

Workplace based assessment in clinical radiology – diverse practice and its implications

W. H. Ramsden and T. E. Roberts

Corresponding author: Dr. W. H. Ramsden ramsdendanburn@blueyonder.co.uk

Department: Consultant Paediatric Radiologist, Clarendon Wing X-Ray Department
The General Infirmary at Leeds, Belmont Grove, Leeds, LS2 9NS, United Kingdom

Received: 11/09/2015

Accepted: 14/09/2015

Published: 22/09/2015

Abstract

Objectives: In 2010 the Royal College of Radiologists (RCR) introduced workplace based assessment (WPBA) into clinical radiology training. The study purposes were firstly to investigate the implementation of WPBA by examining its day-to-day realisation in practice and to relate these findings to the guidance provided by the RCR. The second aim was to explore where the usage of WPBA differed from the RCR's declared intention and to examine the reasons for this deviation.

Methods: Semi-structured interviews were undertaken with 20 radiologists (12 trainers, 8 trainees) to explore their opinions on the implementation and usage of WPBA in radiology. The interview data were analysed using thematic analysis.

Results: The themes derived from the analysis were amending the assessments' purpose, altering WPBA's usage depending upon trainees' seniority or time in an attachment, amending the recording of assessments, and assessing trainees without their knowledge. In terms of amending the assessments' purpose, all interviewees thought that WPBA had a formative ethos but some assessors felt it could be used summatively in certain circumstances. The day-to-day realisation of WPBA varied, and although this might be seen as inevitable due to differing circumstances, there was evidence of some trainers consciously amending the process.

Conclusions: There was variation in the realisation of WPBA beyond that caused by their timing or context related to conscious adaptation of the process by trainers. Although such diversity of practice may or may not have impacted adversely upon the process, there is potential to introduce unfairness into WPBA and thereby disadvantage trainees.

Keywords: Workplace based assessment, implementation and adaptation.

Article

Introduction

Traditionally, formal assessment in post graduate medical education is summative in nature (1) and based upon professional examinations undertaken away from the workplace. In radiology trainees are examined in basic sciences, clinical knowledge, and interpretation of radiological examinations. Allied to these professional examinations are less formal assessments such as trainers' reports of trainees' progress following clinical attachments or placements. If these are all satisfactorily completed, they lead to certification that a doctor has finished their training and reached the standard required for safe independent practice.

Perceived weaknesses in these forms of assessment formed the main stimulus for the introduction of workplace based assessment (WPBA). These weaknesses included concerns that the system of formal examinations did not address trainees' professional practice in the clinical context (2), and the potential for bias and frequent lack of specificity in trainers' comments on trainees' ability. In addition the introduction of the European Working Time Directive (EWTD) leading to measures such as trainees working shifts, has reduced interaction between individual trainers and trainees (3), limiting the shared experience upon which to base such comment.

A further stimulus to the development of new forms of assessment has been the introduction of competency based curricula into postgraduate medical training, with increasing emphasis on the assessment of trainees' performance in the workplace in order to map their achievements to an overall competency framework (4). Assessment of competency has been addressed by the introduction of WPBA, with individual tools developed to allow the evaluation of different facets of trainees' performance.

The tools introduced by the Royal College of Radiologists (RCR) in 2010 include a Mini-Imaging Interpretation Exercise (mini-IPX) and Radiology Direct Observation of Procedural Skills (Rad-DOPS) which are episodic assessments of trainees' everyday clinical practice, and the utility of such methods in evaluating routine performance is endorsed by several authors (5, 6), particularly following multiple encounters.

Guidelines for the conduct of WPBA were issued by the RCR (7, 8). In conjunction with the College Curriculum (9), they provide fairly detailed information on the expected format of assessments (as below), and therefore guide how the RCR feel WPBA should be enacted during day to day practice. A standard form for recording assessments was also provided (10).

The formative ethos of the assessments is clearly stated in the College Curriculum, which describes WPBA (beneath the heading Formative Assessment) as; '*the cornerstone of assessment for day-to-day practice*' (9). The curriculum mandates that the assessor be trained in giving feedback and that this must be delivered after the assessment. It adds that the trainee is instructed to agree the case for assessment, its timing and who will assess them, with the caveat that assessors may initiate unscheduled assessments themselves.

