Comparing the accuracy of initial head CT reporting by radiologists, radiology trainees, neuroradiographers and emergency doctors

^{1,2,3}F A GALLAGHER, PhD, MRCP, FRCR, ¹K Y TAY, FRCR, ^{3,4}S L VOWLER, MSc, ¹H SZUTOWICZ, HDCR(D), HPC, ¹J J CROSS, FRCR, ⁵D J MCAULEY, MRCP, FRCS(A&E), FCEM and ¹N M ANTOUN, FRCR

¹Department of Radiology, Addenbrooke's Hospital, Cambridge, UK, ²Department of Radiology, University of Cambridge, Cambridge, UK, ³Cancer Research UK Cambridge Research Institute, Cambridge, UK, ⁴Centre for Applied Medical Statistics, Department of Public Health and Primary Care, Institute Of Public Health, Cambridge, UK, and ⁵Department of Emergency Medicine, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Objectives: Demand for out-of-hours cranial CT imaging is increasing and some departments have considered addressing this shortfall by allowing non-radiologists to provisionally report imaging studies. The aim of this work was to assess whether it is appropriate for non-radiologists to report head CTs by comparing the misreporting rates of those who regularly report head CTs with two groups of non-radiologists who do not usually report them: neuroradiographers and emergency doctors.

Methods: 62 candidates were asked to report 30 head CTs, two-thirds of which were abnormal, and the results were compared by non-parametric statistical analysis.

Results: There was no evidence of a difference in the score between neuroradiographers, neuroradiologists and general consultant radiologists. Neuroradiographers scored significantly higher than senior radiology trainees, and the emergency doctors scored least well.

Conclusion: The results of this preliminary study show that appropriately trained neuroradiographers are competent at reporting the range of abnormalities assessed with this test and that their misreporting rates are similar to those who already independently report these studies.

Received 16 October 2009

Revised 8 February 2010

Accepted 1 April 2010

DOI: 10.1259/bjr/24581602

© 2011 The British Institute of Radiology

The workload of radiology departments has increased in recent years and the availability of trained radiologists is not always sufficient to meet this demand, particularly in the context of out-of-hours emergency imaging [1, 2]. Some departments have considered addressing this shortfall by allowing referring clinicians or neuroradiographers to provisionally report imaging studies. A major concern with employing non-radiologists to report such studies is the potentially serious and costly consequences of an incorrect report; therefore, for this approach to be widely used it would be necessary to assess an individual's performance using a validated testing procedure. Furthermore, an acceptable level of reporting error would have to be established, and it would be reasonable to assume that this misreporting level could be arrived at by assessing radiologists who regularly report such films. Validated tests are also required to evaluate new teaching methods in radiological training as well as in continuing medical education.

A specific example of reporting by non-radiologists is for out-of-hours cranial CT, the demand for which is increasing: in some centres requests for these are vetted by neuroradiographers (neurotechnologists) and provisionally reported by clinicians who are not radiologists [3, 4]. A few studies have addressed the head CT misreporting rates of radiology trainees (residents or specialist registrars) and emergency doctors but these have recruited only small numbers of candidates and no study has accurately assessed the ability of neuroradiographers to report head CTs. Furthermore, in all of these studies the standard against which candidates are judged has been the report of the consultant (or attending) neuroradiologist rather than the true gold standard diagnosis made from the cumulative results of clinical findings, blood tests and radiological findings. Consequently, no study has assessed the accuracy of consultant neuroradiologists in reporting head CTs.

The purpose of this study was to compare the misreporting rates for cranial CT between consultant neuroradiologists, radiology trainees, general radiology consultants who regularly report these studies, neuroradiographers and emergency doctors. In particular, this study aimed to assess whether these results provide evidence for whether it is appropriate for non-radiologists to provisionally report on-call emergency head CTs.

Address correspondence to: Dr F Gallagher, Department of Radiology, University of Cambridge, Box 218, Level 5, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK. E-mail: fag1000@cam.ac.uk

Methods and materials

Head CT cases

30 admission head CTs were chosen from patients who presented to the emergency department of a large teaching hospital over a 3 month period. 20 of these films showed cranial abnormalities and these are listed in Table 1. In order to maximise the difference in the outcome of the test between the various groups of candidates, some difficult films were deliberately selected. The films were chosen by a senior consultant neuroradiologist and reviewed by two radiology trainees who did not participate as candidates. Films were only included in the study if the abnormality was deemed unequivocal on CT; only films reported by this senior consultant neuroradiologist were chosen to avoid recall bias by one of the candidates in the test. Importantly, all abnormal films were chosen to demonstrate conditions that could result in significant clinical adverse effects if missed on imaging; some representative images are shown in Figure 1. Before inclusion in the study, the diagnosis in each case was confirmed using subsequent imaging, other clinical features, *e.g.* surgical findings, and laboratory results where available. The remaining 10 films were included in the test as "normal" cases but were only considered for this category if patients had no significant neurological deficit, the films were deemed unambiguously normal by a senior neuroradiologist and the patients had no further neurological admissions or imaging over a 6 month period; such studies were performed for minor head trauma or headache.

