How accurate is cancer scan reporting?

Richard FitzGerald, Rakesh Mehra

The wide recognition of advances in medical imaging need to be matched by an awareness of the significant error rates in radiological interpretation of cancer scans. Common errors, their aetiology and potential solutions are discussed. Internet awareness of these errors has appeared.

Scans play a major part in the detection and staging of malignancy, as well as in monitoring the response to therapy. The initial presentation of a patient with cancer is frequently non-specific and the radiologist is often the first to register the likelihood of malignancy. Accurate staging of scan reports can save patients the morbidity and possible mortality of surgery or unpleasant side-effects from radiotherapy and/or chemotherapy. Modern cancer drug regimens may cost thousands of pounds for an individual patient, not to mention the human cost, for example side-effects or time 'lost' while on therapy. Scans provide a major indicator of whether to stop or continue such chemotherapy.

EXTENT OF THE PROBLEM

Robinson (1997) has pointed out radiology's Achilles' heel, namely 'error and variation in the interpretation of the Rontgen image'. As he observed, this can be caused by problems of technique, perception, knowledge or judgment. This limitation of radiology has already reached the public domain in the field of mammography and, indeed, Brown et al (1996) showed that breast cancer detection rates were increased by 13% with consensus double reading. The Wessex Regional Radiology Audit on barium enema detection of colorectal carcinoma (Thomas et al, 1995) found that although carcinoma demonstration rate was 97% the carcinoma reporting rate was only 85%.

It is now becoming clear that similar error rates apply to the interpretation of cancer scans. While ultrasound and nuclear medicine are also used in the imaging of cancer patients, this review will mainly focus on computed tomography (CT) and magnetic resonance imaging (MRI). The general principles apply to all scan modalities. In a recent American review of the clinical importance of reinterpretation of body CT scans reported elsewhere in patients referred for care at a tertiary cancer centre, Gollub et al (1999) found 'major disagreement' with the original report in 17%.

Another American study by Kalbhen et al (1998) found that reinterpretation of outside abdominal CT scans of patients with pancreatic cancer by subspecialist radiologists at a tertiary centre reclassified 32% of patients originally considered resectable on the basis of the first scan report as not resectable. However, even the specialist reinterpretation reports were only correct in 94% of cases in predicting resectability.

British experience has been similar, with Loughrey et al (1999) changing radiological staging in 19% of cases reviewed. However, they admit that in 4 of the 124 cases reviewed by one of their specialist radiologists, there was a serious error of interpretation in the specialist review analysis detected subsequently by a second independent specialist radiologist.

PATTERNS OF ERRORS

Bechtold et al (1997), Gollub et al (1999) and Loughrey et al (1999) all found that the most common error in the reporting of CT cancer scans was lymphadenopathy.

The limitations of MRI in detecting metastatic cervical lymphadenopathy have been made clear by Wide et al (1999) who showed that reporting of lymphadenopathy on MRI scans of patients with oral squamous cell carcinoma had a false-negative rate of 20% and a false-positive rate of 47%. The overall accuracy for this feature lacked the sensitivity and specificity necessary to

Dr Richard FitzGerald is Consultant Radiologist and **Dr Rakesh Mehra** is Consultant Medical Oncologist in New Cross Hospital, Wolverhampton WV10 0QP

Correspondence to: Dr R FitzGerald

replace elective neck dissection for both staging and prognostic purposes.

Fundamentally the problem with all types of scans is that large lymph nodes may be benign but large simply because of infection or inflammation, whereas small lymph nodes may harbour metastatic disease. There is variability in the interpreter's measurements, some using cursors, others simply 'eye-balling'. The use of a single-size criterion for lymph nodes at various anatomical locations is simplistic. There is, however, justifiable subjective variation in radiological 'calling' of lymphadenopathy based on location, appearance and multiplicity of nodes (Gollub et al, 1999).

The limitations of scans in predicting the malignancy or otherwise of multiple pulmonary nodules has been analysed by Ginsberg et al (1999), who found that even in patients who were known to have malignancy and had multiple pulmonary nodules dissected at video-assisted thoracoscopic surgery, all resected nodules were benign in 36% of such patients. Size criteria increase accuracy in predicting malignancy but there is a differential diagnosis of small pulmonary nodules, e.g. hamartomas, intra-pulmonary lymph nodes, scars, granulomata, sarcoid-like reactions, sterilized metastases.