In terms of undertaking the assessments, the guidance for conducting WPBA states that patients should be aware an assessment is occurring and the trainee should be directly observed whilst being assessed. Aside from the conduct of the assessment, the guidance is concerned with how to fill in the assessment form and rating of the trainee's performance.

Despite this central guidance, it is possible that there will be variation in how WPBA is realised by users, due to both differences in local fixed circumstances and other more dynamic factors relating to the timing, context and construction of the assessment episodes by the participants.

Addressing the diversity of assessment beyond that due to differing local circumstances, Delandshere and Petrosky (11) state that important aspects of performance vary according to the time and the context in which they are assessed. The theme was further developed by Govaerts et al (12) who argued that assessors of performance are not passive measurers of it, but instead more active participants whose behaviour is affected by both the assessment's context and their interaction with the trainee, suggesting that the two parties construct the assessment episode together.

Govaerts and Van der Vleuten (13) concur with Delandshere and Petrosky's (11) view that assessee's performance in WPBA may not be a stable situation, and is highly dependent upon the environment which surrounds the assessment episode. Assessee's performance is felt to be dynamic, affected by environmental factors, and based upon social and cultural circumstances, as well as the assessment's context. The authors challenge the assumption that competence is a fixed, permanent attribute and both it and performance are seen as variable and dependent upon circumstances which may be unique to particular situations, making it difficult to assume a trainee's presumed competence in one context is automatically transferable to another.

Allied to such variation are the conscious adaptations end users may make to WPBA and the enactment of assessment episodes. The tension between such actions and central guidance is the main issue investigated by this study. The interviews with participants allowed me to explore the nature of their conscious actions in producing variations in how the assessments were used and how they affected the process.

Method

The study was undertaken in a large teaching hospital Trust in Northern England, and participants were drawn both from this Trust and also (in the case of trainees) from surrounding acute hospitals included within a regional radiology training scheme. Ethical approval for the study was granted by the Leeds East Research Ethics Committee (reference 10/H1306/69). Following this, an invitation letter was sent to all the Trust's consultant radiologists and all trainees beyond the earliest stages of training, as these groups would have had the greatest experience of WPBA in radiology.

Of the 79 invitees, 20 radiologists (12 trainers, 8 trainees) agreed to join the study. The trainers recruited represented a wide range of radiological specialties and the trainees were spread between years 2 and 5 of training (of a 5 year programme). Of the trainers, 7 were male, as were 6 of the trainees. The ratios of trainer and trainee interviewees and male to female participants were both proportionate to those of the original group of invitees.

All subsequently gave informed consent to be involved according to National Research Ethics Service guidelines (14), and a single semi structured interview was undertaken with each participant. This method was chosen as best suited to obtaining individual's views regarding areas which particularly interested or concerned them whilst allowing some standardisation of the interview data. Following a small pilot study to refine the schedule, the same prompts and interview structure were used throughout the study.

WPBA and the new curriculum had been introduced prior to the interviews, and all participants had subsequently been involved with the process. The interviews concerned their experiences of the two

episodic assessments outlined in the introduction. During the interviews, particular emphasis was placed upon interviewees' perceptions of WPBA's purpose, its day to day realisation, and how participants might adapt the process. The interview data were subsequently transcribed and anonymised. Thematic analysis was employed to draw out areas of commonality and difference between and within the two groups of interviewees.

Results

The main finding of the interviews was the difference in views of WPBA's purpose as a formative or summative process and how this influenced the use of the assessments. This difference is exemplified by the following quotations from a trainee and a trainer, the former describing a formative purpose, the latter more summative usage;

'A teaching tool. At the end of the day that's what I believe it is. It's an assessment tool. I learn from my mistakes and it documents the progress that I'm making.' (Trainee 2)

Their prime function is about learning and becoming competent and I think mostly we learn and become competent by having some sort of test.' (Trainer 12)

The other key aspects of the results due to their significant effects on the process were the altering of WPBA's usage depending upon trainees' seniority or time spent in a clinical attachment, amending the recording of assessments, and undertaking them without trainees' knowledge.

