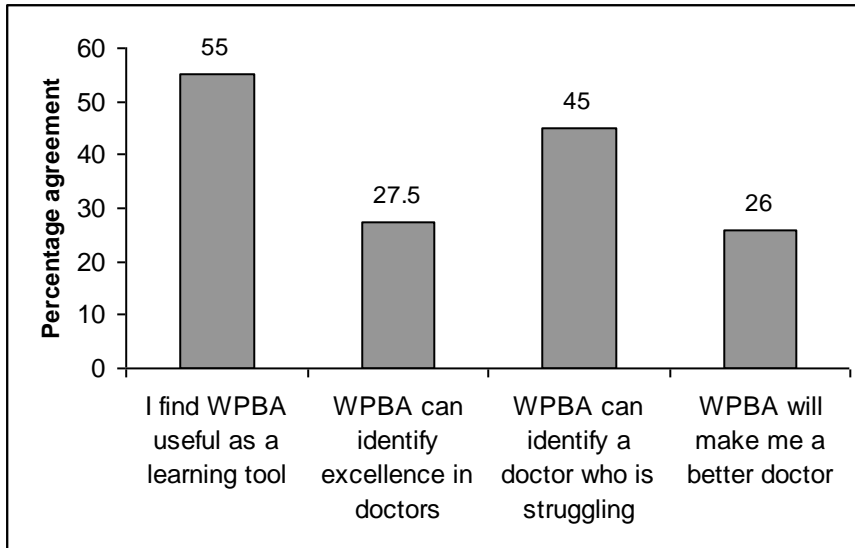


Only 53% agreed that assessors understood WPBA assessments. Trainees were divided as to whether assessors took sufficient time over assessments, with 47% agreeing that they did. A clear majority of GP STs felt that assessors were good at giving verbal feedback (72.5%), while only 39% of trainees agreed that assessors were good at giving written feedback. This stands in marked contrast to the finding reported above that, in principle at least, trainees found written feedback useful.

4. Is WPBA perceived to be useful and valid overall?

Trainees were asked to indicate their agreement with four statements about the usefulness and value of WPBA overall (Figure 4).

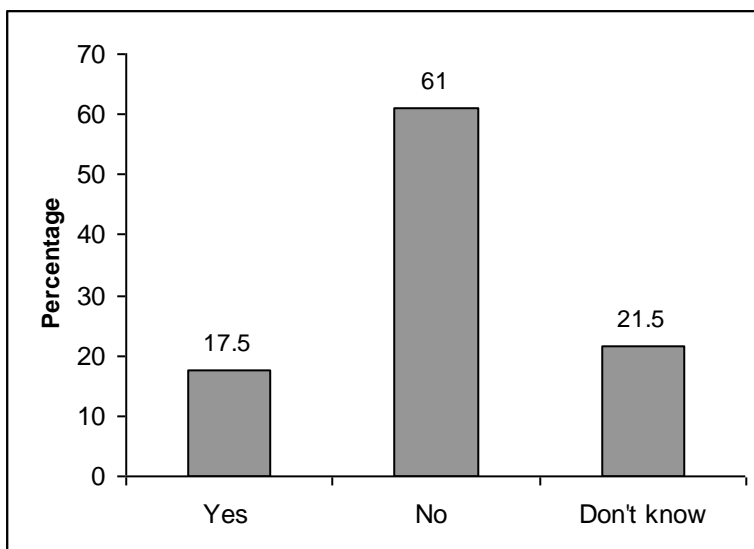
Figure 4. Trainees' views of WPBA overall (% agreement)



55% of trainees agreed that they found WPBA useful as a learning tool.

A minority (27%) of trainees agreed that WPBA could identify excellence in doctors and a similar proportion (26%) agreed that WPBA would make them better doctors. Just below half (45%) agreed that the system could identify a doctor who was struggling.

Figure 5. Do trainees agree that WPBA is a valid judgement of competency?(%)



Overall, only 17.5% of the trainees agreed that WPBA was a valid judgement of their competency (Figure 5).

Phase 2

The main themes arising in Phase 2 will be discussed, with quotes selected to illustrate typical views of trainees. Gender, year of training and training base (A=smaller centre, B=larger centre) are indicated in brackets after each quote.

