Accuracy of detection of radiographic abnormalities by junior doctors

C. A. VINCENT,1 P. A. DRISCOLL,2 R. J. AUDLEY1 & D. S. GRANT3

1Department of Psychology, 2Accident and Emergency Department and 3Department of Radiology, University College Hospital, London, England

SUMMARY

This study assessed the ability of junior doctors in accident and emergency to detect radiographic abnormalities. Their assessments of 505 radiographs taken at nights and weekends over a period of 8 months and showing abnormalities were examined. Each assessment by a senior house officer (SHO) was compared with the subsequent diagnosis of a radiologist of senior registrar or consultant status. An error rate of 35% was found. For abnormalities with clinically significant consequences the error rate was 39%. Although this error rate appears high the results are consistent with those of earlier studies in that missed positive radiographs constitute 2-8% of the total number of radiographs taken in the period. It is considered that the proportion of missed abnormalities gives a truer index of SHOs' abilities. No improvement in performance was evident over the 6-month period of the SHOs tenure of post. It is argued that it is unrealistic to expect accident and emergency SHOs to acquire this complex skill simply through experience and that more formal training and guidance is needed.

INTRODUCTION

A number of previous studies have reviewed or re-reviewed radiographs originally examined by junior doctors in the course of their duties. The results of these studies are summarized in Table 1. A major aim of most of these studies has been to assess the usefulness of radiologists with heavy workloads, re-examining a large number of mostly normal radiographs to detect a relatively small number of errors.

All the studies are consistent in showing an error rate of between 3-9% and 7% overall. For junior doctors, the error rate for missing positives (i.e. missed abnormali-


**Table 1 Summary of previous studies on X-ray interpretation**

<table>
<thead>
<tr>
<th>Study</th>
<th>Rate of error (%)</th>
<th>Rate of missed positives (%)</th>
<th>Rate of significant error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Lacey et al. (1980)</td>
<td>7.0</td>
<td>3.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Carew-McColl (1983)</td>
<td>3.9</td>
<td>3.3</td>
<td>-</td>
</tr>
<tr>
<td>Berman et al. (1985)</td>
<td>6.8</td>
<td>4.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Wardrope &amp; Chennels (1985)</td>
<td>6.2</td>
<td>-</td>
<td>1.1</td>
</tr>
<tr>
<td>Swain (1986)</td>
<td>3.9</td>
<td>2.3</td>
<td>-</td>
</tr>
<tr>
<td>Gleadhill et al. (1987)</td>
<td>4.9</td>
<td>2.7</td>
<td>1.2</td>
</tr>
</tbody>
</table>

*Errors assessed as proportion of total X-rays; figures either quoted or derived from data presented.*

ties) where this was assessed is between 2.3% and 4.2%. As many of the errors are not of clinical significance, in that they would not alter the course of treatment, they are then further classified as clinically significant or insignificant. The rate of clinically significant error is between 1.1% and 2.5% of the total number of X-rays. Most of these studies were primarily designed to assess the usefulness of radiologists’ involvement in accident and emergency (A&E) procedures and their results are clearly important in this regard. De Lacey et al. (1980) drew attention to the deleterious consequences of not having a review of radiographs. Wardrope & Chennels (1985) classified the errors by area of the body and suggested that those areas which seldom produced errors might be eliminated from any reviewing procedure.

The figures quoted above are sometimes also considered to be an accurate reflection of A&E SHOs’ ability to interpret radiographs. For example, Wardrope & Chennels (1985) commenting on an overall error rate of about 6%, concluded that A&E SHOs interpret radiographs with reasonable accuracy. For this purpose, however, the authors consider that it gives a false picture to present the errors as a percentage of the total number of X-rays. A more appropriate index of ability is the percentage of abnormalities that are missed. In fact, these two ways of presenting the data can be confused. De Lacey et al. (1980) write that, ‘just over 2% of clinically significant abnormalities were not positively recognized by casualty officers’, whereas the figure actually refers to the 2% of missed abnormalities relative to the total number of radiographs studied. This is not a trivial statistical distinction as is shown by the results reported here, where the error rate is assessed as a proportion of the total number of positive findings.