Candidates

A total of 62 candidates were recruited: 4 consultant neuroradiologists working in a tertiary referral centre; 18 general radiology consultants from 3 local district general hospitals who routinely report head CTs; 7 neuroradiographers who specialise in neuroimaging (6 from a teaching hospital and 1 from a local district general hospital); 14 senior radiology trainees from a local radiology training scheme who had at least 18 months of radiological experience (second- to fifth-year trainees); 11 first-year radiology trainees with 6 months' experience in radiology; and 8 emergency doctors who had no specific radiological training (3 consultants and 5

trainees). There was 100% recruitment of the available neuroradiologists and neuroradiographers, 78% recruitment of general radiologists in the 3 local hospitals, 88% recruitment of senior trainees and 79% recruitment of first-year trainees from the training scheme. Emergency doctors were recruited following a weekly teaching session and 73% of those attending participated. For those groups in which 100% recruitment was not achieved, potential candidates were either unavailable or declined to take part.

The consultant neuroradiologists had at least 10 years' experience in the subspecialty and all the consultant radiologists had 5 or more years' experience in radiology. Of the emergency doctors, the 3 consultants had at least 5 years' experience in the emergency department and the 5 trainees all had less than this. The neuroradiographers had been practising for at least 1 year as a specialist neuroradiographer with a median experience of 7 years (range 1–28 years). Six of the seven radiographers who took part worked full time in neuroradiological imaging, which mostly involved CT. The final neuroradiographer worked in a local district general hospital as a CT radiographer, had trained as a neuro-radiographer, and a substantial proportion of this neuro-radiographer's job involved head CT imaging. One of the neuroradiographers had a qualification in adult head CT reporting. The more junior neuroradiographers had worked alongside their senior colleagues in a continuous learning environment in CT. All of the neuroradiographers and the emergency doctors in the study had had a lecture from a consultant neuroradiologist before the test took place, but they had not been shown the cases used in the trial.

Both groups of consultant radiologists independently reported head CTs, and films were not routinely double reported in any of the centres included in the study. 56% of the senior radiology trainees reported on-call emergency head CTs, although all studies were subsequently reviewed the next morning by a consultant neuroradiologist. The first-year radiology trainees and emergency doctors did not routinely report head CTs.

Head CT test

A single sheet of printed images for each case was shown to the candidates in appropriate viewing conditions. Only a single-window setting was printed for each

Table 1. Abnormalities present in the cranial CT test series and their frequency. The total number of cases was 30 and 1 case demonstrated 2 abnormalities: a tentorial subdural haemorrhage and a subarachnoid haemorrhage

Abnormality	Number of cases
None: normal cranial CT	10
Malignancy: primary/metastases	4
Cerebral infarct	3
Subarachnoid haemorrhage	3
Pituitary/suprasellar mass	2
Unilateral subdural haemorrhage	2
Middle cerebral artery (MCA) thrombosis (dense MCA)	2
Tentorial subdural haemorrhage	2
Bilateral subdural haemorrhage	1
Diffuse cerebral oedema (pseudo-subarachnoid)	1
Herpes simplex encephalitis	1

Figure 1. Four representative images from the abnormal series of head CT cases. (a) Bilateral isodense subdural haemorrhage. (b) Left middle cerebral artery infarct. (c) Corpus callosal mass. (d) Dense left middle cerebral artery.

case, but this was optimised so that the abnormality was visible at the presented setting. The only clinical details provided for each group was the age of the patient (age range 1–84 years) and the candidates were asked to review the films within 1 h. They were instructed that between one-third and two-thirds of the cases were normal. If the candidate identified a case as abnormal, he or she was asked to state in a few words what the abnormality or diagnosis was. Full marks were awarded for a correct diagnosis; half marks were given for a partially correct diagnosis, *e.g.* identifying the abnormality but misinterpreting it; and no marks were awarded for an incorrect answer, or if the case was listed as abnormal without stating the abnormality. The maximum score on the examination was 30.