Feedback

Phase 2 revealed that the trainees valued feedback for providing reassurance and giving a feeling of being appreciated. However, value was lost as feedback often came too late in the placement, especially in busier posts, leaving no time to act on it and giving assessors little reason to comment on what should be improved:

"[feedback given is] very little use because it's finished, it doesn't really matter" (F/ ST1/A)

A theme also evident was of poor quality feedback, especially in mini-CEX and CbD with quality being defined in terms of detail and individuality. GP STs reported receiving many similar comments across assessors and specialties, with lack of specific detail about performance:

"It's all just platitudes" (F/ST2/B)

"They'll write something a bit vague and a bit meaningless" (M/ST2/B).

"..'Needs more clinical experience' is generally what everyone writes"(F/ST1/A)

Those who provided more individual feedback were seen to stand out:

"..occasionally you will get registrars who will say 'oh you know, shall we do an assessment on that case?' And you think wow! And they do it and they will often write much more constructive comments that are more specific to you than the people that you go and nag essentially and those people are more few and far between." (F/ST2/A)

A case shared by one trainee, which did not surprise other participants, gives cause for concern about attitudes of some assessors, further undermining feedback. The assessor stated that all trainees would be rated as competent,

however well they did, and the trainee was signed off after only a month. Some trainees admitted to not actually being observed during the DOPS procedures. Again, this was not refuted by other participants:

“It’s very rare that somebody watches you for all the bits that actually they’re supposed to assess you for.” (F/ST2/B),

Related to this was the strong theme that trainees were doubtful about the level of honesty of feedback, particularly when given face-to-face in the mini-CEX and CbD:

“If everything was done face-to-face all you’ll get are usually people praising you.” (M/ST2/B)

The anonymous feedback provided by the MSF was felt to be more honest and, because of the larger number of people giving feedback, was perceived as more useful and valid. Putting more emphasis on the MSF as a *“more thorough review”* (M/ST2/A), and perhaps reducing the number of other assessments, were suggested as ways to improve quality in WPBA.

Bias

There was a perception that bias in assessment arose from trainees being able to choose their assessors. Participants openly admitted to choosing friends and this was felt to promote positive feedback, because of the need to preserve relationships among colleagues. This damaged the perceived validity of WPBA and some robust remarks were made in this regard:

“The whole tool is completely flawed because you choose your assessors.” (F/ST2/A)

“If you do just get your mates to do it, then it’s a waste of time.” (F/ST2/A)

A similar problem arose with CbD as trainees could choose the cases on which they were assessed. This resulted in trainees choosing positive cases rather than those they had found challenging. Having to select cases that have gone less well, or getting consultants to choose cases, would have made these tools more useful and put the emphasis on learning rather than assessing:

“It would then be their teaching agenda...it would work a lot better if the consultants picked the cases.” (M/ST2/B)

Use of ratings/scores in assessments.

A clear theme emerged relating to the inconsistent use of the rating scales, with some assessors judging the trainee in relation to their progress to date (i.e. for an ST1 or an ST2) and others in relation to the end-point of GP training – the latter being how the scales were intended to be used. This made it difficult to establish progress but also impacted on perceived validity:

“So you get these grades of excellent, but they’re a bit meaningless because they haven’t understood what’s behind it.” (F/ST2/B).

Some assessors appeared not to have understood the scales, which the trainees acknowledged may have resulted from hospital consultants having to judge expectations in relation to a fully qualified GP. Comments were made about being graded as ‘*excellent*’ against criteria which could not be assessed in hospital, for example against the criteria ‘*Primary care admin and IMT*’ and ‘*Community Orientation*’. A better match between assessment and placement is needed:

“I know we’re meant to be on a GP training scheme but at the end of the day you’re doing a hospital post and you’re seeing a hospital patient and it’s a hospital-based problem.” (F/ST2/A).

Suggestions included having separate forms for use in hospital and GP placements, and moving DOPS to be completed in the practice placements,

which were recognised to give greater opportunities to be observed on GP-orientated skills:

"[A GP is] "in a much better position to actually watch you and say what you can and can't do." (F/ST1/B)

Training was also implicated here with references to GP trainers, who were fully trained in WPBA assessment methods, using the scales correctly and giving more reliable feedback.