**METHOD**

The study was carried out in the A&E department, University College Hospital, London, England. From 9 a.m. to 5 p.m. during the week, SHOs in A&E automatically obtain a radiological opinion of any X-ray. As this study is only concerned with the accuracy of the SHOs’ interpretations, and not with the performance of the department as a whole, the authors have only studied those films for which they did not have immediate feedback from radiologists; in practice, this means those taken at nights and weekends. These radiographs were subsequently reviewed by radiologists of senior
Detection of radiographic abnormalities

registrar or consultant status. Those films showing abnormalities not detected by the casualty officer are routinely passed to the senior registrar or consultant of the A&E department. For the purposes of this study, however, all X-rays showing traumatic abnormalities were noted, whether or not these abnormalities had been detected by the SHO. All X-rays taken at nights and weekends over an 8-month period were studied. Thirty-two SHOs were working at some time in the 8-month period but only 18 of these were regular members of the department.

The following information was recorded from the casualty card: date and time of registration, age and sex of patient, presenting complaint and SHO’s assessment of the radiograph. The experience of the SHO (months in A&E) was also recorded. The radiologist’s report was then compared with the SHO’s assessment and discrepancies between them recorded. Abnormalities, according to the radiologist’s report, were subsequently graded as clinically significant or insignificant by one of the authors (PD). Radiographic abnormalities were graded as clinically significant if a failure to detect them would mean that the patient would need to be recalled to the hospital for treatment. The classification system is presented in appendix I.

RESULTS

Errors expressed as proportion of total X-rays

In the 8-month period, 6448 X-rays were taken at nights and weekends. Five hundred and five showed abnormalities and, of these, 178 were missed by the SHOs. Therefore, missed positives constitute 2.8% of total X-rays, which corresponds with the results of earlier studies.

Errors in the detection of abnormalities

Table 2 shows that when the focus is on the SHO’s ability to detect abnormal X-rays, a quite different picture emerges. Of radiographic abnormalities, 35% were missed by A&E SHO’s and 39% of clinically significant positives were missed.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Accuracy of detection of abnormalities on X-ray</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total detected</td>
</tr>
<tr>
<td>All positive</td>
<td>505</td>
</tr>
<tr>
<td>Clinically significant positive</td>
<td>200</td>
</tr>
</tbody>
</table>

Site of X-ray

The figures in Table 3 show the error rates for individual sites. They vary from 13% for X-rays of longbones to 60% for skull X-rays. The primary use of these figures must be to identify the classes of X-rays that need to be emphasized in training in radiographic
interpretation. Several areas have an error rate of over 40% while the X-rays of the ankle and skull each have an error rate of about 60%. Thus, a patient with a fracture of the ankle or skull has a less than even chance of having it detected by an unassisted SHO.

These results can be grouped and further analyses performed. For instance, X-rays of limbs are significantly better interpreted than other X-rays (Chi-squared = 4.37, \(P<0.04\)), the error rates being 32% and 42%, respectively. X-rays of joints and of longbones are, however, interpreted equally well.

### Table 3: Accuracy of detection by site

<table>
<thead>
<tr>
<th>Site</th>
<th>Total</th>
<th>Total incorrect</th>
<th>Percentage incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>27</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>Skull</td>
<td>10</td>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>Spine</td>
<td>19</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td>Shoulder</td>
<td>20</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Chest</td>
<td>81</td>
<td>37</td>
<td>46</td>
</tr>
<tr>
<td>Abdomen</td>
<td>14</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Pelvis</td>
<td>4</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Knee</td>
<td>14</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Ankle</td>
<td>22</td>
<td>13</td>
<td>59</td>
</tr>
<tr>
<td>Foot</td>
<td>43</td>
<td>17</td>
<td>40</td>
</tr>
<tr>
<td>Toes</td>
<td>33</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Elbow</td>
<td>23</td>
<td>8</td>
<td>35</td>
</tr>
<tr>
<td>Wrist</td>
<td>33</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>Hand</td>
<td>46</td>
<td>19</td>
<td>41</td>
</tr>
<tr>
<td>Fingers</td>
<td>108</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>Longbones</td>
<td>8</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>505</td>
<td>178</td>
<td>35</td>
</tr>
</tbody>
</table>

*Individual differences in SHO's abilities*

Table 4 shows the error rates for the 10 SHOs who had seen 20 or more of the radiographs. Error rates range between 19% and 50%, but the individual variation is not statistically significant (Chi-squared = 10.9, \(P<0.28\)). However, this study was not designed to investigate individual differences in ability; it would be unwise to conclude that the differences seen here were completely due to chance variation.