Schwartz et al (1999) found that multiple small hepatic lesions in patients with known cancer represented metastases in only 12% of the 2 978 cancer patients they studied. Small hepatic lesions are often cysts or biliary hamartomas.

Other non-malignant conditions need to be borne in mind by radiologists and clinicians when reviewing scans of patients with cancer. Abnormalities may be the result of surgery, radiotherapy or chemotherapy, as well as opportunistic infections. Gosselin et al (1998) found that 9% of inpatients with cancer have unsuspected pulmonary emboli.

Moreover, simply having cancer does not protect one against the spectrum of other illnesses that one can develop during life. Yue et al (1997) performed MRI scans of the brain in a population-based longitudinal study of cardiovascular disease and stroke in asymptomatic men and women aged 65 years and older; 0.5% were unexpectedly found to have meningiomas. Even more frequent comorbidity, for example with cardiovascular disease, cerebrovascular disease or musculoskeletal problems, can be expected in the age group in which patients typically have cancer.

Survival from a first tumour allows one to live long enough to develop a second different type of tumour. Also some patients treated for one type of cancer/tumour are at increased risk of developing a second different type of tumour as a result of treatment for the first malignancy. Second malignant neoplasms are a consequence of the long-term survival now being seen in children and some adults who have been treated for cancer. While a peak incidence of acute leukaemia is reached after about 4–5 years, the incident of solid tumours, e.g. lymphoma or sarcoma, continues to rise over time (Malpas, 1998).

Radiology, even sophisticated scanning, is neither histology nor microbiology. The medico-legal implications of the patient who died following radiotherapy for a suspected brain tumour which, in fact, was an abscess has been described (Berlin and Saxena, 1999). They point out that radiological imaging establishes a presumptive but not necessarily definitive diagnosis.

TECHNIQUE ERRORS

Spiral CT has been shown to detect 40% more lung nodules than conventional CT (Remy-Jardin et al, 1993) and further increase in sensitivity can be expected from the latest multislice CT scanners.

However, if in addition one wished to confirm or exclude the possibility of pulmonary emboli in a cancer patient, a significant and more time-consuming alteration to the CT protocol would be required (CT pulmonary angiography perhaps combined with CT venography) (Loud et al, 2000).

Contrast-enhanced CT is more sensitive than plain CT in detecting hepatic masses. Multiple factors, such as the volume and concentration of contrast, rate of administration of contrast, venous access, timing of scan, type of scanner, cardiac output, hydration status, renal function, body weight or time since last meal, may affect both conspicuity and apparent size of liver metastases. Indeed, these myriad factors have led Nazarian et al (1999) to suggest the use of non-enhanced CT scans for the liver at each examination as the most reproducible method of comparison, with post-contrast scanning being performed to optimize the detection of new lesions. Even where lesions are known to be malignant and are being followed up for response to chemotherapy, there is considerable inter-observer and, indeed, intra-observer variability in CT tumour measurements. Hopper et al (1996) found inter-observer variability of 15%, which was worse for poorly defined and irregular lesions.

PERCEPTUAL ERRORS

Perception of abnormalities on scans depends on satisfactory scan technique, adequate (if necessary corrected) visual acuity, good viewing facilities with optimal lighting, temperature, noise insulation and ideally minimal interruption.

The first scan of a patient with cancer will not always be labelled as a 'cancer scan' with a 'red flag'. The radiological level of suspicion and detection will increase with the gravity of clinical presentation as outlined in the information given by the referring doctor. Indeed, the radiologist can be pointed away from a diagnosis of cancer. In a recent study of 624 patients referred for spiral CT angiography for vascular disease six had previously unknown cancers, of which only four were detected prospectively (Katz et al, 1999).

Tumour detection may be hampered by surrounding non-malignant disease, e.g. lung fibrosis. Moreover, some tumours are poorly demonstrated by scanning, e.g. early gastric or bladder malignancy.

Cancer detection on scans requires concentration and careful systematic analysis of sometimes hundreds of images of individual patients, as well as comparison with relevant previous scans or radiographs.

