RadBench : Benchmarking Image Interpretation Performance

WRIGHT, Chris

Benchmarking Image Interpretation Performance

Dr Chris Wright PhD, MSc, HDCR, CertEd

Chris Wright is Senior Lecturer and Principal Researcher (RadBench) at Sheffield Hallam University, England (chris.wright@shu.ac.uk Direct: 0114 225 5488)

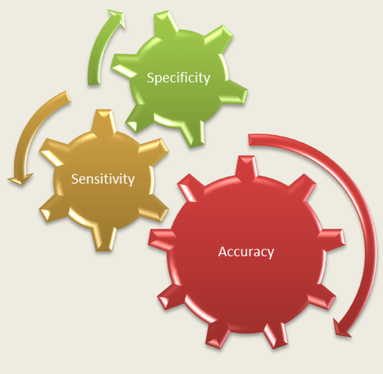
INTRODUCTION

Image interpretation has three distinct forms. The most basic form is the **Red Dot**, practised in 90% of UK hospitals (Snaith & Hardy, 2007). **Commenting** on trauma images was advocated by the CoR in 2006 and expected in the latest 2013 report. **Reporting** offers the potential to greatly reduce interpretation errors (Hardy, Snaith, & Scally, 2013) particularly if done at the time of imaging. Over the past decade numerous authors have carried out a wide range of studies to investigate the performance of different professions. Gold-standard accuracy of 95% is based on that of experienced consultant radiologists (Paterson, 2004; Nightingale, 2008). The vast majority of image interpretation studies have been on a small scale. This research aims to benchmark anyone taking the same test(s) regardless of profession, and provides a unique characterisation of an individuals image interpretation performance. Analysing error enables a development plan to be created as part of CPD, either individually or as part of a common theme within an organisation. Performance is typically tracked annually and certified. Respondents are able to benchmark themselves against others as desired selecting from a range of variables e.g. by hospital, geography, University, years qualified, profession etc

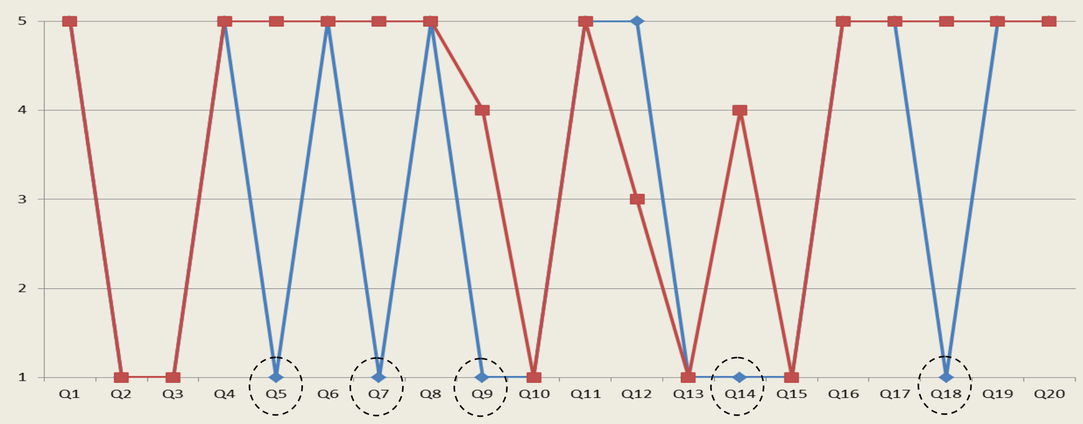
Several applications of RadBench have been tested; 1) Individual annual performance check as part of CPD, 2) Training planning for managers, 3) Selection differentiator for job interviews, 4) Selection for UCAS applications, 5) Anonymised global database to support other research projects. Further research is on-going.

Methodology

A quantitative approach is adopted using a series of image data banks (primarily MSK1.5 & MSK2.5). All data is anonymised. Respondents view a series of 20 images and select from a series of 5 decision states: additionally commenting on appearances. Performance is measured in terms of Accuracy, Sensitivity and Specificity. Additional analysis provides PPV, NPV and ROC.