Box-ticking

The attitude of some assessors described above, together with impressions of poor understanding of scales and criteria, contributed to the view that WPBA was a 'box-ticking exercise'. Remarks about assessors ticking one category, such as 'meets expectations', all the way down the form were typical:

"The assessments say so much more about them than you really..like 'haven't got time to do this for you, I'll just tick the boxes.'"(F/ST2/B)

Some cited experiences of assessors wandering off half-way through assessments, with reluctance and time-pressures being blamed, leading to assessments based on just *"their instinct about you"*:

"It's difficult enough to get them even if they are just tick-box exercises, sometimes it takes five minutes to get somebody to do it properly, it's a lot of time." (F/ST2/B)

This fuelled a similar attitude among trainees:

"Everyone plays it down...and it just becomes a complete tick-box exercise...and nobody will ever look at it again" (F/ST2/B)

This was reinforced by the emphasis on scores and ratings on the forms, rather than feedback, particularly evident in the 'once only' basis of DOPS assessment:

"There's a million things in medicine that you invariably get better the more you do and you should really only be signed off when you're good enough to do it, not on your first one." (M/ST2/B).

Trainees consistently proposed fewer assessments in hospital posts with more opportunity for "qualitative" feedback that would lead to greater learning:

"You'd get a lot more comments about the things that you might be able to improve on without having to have said 'fail' or 'bad'." (F/ST2/B)

Discussion

Summary of main findings

The STs surveyed in our study make a clear distinction between those parts of the WPBA process they find credible and those that they do not. WPBAs enable regular, structured discussions about performance to take place between GP STs and hospital clinicians. Many of the trainees here rate such face-to-face contact and verbal discussions as being useful. However, they place a low value on rating scale scores, they hold negative views of Mini-CEX, DOPS and, to a lesser extent, CbD, and acknowledge the bias from being able to choose assessors. Overall, the data expose a clear message about a lack of honesty in assessments. This is compounded by perceived poor practice among assessors in written feedback, use of ratings, time taken over assessments, and evidence of cursory box-ticking judgements; as well as among trainees in selecting cases that by their own admission are not the ones from which they best learn. The assessments thus become a set of hoops to jump through, rather than a robust system for feedback and learning.

These findings have the potential to undermine the credibility of WPBA and to weaken a key assumption that WPBA leads to improvement in performance.

A recent systematic review by Miller and Archer (2010) of the impact of WPBA on education and performance exposes this weakness further, finding a lack of conclusive evidence that mini-CEX, DOPs or CbD result in improved performance and that doctors have a variable willingness to make changes as a result of MSF.²² Factors such as the context of feedback, facilitation and 'accurate and credible' feedback as highlighted by our trainees, are implicated in underpinning the link to performance improvement.

A further assumption in the move to WPBA was that this would bring objectivity in performance assessment but our data suggest that the qualitative elements may be more highly valued among trainees. Since this study, the shortcomings of WPBA have been more widely acknowledged with the Academy of Medical Royal Colleges (AoRMC) recently referring to a 'crisis in assessment', and highlighting the same failings in the objective elements of WPBA through 'reductive tick-boxing approaches'.²³ There is support by the AoMRC for a 'move from use of numerical values'²³ to a professional judgement of a doctor's competence. Van der Vleuten and Schuwirth (2005) make a similar call for assessment methods that rely on qualitative information and thus professional judgement.²⁴ Such change may help to capture those abilities which Rughani (2008) rightly points out are at the heart of being a GP for which 'counting is not enough'.²⁵

Strengths and limitations of the study

The good response rate strengthens the study. The inclusion of a qualitative phase of data collection, in which a high level of consistency was found among trainees, strongly endorsed the questionnaire data as well as adding depth. However, as the study consulted GP ST participants from two areas within one deanery, the findings are not necessarily generalisable to other geographical areas or other specialities. Furthermore, the study relies on the opinions of the STs themselves, and may not reflect those of other stakeholders.

