EDITORIAL

Benchmarking in radiology: apples and oranges?

DANIEL BIRCHALL

Neuroradiology Department, Regional Neurosciences Centre, Newcastle General Hospital, Westgate Road, Newcastle-upon-Tyne NE4 6BE, UK

Received 23 September 2009
Revised 6 October 2009
Accepted 8 October 2009

DOI: 10.1259/bjr/31880882

© 2010 The British Institute of Radiology

There is little formal coordination or understanding between UK radiology departments in terms of standards relating to departmental work output or consultant workload. As a result, there are likely to be significant variations in work patterns, individual consultants' workload and work intensity in different units around the country.

To some extent, these differences will reflect variations in centre size, nature of clinician base and type of imaging performed. However, it may well be that work patterns are more optimised in some centres than in others. In such a situation, it is in the interest of individual radiology services, trusts and consultants that 'best practice' be identified and aspired to [1].

Benchmarking provides a means for contextualising an individual department's workloads and work patterns, and in doing this, it can assist with operational management and strategic planning. This is particularly important, given the increased demand for services arising from new imaging imperatives (for example, providing imaging for acute stroke), the loss of support from independent imaging providers and from internal pressures to enhance efficiency.

Benchmarking is a central part of organisational strategic management and an effective means of optimising operational efficiency and strategic planning. It is a process that compares business activities within similar organisations questions how organisations operate, identifies opportunities for improvement and provides the momentum necessary for implementing positive change.

Most benchmarking compares key performance indicators and tends to focus on objective measures such as productivity and efficiency. Although these terms are not naturally associated with clinical activity, the key activities in radiology departments can be identified

and measured to develop quantitative benchmarking comparisons. These include average output per consultant for diagnostics and interventions, consultant workload, and average output per session – consultant work intensity.

These measures are fundamental to strategic planning and to the optimal management of operations within radiology departments. They are quantifiable, and potentially suitable for benchmarking, as part of a strategic drive for service improvement and optimisation.

Benchmarking of service productivity has been used extensively within the health service – for example, in the assessment of outpatient clinic throughput and operating theatre utilisation. Latterly, some cross-trust benchmarking has been applied to clinician 'productivity', and this has included relating output figures to sessional salary.

Although internal comparative 'performance data' is available in most trusts, there are few mechanisms for comparing the performance differences between different hospitals and trusts. In most cases, such cross-trust clinical benchmarking has been facilitated by management consultant houses. This may have the benefit of impartiality, but the lack of direct clinician involvement remains a significant issue for the validity of such studies.

A crucial component of any benchmarking must be to ensure, as far as possible, that performance data is comparable between organisations – that apples are not being compared to oranges. In many, and possibly all, instances of clinical benchmarking, a high degree of clinical insight is necessary to ensure that comparative observations and conclusions are valid. This also includes ensuring that the complexity of the individual clinical settings is fully appreciated.

It is arguable that benchmarking studies of clinician performance are likely to have limited validity unless there is a high level of clinician involvement. Similarly, the extent to which clinicians will engage with the conclusions of benchmarking studies is likely to depend on the credibility of the reports' authors.

Address correspondence to: Daniel Birchall, Neuroradiology Department, Regional Neurosciences Centre, Newcastle General Hospital, Westgate Road, Newcastle-upon-Tyne NE4 6BE, UK. E-mail: Daniel.birchall@nuth.nhs.uk

The current financial climate emphasises to an even greater degree the necessity of incorporating 'lean thinking' into all aspects of medical practice. That is to say, to eliminate as far as possible inefficient work practices and to optimise workflow processes [1, 2].

Given that radiological equipment makes up some of the highest capital equipment costs within hospital environments, it is likely that the pressure to demonstrate optimal use of imaging hardware will increase and that a high level of political capital will rest on this optimisation [3]. However, there is a relative lack of meaningful performance benchmarking between radiology departments.

There are certainly high profile publications that assess or 'benchmark' gross departmental function. For example, departmental output for individual modalities, operational hours and 'examinations per annum per full time equivalent radiologist' [4, 5]. But there is very little comparative information on how effectively and efficiently each 'work hour' is used within radiology departments. Perhaps this could be termed 'benchmarking effectiveness'.

There are undoubtedly major issues to take into account in proposing that the effectiveness and efficiency of radiology services should be benchmarked. These centre on the comparability of data. Radiologists operate within a highly complex environment, involving speciality and sub-speciality practice variations, a variety of hospital settings and clinical pressures, and individually varying levels of 'extra-core' commitment. These aspects have reduced the validity of previous attempts to provide meaningful benchmarking data between departments. For example, relating to Körner data and to efforts within the USA, Canada, Australia and New Zealand to assess radiologist 'productivity' with RVUs (relative value units) [6, 7].

This complexity makes it imperative that any comparative study should necessitate a very high level of clinical radiologist involvement. This contrasts to the consultancy house model currently being used in the UK – not only to ensure validity, but also to maximise acceptance of study conclusions within the medical body.

Such issues are stressed in a recent report by the British Society of Cardiovascular Imaging (BSCI) [8], in which they state that "the profession should lead this process to ensure a fair and equitable benchmarking method". The BSCI also go on to suggest a system for weighting examinations, in an analogous way to the RVU system.