Trainers' and trainees' views were compared during the data analysis, and as illustrated above, there were often contrasts between each group's views of the process. All the trainees perceived WPBA as a solely developmental (the more common term interviewees used) or formative exercise, a further example of their view being demonstrated by the following;

'Progress, development of the trainee, the basic skills and knowledge of a trainee' (Trainee 8)

Trainers mentioned a formative purpose less frequently, although they all acknowledged that WPBA had some developmental intent, and should be used to deliver feedback to help trainees improve and to direct their training. A representative example of this view was;

'It's primarily for me a method by which I can feed back to my trainee each week, every day hopefully. Feed back to my trainee, how did that go, how can we do that better, and next time we do this next week we'll do it this way.' (Trainer 4)

However, six of the trainers held the opinion that the assessments could also represent a test which could be passed or failed, positioning WPBA as a summative assessment of competence and limiting its developmental potential. Some stated that this purpose increased towards the end of a clinical attachment, as demonstrated by the following quotation;

'I suspect, near the end of the attachment, it does become more of an assessment, whereas in the beginning of the attachment, I think it's more teaching.' (Trainer 2)

When asked why they might use the assessments summatively, some trainers emphasised the need to ascertain trainees' level of performance, particularly those who might be performing poorly, two stating;

'I think they (workplace based assessments) are only useful if they identify someone who is failing to come up to a standard that you anticipate.' (Trainer 11)

and

'You know, you're really using this as an early-warning system for trainees that are struggling, or failing, and, in my experience, that may be 10% of trainees.'

(Trainer 1)

Some trainers also made the point that the nature of the marking schedule used to record the assessments influenced their decision to use them summatively, one saying;

'You can pass or fail, because the system of marking actually says, is this person performing at above or below the level that they should be expected to?'

(Trainer 11)

In addition to altering their use of assessments depending upon time spent in a clinical attachment (as above), some trainers amended their purpose depending upon trainees' seniority, as exemplified below;

'You see, again for the procedure I think is to have a competency to actually say if they can perform it on their own, at least for the older students. I think for the younger ones who haven't done many perhaps we just want to assess at what stage they are and see what needs they have.' (Trainer 8)

Another trainer who also pursued a similar approach explained that she expected greater depth of understanding and demonstration of their fitness to practice from senior trainees, saying;

'I expect junior trainees to know the basic, the core, aspects of the subject. Whereas senior trainees I expect them to know a more detailed aspect of exactly the same subject... That's when you look at the training more holistically, from all angles, and it's not just as a trainee who's come in "Hey well you've done this, and I've ticked your six boxes and finished."'

(Trainer 6)

There were instances where trainers would either not record the results of assessments, or where the written record might not truly represent what had occurred, as illustrated by the following;

'If a trainee's done a case very badly, I will say, you know, "Could do better. This is what he's done well; this is what he hasn't done very well." So I may verbally say to them, "Well, that was actually a disaster." But I won't write down on paper that that was a disaster. I'll write down what they did well and what they didn't do so well, and how they can improve.'

(Trainer 2)

This represents a different amendment from alteration of WPBA's purpose, and trainers who chose not to record assessments, or recorded them in a modified form, justified this by saying that they would only apply this to trainees' preliminary efforts which may not have gone well, as they did not wish to damage learners' confidence during early stages of training.

Another important issue was that other trainers sometimes undertook assessments without trainees' knowledge as they felt they could obtain a more naturalistic view of their work, one justifying this during the following exchange;

A: *'Yes, because tomorrow when the trainee goes in and does the cases on their own, every case will be different and they have to independently make all the decisions which are needed through the whole of the procedure.'*

Q: *'When will they know it's an assessment?'*

A: *'Afterwards; after the case has finished irrespective of the outcome.'* (Trainer 6)

This and the other amendments described represent the key areas in which there was conscious adaptation of WPBA. It was clear that the amendments were instituted by trainers, which is unsurprising as it was they who had control of the episodes. It should be emphasised that aside from amending the assessments' purpose, the adaptations were instituted by a minority of trainers, but the examples quoted illustrate the full range of amendment which occurred.

Discussion

The amendments to WPBA described in the results go beyond altering the assessments' purpose, and perhaps the most radical of all is the assessment of trainees without their knowledge. Beyond the implication that such changes may introduce unfairness into the process, there are the potentially more serious effects of disadvantaging trainees by affecting their records and impairing progression, which are discussed later.

Aside from the situations described in the results, and as highlighted in the introduction, there are many other factors contributing to variation in the realisation of WPBA, and it would be naive to think that the RCR would have been unaware of this. This is a likely reason for their issuing guidelines rather than rules for the assessments' conduct, as it is not possible to legislate for all circumstances in which assessment might occur. An example would be the differing sizes of, and facilities offered by, training schemes, where the availability (or not) of specialist faculty and assessment cases might greatly affect the conduct of WPBA.

