Added value of double reading in diagnostic radiology, a systematic review

Håkan Geijer 1 · Mats Geijer 1,2

Received: 2 November 2017 / Revised: 10 January 2018 / Accepted: 15 January 2018 / Published online: 28 March 2018
© The Author(s) 2018

Abstract
Objectives Double reading in diagnostic radiology can find discrepancies in the original report, but a systematic program of double reading is resource consuming. There are conflicting opinions on the value of double reading. The purpose of the current study was to perform a systematic review on the value of double reading.

Methods A systematic review was performed to find studies calculating the rate of misses and overcalls with the aim of establishing the added value of double reading by human observers.

Results The literature search resulted in 1610 hits. After abstract and full-text reading, 46 articles were selected for analysis. The rate of discrepancy varied from 0.4 to 22% depending on study setting. Double reading by a sub-specialist, in general, led to high rates of changed reports.

Conclusions The systematic review found rather low discrepancy rates. The benefit of double reading must be balanced by the considerable number of working hours a systematic double-reading scheme requires. A more profitable scheme might be to use systematic double reading for selected, high-risk examination types. A second conclusion is that there seems to be a value of sub-specialisation for increased report quality. A consequent implementation of this would have far-reaching organisational effects.

Key Points
- In double reading, two or more radiologists read the same images.
- A systematic literature review was performed.
- The discrepancy rates varied from 0.4 to 22% in various studies.
- Double reading by sub-specialists found high discrepancy rates.

Keywords Diagnostic errors · Observer variation · Diagnostic imaging · Review · Quality assurance, healthcare

Introduction
In the industrialised world, there is an increasing demand for radiology resources with an increasing number of images being produced, which has led to a relative scarcity of radiologists. With limited resources, it is important to question and evaluate work routines, to provide settings for high-quality output and high cost-effectiveness, but at the same time keep medical standards high and avoid costly lawsuits. One way to increase the quality of radiology reports may be double reading of studies between peers, i.e. two radiology specialists of similar and appropriate experience reading the same study.

Most radiologists hold a very firm view on the concept of double reading—either for or against. Arguments for are that it reduces errors and increases quality in radiology. Arguments against are that it does not increase quality significantly, is time-consuming, and wastes time and resources. Despite these firm beliefs, there is comparatively scant evidence supporting either view, and both systems are widely practiced [1]. In some radiology departments or department sections, it is accepted that no systematic double reading is performed.
between specialists of a similar or above a certain degree of expertise. In other departments, such double reading between peers is mandatory. A survey among Norwegian radiologists reported a double reading rate of 33% of all studies [1], which is consistent with a previous Norwegian survey [2].

The concept of observer variation in radiology was introduced in the late 1940’s when tuberculosis screening with mass chest radiography was evaluated [3, 4]. In a comparison between four different image types (35-mm film, 4 × 10-inch stereophotofluorogram, 14 × 17-inch paper negative, 14 × 17-inch film), it was discovered that the observer variation was greater than the variation between image types [3]. The authors recommended that “In mass survey work … all films be read independently by at least two interpreters”. Double reading in mammography and other types of radiologic screening is, however, not the purpose of the current study since the approach of the observer in screening work is different from that in clinical work. In screening, the focus leans towards finding true positives and avoiding false negatives, whereas in clinical work also false positive and true negative findings are of importance. Neither is the purpose of the current study the evaluation of double reading in a learning situation, such as the double reading of residents’ reports by specialists in radiology. In such cases, the report and findings of a resident are checked by a more experienced colleague. This has an educational purpose and serves to improve the final report to provide better healthcare, with a better patient outcome in the end. The value of such double reading is hardly debatable.

Double reading can be broadly divided into three categories: (1) both primary and secondary readings by radiologists of the same degree of sub-specialisation, in consensus, or serially with or without knowledge of the contents of the first report; (2) secondary reading by a radiologist of a higher level of sub-specialisation; (3) double reading of resident reports [5].

The concept of double reading is at times confusing and can apply to several practices.

In screening, the concept of double reading implies that if both readers are negative, the combined report is negative. If one or both readers are positive, the report is positive (i.e. the “Or” rule or “Believe the positive”). In dual reading, the two readers reach a consensus over the differing reports [6].

Some studies use arbitration: with conflicting findings, a third reader considers each specific disagreement and decides whether the reported finding is present or not. Similar to this is pseudo-arbitration: with conflicting findings, the independent and blinded report of a third reader casts the deciding “vote” in each dispute between the original readers. In contrast to the “true arbitration” model, the third reader is not aware of the specific disagreement(s) [7]. These concepts are summarised in Table 1.

Considering the paucity of evidence either for or against double reading among peers in clinical practice, the purpose of the current study was to, through a systematic review of
available literature, gather evidence for or against double reading in imaging studies by peers and its potential value. A secondary aim was to evaluate double reading with the secondary reading being performed by a sub-specialist.

Materials and methods

The study was registered in PROSPERO International prospective register of systematic reviews, CRD42017059013. The inclusion criterion in the literature search was: studies calculating the rate of misses and overcalls with the aim of establishing the added value of double reading by human observers. The exclusion criteria were: (1) articles dealing solely with mammography; (2) articles dealing solely with screening; (3) articles dealing solely with double reading of residents; (4) articles not dealing with double reading; (5) reviews, editorials, comments, abstracts or case reports; (6) articles without abstract; (7) article not written in English, German, French or the Nordic languages; (8) duplicate publications of the same data.

Literature search

A literature search was performed on 26 January 2017 in PubMed/MEDLINE and Scopus. The search expressions were a combination of “radiography, computed tomography (CT), magnetic resonance imaging (MRI) and double reading/reporting/interpretation” (Appendix 1).

Both authors read all titles and abstracts independently. All articles that at least one reviewer considered worth including were chosen for reading of the full text. After independent reading of the full text, articles fulfilling the inclusion criteria were selected. Disagreements were solved in consensus. The material was stratified into two groups depending on whether the double reading was performed by a colleague of similar or higher sub-speciality.

Results

The literature search resulted in 1,610 hits. Another eight articles were added after manual perusal of the reference lists. Of these, 165 articles were chosen for reading of the full text. Forty-six of these that fulfilled the inclusion criteria and did not comply with the exclusion criteria were selected for final analysis. The study flow diagram is shown in Fig. 1. Study characteristics and results are shown in Table 2. Excluded articles are shown in Appendix 2.

When perusing the material, it was found that there were not sufficient data to perform a meta-analysis. Instead, a verbal summary was performed. In the results, two distinct groups of studies appeared: studies reporting double reading by peers of similar competence level and studies reporting the second reading performed by a sub-specialist, often performed at a referral hospital.

Double reading by peers of similar degree of sub-specialisation

Fifteen articles evaluated double reading in CT.