Bechtold et al (1997) found that error rates in reporting abdominal CT scans more than doubled when a radiologist reported more than twenty studies a day. Burnout, professional isolation, interpersonal problems at home or at work, or illness can all seriously compromise performance.

LACK OF KNOWLEDGE

The explosive growth in radiological knowledge and techniques over the past 20 years means that many radiologists spend the majority of their time on activities that did not exist in clinical radiological practice before 1980. This revolution shows no sign of slackening. Despite increasing specialization by radiologists it is not economically affordable in most health-care settings for all cancer scans to be reported by super-specialists.

JUDGMENT ERRORS

Wood (1999) has pointed out that errors and decision making by inexperienced radiologists usually take the forms of 'pseudodiagnosticity' or premature diagnostic conclusions. She considers pseudodiagnosticity a situation where observations are interpreted on the basis of a single hypothesis in the belief that a high true positive rate of supportive evidence is diagnostic in itself and ignoring false-positive results. In adjudicating the scan abnormalities multiple hypotheses, e.g. benign disease, need to be considered.

Experienced radiologists are not immune to errors of judgment and Wood points out that these radiologists may rely too heavily on 'rules of thumb' and 'subjective probability estimates'. While these may work in uncomplicated diagnostic situations and be rather time efficient, they may be inaccurate.

The wording of a radiological report is analagous to the slant a journalist uses in reporting news. Berlin and Saxena (1999) have emphasized how the use of the word 'mass' can prejudice clinicians towards believing that malignancy is present.

WHAT CAN BE DONE ABOUT IT ALL?

Staffing issues

Specialist reporting and double reading of scans might seem obvious solutions to improve cancer scan accuracy. However, these have enormous cost implications.

Even specialist reporting is fallible (Kalbhen et al, 1998; Loughrey et al, 1999). Moreover, the improved performance of specialist radiologists reviewing cancer scans done elsewhere may reflect:

1. Greater clinical information than given to the original radiologist
2. The review being done in a different clinical context, i.e definite diagnosis of malignancy *vs* non-specific clinical presentation at referral hospital
3. The reviewing radiologist may well have access to the original radiological report, i.e double reporting.

British radiologists will marvel at the resources available in other countries, for example as described by Seltzer et al (1997) in Brigham and Women's Hospital, Boston, where at the time of their paper there were 54 consultant radiologists and 48 registrars.

Site specialization and participation by British radiologists in multidisciplinary cancer teams have been welcome developments arising from the implementation of the Calman–Hine cancer report. However, their full potential, much less the thought of double reporting, is constrained by overall staffing problems in British radiology with scores of unfilled consultant radiologist posts.

Cancer scan technique can only be optimized by having an adequate number of diagnostic radiographers. The current UK shortage of diagnostic radiographers impacts on the timeliness and quality of scanning, especially as the latest scanners of all modalities have more sophisticated post processing.

Moreover, cancer imaging while very important is only one of many equally important roles of a radiology department. Improved radiological performance is also expected, quite reasonably, in these areas.

Increased manpower even with double reporting does not eliminate cancer scan errors as the international literature shows.

In recent years teleradiology (Wunderbaldinger et al, 1999) has facilitated second opinions by specialist radiologists on 'difficult' cancer scans.

While computer-aided diagnosis still has a long way to go in its development, its already proven ability to detect pulmonary nodules on CT scans illustrates its potential to complement radiological perception (Armato et al, 1999; Tourassi, 1999).

Similarly, the work already done on applying artificial neural networks, e.g. to the prediction of ovarian malignancy on colour Doppler ultrasound, shows another potential role of information technology in improving cancer scan accuracy (Biagiotti et al, 1999).

These information technology developments also emphasize the current limitations of human performance in cancer scan reporting.

Equipment issues

The Royal College of Radiologists has produced recommendations relating to equipment replacement, recognizing that all electronic equipment has a finite life and that in some cases this is shorter than expected because of improvements in technology. It recommends replacing CT scanners, MRI scanners, radionuclide scanners at 7 years and ultrasound scanners at 5 years (Royal College of Radiologists, 1999). Providing adequate numbers of up to date scanners has enormous capital expenditure implications, and a rolling programme of planned replacement is required.

The improved performance of recent MRI scanners can be seen with total body echo planar imaging. Six minutes of such scanning has been shown to produce 180 images from the cranial vertex through to the feet (Horvath et al, 1999). Latest multislice CT scanners can capture 8 images per second in a single breathhold (Sandrick, 1999).