Data analysis

Graphical representation of the data and statistical analysis were carried out using the SPSS software

package (V13.0; SPSS Inc., Chicago, IL). Median marks were compared between groups and were plotted graphically using boxplots. Statistics were performed by a medical statistician. Comparison between groups was made using non-parametric methods as the group sizes were small: the Mann–Whitney *U*-test was used to compare two groups and the Kruskal–Wallis test to compare more than two groups. A Jonckheere–Terpstra trend test was performed to assess for a trend between the outcome on the test and the level of radiological experience for each group.

Results

The median results (with ranges) were: neuroradiologists, 24 (22–26); neuroradiographers, 23 (19.5–27.5); general consultant radiologists, 23 (15.5–29); senior radiology trainees, 15 (9.5–23); first-year trainees, 12 (6.5–15); emergency doctors, 8 (6–10.5). The median scores for each group with the corresponding interquartile ranges are

shown as a boxplot in Figure 2. Using Kruskal–Wallis analysis, there was an overall highly significant difference in head CT scores between the six groups; given this, it was possible to do *post hoc* pairwise tests to look for differences between the groups.

Although the consultant neuroradiologists had the highest median score, there was no evidence of a difference between the neuroradiologists, general radiologists and neuroradiographers ($\chi^2_{\text{kw}}=0.36$; degrees of freedom=2; $p=0.83$). Therefore, there was no evidence of a difference between the neuroradiographers, who do not routinely report cranial CTs, and the consultant radiologists, who do. Using Mann–Whitney *U*-test analysis, the neuroradiographers scored significantly higher than the senior radiology trainees ($U=5.0$, $p=0.0003$). There was a 5% chance that a neuroradiographer would have a head CT score lower than a senior trainee. Even if only those trainees who independently reported head CTs on-call were included (8 of the 14 senior radiology trainees), the neuroradiographers still tended to score higher ($U=4.0$, $p=0.029$). Therefore, the neuroradiographers performed better on the test than a group of senior trainees who already report head CTs on-call.

The scores for the radiologists on the test corresponded well with their level of seniority in radiology: the general radiology consultants scored higher than the senior radiology trainees ($U=17.0$, $p<0.0001$), who, in turn, scored higher than the first-year trainees ($U=25.0$, $p=0.003$). There was a 16% chance that a senior trainee would have a lower head CT value than a first-year trainee. Even if a Bonferroni correction was used to correct for the multiple comparisons performed, this did not change the study outcome or which results were

considered significant. Jonckheere–Terpstra analysis, which assesses the significance of increasing median scores, showed that there was a significant trend for an increasing median mark over the groups: first-year trainees, through senior trainees to consultant radiologists ($JT=562$, $p<0.0001$). This, therefore, also confirmed that the outcome on the test is associated with experience in radiology.

The emergency doctors scored least well, and Mann–Whitney *U*-test analysis showed that there was a significant difference in the head CT scores between these candidates and all the other groups combined ($U=25.0$, $p<0.0001$). When the results for the emergency doctors were compared with the first-year radiology trainees there was only a borderline significant difference ($U=23.5$, $p=0.09$), but this became highly significant when compared with the more senior radiology trainees ($U=1.5$, $p<0.0001$). There is a 1% chance that a radiology trainee with at least 1 year of experience in radiology will have a lower head CT score than an emergency doctor. However, it should be noted that there was only a small number of emergency doctors sampled in this study.

No individual scored full marks on the examination, but there was no single film that every individual misreported. In particular, there was no case that even all four of the consultant neuroradiologists reported completely incorrectly.

Discussion

Many studies have assessed the competence of non-radiologists to report various imaging investigations. These have usually been undertaken to determine whether it may be appropriate for such individuals to provisionally report radiological studies and therefore help to meet the increasing demand for on-call radiology. For instance, several studies have evaluated the discrepancy rates between radiologists and emergency doctors in reporting trauma radiographs, demonstrating misreporting rates in the region of 1–3% [5–7]. In the context of cranial CT, previous studies have assessed the clinically significant non-concordance rates between the reports of radiology trainees and neuroradiologists, showing these to be between 1.7% and 5.7% [8–10]. Higher rates of non-concordance have been demonstrated between the provisional reports of consultants in the emergency department and the final report: 13.3–24% [9, 11, 12]. A single study assessing the difference between the reports of board-certified general radiologists and neuroradiologists showed a discrepancy in 5%; 40% of these were deemed clinically significant [13]. Such discrepancies may have serious clinical implications given that these reports, unlike those of radiology trainees and emergency doctors, are not subsequently checked by a neuroradiologist. Finally, despite the recent interest in cranial CT reporting by neuroradiographers [3], few studies have assessed report discrepancy rates in this group: 1 study assessed 2 radiographers' reports for paediatric head CT and found the accuracy in detecting significant abnormalities to be between 49% and 75% [14]; the only other study has assessed a single radiographer and shown a report sensitivity of 85.4%, which compared favourably with the sensitivity for senior trainees: 87.5% [15].