- In trauma CT, three papers found initial discordant readings of 26–37% [13–15]. However, in one of these articles patient care was changed in only 2.3% by a non-blinded second reader [13]. Eurin et al. [16] reported a high rate of missed injuries initially, predominantly minor and musculoskeletal injuries.
- In abdominal CT, a discrepancy rate of 17% resulted in 3% treatment change when reviewed by a non-blinded second reader [12]. Five articles evaluated sensitivity and specificity. In CT of ovarian cancer and CT colonography, there was a non-significant trend towards higher sensitivity in double reading [18, 19], but double reading increased the false-positive rate [20].
- In chest CT for pulmonary nodules, double reading increased sensitivity [8, 22, 23], but computer-aided diagnosis (CAD) was even more beneficial [8, 22]. Another article found clinically important changes in 9% of cases [24].

Eight articles evaluated double reading in radiography.

- Two articles found negligible improvement by double reading in small-bowel and large-bowel barium studies, one study even reported increased false positives with double reading [27, 28].
- In chest radiography, Hessel et al. [7] combined independent readings by eight radiologists. Using a third independent interpretation to resolve disagreements between pairs of readers (pseudo-arbitration) was the most effective method overall, reducing errors by 37%, increasing correct interpretations by 18%, and adding 19% to the cost of an error-free interpretation.
- Quekel et al. [6] reported that double or dual reading increased sensitivity, at the same time reducing specificity.
- Two articles quoted 3–9% disagreement between observers in general radiography [30, 31].

Mixed modalities.

- Siegle et al. [33] evaluated general radiology in six departments, and found a mean rate of disagreement of 4.4%.
- In another large study, 11,222 cases (3.3% of the total production) underwent randomised peer review using a consensus-oriented group review with a rate of discordance (“report should change”) of 2.7% [37].
Babiarz and Yousem [35] found 2% disagreement when 1,000 neuroradiology cases were double read by another neuroradiologist, all working in the same institution.

In breast MRI, double reading increased sensitivity from 80 to 91%, while reducing specificity from 88 to 81% [34].

Agrawal et al. [36] performed parallel dual reporting in teleradiology emergency radiology which resulted in 3.8% disagreements. The authors suggested that abdominal CT and head/spine MRI were the most common error sources and that a focused double reading of error-prone case types may be considered for optimum utilisation of resources.

Second reading by a sub-specialist

Six articles reported on abdominal imaging, five of these for distinct conditions, usually malignancy. The discrepancy rates for these varied from about 12% up to 50% [5, 38, 39, 41, 42].

- Bell and Patel [40] reported on 1,303 cases of body CT with the primary report from non-sub-specialised radiologists and found a higher frequency of clinically relevant discrepancies in the 742 cases that were double read by radiologists with a higher degree of sub-specialisation.
- In chest radiography, a statistically significantly higher rate of seemingly obvious misdiagnoses was found for non-chest speciality radiologists [43], while a thoracic radiologist had higher sensitivity and reported fewer indeterminate nodules in chest CT for colorectal cancer [44].
- In neuroradiology, two articles demonstrated the benefit from sub-specialist second opinion [46, 47], while two did not [45, 48].
- In paediatric radiology, Eakins et al. [49] found a high rate of discrepancies in neuroimaging and body studies, while discrepancies were much rarer in extremity radiography [50]. In abdominal trauma CT, 12 new injuries were found in 98 patients [51].

Discussion

This systematic review found a wide range of significant discrepancy rates, from 0.4 to 22%, with minor discrepancies being much more common. Most of this variability is probably due to study setting. Double reading generally increased sensitivity at the cost of decreased specificity. One area where double reading seems to be important is in trauma CT, which is not surprising considering the large number of images and often stressful conditions under which the primary reading is performed. Thoracic and abdominal CT were also associated with more discrepancies than head and spine CT [54]. Higher rates of discrepancy can be expected in cases with a high probability of disease with complicated imaging findings [5].

More surprising was the fact that double reading by a sub-specialist almost invariably changed the initial reports to a high degree, although the second reader was also the reference standard for the study, which might have introduced bias. This leads to the conclusion that it might be more efficient to strive for sub-specialised readers than to implement double reading. It might also be more cost-efficient considering the fact that in one study, double reading of one-third of all studies consumed an estimated 20–25% of all working hours in the institutions concerned [1]. In modern digital radiology it is easy to send images to another hospital, and it should thus be possible to include even small radiology departments in a large virtual department where all radiologists can be sub-specialised. However, even a sub-specialised reader is subject to the same basic reading errors and this needs further study comparing outcomes from various reading strategies.

