ORIGINAL ARTICLE

AUDIT TO ESTIMATE THE PROPORTION OF RADIOLOGY REPORTS CHANGED DURING DOUBLE READING OF IMAGES AND TO ASSESS THEIR POTENTIAL CLINICAL IMPACT

Ummara Siddique Umer, Aruba Nawaz Khattak, Mahwish Jabeen, Abdullah Safi, Hadia Abid, Anisa Sundal

Department of Radiology, Rehman Medical Institute (RMI), Peshawar, Pakistan.

PJR October - December 2022; 32(4): 203-208

ABSTRACT

OBJECTIVES: The objectives of this study were to estimate the proportion of radiology reports that were changed during double reading and to assess the potential clinical impact of these changes. **MATERIAL AND METHODS:** Rates of double reading in radiology department were explored in survey issued to consultant radiologists, covering practice of double reading, department guidelines and quality improvement work. The responses of consultant radiologists, grouped according to workplace, were used to validate management responses about working hours consumed by double reading. The clinical importance of changes to radiology reports was estimated retrospectively. We collected pairs of preliminary and final reports from 1500 double read examinations (CT, MRI, X ray, ultrasound). Exploratory analysis of associations between clinically important changes and characteristics of patients, examinations, and readers was performed with multivariate logistic regression. We also constructed two random effects models to test for clustering of clinically important report changes in separate examinations read by the same radiologist. **RESULTS:** We found double reading rate of 21% for CT and MRI, 17% for X ray images, 15% for ultrasound and 5% for mammograms. All exams read by consultants, consuming an estimated 20-25% of consultant working hours. By modality double reading rates were highest for Magnetic Resonance Imaging (MRI) (47%) and CT (33%), intermediate for X-ray (24%) and fluoroscopy (23%), and lowest for ultrasonography (16%) and intervention (16%). Chest radiologists and sonologists made more clinically important changes than other second readers. The severity of the radiological findings was increased in 30% of the clinically important changes. Double reading caused upto 25% increase in time consumed for reporting. **CONCLUSION:** Double reading has a major impact on workflow and output directly by consuming working hours. The rates of clinically important changes to radiology reports following double reading indicate that some quality assurance of radiological interpretation is warranted.

Keywords: Radiology reports; impact; Double reading

Introduction

To assess the world around them, humans rely on their eyes along with other senses. Radiologists transmit their visual impression of images by non-visual means (the report) to the clinicians.[1] Garland first noticed and outlined errors in medical imaging in 1959.[2] Over a course of 61 years these errors have persisted and even today are a hindrance to the concept of perfect reporting . These errors and dis-

Correspondence : Dr. Ummara Siddique
Department of Radiology,
Rehman Medical Institute (RMI),
Peshawar, Pakistan.
Email: ummara_81@hotmail.com

Submitted 9 October 2022, Accepted 15 October 2022

crepancies in the radiology practice have an estimated day to day rate of 3-5 % in studies reported, the number being higher for targeted studies.[3] Currently, despite multiple advances in the imaging technology, perceptual abilities of human eye and brain are a limitation.[4] Given the situation, throughout the world multiple solutions have been devised to overcome this hurdle.[5] Among many methods, one way to increase the quality of reporting is double reading of imaging studies between peers.[6] Double reading is basically a practice in which two readers read and interpret the same imaging examination and help each other in reducing error, thus increasing sensitivity of reporting.[7] The concept has originated from inter observer variation in radiology, first introduced in late 1940 s, when mass chest radiography for tuberculosis screening was evaluated.[8,9] This practice can be conducted in several ways. Applied prospectively, it may be used for quality assurance of radiology reports, and it is routine in the education of residents.[9]

It has been suggested that 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.Although the concept is simple, double reading can be conducted in several ways. There are large variations in the reported effect of double reading in different settings, and the cost effectiveness is not well established. Double reading can be broadly divided into three categories: (1) both primary and secondary reading 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.

Final comments made in a radiology report have a major impact on direction of patient s management. Double reading reduces error rate. The objectives of this study were to estimate the proportion of radiology reports that were changed during double reading and to assess the potential clinical impact of these changes.

Material and Methods

This is a cross sectional observational study conducted at Radiology Department of Rehman Medical Institute Peshawar. Rates of double reading in radiology department were explored in survey issued to consultant radiologists, covering practice of double reading, department guidelines and quality improvement work. The responses of consultant radiologists grouped according to workplace were used to validate management responses about working hours consumed by double reading. The clinical importance of changes to radiology reports was estimated retrospectively.