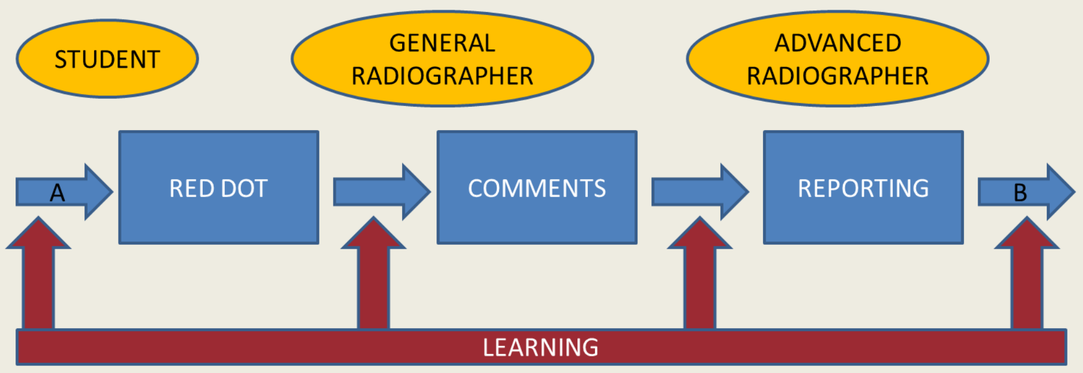
A performance grid of individual response against the expected provides a visual picture decision making performance which when cross referenced to the image database highlights development needs by anatomical area.



Re-test is typically done annually as part of CPD, but is also useful after a training intervention.

RESULTS & DISCUSSION

The majority of Radiographers fall into the 'non-reporting' category, yet are often expected to be able to express accurate opinions on the images they produce every day of their working lives. RadBench facilitates progression from Red Dot > Commenting > Reporting by identifying key talent and identifying training needs. It supports the DoH Return on Investment (ROI) strategy by benchmarking an individual pre and post training event.



The evidence highlights the benefits of benchmarking new graduates such that any development needs can be identified early in their career. Implementing the RadBench test as part of recruitment, ensures that managers have full visibility of current performance within the interviewees, and also how they perform within a global cohort with a similar level of experience.

This research demonstrates the idealistic nature of the generic gold-standard accuracy performance of 95%. Whilst the reporting radiographers and radiologists tested so far have on the whole met this standard, performance differs depending on the image set. Differences become emphasised with lesser experience. Different image sets can result in a difference in performance of up to 7%, even for experienced readers. The table below summarises the results from MSK1.5 (n=1876):

Mean Values	New Graduates	Experienced Radiographers	Reporting Radiographers
SENSITIVITY	85%	84%	95%
SPECIFICITY	69%	92%	95%
ACCURACY	77%	88%	95%

The typical emerging pattern is that without further training, respondents tend to be stronger at spotting abnormality (TP) than confirming normality (TN). This finding is consistent across multiple image banks. Image interpretation skills develop over time and can become highly tuned with focussed support. The first critical step is to be able to make the correct decision. Next is to increase confidence in decision making. Then develop written commenting and reporting. It is unrealistic in the short-term to expect all radiographers to be able to issue accurate reports, although this staged development process goes some way to ensuring that more accurate and reliable image interpretation decisions are made at all levels, which adds value and further credibility to the radiography profession.

Results from the testing of undergraduate UCAS applicants , with a dedicated image bank suited to novice level, demonstrated a positive correlation between RadBench score and selection event (interview) performance, providing a useful aid to student radiographer recruitment. For 2014 entry, all applicants will have the opportunity to take a specific RadBench test on-line and use their result to inform their personal statements. A similar approach could also be desirable for entry into other professions.

CONCLUSION

RadBench provides a reliable objective measure of image interpretation performance. A key contribution to knowledge is recognising variation in performance by image bank and the importance of benchmarking linked to a prescribed development pathway through CPD. RadBench has implications for practice at all stages of professional development. The results are of value to the individual, to organisations, governments, and to professional bodies. With increasing adoption, a life-long profile can be generated which will help inform clinical training and practice on a global scale. Partnering with other institutions enables the rich data source to generate and support further research. The on-line product enables wide reach to all imaging professions regardless of geography. A multi-lingual version is in development.

REFERENCES

College of Radiographers (2006) 'Medical image interpretation & clinical reporting by non-radiologists: the role of the radiographer'. SCoR. London

Hardy,M. Snaith,B. & Scally,A. (2013) 'The impact of immediate reporting on interpretative discrepancies and patient referral pathways within the emergency department'. BJR. 86. 20120112. 1-8

Nightingale,J. (2008) 'Developing protocols for advance practice'. Radiography. 14. 55-60

Paterson,A. (2004) 'Reporting by radiographers: a policy and practice guide'. Radiography. 10. 205-212

Snaith,B. & Hardy,M. (2007) 'Radiographer abnormality detection schemes in the trauma environment - an assessment of current practice'. Radiography. 1. 277-287

Society of Radiographers (2013) 'Preliminary clinical evaluation and clinical reporting by radiographers'. SCoR. February. London.

RadBench is a joint venture initiative between Sheffield Hallam University and Papaya UK

now available on-line at **www.radbench.org**