Figure 2. Boxplots of the head CT scores for each of the six groups. The line represents the median score out of a total mark of 30 (30 films were used); the box indicates the interquartile range and the whiskers show the spread of results. The number of candidates in each group is provided. *p*-values are shown for the non-parametric statistical comparison of adjacent groups (see text for details).

All these studies have recruited only small numbers of candidates: the maximum number of participants in previous studies assessing head CT reporting has been 18 [8]. Furthermore, no study has simultaneously assessed all the appropriate groups using one test. Finally, previous studies have assumed that the neuro-radiologist's report is the gold standard. This study has addressed these issues by recruiting a very large number of candidates from six groups of individuals who are directly involved in the requesting, acquisition and reporting of cranial CT studies. However, similar numbers could not be obtained in each group owing to the availability of candidates within the local area. The simultaneous evaluation of these groups against a clinical gold standard has allowed the direct comparison of all the relevant categories. Recruiting large numbers of candidates has been made possible by using a relatively small number of cases, thus reducing the time commitment required for the participating individuals. Consequently, some difficult abnormalities were selected to allow discrimination between the groups of candidates. This differs from many of the previous studies which used larger numbers of randomly selected or consecutive cases; it is likely that our selection of CTs is similar to the relatively small number of missed cases found in previous studies. This would also explain the apparently higher misreporting rates for the candidates in our study than much of the previous work in this area.

To assess how well a neuroradiographer or an emergency doctor may report head CTs, it is necessary to validate the test used by demonstrating that the outcome is related to the level of experience in radiology for the other groups sitting the examination. Here we show that there is an association between scores on this test and seniority in radiology: consultant radiologists scored significantly higher than senior radiology trainees, who in turn scored higher than first-year trainees. Although neuroradiologists had the highest median score, this showed no evidence of being higher than that for the general consultant radiologists.

No evidence of a difference in score was noted between radiographers specialising in neuroimaging and consultant radiologists who regularly report head CT studies independently (both neuroradiologists and general radiologists). This preliminary study provides supporting evidence for the role of appropriately trained neuroradiographers in provisionally reporting head CTs on-call; such reports could be checked the following day by a neuroradiologist in a similar way to the process undertaken in many centres for radiology trainees. In fact, the neuroradiologists scored significantly higher in this study than the trainees who are already given the responsibility of provisionally reporting head CTs out of hours. The performance of the neuroradiographers may partially be accounted for by the lecture on emergency head CT imaging they received a few weeks prior to the study. However, the emergency doctors undertook a similar lecture and scored significantly lower. It is likely that the competence of an individual radiographer is related to his or her level of experience, which ranged widely in this study; despite this, the range of scores was relatively narrow, suggesting that even a single year of experience as a neuroradiographer may produce a sufficient level of competence for reporting. Even though

the first-year trainees had a similar level of radiological experience, it is likely that they did not perform so well because of the range of modalities and body areas that they have to gain experience in, as opposed to the more focused approach of training for neuroradiographers. A future trial could assess the reporting accuracy of neuroradiographers from provisional reports of routine head CTs performed in a department over a long period of time. Ideally, this would be a multicentre study and would enable the assessment of the incidence of clinically significant errors.

The small number of emergency doctors recruited performed least well on this test, which is not unexpected given their limited exposure to the specialty of radiology. With appropriate training it may be possible for this group to safely report cranial CTs in conjunction with a neuroradiologist.

Finally, even the neuroradiologists did not score 100% on the test. The reasons for this may be related to the limitations of the study, which include the lack of clinical details provided, inability to manipulate the display settings of the images, lack of further imaging (e.g. following contrast medium), no access to previous films or insufficient time to report the studies. Although these may have contributed to the result, no single film was completely misreported by all four consultant neuroradiologists. Such an outcome suggests that it is inappropriate to use a single individual as a reporting gold standard, as has been used in some previous studies. Furthermore, it provides some evidence that "double reporting" head CTs by two consultant radiologists may be beneficial. This approach has been suggested for other radiological investigations [16], and, although it may increase the workload in a radiology department, it is likely to improve diagnostic accuracy. Further studies are required to assess the incidence of clinically significant errors in the reporting of head CTs, its impact on patient management and the cost-effectiveness of double reporting [17].