*Experience of SHOs*

For every radiograph interpreted by a regular member of the department, rather than a locum, the experience of the SHO (in months) in A&E at that time was recorded. Table 5 groups the abnormal radiographs studied according to experience.

Table 5 shows that, although the highest error rate is in the first month, the SHOs do not improve significantly thereafter (Chi-squared = 5.12, \(P=0.4\)); various tests of trends also showed no significance. This relatively stable error rate is further illustrated in Fig. 1 which plots the cumulative number of errors over time against the cumulative total of X-rays with positive findings. The strong linear relationship between the two
Detection of radiographic abnormalities

Table 4 Individual differences in error rates

<table>
<thead>
<tr>
<th>SHO</th>
<th>Total</th>
<th>Incorrect</th>
<th>Error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>11</td>
<td>44</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>18</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>10</td>
<td>36</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>9</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>26</td>
<td>13</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>35</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>10</td>
<td>37</td>
<td>14</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 5 Accuracy of detection by month

<table>
<thead>
<tr>
<th>Month</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>65</td>
<td>67</td>
<td>52</td>
<td>71</td>
<td>55</td>
<td>72</td>
<td>382</td>
</tr>
<tr>
<td>Incorrect</td>
<td>29</td>
<td>20</td>
<td>15</td>
<td>26</td>
<td>23</td>
<td>26</td>
<td>139</td>
</tr>
<tr>
<td>Percentage incorrect</td>
<td>45</td>
<td>30</td>
<td>29</td>
<td>37</td>
<td>42</td>
<td>36</td>
<td>36</td>
</tr>
</tbody>
</table>

Fig. 1 Cumulative errors against cumulative abnormal X-rays over 6 months.

Variables illustrates clearly that the error rate does not alter over 6 months. Plots for individual SHOs present a similar picture. Only 382 radiographs were interpreted by regular SHOs. Locums saw 107 radiographs and missed 36, an error rate of 34%. Locums and regular SHOs interpreted radiographs equally accurately. (The remaining 16 radiographs were seen by regular SHOs who had worked beyond 6 months.)
DISCUSSION

The results show that A&E SHOs miss 35% of radiographic abnormalities and 39% of clinically significant abnormalities. These high error rates, however, are consistent with the results of previous investigations. The authors’ own finding that missed abnormalities constitute 2.8% of all X-rays is slightly lower than the figure in most other studies. They believe that the SHOs under study were of average or better ability. All these figures assume that the radiologist’s interpretation is correct. De Lacey’s et al. (1980) study showed that the reviewing radiologists will miss a proportion of the abnormalities missed by casualty officers. This suggests that the figures reported here are, if anything, an underestimate of the SHO’s true error rates.

An ideal assessment of performance would include both the percentage of abnormalities missed and the percentage of normal cases misdiagnosed as abnormalities. This would enable both the relative detectability of various classes of abnormalities and the bias towards positive or negative decisions to be assessed. Some investigators (De Lacey et al. 1980) have reported the total number of normal radiographs misdiagnosed as abnormal, but only one study (Gleadhill et al. 1987) has yet considered the proportion of normal radiographs misdiagnosed or studied this type of error with respect to the site of the X-ray.

The figure of greatest interest for anyone as a prospective patient is surely the probability of a fracture or other abnormality being detected. The results of the present study indicate that this probability is about one in three if they are seen by an unassisted SHO. For some classes of X-rays, the chance of detection appears to be considerably lower. However, it is important to stress that all these abnormalities were subsequently detected during the routine radiological review. It was only possible to assess the SHOs’ abilities because of the existence of an effective review system. The fact remains that 39% of patients with significant radiographic abnormalities were initially diagnosed incorrectly and usually wrongly discharged, although subsequently recalled. It can be seen that radiologists play a crucial role in the management of patients who show abnormalities on X-ray (i.e. about 10% of all radiographs). In addition, radiographers can help by marking suspicious films (Berman et al., 1985).