These innovations bring significant challenges for radiologists to interrogate on workstations literally hundreds of extra images per patient. Issues of image storage and how much imaging data clinicians will have access to have yet to be resolved. In short, these newer scanners have yet to be 'tamed'. It is not hard to imagine reporting errors continuing with the dramatic increase in data to be analysed.

New, often costly contrast media or labelled tracers in ultrasound, nuclear medicine and MRI all can improve accuracy of cancer scans, e.g. MR lymphangiography is a major imaging advance in predicting lymph nodes as benign or malignant (Harisinghani et al, 1999).

Positron emission tomographic (PET) scans have a proven role in helping to distinguish malignancy from benign abnormalities (e.g. post-treatment change). However, their cost limits their application (Cook and Maisey, 1996).

Hospital-wide picture archiving and communications systems (PACS) enable radiologists and clinicians to have 'immediate' access to all current and prior imaging of patients. This is particularly important in cancer radiology where comparison with previous scans is vital to assess response or otherwise to therapy. PACS, when integrated with electronic patient records, empowers reporting radiologists to have immediate access to much more patient clinical information. However, PACS is not cheap, with a hospital-wide system typically costing several million pounds. While substantial savings are made, e.g. in respect of film, the net effect of implementing PACS is to increase radiology costs significantly by up to several hundred thousand pounds a year in a large hospital-wide system (Bryan et al, 1999).

Education issues

The internet gives radiologists an enormous knowledge resource with radiology teaching files, journals, textbooks and bulletin boards already on line (Wunderbaldinger et al, 1999).

The advent of revalidation will bring the issue of continuing medical education and professional development higher up radiologists' agendas. Increased time spent on these areas will inevitably make inroads on radiologists' time. Quality of reporting should improve but quantity may decrease.

Clinical governance issues

Clinical governance brings greater emphasis on standard setting and monitoring. The intrinsic uncertainty in cancer scan reporting (as well as in all of medicine) is not a defence for poor quality radiology. Some cancer patients suffer as a result of errors in radiologists' reports. The General Medical Council explicitly places a duty on doctors to act quickly to protect patients from risk if they have good reason to believe that they themselves or colleagues may not be fit to practice (General Medical Council, 1998).

However, distinction must be made between best practice, average practice and minimum acceptable standards of practice. Factors outwith radiological control, e.g. equipment and staffing levels, have to be considered along with features clearly the personal responsibility of radiologists, e.g. application of due care and concentration. Ultimately continuing medical education will largely be the duty of the individual radiologist but his/her employer will need to provide the time and other resources necessary for his/her continuing training, e.g. in mastering new technology.

Analysis of malfunctioning hospital specialists of various branches of medicine in North Holland (Lens and van der Wal, 1997) found that lack of social skills/inability to work in a team/incompatability of temperament, far outnumbered cases of medical and technical errors of judgment. Professor Donaldson, now Chief Medical Officer for England, has written that when an apparent problem arises in respect of a doctor's practice: 'it is not infrequent to find matters being manipulated by someone who has a vested interest in making trouble for the doctor concerned' (Donaldson, 1997).

Mistakes made by a radiologist need to be brought to his/her attention in a constructive fashion. Good team working both among radiology department staff and site-specific multidisciplinary groups is an essential prerequisite for optimal standards in cancer radiology. Cancer scan errors need to be put in context, e.g. did any patient harm ensue, might other radiologists have made the same mistake given identical clinical information? The fact that cancer scans are an important part but not the whole picture of a patient with cancer needs to be borne in mind. Discussion of mistakes should be done not in the form of a witchhunt but in a fashion where the individual radiologist and colleagues can learn lessons from the experience. Marginalized radiologists can not perform to their full potential.

CONCLUSIONS

Medicine and society are in the process of renegotiating their relationship with each other. Claire Rayner of The Patients' Association, recently observed with characteristic wisdom that patients should not have a childlike trust in their doctors, but rather a relationship more appropriate to adults.