The influence of the timing of WPBA and its context upon varying trainee performance should not be underestimated, but the importance of assessors being regarded as active participants (12) in the process and their significant role in the construction (and amendment) of assessment episodes is supported by the data. This particularly relates to the purpose of WPBA, its varying usage with different trainees, its recording and unannounced assessment.

The concept of assessors being active participants during assessment brings into focus their potential to reinterpret policy according to their own priorities and it has been argued that this is almost

inevitable due to varying local circumstances as articulated by Darling-Hammond (15) in the context of instituting teaching policy in schools, stating;

'local leadership and motivation for change are critical for policy success; that local ideas and circumstances always vary (therefore local agencies must adapt policies rather than adopting them).'

These numerous factors help to explain some of the inevitable heterogeneity apparent in the enactment of WPBA by end users, and the impossibility of standard interactions. In addition, the study was undertaken soon after the introduction of WPBA in radiology, when all assessors may not have been fully trained (another potential source of variation), although initial local training had been offered to all and guidance in the assessments' use was freely available on the RCR website. Despite possible limitations in their training, it is important to recognise that some assessors' beliefs and actions may have impacted upon central tenets of the scheme, such as its formative ethos.

The amendments to WPBA highlighted by the study all represent discretionary actions by trainers rather than enforced alteration to the process to accommodate outside circumstances. The data suggest that assessors' motives for undertaking such actions are variable, and whilst some may seek to enhance WPBA, others may use it for purposes for which it was not intended.

The differing perspectives of trainers (or those who administer training) and trainees may help explain why some trainers adapt the process. An example such as trainers using assessments summatively to identify underperformers, might be partly explained by trainers' responsibility for many trainees and perception of their role to uphold standards of clinical work, whilst trainees need only be concerned with their own progress.

A tendency for some trainers to utilise WPBA in the manner they see fit may also be derived from some medical staff (who will represent the vast majority of assessors) valuing autonomy, being resistant to change, and excluding themselves from policy implementation (16). Whatever the underlying motive, it is the trainer who leads the episode and adapting it carries the potential risk of unfairness to those being assessed and misinforming those who utilise the assessments' results.

These effects might be mitigated were WPBA always formative exercises without the recording of individual episodes, with any procedural variation agreed in advance by both trainer and trainee. However, the results are recorded in trainees' portfolios and become part of their permanent record. These records are reviewed by their trainers and supervisors when deciding if a trainee may progress, and they may be accessed by third parties should the trainee encounter difficulties or move between schemes.

It is apparent that the issues raised in the discussion have significant implications for the reliability and equity of trainee assessment and progression. How reliable are the findings? The limitations of the study include the fact that the interviewer had some managerial responsibility for the trainees, and this might have impacted upon their candour during the interviews, although they did not appear inhibited from criticising the assessment process. In addition, as participation was voluntary, the study sample was likely to have been skewed towards participants who were successful and enthusiastic, rather than those who were disengaged or performing poorly.

In terms of the generalisability of the study's findings, it is important to acknowledge that the study was undertaken in one training scheme. However, the guidelines were common to all schemes, and none of the amendments described appeared due to local circumstances. Thus although it is probable

that the study's findings are likely to be generalisable, this issue merits attention on a larger scale in the future.

Conclusions

There are numerous unavoidable variables impacting upon the realisation of WPBA related to both the participants and surrounding environment that make it difficult to deliver a standardised process. However, this study has demonstrated that superimposed upon these variables are instances in which trainers may undertake discretionary adaptation of WPBA to the potential disadvantage of trainees. These comprise adapting the assessments' purpose, amending the recording of WPBA and assessing trainees without their knowledge.

If some of these amendments were incorporated into the future development of WPBA they might not necessarily harm the scheme, but (if agreed) serve to enhance it. An example of this is the potential division of WPBA into summative procedural based assessments (PBAs) and completely formative supervised learning events (SLEs) which cannot be scored, ranked, passed or failed (17). Both of these have been considered by the RCR, and if instituted would serve to accommodate both formative and summative usage of WPBA without the ambiguity of purpose highlighted by some trainers' responses in the study.