There has recently been a Neuroradiology Benchmarking Project in the UK comparing the efficiency of selected neuroradiology units [9]. This study was driven by clinical neuroradiologists and was stimulated, in part, by the inability of consulting houses to provide meaningful comparisons in the field, such studies have previously had poor uptake, poor data and poor insight into radiology service complexities.

The project's objectives were to establish a group of neuroradiology centres motivated to exchange information as part of a benchmarking process; to collate, analyse and report benchmarked information among the group. In doing so, this would assist individual departments in their operations and strategic planning.

The critical aspect of the project centred on 'data capture'. Several units were unable to supply any data on performance and productivity, indicating a deficiency

in basic information reporting. Among the remaining units, there was variety in the way in which activity is recorded – for example, as 'attendances' or as 'examinations'. The degree of complexity of data recording also varied – for example, whether post-contrast CT or MRI scans were counted as one or two examinations.

Therefore, considerable effort was made to ensure that, as far as possible, like was being compared with like in the way in which activity was recorded. Similarly, it was evident that there is considerable variation in the way in which job plans and consultant work patterns are structured, which further complicated the measurement of 'consultant effectiveness'.

Nevertheless, it was possible to co-ordinate a group of units with comparable ways of recording activity and with similar consultant work patterns, and to derive comparative information relating to average consultant workload and work intensity within individual units. This information was qualified to a significant extent by the recognition of the inherent complexities within the speciality. For example, the study did not take into account the complexity of examination type; the existence of other clinical or professional commitments; the practice of 'multi-tasking' or flexible work patterns; the presence of modern, optimally efficient scanner hardware; the reliability and speed of PACS and RIS support; and the level of support from ancillary professional groups. But perhaps the most important variable that was not included, and probably the one most difficult to define, was quality.

However, the project succeeded in generating a broad-brush comparative study that had a level of validity and acceptability which was greater than that available from other sources. This initial study did not identify the processual reasons for work pattern variations (*i.e.* why things operate the way they do) and this remains an important next step in the analytical process. Indeed, this should be an important goal of any effort to benchmark 'effectiveness' and 'efficiency' in radiological practice.

Although benchmarking studies are often directed to identifying 'best practice', this term will have different meanings for different departments (*e.g.* depending on size and specialities supported). Furthermore, 'best practice' will be interpreted differently according to the perspective of the observer.

For example, 'best practice' to an individual consultant may be to have a relatively low figure for work intensity, so that sufficient time is allowed for thoughtful interpretation and reporting. Alternatively, an administrator may see 'best practice' as being as high a figure as possible, so that throughput is maximised within a resource-limited environment. A metric 'examinations reported per hour' could therefore be termed as 'work intensity' or alternatively as 'efficiency'. There are no standards for 'best practice' in this respect, but it could be argued that mid-range work intensity would be an optimal level, taking into account the balance between service needs and individual consultant workloads.

It is very likely that individual radiology departments will come under increasing pressure to demonstrate that they are providing 'value for money'. Clinical radiologists need to take the lead in optimising their services and a necessary part of this process is to benchmark services with other similar departments. It may well be that this is

best achieved through sub-speciality interest groups. However, there is a danger that this process could be driven without significant clinical input, which runs the risk of generating conclusions that lack insight and validity, but which, nevertheless, carry political clout.

It is important that clinical radiologists are aware of the medico-political issues that face the profession – particularly in a period of increasing demand and potentially decreasing resource. It is also important that they address these issues proactively and maintain high levels of involvement to promote a constructive and inclusive approach to workforce planning and process optimisation.

References

1. Boland GW. Government reform of the National Health Service: implications for radiologists and diagnostic services. *Br J Radiol* 2006;79:861–865.
2. Department of Health. Healthcare output and productivity: accounting for quality change. Department of Health, London, 2005.
3. NHS Scotland Diagnostics Steering Group. A report from the Diagnostics Steering Group setting out an approach to benchmarking radiology services. The Scottish Government, Edinburgh, 2009. <http://www.scotland.gov.uk?Publications/2009/09/02090823/0>
4. Board of the Faculty of Clinical Radiology, The Royal College of Radiologists. Clinical Radiology: a workforce in crisis. Royal College of Radiologists, London, 2002.
5. Board of the Faculty of Clinical Radiology, The Royal College of Radiologists. How many radiologists do we need? A guide to planning hospital radiology services. Royal College of Radiologists, London, 2008.
6. Pitman AG, Jones DN. Radiologist workloads in teaching hospital environments: measuring the workload. *Australas Radiol* 2006;50:12–20.
7. Bhargavan M, Sunshine JH. Workload of radiologists in the United States in 1998–1999 and trends since 1995–1996. *AJR* 2002;179:1123–1128.
8. British Society of Cardiovascular Imaging. Benchmarking in Cardiovascular Imaging. 2008 http://www.bsci.org.uk/downloads/cat_view/45-benchmarking
9. Birchall D, Straiton J. Neuroradiology Benchmarking Project. Unpublished data, 2009.