CancerNet, the website of the USA National Cancer Institute (http://cancernet.nci.nih.gov/testing.html) spells out the false-negative and false-positive rates of screening mammography. In April 2000 it added an additional section on spiral CT screening for lung cancer which points out that 'non cancerous changes in the lungs can mimic tumors on CT scans, challenging the radiologists who read them. Interpretation of the scans can vary leading to confusion about recommendations for follow-up care'. This is the first entry on this influential website that raises the issue of inaccuracies in cancer scan reporting. Many cancer patients would prefer not to know such information. However, we do believe that problems in cancer scan accuracy need to be more widely appreciated than they are. The causes are multifactorial and there is no quick fix solution.

It is appropriate to conclude with a cancer patient's perspective (Anonymous, 2000) written by a 42-year-old patient whose 'very obvious' sacral chordoma was missed by radiologists reporting not only his radiographs, a subsequent myelogram but also his MRI scan. 'If I had to leave doctors with one message from my experience it would be to emphasise the importance of listening to the patient and querying reports of investigations and examinations when symptoms and signs do not match the results'. That 'querying' must lead to clinicians and radiologists discussing and hopefully resolving such diagnostic dilemmas sooner. **HM**

Conflict of interest: none.

Anonymous (2000) Why a massive tumour went undetected. *Br Med J* **320:** 523

Armato SG, Giger ML, Moran CJ, Blackburn JT, Doi K, MacMahon H (1999) Computerised detection of pulmonary nodules on CT scans. *Radiographics* **19:** 1303–11

Bechtold RE, Chen MYM, Ott D et al (1997) Interpretation of abdominal CT: analysis of errors and their causes. *J Comput Assist Tomogr* **21:** 681–5

Berlin L, Saxena VA (1999) Radiation therapy of suspected brain tumour without definitive diagnosis. *Am J Roentgenol* **172:** 1477–80

Biagiotti R, Desii C, Vanzi E, Gacci G (1999) Predicting ovarian malignancy: application of artificial neural networks to transvaginal and colour doppler flow US. *Radiology* **210:** 399–403

Brown J, Bryan S, Warren R (1996) Mammography screening: an incremental cost effectiveness analysis of double *vs* single reading of mammograms. *Br Med J* **312:** 809–12

Bryan S, Weatherburn G, Buxton M, Watkins J, Keen J, Muris N (1999) Evaluation of a hospital picture archiving and communication system. *J Health Serv Res Policy* **4:** 204–9

Cook GJR, Maisey MN (1996) Review: The current status of clinical PET imaging. *Clin Radiol* **51:** 603–13

Donaldson L (1997) Doctors with problems in a hospital workforce. In: Lens P, van der Wal G, eds. *Problem Doctors, A Conspiracy of Silence.* IOS Press, Amsterdam: 99–109

General Medical Council (1998) *Good Medical Practice.* General Medical Council, London