In summary, this preliminary study shows a relationship between misreporting rates and seniority in radiology. No evidence of a difference is noted between neuroradiographers and those who routinely and independently report cranial CTs, providing support to suggest a role for them in provisionally reporting cranial head CTs on-call after appropriate training. No individual reported all the films correctly, and further work is required to assess the clinical impact of such errors on patient outcome; if possible, double reporting of head CTs by two radiologists should be considered. Furthermore, this work has the potential to be used in continuing medical education, to assess an individual prior to reporting on-call studies or as a tool to assess the efficacy of different teaching methods used in training radiologists.

Conclusion

This study suggests that appropriately trained neuroradiographers have similar detection rates to consultant radiologists for the limited range of abnormalities assessed in this study; neuroradiographers also outperformed radiology registrars and emergency doctors.

A future large multicentre trial would be helpful in evaluating these findings further.

Acknowledgments

We are grateful to all the radiologists, emergency doctors and neuroradiographers who participated in this study. The centre has funding from the National Institute for Health Research, Cambridge Biomedical Research Centre.

References

1. Harrigal CL, Erly WK. On-call radiology: community standards and current trends. *Semin Ultrasound CT MR* 2007;28:85–93.
2. Saket DD. The provision of emergency radiology services and potential radiologist workforce crisis: is there a role for the emergency-dedicated radiologist? *Semin Ultrasound CT MR* 2007;28:81–4.
3. Friedenbergr RM. The role of the supertechnologist. *Radiology* 2000;215:630–3.
4. Mukerji N, Wallace D, Mitra D. Audit of the change in the on-call practices in neuroradiology and factors affecting it. *BMC Med Imaging* 2006;6:13.
5. Klein EJ, Koenig M, Diekema DS, Winters W. Discordant radiograph interpretation between emergency physicians and radiologists in a pediatric emergency department. *Pediatr Emerg Care* 1999;15:245–8.
6. Thomas HG, Mason AC, Smith RM, Fergusson CM. Value of radiograph audit in an accident service department. *Injury* 1992;23:47–50.
7. Williams SM, Connelly DJ, Wadsworth S, Wilson DJ. Radiological review of accident and emergency radiographs: a 1-year audit. *Clin Radiol* 2000;55:861–5.
8. Erly WK, Berger WG, Krupinski E, Seeger JF, Guisto JA. Radiology resident evaluation of head CT scan orders in the emergency department. *AJNR Am J Neuroradiol* 2002;23:103–7.
9. Mills TD, Jones JO, Evanson J. Accuracy of interpretation of emergency cranial computed tomography scans by radiology residents and the senior attending member of the requesting clinical team. *Radiology* 2001;221(RSNA abstract suppl.):440.
10. Wysoki MG, Nassar CJ, Koenigsberg RA, Novelline RA, Faro SH, Faerber EN. Head trauma: CT scan interpretation by radiology residents versus staff radiologists. *Radiology* 1998;208:125–8.
11. Alfaro D, Levitt MA, English DK, Williams V, Eisenberg R. Accuracy of interpretation of cranial computed tomography scans in an emergency medicine residency program. *Ann Emerg Med* 1995;25:169.
12. Mucci B, Brett C, Huntley LS, Greene MK. Cranial computed tomography in trauma: the accuracy of interpretation by staff in the emergency department. *Emerg Med J* 2005;22:538–40.
13. Erly WK, Ashdown BC, Lucio II, RW, Carmody RF, Seeger JF, Alcala JN. Evaluation of emergency CT scans of the head: is there a community standard? *AJR Am J Roentgenol* 2003;180:1727–30.
14. Brandt A, Andronikou S, Wieselthaler N, Louw B, Kilborn T, Dekker G, et al. Accuracy of radiographer reporting of paediatric brain CT. *Pediatr Radiol* 2007;37:291–6.
15. Craven CM, Blanshard KS. Computed tomography (CT) head scans reported by an experienced CT radiographer. *Radiography* 1997;3:105–11.
16. Markus JB, Somers S, O'Malley BP, Stevenson GW. Double-contrast barium enema studies: effect of multiple reading on perception error. *Radiology* 1990;175:155–6.
17. Brown J, Bryan S, Warren R. Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms. *BMJ* 1996;312:809–12.