The primary goal of the current study was to evaluate double reading in a clinically relevant context, i.e. where the
| First author, country | Year | Clinical setting | Method | Total number of cases | Results | Conclusion |
|-----------------------|------|------------------|--------|-----------------------|---------|------------|
| **Double reading by peers; CT** | | | | | | |
| Yoon LS, USA [13] | 2002 | Abdominal and pelvic trauma CT | Original report reviewed by a second non-blinded reader | 512 | 30% discordant readings, patient care was changed in 2.3% | Most discordant readings do not result in change in patient care |
| Agostini C, France [14] | 2008 | CT in polytrauma patients | Official interpretation reviewed by two radiologists | 105 | 280 lesions out of 765 (37%) were not appreciated during first reading, of these 31 major | Double reading is recommended in polytrauma patients |
| Sung JC, USA [15] | 2009 | Trauma CT from outside hospital | Re-interpretation by local radiologist | 206 | 12% discrepancies, judged as perceptual in 26% and interpretive in 70% | Double reading is beneficial |
| Eurin M, France [16] | 2012 | Whole-body trauma CT | Scans were re-interpreted for missed injuries by second reader, blinded to initial data | 177 | 157 missed injuries in 85 patients (48%), predominantly minor and musculoskeletal | Double reading is recommended |
| **Bechtold RE, USA [17]** | 1997 | Abdominal CT | Clinical report compared with reference standard from a consensus panel | 694 | 56 errors in 694 patients | 7.6% errors in CT abdomen, 2.7% clinically significant |
| Fultz PJ, USA [18] | 1999 | CT of ovarian cancer | Four independent readers tested single, single with checklist, paired consensus, and replicated readings | 147 | Sensitivity for single reader, checklist, paired and replicated readings were 93 to 94% with specificities 79, 80, 82 and 85%, almost all non-significant | The diagnostic aids did not lead to an improved mean observer performance, however an increase in the mean specificity occurred with replicated readings |
| Gollub MJ, USA [12] | 1999 | CT abdomen and pelvis in cancer patients | Original report and re-interpretation report by a non-blinded reader in another hospital was retrospectively compared | 143 | Major disagreement in 17%, treatment change in 3% | Reinterpretation of body CT scans can have a substantial effect on the clinical care |
| Johnson KT, USA [19] | 2006 | CT colonography with virtual dissection software | Single reading compared with double reading, no consensus | 20 | Sensitivity/specificity single reading 78–85/80–100%, sensitivity double reading 75–95% | 5 mm polyps and larger. No significant increase in sensitivity with double reading |
| Murphy R, UK [20] | 2010 | CT colonography with minimal preparation | Independent and blinded double reading | 186 | Single reading found 11 cancers and double reading 12, at the expense of 5 false positives for single and 10 for double reading, giving positive predictive values of 69% and 54%, respectively | There is some benefit of double reporting; however, with major resource implications and at the expense of increased false-positives |
| Lauritzen PM, Norway [21] | 2016 | Abdominal CT | Double reading, peer review | 1,071 | Clinically important changes in 14% | Primary reader chose which studies should be double-read, thus probably more difficult cases. Important changes were made less frequently when abdominal radiologists were first readers, more frequently when they were second readers, and more frequently to urgent examinations |
| First author, country | Year | Clinical setting | Method | Total number of cases | Results | Conclusion |
|-----------------------|------|------------------|--------|-----------------------|---------|------------|
| Wormanns D, Germany [8] | 2004 | Low-dose chest CT for pulmonary nodules | Independent double reading | 9 patients with 457 nodules | Sensitivity of single reading, 54%; double reading, 67%; single reader with CAD, 79%. False positives, 0.9–3.9% for readers, 7.2% for CAD | Double reading and CAD increased sensitivity, CAD more than double reading, at the cost of more false positives for CAD |
| Rubin GD, USA [22] | 2005 | Pulmonary nodules on CT | Independent reading by three radiologists, reference standard by two thoracic radiologists + CAD | 20 | Sensitivity single reading 50%, double reading 63%, single reading + CAD 76–85% | Double reading increased sensitivity slightly. Inclusion of CAD increased sensitivity further |
| Womanns D, Germany [23] | 2005 | Chest CT for pulmonary nodules | Independent double reading of low- and standard-dose CT | 9 patients with 457 nodules | Sensitivity of single reading, 64%; double reading, 79%; triple reading, 87% (low-dose CT) | Double reading significantly increased sensitivity |
| Lauritzen PM, Norway [24] | 2016 | Chest CT | Double reading, peer review | 1,023 | Clinically important changes in 9% | Primary reader chose which studies should be double-read, thus probably more difficult cases. More clinically important changes were made to urgent examinations, chest radiologists made more clinically important changes than the other consultants |
| Lian K, Canada [25] | 2011 | CT angiography of the head and neck | Blinded double reading by two neuroradiologists in consensus, compared with original report by a neuroradiologist | 503 | 26 significant discrepancies were found in 20 cases, overall miss rate of 5.2% | Double reading may decrease the error rate |
| Double reading by peers; radiography | | | | | | |
| Markus JB, Canada [26] | 1990 | Double-contrast barium enema | Double and triple reporting, colonoscopy as reference standard | 60 | Sensitivity/specificity of single reading, 68/96%; double reading, 82/91% | Double reading increased sensitivity and reduced specificity slightly |
| Tribl B, Austria [27] | 1998 | Small-bowel double contrast barium examination in known Crohn’s disease | Clinical report double read by two gastrointestinal radiologists; ileoscopy as reference standard | 55 | Sensitivity/specificity of single reading, 66/82%; double reading, 68/91% | Negligible improvement by double reading |
| Canon CL, USA [28] | 2003 | Barium enemas, double- and single-contrast | Two independent readers, final diagnosis by consensus. Endoscopy as reference standard | 994 | Sensitivity/specificity of single reading, 76/91%; simultaneous dual reading, 76/86% | Dual reading led to an increased number of false positives which reduced specificity. No benefit in sensitivity |
| Marshall JK, Canada [29] | 2004 | Small-bowel meal with pneumocolon for diagnosis of ileal Crohn’s disease | Double reading of clinical report by two gastrointestinal radiologists with endoscopy as reference standard | 120 | Sensitivity/specificity of single reading, 65/90%; double reading, 81/94% | Possibly increased sensitivity with double reading, however unclear information on how study was performed |
| First author, country    | Year | Clinical setting                      | Method                                                                 | Total number of cases | Results                                                                 | Conclusion                                                                                                                                 |
|-------------------------|------|---------------------------------------|------------------------------------------------------------------------|-----------------------|------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| Hessel SJ, USA [7]      | 1978 | Chest radiography                     | Independent reading by eight radiologists, combined by various strategies | 100                   | Pseudo-arbitration was the most effective method overall, reducing errors by 37%, increasing correct interpretations 18%, and adding 19% to the cost of an error-free interpretation |                                                                                                                                           |
| Quekel LGBA, Netherlands [6] | 2001 | Chest radiography                     | Independent and blinded double reading as well as dual reading in consensus | 100                   | Sensitivity/specitivity of single reading, 33/92%; independent double reading, 46/87%; simultaneous dual reading, 37/92% | Double or dual reading increased sensitivity and decreased specificity, altogether little impact on detection of lung cancer in chest radiography |
| Robinson PJA, UK [30]   | 1999 | Skeletal, chest and abdominal radiography in emergency patients | Independent reading by three radiologists                              | 402                   | Major disagreements in 5–9% of cases                                  | The magnitude of interobserver variation in plain film reporting is considerable                                                        |