However, other recent research offers much support for our findings. A recent study reporting on consultants' attitudes to WBPA for GP STs shows many

parallels including their acknowledgement of issues with time, the challenge of negative feedback and a need for training.²⁶ Archer et al (2010) recently identified biases among assessors, in particular relating to assessor seniority and length of working relationship, and agree that even the perception of bias can undermine assessment. Our work reinforces their conclusion that 'self-selection of assessors is no longer supportable.'²⁷ Johnson et al (2008) found poor assessor performance within a pilot of WPBA in core medical training, partly due to a lack of time, as seen here.²⁸ A survey of UK dermatology trainees' views of WBPA tools revealed problems with securing assessors and 'rushed' assessments but again, found the MSF was valued for the insight it provided.²⁹

Conclusions

WPBA may not deliver consistently high quality, objective and constructive feedback on performance from which trainees can learn and continually improve their practice. Our findings add weight to a recent call for 'a robust framework for quality assurance' within WPBA.³⁰ We have offered suggestions which may improve the effectiveness of WBPA as a formative tool and increase the trust of the profession in a significant element of the assessment process for GP training.

Box 2: Possible ways to improve the effectiveness of WPBA

- Suitable training for supervisors in secondary care, as called for in Lord Patel's recent review.³¹
- Re-designing some of the tools to make them more self-explanatory, so that training is less urgently required.
- Having generic assessment and software systems across the specialities.
- Ending the self-selection of assessors for multi-source feedback
- Involving consultants in the selection of more challenging cases for CbD.

- Identifying dedicated trainers who can become expert in completing the six-monthly reviews and be given dedicated time to review the performance of a larger number of STs.
- Reducing the number of assessments in hospital posts with more emphasis on assessments that are valued for qualitative and honest feedback, such as the MSF.
- Adjustments to tools so that there is less use of tick-boxes and rating scales and more use of questions like '*what went well?*' or '*what could have been done better and how?*' to encourage more useful feedback.
- Modifying WPBA forms for hospital posts, either so that these do not refer to criteria that can less commonly be assessed in a hospital post, or so that those criteria are labelled as 'optional'.

Ethics approval

The study was approved by the UWE Faculty Ethics Sub-Committee, ref HSC/08/10/78/. NHS ethics approval was confirmed as not being required for an educational evaluation.

Acknowledgements

We are very grateful to postgraduate centre staff for helping contact participants. We also wish to thank all the trainees who gave their time to the project, in particular the trainee who helped with Phase 2 recruitment. The contribution of Drs Kathy Pollard and Pam Moule to design and data analysis is also acknowledged.

Source of funding

Funding was provided by Severn Deanery.

Conflict of interest

None.

References

1. Royal College of General Practitioners (2007) *The RCGP Curriculum Statements*. At: www.rcgp-curriculum.org.uk/rcgp_gp_curriculum_documents.aspx. Accessed: 05/03/10.
2. Postgraduate Medical Education and Training Board (2008) *Standards for Curricular and Assessment Systems*. London: PMETB.
3. Davies, H., Archer, J., Heard, S., Southgate, L. (2005) Assessment tools for foundation programmes – a practical guide. *British Medical Journal*, vol.330: 195-196.
4. Jolly, B. (2007) The new curriculum framework and assessment practices: current challenges for postgraduate years 1 and 2. *Medical Journal of Australia*, vol. 186 (7 Suppl): S33-36.
5. Wilkinson, T (2007) Assessment of clinical performance: gathering evidence. *Internal Medicine Journal*, vol. 37: 631-636.
6. Hays, R. B., Davies, H. A., Beard, J. D. et al. (2002) Selecting performance assessment methods for experienced physicians. *Medical Education*, vol. 36: 910-917.
7. Rethans, J-J., Norcini, J. J., Baron-Maldonado, M. et al. (2002) The relationship between competence and performance: implications for assessing practice performance. *Medical Education*, vol. 36: 901-909.
8. Southgate, L., Cox, J., David, T. et al. (2001) The General Medical Council's Performance Procedures: peer review of performance in the workplace. *Medical Education*, vol.35 (1. Suppl): 9-19.
9. Norcini, J., Blank, L., Duffy, D., Fortna, G.S. (2003) The mini-CEX: a method for assessing clinical skills. *Annals of Internal Medicine*, vol. 138: 476-481.
10. Holmboe, E.S., Huot, S., Chung, J., Norcini, J., Hawkins, R.E. (2003) Construct validity of the Mini Clinical Evaluation Exercise (MiniCEX). *Academic Medicine*, vol. 78 (8): 826-830.
11. Wilkinson, J., Benjamin, A., Wade, W. (2003) Assessing the performance of doctors in training. *British Medical Journal Career Focus*, vol. 327: S91-92.