It is commonly held that A&E SHOs will ‘pick things up’ and acquire knowledge and expertise by experience alone. These results show, however, that there is no improvement in the ability to detect radiographic abnormalities over the 6-month tenure of the post. Swain (1986) also found no change between the first, fourth and sixth months. Currently, formal training consists only of a small number of introductory lectures though there are many other opportunities for the SHOs to learn to interpret radiographs. Firstly, during the day there is direct reporting of films by radiologists; secondly, they are encouraged to attend the regular reviews of radiographs by senior radiological staff; thirdly, a library of abnormal radiographs is available for study and reference; and, fourthly, when an SHO has missed a significant abnormality a senior member of the A&E department reviews the film with him. This system of training however, appears insufficient to enable them to learn the specialized skill of radiographic interpretation; the authors suggest that more formal training and guidance is needed.

Chronic shortage of staff means that A&E SHOs cannot usually be spared to visit radiology departments to study radiographs or undergo extensive formal training. It
Detection of radiographic abnormalities

should be possible, however, to collect some of the radiographic errors from one generation of SHOs and use them as guidance for the next. It is obviously important to identify sites where errors are frequent (Wardrope & Chennels, 1985; Gleadhill et al., 1987) but the authors also think it important to single out those areas where the rate of error is high, even though X-rays of that area are less frequent. For example, abnormal radiographs of the skull are infrequently encountered but are poorly interpreted when they do occur. A good solution would be to develop a standard series of radiographs showing frequent errors, possibly in the context of a computer-assisted learning programme.

Simply viewing training material available is unlikely to improve the SHO’s rate of detection of abnormalities. An important aspect of learning any skill is feedback about performance (Annett, 1969). In some A&E departments, such as University College Hospital, SHOs do receive feedback on their interpretation of X-rays performed out of hours, but this is necessarily some time after they have made their decisions on patient management. It is true that SHOs can get immediate radiological opinions during the day and out of hours for emergencies on an on-call basis, but it is doubtful in these circumstances if they trouble to form their own judgements about the film first. The authors propose, based on the psychology of training (Stammers & Patrick, 1975), that:

1. SHOs, both in formal sessions and during their work, must make their own decisions about an X-ray before they are given feedback;
2. they need to be shown examples of both normal and abnormal X-rays; and
3. they should be given regular feedback on their performance during their work. The authors are currently carrying out some preliminary studies to assess the usefulness of these proposals.

ACKNOWLEDGEMENTS

We thank Dr H. Baderman and Ms F. Moore for permission to conduct the study in their department.

C. A. Vincent is funded by the Medical Protection Society, on a project investigating avoidable mishaps in medicine, under the direction of Professor Audley.

REFERENCES

108 C. A. Vincent et al.


APPENDIX I

Classification scheme

Clinically significant abnormalities are those that, if undetected, would require recall to the A&E department or other hospital treatment. Cases requiring only a letter to the patient’s general practitioner are rated insignificant. The scheme that follows must only be considered a guide as each case was considered individually; nevertheless, almost all cases were eventually classified as stated. Fractures and dislocations are rated similarly unless otherwise specified.

Significant abnormalities

Foot, if dislocated metatarsal with phalanx or talus.
Talus and other talar bones.
Ankle, unless just soft tissue swelling.
Tibia.
Knee.
Femoral shaft.
Hip, lesser trochanter and pelvis.
Ribs, if two fractured on same side or with paradoxical movement.
Lumbar/thoracic vertebrae: subluxation/fracture/dislocation.
Fingers: subluxation/avulsion/fractures excluding terminal phalanx.
Thumb and first metacarpal.
Wrist.
Scaphoid.
Colles.
Smiths.
Radius/Ulna shaft.
Elbow: radial head, unless impacted.
Humoral shaft, fracture or comminuted fracture.
Shoulder: dislocated acromioclavicular joint, not clavicle.
Skull: facial bones, jaw or vault.

Non-significant abnormalities

There will be some variation in the individual case in the following but, in general:
Soft tissue swelling, foreign bodies, flake fractures.
Toes, whether simple or compound fracture.
Hallux whether displaced or not.
Fibula, excluding ankle.
Rib, if single fracture.
Vertebrae: old scoliosis or reduced space.
Finger, terminal phalanx.
Metacarpal, second to fifth.
Elbow, impacted radial head.
Clavicle.
Nose.
Accuracy of detection of radiographic abnormalities by junior doctors.

C A Vincent, P A Driscoll, R J Audley and D S Grant

doi: 10.1136/emj.5.2.101

Updated information and services can be found at:
[http://emj.bmj.com/content/5/2/101](http://emj.bmj.com/content/5/2/101)

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
[http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to:
[http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to:
[http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)