| Sofía DJ, USA [31]      | 2004 | General radiography                   | Independent double reading by two radiologists                        | 3,763                 | Significant disagreement in 3%                                       | Part of a quality assurance program                                                                                                        |
| Siegle RL, USA [33]     | 1998 | General radiology in six departments, including CT, nuclear medicine and ultrasound | Double reading by a team of QC radiologists                           | 11,094                | Mean rate of disagreement 4.4% in over 11,000 images                 | Rates of disagreement lower than previously reported                                                                                     |
| Wakeley CJ, UK [32]     | 1995 | MR imaging                            | Double reading by two radiologists. Arbitration in case of disagreement | 100                   | 9 false-positive, 14 false-negative reports in 100 cases              | The study promotes the benefits of double reading MRI studies                                                                           |
| Warren RM, UK [34]      | 2005 | MR breast imaging                     | Blinded and independent double reading by two observers, 44 in total!  | 1,541                 | Sensitivity/specitivity of single reading, 80/88%; double reading, 91/81% | Double reading increased sensitivity at the cost of decreased specificity                                                                   |
| Babiarz LS, USA [35]    | 2012 | Neuroradiology cases                  | Original report by neuroradiologist, double reading by another neuroradiologist | 1,000                | 2% rate of clinically significant discrepancies                       | Low rate of disagreements, but all worked in the same institution                                                                       |
| Agrawal A, India [36]   | 2017 | Teleradiology emergency radiology     | Parallel dual reporting                                                | 3,779                 | 3.8% error rate, CT abdomen and MRI head/spine most common error sources | Focused double read of pre-identified complex, unfamiliar or error-prone case types may be considered for optimum utilisation of resources |
| Harvey HB, USA [37]     | 2016 | CT, MRI and ultrasound                | Peer review using consensus-oriented group review                      | 11,222                | Discordance in 2.7%, missed findings most common                      | Highest discordance rates in musculoskeletal and abdominal divisions                                                                   |
| Kalbhen CL, USA [38]    | 1998 | Abdominal CT for pancreatic carcinoma | Original report reviewed by sub-specialty radiologists                | 53                    | 32% discrepancies in 53 patients, all under-staging                  | Reinterpretation of outside abdominal CT was valuable for determining pancreatic carcinoma resectability                                  |
| First author, country | Year | Clinical setting | Method | Total number of cases | Results | Conclusion |
|----------------------|------|------------------|--------|-----------------------|---------|------------|
| Tilleman EH, Netherlands [39] | 2003 | CT or ultrasound in patients with pancreatic or hepatobiliary cancer | Reinterpretation by sub-specialised abdominal radiologist | 78 | 48% of ultrasound and 30% of CT studies were judged as not sufficient for reinterpretation, Major discordance in 8% for ultrasound, 12% for CT | Change in treatment strategy in 9%. Many initial reports were incomplete |
| Bell ME, USA [40] | 2014 | After-hours body CT | Abdominal imaging radiologists reviewed reports by non-sub-specialists | 1,303 | 4.4% major discrepancies in 742 cases double read by primary members of the abdominal imaging division, 2.0% major discrepancies in 561 cases double read by secondary members | The degree of sub-specialisation affects the rate of clinically relevant and incidental discrepancies |
| Lindgren EA, USA [5] | 2014 | CT, MR and ultrasound from outside institutions submitted for secondary interpretation | Second opinion by sub-specialised GI radiologist | 398 | 5% high clinical impact and 7.5% medium clinical impact discrepancies | The second reader had 2% medium clinical impact discrepancies. There was a trend towards overcalls in normal cases and misses in complicated cases with pathology |
| Wibmer A, USA [41] | 2015 | Diagnosis of extracapsular extension of prostate cancer on MRI | Second-opinion reading by sub-specialised genitourinary oncological radiologists | 71 | Disagreement between the initial report and the second-opinion report in 30% of cases, second-opinion correct in most cases | Reinterpretation by sub-specialist improved detection of extracapsular extension |
| Rahman WT, USA [42] | 2016 | Abdominal MRI in patients with liver cirrhosis | Re-interpretation by sub-specialised hepatobiliary radiologist | 125 | 10% of subjects had a discrepant diagnosis of hepatocellular cancer, and 10% of subjects had discrepant Milan status for transplant | Reinterpretations were more likely to describe imaging findings of cirrhosis and portal hypertension and more likely to make a definitive diagnosis of HCC 50% change in management |
| Double reading by sub-specialist; chest | | | | | | |
| Cascade PN, USA [43] | 2001 | Chest radiography | Performance of chest faculty and non-chest radiologists was evaluated | 485,661 | No difference in total rate of incorrect diagnoses, but non-chest faculty had a statistically significant higher rate of seemingly obvious misdiagnoses | There are several potential biases in the study which complicate the conclusions |
| Nordholm-Carstensen A, Denmark [44] | 2015 | Chest CT in colorectal cancer patients, classification of indeterminate nodules | Second opinion by sub-specialised thoracic radiologist | 841 | Sensitivity/specificity primary reading 74-99%, sub-specialist 92/100% | Higher sensitivity for the thoracic radiologist with fewer indeterminate nodules |
| Double reading by sub-specialist; neuro | | | | | | |
| Jordan MJ, USA [45] | 2006 | Emergency head CT | Original report reviewed by sub-specialty neuroradiologists | 1,081 | 4 (0.4%) clinically significant and 10 insignificant errors | Double reading of head CT by sub-specialist appears to be inefficient |
| First author, country | Year | Clinical setting | Method | Total number of cases | Results | Conclusion |
|----------------------|------|------------------|--------|-----------------------|---------|------------|
| Briggs GM, UK [46]   | 2008 | Neuro CT and MR  | Second opinion by sub-specialised neuro-radiologist | 506 | 13% major discrepancy rate | The benefit of a formal specialist second opinion service is clearly demonstrated |
| Zan E, USA [47]      | 2010 | Neuro CT and MR  | Reinterpretation by sub-specialised neuroradiologist | 4,534 | 7.7% of clinically important differences | Double reading is recommended |
| Jordan YJ, USA [48]  | 2012 | Head CT, stroke detection | Original report reviewed by sub-specialty neuroradiologists | 560 | 0.7% rate of clinically significant discrepancies | Low rate of discrepancies and double reading by sub-specialist was reported as inefficient. However the study was limited to ischaemic non-haemorrhagic disease |
| Eakins C, USA [49]   | 2012 | Paediatric radiology | Cases referred to a children’s hospital were reviewed by a paediatric sub-specialist | 773 | 22% major disagreements | Interpretations by sub-specialty radiologists provide important clinical information |
| Bisset GS, USA [50] | 2014 | Paediatric extremity radiography | Official interpretation reviewed by one paediatric radiologist, blinded to official report. Arbitration by a second radiologist when reports differed | 3,865 | Diagnostic errors in the form of a miss or overcall occurred in 2.7% of the radiographs | Diagnostic errors quite rare in paediatric extremity radiography. Clinical significance of the discrepancies was not evaluated |
| Onwubiko C, USA [51]| 2016 | CT abdomen in paediatric trauma patients | Re-review of images by paediatric radiologist | 98 | 12.2% new injuries identified, 3% had solid organ injuries upgraded, and 4% downgraded to no injury | Clear benefit to having referring hospital trauma CT scans reinterpreted by paediatric radiologists |
| Loevner LA, USA [52] | 2002 | CT and MR in head and neck cancer patients | Second opinion by sub-specialised neuroradiologist | 136 | Change in interpretation in 41%, TNM change in 34%, mostly up-staging | Sub-specialist increases diagnostic accuracy |
| Kabadi SJ, USA [53]  | 2017 | CT, MR and ultrasound from outside institutions submitted for formal over-read | Retrospective review | 362 | 12.4% had clinically significant discrepancies | 64% perceptual errors Strategies for reducing errors are suggested |