12. Ramsey, P.G. Wenrich, M.D., Carline, J.D., Inui, T.S., Larson, E.B., LoGerfo, J.P. (1993) Use of peer ratings to evaluate physician performance. *Journal of the American Medical Association*, vol. 269 (13): 1655-1660.
13. Norcini, J. (2003) Peer assessment of competence. *Medical Education*, vol. 37: 539-543.
14. Swanwick, T. (2009) Workplace-based assessment. *British Journal of Hospital Medicine*, vol. 70 (5): 290-293.
15. Wilkinson, J.R., Crossley, J. G. M., Wragg, A., Mills, P., Cowan, G., Wade, W. (2008) Implementing workplace-based assessment across the medical specialties in the United Kingdom. *Medical Education*, vol. 42: 364-373.
16. Murphy D.J., David. A. B., Mercer, S.W., Eva, K.W. (2009) The reliability of workplace-based assessment in postgraduate medical education and training: a national evaluation in general practice in the United Kingdom. *Advances in Health Sciences Education*, vol. 14: 219-232.
17. Postgraduate Medical Education and Training Board (2009) *Joint Statement from PMETB, AORMC, GMC, COPMED and NHS Employers in relation to the role and recognition of supervisors of postgraduate medical education and training*. London:PMETB.
18. Sabey, A., Feest, K., Gray, S. (2010) Lessons from the UK: Doctors' views of changes in postgraduate training. *Focus on Health Professional Education*, vol.11 (3): 42-51.
19. Green, J. and Thorogood, N. (2004) *Qualitative Methods for Health Research*. London: Sage.
20. Bowling, A. (2009) *Research Methods in Health: Investigating Health and Health Services (3rd Edition)*. Maidenhead: Open University Press.
21. Barbour, R. (2007) *Doing Focus Groups*. London: Sage.
22. Miller, A. and Archer, J. (2010) Impact of workplace based assessment on doctors' education and performance: a systematic review. *British Medical Journal*, vol.341:c5064
23. Academy of Medical Royal Colleges (2009) *Improving Assessment*. London: AoMRC.

24. Van der Vleuten, C.P.M. and Schuwirth, L.W.T. (2005) Assessing professional competence: from methods to programmes. *Medical Education*, vol. 39: 309-317.
25. Rughani, A. (2008) Workplace-based assessment and the art of performance. *British Journal of General Practice*, vol.58 (553): 582-584
26. Makris, J., Curtis, A., Main, P., Irish, B. (2010) Consultants' attitudes to the assessment of GP specialty trainees during hospital placements. *Education for Primary Care*, vol.21: 236-242.
27. Archer, J., McGraw, M., Davies, H. (2010) Republished paper: Assuring validity of multisource feedback in a national programme. *Postgraduate Medical Journal*, vol. 86: 526-531.
28. Johnson, G., Barrett, J., Jones, M., Parry, D., Wade, W. (2008) Feedback from educational supervisors and trainees on the implementation of curricula and the assessment system for core medical training. *Clinical Medicine*, vol. 8 (5): 484-489.
29. Cohen, S.N., Farrant, P.B.J., Taibjee, S.M. (2009) Assessing the assessments: U.K. dermatology trainees' views of the workplace assessment tools. *British Journal of Dermatology*, vol. 161: 34-39.
30. Mamelok, J. (2009) Workplace-based assessment (WPBA) portfolios in licensing for general practice specialty training. *Education for Primary Care*, vol. 20: 139-142.
31. Postgraduate Medical Education and Training Board (2010) *Final Report of the Education and Training Regulation Policy Review: Recommendations and Options for the Future Regulation of Education and Training*. London: PMETB.