Ginsberg MS, Griff SK, Go BD, Yoo HH, Schwartz LH,

Panicek DM (1999) Pulmonary nodules resected at video-assisted thoracoscopic surgery: aetiology in 426 patients. *Radiology* **213:** 277–82
Gollub MJ, Panicek DM, Bach AM, Penalver A, Castellino RA (1999) Clinical importance of reinterpretation of body CT scans obtained elsewhere in patients referred for care at a tertiary cancer centre. *Radiology* **210:** 109–12
Gosselin MV, Rubin GD, Leung AN, Huang J, Rik NW (1998) Unsuspected pulmonary embolism: prospective detection on routine helical CT scans. *Radiology* **208:** 209–15
Harisinghani MG, Saini S, Weissleder R et al (1999) MR lymphangiography using ultra-small super paramagnetic iron oxide in patients with primary abdominal and pelvic malignancies: radiographic - pathologic correlation. *Am J Roentgenol* **172:** 1347–51
Hopper KD, Kasales CJ, Van Slyke MA, Schwartz TA, TenHave TR, Joefiak JA (1996) Analysis of inter-observer and intra-observer variability in CT tumour measurements. *Am J Roentgenol* **167:** 851–4
Horvath LJ, Burtness BA, McCarthy S, Johnson KM (1999) Total body echo-planar MR imaging in the staging of breast cancer: comparison with conventional methods — early experience. *Radiology* **211:** 119–28
Kalbhen CL, Yetter EM, Olson MC, Posniak HV, Aranha GV (1998) Assessing the resectability of pancreatic carcinoma: the value of reinterpreting abdominal CT performed at other institutions. *Am J Roentgenol* **171:** 1571–6
Katz DS, Jorgensen MJ, Rubin GD (1999) Detection and follow-up of important extra arterial lesions with helical CT angiography. *Clin Radiol* **54:** 294–300
Lens P, van der Wal G (1997) Malfunctioning of specialists in Dutch hospitals. In: Lens P, van der Wal G, eds. *Problem Doctors, A Conspiracy of Silence.* IOS Press, Amsterdam: 191–9
Loud PA, Katz DS, Klippenstein DL, Shah RD, Grossman ZD (2000) Combined CT venography and pulmonary angiography in suspected thromboembolic disease. *Am J Roentgenol* **174:** 61–5
Loughrey GJ, Carrington BM, Anderson H, Dobson MJ, Lo Ying Ping F (1999) The value of specialist oncological radiology review of cross-sectional imaging. *Clin Radiol* **54:** 149–54
Malpas J (1998) Second malignancies. In: Husband JES, Reznek RH, eds. *Imaging in Oncology*. Isis medical media, Oxford: 945–52
Nazarian LN, Park JH, Halpern EJ et al (1999) Size of colorectal liver metastases at abdominal CT: comparison of pre-contrast and post-contrast studies. *Radiology* **213:** 825–30
Remy-Jardin M, Remy J, Giraud F, Marquette CH (1993) Pulmonary nodules: detection with thick section spiral CT *vs* conventional CT. *Radiology* **187:** 513–20
Robinson PJA (1997) Radiology's Achilles' heel: error and variation in the interpretation of the Rontgen image. *Br J Radiol* **70:** 1085–98
Royal College of Radiologists (1999) *Good Practice Guide for Clinical Radiologists.* Board of the Faculty of Clinical Radiology, The Royal College of Radiologists, London
Sandrick K (1999) Multi-slice CT gains wider acceptance in Europe. *Diagn Imaging Eur* **12:** 23–5
Schwartz LH, Gandras EJ, Colangelo SM, Ercolani MC, Panicek DM (1999) Prevalence and importance of small hepatic lesions found at CT in patients with cancer. *Radiology* **210:** 71–4
Seltzer SE, Kelly P, Adams DF et al (1997) Expediting the turnaround of radiology reports in a teaching hospital setting. *Am J Roentgenol* **168:** 889–93
Thomas RD, Fairhurst JJ, Frost RA (1995) Wessex Regional Radiology Audit: barium enema in colorectal carcinoma. *Clin Radiol* **50:** 647–50
Tourassi GD (1999) Journey toward computer aided diagnosis: role of image texture analysis. *Radiology* **213:** 317–20
Wide JM, White DW, Woolgar JA, Brown JS, Vaughan ED, Lewis-Jones HG (1999) Magnetic resonance imaging in the assessment of cervical nodal metastasis in oral squamous cell carcinoma. *Clin Radiol* **54:** 90–4
Wood BP (1999) Decision making in radiology. *Radiology* **211:** 601–3
Wunderbaldinger P, Schima W, Turetschek K, Helbich TH, Bankier AA, Herold CJ (1999) Worldwide web and internet: applications for radiologists. *Eur Radiol* **9:** 1170–82
Yue NC, Longstreth WT, Elster AD, Jungreis CA, O'Leary DH, Poirier VC (1997) Clinically serious abnormalities found incidentally at MR imaging of the brain: data from the Cardiovascular Health Study. *Radiology* **202:** 41–6

KEY POINTS

- International literature review indicates a significant cancer scan reporting error rate.
- Suboptimal scan technique, perceptual errors by radiologists, deficient knowledge in a rapidly changing field and errors in radiological judgment may all be responsible.
- The scan technique applied depends on the disease suspected clinically. 'Wonderscans' revealing all pathology do not exist. Each scan technique has applications and limitations.
- Radiology is not histology. Similar scan appearances can be caused by different diseases.
- Accuracy of radiological reports is dependent on clinical information given.
- Good team working between radiologists and clinicians is an essential prerequisite for optimal cancer radiology.
- Ideal scanner provision, staffing levels and information technology support has considerable resource implications.
- Cancer scanning is competing with other important roles for radiology in a setting of limited resources.