CAD computer aided diagnosis, HCC hepatocellular cancer
second reader double-reads the case in a non-blinded context before the report is finalised. Only two studies used a method approaching this [12, 13]. Reinterpretation of body CT in another hospital was beneficial [12] but double reading of abdominal and pelvic trauma CT resulted in only 2.3% changes in patient care [13].

One method for peer review of radiology reports is error scoring such as is practiced in the RadPeer program [55]. This differs from clinical double reading in that it does not confer direct benefit for the patient at hand. The use of old reports can also be seen as a form of second reading [56].

Double reading has been evaluated in a recent systematic review which dedicated much space to mammography screening [57]. This review suggested further attention to other common examinations and implementation of double reading as an effective error-reducing technique. This should be coupled with studies on its cost-effectiveness. The literature search in the current study resulted in some additional articles and a slightly different conclusion, which is not surprising considering the wide variety of studies included. In a systematic review on CT diagnosis, a major discrepancy rate of 2.4% was found, even lower when the secondary reader was non-blinded [54]. There is also a Cochrane review on audit and feedback which borders on the subject in the current study, even though no radiology-specific articles were included [58]. Errors and discrepancies in radiology have been covered in a recent review article [59].

Observer variation analysis is now customary when evaluating imaging modalities or procedures, or when starting studies on larger image materials [60–62], and it is well known that observer variation can be small or large between observers, due to differences in experience and variations in image quality or ease of detection and characterisation of a lesion.

A quality assessment of the individual evaluated articles was not performed in the current study. It was judged to be not feasible to get any meaningful results out of this, due to the wide variability in subject matter and methods.

Limitations of the study are the widely varying definitions of what is a clinically important discrepancy, which makes a meaningful meta-analysis impossible. In studies with a sub-specialised second reader there is a risk that the discrepancy rate is inflated since the second reader decides what should be included in the report.

In conclusion, the systematic review found, in general, rather low discrepancy rates when double-reading radiological studies. The benefit of double reading must be balanced by the considerable number of working hours a systematic double reading scheme requires. A more profitable scheme might be to use systematic double reading for selected, high-risk examination types. A second conclusion is that there seems to be a value in sub-specialisation for increased report quality. A consequent implementation of this would have far-reaching organisational effects.