However, other adaptations to the process such as amending the recording of WPBA and assessing trainees without their knowledge are far more difficult to accept due to their effect upon the assessments' fairness and potential to disadvantage trainees. These amendments have the potential to reduce users' trust in the recorded process, and as such, should be discouraged.

Acknowledgements: None

Funding: None

Conflicts of Interest: None

References

1. Burkill G J C. Work-based assessment for trainees – more than just a few new tools? *Clinical Radiology*. 2008; 63: 12-14.
<http://dx.doi.org/10.1016/j.crad.2007.07.014>
2. Harding J and McCoubrie P. Competency-based training versus traditional experience in radiology; how best to educate the radiologists of the future? *Clinical Radiology*. 2009; 64: 569-573.
<http://dx.doi.org/10.1016/j.crad.2009.02.006>
3. Temple J. Time for training: a review of the EWTD on the quality of training. Department of Health, London. 2010.
4. General Medical Council. Workplace Based Assessment: A guide for implementation. General Medical Council, London. 2010.
5. Wilkinson J R, Crossley J G M, Wragg A, Mills P, Cowan G and Wade W. Implementing workplace-based assessment across the medical specialties in the United Kingdom. *Medical Education*. 2008; 42: 364-373.
<http://dx.doi.org/10.1111/j.1365-2923.2008.03010.x>
6. Setna Z, Jha V, Boursicot K A M and Roberts T E. Evaluating the utility of workplace-based assessment tools for specialty training. *Best Practice & Research Clinical Obstetrics and Gynaecology*. 2010; 24: 767-782.
<http://dx.doi.org/10.1016/j.bpobgyn.2010.04.003>
7. The Royal College of Radiologists. mini-IPX Guidance for Assessors. The Royal College of Radiologists, London. 2010. Available at: www.rcr.ac.uk/docs/radiology/other/mini-IPX%20Guidance%20for%20Assessors%20v5.0.pdf (accessed 22/2/15)
8. The Royal College of Radiologists. Rad-DOPS Guidance for Assessors. The Royal College of Radiologists, London. 2010. Available at: www.rcr.ac.uk/docs/radiology/other/Rad-DOPS%20Guidance%20for%20Assessors%20v5.0.pdf (accessed 22/2/15)
9. The Royal College of Radiologists (Faculty of Clinical Radiology). Specialty Training Curriculum for Clinical Radiology. The Royal College of Radiologists. London. 2010.
10. The Royal College of Radiologists. mini-IPX and Rad-DOPS forms. The Royal College of Radiologists, London. 2010. Available at: www.rcr.ac.uk/docs/radiology/other/mini-IPX%20v5.0.doc and www.rcr.ac.uk/docs/radiology/other/Rad-DOPS%20v5.0.doc (accessed 22/2/15)
11. Delandshere G and Petrosky A R. Assessment of Complex Performances: Limitations of Key Measurement Assumptions. *Educational Researcher*. 1998; 27(2): 14-24.
<http://dx.doi.org/10.3102/0013189X027002014>

12. Govaerts M J B, Van der Vleuten C P M, Schuwirth L W T and Muijtjens A M M. Broadening Perspectives on Clinical Performance Assessment: Rethinking the Nature of In-training Assessment. *Advances in Health Sciences Education*. 2007; 12: 239-260.
<http://dx.doi.org/10.1007/s10459-006-9043-1>
13. Govaerts M and Van der Vleuten CPM. Validity in work- based assessment: expanding our horizons. *Medical Education*. 2013; 47: 1164-1174.
<http://dx.doi.org/10.1111/medu.12289>
14. National Research Ethics Service. Information Sheets & Consent Forms. Guidance for Researchers and Reviewers. Version 3.5. National Research Ethics Service, London. 2009.
15. Darling-Hammond L. Instructional Policy Into Practice: "The Power of the Bottom Over the Top". *Educational Evaluation and Policy Analysis*. 1990; 12(3): 339-347.
<http://dx.doi.org/10.3102/01623737012003339>
16. O'Connor N, Kotze B, Vine R, Patton M and Newton R. The emperor's edicts stop at the village gate. *Australasian Psychiatry*. 2012; 20 (1): 28-34.
<http://dx.doi.org/10.1177/1039856211432460>
17. Kessel D, Jenkins J and Neville E. Workplace based assessments are no more. *BMJ Careers* 29/9/2012: 5.