Acknowledgements Many thanks to Birgitta Eriksson at the Medical Library at Örebro University for assistance with literature searches.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Lauritzen PM, Hurlen P, Sandbaek G, Gulbrandsen P (2015) Double reading rates and quality assurance practices in Norwegian hospital radiology departments: two parallel national surveys. Acta Radiol 56:78–86
2. Husby JA, Espeland A, Kalyanpur A, Brocker C, Haldorsen IS (2011) Double reading of radiological examinations in Norway. Acta Radiol 52:516–521
3. Birkelo CC, Chamberlain WE et al (1947) Tuberculosis case finding: a comparison of the effectiveness of various roentgenographic and photofluorographic methods. J Am Med Assoc 133:359–366
4. Garland LH (1949) On the scientific evaluation of diagnostic procedures. Radiology 52:309–328
5. Lindgren EA, Patel MD, Wu Q, Melikian J, Hara AK (2014) The clinical impact of subspecialized radiologist reinterpretation of abdominal imaging studies, with analysis of the types and relative frequency of interpretation discrepancies. Abdom Imaging 39:1119–1126
6. Quekels LG, Goei R, Kessels AG, van Engelshorn JM (2001) Detection of lung cancer on the chest radiograph: impact of previous films, clinical information, double reading, and dual reading. J Clin Epidemiol 54:1146–1150
7. Hessel SJ, Herman PG, Swensson RG (1978) Improving performance by multiple interpretations of chest radiographs: effectiveness and cost. Radiology 127:589–594
8. Wormald D, Beyer F, Diederich S, Ludwig K, Heindel W (2004) Diagnostic performance of a commercially available computer-aided diagnosis system for automatic detection of pulmonary nodules: comparison with single and double reading. Råö 176:953–958
9. Law RL, Slack NF, Harvey RF (2008) An evaluation of a radiographer-led barium enema service in the diagnosis of colorectal cancer. Radiography 14:105–110
10. Garrett KG, De Cecco CN, Schoepf UJ et al (2014) Residents’ performance in the interpretation of on-call “triple-rule-out” CT studies in patients with acute chest pain. Acad Radiol 21:938–944
11. Guerin G, Jamali S, Soto CA, Guilbert F, Raymond J (2015) Interobserver agreement in the interpretation of outpatient head CT scans in an academic neuroradiology practice. AJNR Am J Neuroradiol 36:24–29
12. 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 center. Radiology 210:109–112
13. Yoon LS, Haims AH, Brink JA, Rabinovici R, Forman HP (2002) Evaluation of an emergency radiology quality assurance program at a level I trauma center: abdominal and pelvic CT studies. Radiology 224:42–46
14. Agostini C, Durieux M, Milot L et al (2008) Value of double reading of whole body CT in polytrauma patients. J Radiol 89: 325–330
15. Sung JC, Sodickson A, Ledbetter S (2009) Outside CT imaging among emergency department transfer patients. J Am Coll Radiol 6:626–632
16. Eurin M, Haddad N, Zappa M et al (2012) Incidence and predictors of missed injuries in trauma patients in the initial hot report of whole-body CT scan. Injury 43:73–77
17. Bechtold RE, Chen MY, Ott DJ et al (1997) Interpretation of abdominal CT: analysis of errors and their causes. J Comput Assist Tomogr 21:681–685
18. Fultz PJ, Jacobs CV, Hall WJ et al (1999) Ovarian cancer: comparison of observer performance for four methods of interpreting CT scans. Radiology 212:401–410
19. Johnson KT, Johnson CD, Fletcher JG, MacCarty RL, Summers RL (2006) CT colonography using 360-degree virtual dissection: a feasibility study. AJR Am J Roentgenol 186:90–95
20. Murphy R, Slater A, Uberoi R, Bungay H, Ferrett C (2010) Reduction of perception error by double reporting of minimal preparation CT colon. Br J Radiol 83:331–335
21. Lauritzen PM, Andersen JG, Stokke MV et al (2016) Radiologist-initiated double reading of abdominal CT: retrospective analysis of the clinical importance of changes to radiology reports. BMJ Qual Saf 25:595–603
22. Rubin GD, Lyo JK, Paik DS et al (2005) Pulmonary nodules on multi-detector row CT scans: performance comparison of radiologists and computer-aided detection. Radiology 234:274–283
23. Womanns D, Ludwig K, Beyer F, Heindel W, Diederich S (2005) Detection of pulmonary nodules at multilayer-detector CT: effectiveness of double reading to improve sensitivity at standard-dose and low-dose chest CT. Eur Radiol 15:14–22
24. Lauritzen PM, Stavnen K, Andersen JG et al (2016) Double reading of current chest CT examinations: clinical importance of changes to radiology reports. Eur J Radiol 85:199–204
25. Lian K, Bharatha A, Avir RI, Symons SP (2011) Interpretation errors in CT angiography of the head and neck and the benefit of double reading. AJNR Am J Neuroradiol 32:2132–2135
26. Markus JB, Somers S, O’Malley BP, Stevenson GW (1999) Double-contrast barium enema studies: effect of multiple reading on perception error. Radiology 175:155–156
27. Tribl B, Tureschek K, Mostbeck G et al (1998) Conflicting results of ileoscopy and small bowel double-contrast barium examination in patients with Crohn’s disease. Endoscopy 30:339–344
28. Canon CL, Smith JK, Morgan DE et al (2003) Double reading of barium enemas: is it necessary? AJR Am J Roentgenol 181:1607–1610
29. Marshall JK, Cawdron R, Zealley I, Riddell RH, Somers S, Irvine AJ (2004) Prospective comparison of small bowel meal with pneumocolon versus ileo-colonoscopy for the diagnosis of ileal Crohn’s disease. Am J Gastroenterol 99:1321–1329
30. Robinson PJ, Wilson D, Coral A, Murphy A, Verow P (1999) Variation between experienced observers in the interpretation of accident and emergency radiographs. Br J Radiol 72:323–330
31. Soffa DJ, Lewis RS, Sunshine JH, Bhargavan M (2004) Disagreement in interpretation: a method for the development of benchmarks for quality assurance in imaging. J Am Coll Radiol 1: 212–217
32. Wakeley CJ, Jones AM, Kabala JE, Prince D, Goddard PR (1995) Audit of the value of double reading magnetic resonance imaging films. Br J Radiol 68:358–360
33. Siegle RL, Baram EM, Reuter SR, Clarke EA, Lancaster JL, McManus CA (1998) Rates of disagreement in imaging interpretation in a group of community hospitals. Acad Radiol 5:148–154
34. Warren RM, Pointon L, Thompson D et al (2005) Reading protocol for dynamic contrast-enhanced MR images of the breast: sensitivity and specificity analysis. Radiology 236:779–788
35. Babiarz LS, Yousen DM (2012) Quality control in neurroradiology: discrepancies in image interpretation among academic neuroradiologists. AJNR Am J Neuroradiol 33:37–42
36. Agrawal A, Koundinya DB, Raju JS, Agrawal A, Kalyanpur A (2017) Utility of contemporaneous dual read in the setting of emergency telereadiology reporting. Emerg Radiol 24:157–164
37. Harvey HB, Alkasab TK, Prabhakar AM et al (2016) Radiologist peer review by group consensus. J Am Coll Radiol 13:656–662
38. 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. AJR Am J Roentgenol 171:1571–1576
39. Tilleman EM, Phoa SS, Van Delden OM et al (2003) Reinterpretation of radiological imaging in patients referred to a tertiary referral centre with a suspected pancreatic or hepatobiliary malignancy: impact on treatment strategy. Eur Radiol 13:1095–1099
40. Bell ME, Patel MD (2014) The degree of abdominal imaging (AI) subspecialization of the reviewing radiologist significantly impacts the number of clinically relevant and incidental discrepancies identified during peer review of emergency after-hours body CT studies. Abdom Imaging 39:1114–1118
41. Wibmer A, Vargas HA, Donahue TF et al (2015) Diagnosis of extracapsular extension of prostate cancer on prostate MRI: impact of second-opinion readings by subspecialized genitourinary oncologic radiologists. AJR Am J Roentgenol 205:W73–W78
42. Rahman WT, Hussain HK, Parikh ND, Davenport MS (2016) Reinterpretation of outside hospital MRI abdomen examinations in patients with cirrhosis: is the OPTN mandate necessary? AJR Am J Roentgenol 191:1–7
43. Cascade PN, Kazerooni EA, Gross BH et al (2001) Evaluation of competence in the interpretation of chest radiographs. Acad Radiol 8:315–321
44. Nordholm-Carstensen A, Jorgensen LN, Wille-Jorgensen PA, Hansen H, Harling H (2015) Indeterminate pulmonary nodules in colorectal-cancer: do radiologists agree? Ann Surg Oncol 22: 543–549
45. Jordan MJ, Lightfoote JB, Jordan JE (2006) Quality outcomes of reinterpretation of brain CT imaging studies by subspecialty experts in neuroradiology. J Natl Med Assoc 98:1326–1328
46. Briggs GM, Flynn PA, Worthington M, Rennie I, McKinstry CS (2008) The role of specialist neuroradiology second opinion reporting: is there added value? Clin Radiol 63:791–795
47. Zan E, Yousem DM, Carone M, Lewin JS (2010) Second-opinion consultations in neuroradiology. Radiology 255:135–141
48. Jordan YJ, Jordan JE, Lightfoote JB, Ragland KD (2012) Quality outcomes of reinterpretation of brain CT studies by subspecialty experts in stroke imaging. AJR Am J Roentgenol 199:1365–1370
49. Ekins C, Ellis WD, Pruthi S et al (2012) Second opinion interpretations by specialty radiologists at a pediatric hospital: rate of disagreement and clinical implications. AJR Am J Roentgenol 199:916–920
50. Bisset GS 3rd, Crowe J (2014) Diagnostic errors in interpretation of pediatric musculoskeletal radiographs at common injury sites. Pediatr Radiol 44:552–557
51. Onwubiko C, Mooney DP (2016) The value of official reinterpretation of trauma computed tomography scans from referring hospitals. J Pediatr Surg 51:486–489
52. Loevner LA, Sonners AI, Schulman BJ et al (2002) Reinterpretation of cross-sectional images in patients with head and neck cancer in the setting of a multidisciplinary cancer center. AJNRM Am J Neuroradiol 23:1622–1626
53. Kabadi SJ, Krishnaraj A (2017) Strategies for improving the value of the radiology report: a retrospective analysis of errors in formally over-read studies. J Am Coll Radiol 14:459–466
54. Wu MZ, McIntnes MD, Macdonald DB, Kielar AZ, Duigenan S (2014) CT in adults: systematic review and meta-analysis of interpretation discrepancy rates. Radiology 270:717–735
55. Jackson VP, Cushing T, Abujudeh HH et al (2009) RADPEER scoring white paper. J Am Coll Radiol 6:21–25
56. Berbaum KS, Smith WL (1998) Use of reports of previous radiologic studies. Acad Radiol 5:111–114
57. Pow RE, Mello-Thoms C, Brennan P (2016) Evaluation of the effect of double reporting on test accuracy in screening and diagnostic imaging studies: a review of the evidence. J Med Imaging Radiat Oncol 60:306–314
58. Ivers N, Jamtvedt G, Flottorp S et al (2012) Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. https://doi.org/10.1002/14651858.CD000259.pub3
59. Brady AP (2017) Error and discrepancy in radiology: inevitable or avoidable? Insights Imaging 8:171–182
60. Collin D, Dunker D, Göthlin JH, Geijer M (2011) Observer variation for radiography, computed tomography, and magnetic resonance imaging of occult hip fractures. Acta Radiol 52:871–874
61. Geijer M, Göthlin GG, Göthlin JH (2007) Observer variation in computed tomography of the sacroiliac joints: a retrospective analysis of 1383 cases. Acta Radiol 48:665–671
62. Ornetti P, Maillefert JF, Paternotte S, Dougados M, Gossec L (2011) Influence of the experience of the reader on reliability of joint space width measurement. A cross-sectional multiple reading study in hip osteoarthritis. Joint Bone Spine 78:499–505
63. Groth-Petersen E, Moller AV (1955) Dual reading as a routine procedure in mass radiography. Bull World Health Organ 12:247–259
64. Griep WA (1955) The role of experience in the reading of photofluorograms. Tubercle 36:283–286
65. Yerushalmy J (1955) Reliability of chest radiography in the diagnosis of pulmonary lesions. Am J Surg 89:231–240
66. Williams RG (1958) The value of dual reading in mass radiography. Tuberbile 39:367–371
67. Discher DP, Wallace RR, Massey FJ Jr (1971) Screening by chest photofluorography in los angeles. Arch Environ Health 22:92–105
68. Felson B, Morgan WKC, Bristol LJ et al (1973) Observations on the results of multiple readings of chest films in coal miners’ pneumoconiosis. Radiology 109:19–23
69. Angerstein W, Oehmke G, Steinbruck P (1975) Observer error in interpretation of chest-radiographs (author’s transl). Z Erkr Atmungsorgane 142:87–93
70. Herman PG, Hessel SJ (1975) Accuracy and its relationship to experience in the interpretation of chest radiographs. Investig Radiol 10:62–67
71. Labruné M, Dayras M, Kalifa G, Rey JL (1976) “Cirrhotic’s lund”. A new radiological entity? 182 CASES (AUTHOR’S TRANSL). J Radiol Electro Med Nucl 57:471–475
72. Stitik FP, Tockman MS (1978) Radiographic screening in the early detection of lung cancer. Radiol Clin N Am 16:347–366
73. Aoki M (1985) Lung cancer screening-its present situation, problems and perspectives. Gan To Kagaku Ryoho 12:2265–2272
74. Gjorup T, Nielson H, Jensen LB, Jensen AM (1985) Interobserver variation in the radiographic diagnosis of gastric ulcer. Gastroenterologists’ guesses as to level of interobserver variation. Acta Radiol Diagn (Stockh) 26:289–292
75. Gjorup T, Nielson H, Bording Jensen L, Morup Jensen A (1986) Interobserver variation in the radiographic diagnosis of duodenal ulcer disease. Acta Radiol Diagn (Stockh) 27:41–44
76. Fukushima K, Matsumoto T, Inuma TA et al (1989) On the assessment of the diagnostic accuracy of imaging diagnosis by ROC and BVC analyses—in reference to X-ray CT and ultrasound examination of liver disease. Nihon Igaku Hoshasen Gakkai Zasshi 49:863–874
77. Stephens S, Martin I, Dixon AK (1989) Errors in abdominal computed tomography. J Med Imaging 3:281–287
78. Shaw NJ, Hendry M, Eden OB (1990) Inter-observer variation in interpretation of chest X-rays. Scott Med J 35:140–141
79. Anderson N, Cook HB, Coates R (1991) Colonoscopy—colorectal cancer missed on barium enema. Gastrointest Radiol 16:123–127
80. Corbett SS, Rosenfeld CR, Laptook AR et al (1991) Intraobserver and interobserver reliability in assessment of neonatal cranial ultrasounds. Early Hum Dev 27:9–17
81. Haug PJ, Clayton PD, Tocino I et al (1991) Chest radiography: a tool for the audit of report quality. Radiology 180:271–276
82. Hopper KD, Rosetti GF, Edmiston RB et al (1991) Diagnostic radiology peer review: a method inclusive of all interpreters of radiographic examinations regardless of specialty. Radiology 180:557–561
83. Slovis TL, Guzzardo-Dobson PR (1991) The clinical usefulness of teleradiology of neonates: expanded services without expanded staff. Pediatr Radiol 21:333–335
84. Matsumoto T, Doi K, Nakamura H, Nakanishi T (1992) Potential usefulness of computer-aided diagnosis (CAD) in a mass survey for lung cancer using photo-fluorographic films. Nihon Igaku Hoshasen Gakkai Zasshi 52:500–502
85. Frank MS, Mann FA, Gillespy T (1993) Quality assurance: a system that integrates a digital dictation system with a computer data base. AJR Am J Roentgenol 161:1101–1103
86. O’Shea TM, Volberg F, Dillard RG (1993) Reliability of interpretation of cranial ultrasound examinations of very low-birthweight neonates. Dev Med Child Neurol 35:97–101
87. Friedman DP (1995) Manuscript peer review at the AJR: facts, figures, and quality assessment. AJR Am J Roentgenol 164:1007–1009
88. Gacinovic S, Buscombe J, Costa DC, Hilson A, Bomanji J, Ell PJ (1996) Inter-observer agreement in the reporting of 99Tcm-DMSA renal studies. Nucl Med Commun 17:596–602
89. Nitowski LA, O’Connor RE, Reese CL (1996) The rate of clinically significant plain radiograph misinterpretation by faculty in an emergency medicine residency program. Acad Emerg Med 3:782–789
90. Filippi M, Barkhof F, Bressi S, Yousry TA, Miller DH (1997) Inter-rater variability in reporting enhancing lesions present on standard and triple dose gadolinium scans of patients with multiple sclerosis. Mult Scler 3:226–230
91. Gale ME, Vincent ME, Robbins AH (1997) Teleradiology for remote diagnosis: a prospective multi-year evaluation. J Digit Imaging 10:47–50
92. Law RL, Longstaff AJ, Slack N (1999) A retrospective 5-year study on the accuracy of the barium enema examination performed by radiographers. Clin Radiol 54:80–83 discussion 83-84
93. Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Doi K (2000) Relative gains in diagnostic accuracy between computer-aided diagnosis and independent double reading. Proc SPIE 3981:10–15
94. Kopans DB (2000) Double reading. Radiol Clin N Am 38:719–724
95. Connolly DJA, Traill ZC, Reid HS, Copley SJ, Nolan DJ (2002) The double contrast barium enema: a retrospective single centre audit of the detection of colorectal carcinomas. Clin Radiol 57:29–32
96. Fidler JL, Johnson CD, MacCarty RL, Welch TJ, Harmsen WS (2002) Detection of flat lesions in the colon with CT colonography. Abdom Imaging 27:292–300
97. Leslie A, Virjee JP (2002) Detection of colorectal carcinoma on double contrast barium enema when double reporting is routinely performed: an audit of current practice. Clin Radiol 57:184–187
98. Murphy M, Loughran CF, Birchenough H, Savage J, Sutcliffe C (2002) A comparison of radiographer and radiologist reports on radiographer conducted barium enemas. Radiography 8:215–221
99. Summers RM, Aggarwal NR, Sneller MC et al (2002) CT virtual bronchoscopy of the central airways in patients with Wegener's granulomatosis. Chest 121:242–250
100. Baarslag HJ, van Beek EJ, Tijssen JG, van Delden OM, Bakker AJ, Reekers JA (2003) Deep vein thrombosis of the upper extremity: intra- and inter-observer study of digital subtraction venography. Eur Radiol 13:251–255
101. Johnson CD, Harmsen WS, Wilson LA et al (2003) Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. Gastroenterology 125:311–319
102. Quekel LGBA, Goei R, Kessels AGH, Van Engelschoven JMA (2003) The limited detection of lung cancer on chest X-rays. Ned Tijdschr Geneeskd 147:1048–1056
103. Borgstede JP, Lewis RS, Bhargavan M, Sunshine JH (2004) RADPEER quality assurance program: a multifacility study of interpretive disagreement rates. J Am Coll Radiol 1:59–65
104. Halsted MJ (2004) Radiology peer review as an opportunity to reduce errors and improve patient care. J Am Coll Radiol 1:984–987
105. Järvenpää R, Holli K, Hakama M (2004) Double-reading of plain radiographs–no benefit with regard to earliness of diagnosis of cancer recurrence: a randomised follow-up study. Eur J Cancer 40:1668–1673
106. Johnson CD, MacCarty RL, Welch TJ et al (2004) Comparison of the relative sensitivity of CT colonography and double-contrast barium enema for screen detection of colorectal polyps. Clin Gastroenterol Hepatol 2:314–321
107. Smith PD, Temte J, Beasley JW, Mundt M (2004) Radiographs in the office: is a second reading always needed? J Am Board Fam Pract 17:256–263
108. Taylor P, Given-Wilson R, Champie J, Potts HW, Johnston K (2004) Assessing the impact of CAD on the sensitivity and specificity of film readers. Clin Radiol 59:1099–1105
109. BamhartHX, Song J, Haber MJ (2005) Assessing intra, inter and total agreement with replicated readings. Stat Med 24:1371–1384
110. Booth AM, Mannion RAJ (2005) Radiographer and radiologist perception error in reporting double contrast barium enemas: a pilot study. Radiology 11:249–254
111. Bradley AJ, Rajashanker B, Atkinson SL, Kennedy JN, Purcell RS (2005) Accuracy of reporting of intravenous urograms: a comparison of radiographers with radiology specialist registrars. Clin Radiol 60:807–811
112. Den Boon S, Bateman ED, Enarson DA et al (2005) Development and evaluation of a new chest radiograph reading and recording system for epidemiological surveys of tuberculosis and lung disease. Int J Tuberc Lung Dis 9:1088–1096
113. Jarvenpaa R, Holli K, Hakama M (2005) Resource savings in the single reading of plain radiographs by oncologist only in cancer patient follow-up: a randomized study. Acta Oncol 44:149–154
114. Peldschus K, Herzog P, Wood SA, Cheema JI, Costello P, Schoepf UJ (2005) Computer-aided diagnosis as a second reader: spectrum of findings in CT studies of the chest interpreted as normal. Chest 128:1517–1523
115. Bimbbaum LM, Filion KB, Joyal D, Eisenberg MJ (2006) Second reading of coronary angiograms by radiologists. Can J Cardiol 22:1217–2221
116. Borgstede J, Wilcox P (2007) Quality care and safety know no borders. Biomed Imaging Interv J 3:e34
117. Foinant M, Lipiecza E, Buc E et al (2007) Impact of computed tomography on patient's care in nontraumatic acute abdomen: 90 patients. J Radiol 88:559–566
118. Freioli F, Bertoletti L, Napoli A et al (2007) Computer-aided detection (CAD) in lung cancer screening at chest MICT: ROC analysis of CAD versus radiologist performance. J Thorac Imaging 22:241–246
119. Capobianco J, Jasinowodolinski D, Szarf G (2008) Detection of pulmonary nodules by computer-aided diagnosis in multidetector computed tomography: preliminary study of 24 cases. J Bras Pneumol 34:27–33
120. Johnson CD, Mandauc A, Fletcher JG et al (2008) Noncathartic CT colonography with stool tagging: performance with and without electronic stool subtraction. AJR Am J Roentgenol 190:361–366
121. Law RL, Titcomb DR, Carter H, Longstaff AJ, Slack N, Dixon AR (2008) Evaluation of a radiographer-provided barium enema service. Color Dis 10:394–396
122. Nellensteijn DR, ten Duis HJ, Oldenziel J, Polak WG, Hulscher AJ (2009) Only moderate intra- and inter-observer agreement between radiologists and surgeons when grading blunt paediatric hepatic injury on CT scan. Eur J Pediatr Surg 19:392–394
123. Brinkjikki W, Kallmes DF, White JB, Lanzino G, Morris JM, Cloft HJ (2010) Inter- and intraobserver agreement in CT characterization of nonaneurysmal perforating subarachnoid hemorrhage. AJNR Am J Neuroradiol 31:1103–1105
124. Liu PT, Johnson CD, Miranda R, Patel MD, Phillips CJ (2010) A reference standard-based quality assurance program for radiology. J Am Coll Radiol 7:61–66
125. Monico E, Schwartz I (2010) Communication and documentation of preliminary and final radiology reports. J Healthc Risk Manag 30:23–25
126. Saurin JC, Pelleau J, Soussan EB et al (2010) Small-bowel capsule endoscopy diagnoses early and advanced neoplasms in asymptomatic patients with lymph syndrome. Endoscopy 42:1057–1062
127. Schu YR, Feder E, Balsim I, Levin VF, Bleicher AG, Branstetter BF (2010) Optimizing radiology peer review: a mathematical model for selecting future cases based on prior errors. J Am Coll Radiol 7:431–438
128. Brook OR, Kane RA, Tyagi G, Siewert B, Kruskal JB (2011) Lessons learned from quality assurance: errors in the diagnosis of acute cholecystitis on ultrasound and CT. AJR Am J Roentgenol 196:597–604
166. Natarajan V, Bosch P, Dede O et al (2017) Is there value in having radiology provide a second reading in pediatric Orthopaedic clinic? J Pediatr Orthop 37:e292–e295

167. O’Keeffe MM, Davis TM, Siminoski K (2016) Performance results for a workstation-integrated radiology peer review quality assurance program. Int J Qual Health Care 28:294–298

168. Olthof AW, van Ooijen PM (2016) Implementation and validation of PACS integrated peer review for discrepancy recording of radiology reporting. J Med Syst 40:193

169. Pedersen MR, Graumann O, Horlyck A et al (2016) Inter- and intraobserver agreement in detection of testicular microlithiasis with ultrasonography. Acta Radiol 57:767–772

170. Verma N, Hippe DS, Robinson JD (2016) JOURNAL CLUB: assessment of Interobserver variability in the peer review process: should we agree to disagree? AJR Am J Roentgenol 207:1215–1222

171. Vural U, Sarisoys HT, Akansel G (2016) Improving accuracy of double reading in chest X-ray images by using eye-gaze metrics. Proceedings SIU 2016—24th Signal Processing and Communication Application Conference, 16-19 May 2016, Zonguldak, pp 1209-1212

172. Steinberger S, Plodkowski AJ, Latson L et al (2017) Can discrepancies between coronary computed tomography angiography and cardiac catheterization in high-risk patients be overcome with consensus reading? J Comput Assist Tomogr 41:159–164

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.