We collected pairs of preliminary and final reports of 1500 consecutive double read examinations (CT, MRI and X-ray) and compared them for changes made. The hospital management and integration system software (HMIS) of Rehman Medical Institute for radio diagnostic reporting includes a separate column for internal notes. We used this column as a means of saving preliminary reports for comparison. On a daily basis a specific time period was allocated for double reading of reports. The time period was decided after generalized consensus among consultant radiologists and set during later hours of routine shift from 1400-1600 hrs. Extended time period expended beyond the set duration was marked as out of working hours. For assessment of clinical outcome a multidisciplinary team was devised constituting specialist doctors in fields of neurology, pulmonology, gastroenterology, urology, gynecology and general surgery with more than 10 years of experience. This team would review the submitted final reports (which were the initially double read reports), every Wednesday to pass a verdict about clinical rating of the changed reports. To facilitate the clinical raters, the patient population was restricted to in and out patients of the parent institute. Out of hospital referrals were excluded. Repeat examinations were not carried out.

For assessment of clinically significant outcome we took the study of Peter Maehre et al[10] as our precedent. The clinical raters were given a roughly drafted scale of 1-5 to rate the changes in radiological reports that signified impact on outcomes. Any score above 2 was taken as a cut off for clinically significant change. Data was collected based on patient s gender, age, in/outpatient status, urgency of examination, clinical notes of referral, identities of first and second readers, time of preliminary and final reports (during working hours of 0800 to 1600, or out of working hours. Preliminary and final reports were then com-

pared (preliminary findings had been saved before making changes). For clinical rating, reports with content change beyond misspelling and layout corrections were submitted. Rating was done on basis of color coded changes in radiology reports, patient s age and gender in keeping with clinical notes of referral. Thus clinically significant changes were grouped according to the issues concerned.

The reporting radiologists were also classified on basis of experience in years and subspecialty.

Exploratory analysis of associations between clinically important changes and characteristics of patients, examinations, and readers was performed with multivariate logistic regression. We also constructed two random effects models to test for clustering of clinically important report changes in separate examinations read by the same radiologist.

Results

We found double reading rate of 21% for CT and MRI, 17% for X ray images, and 5% for mammograms. Overall mean double reading rate was 19%. All procedure findings were read by consultants, consuming an estimated 20-25% of consultant working hours. By modality double reading rates were highest for Magnetic Resonance Imaging (MRI) (47%) and CT (33%), intermediate for X-ray (24%) and fluoroscopy (23%), and lowest for intervention (16%). Chest radiologists made more clinically important

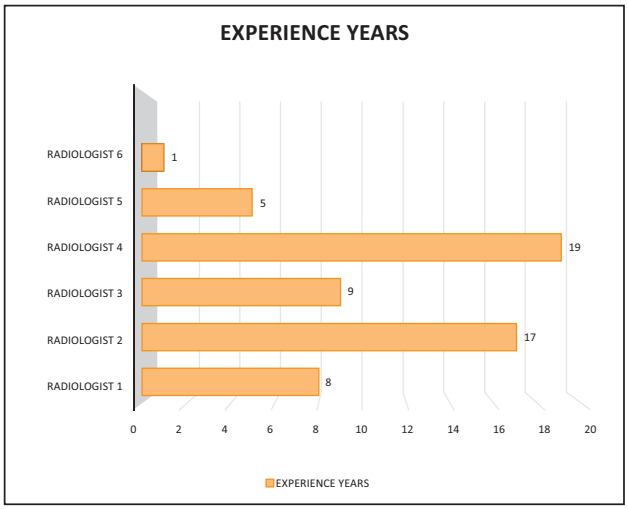

Figure 1: Bar chart showing experience of radiologists assessed for double read effect.

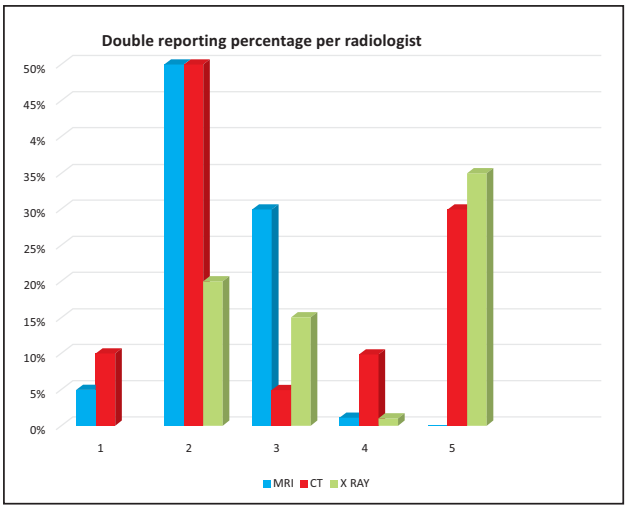

Figure 2: Higher double reading rates were observed at two ends of the spectrum. One end included radiologists with less experience years (<5 years), while the other end included radiologists with most experience, submitting more reports for double reading to improve quality assurance.

1. Double reading (per modality):

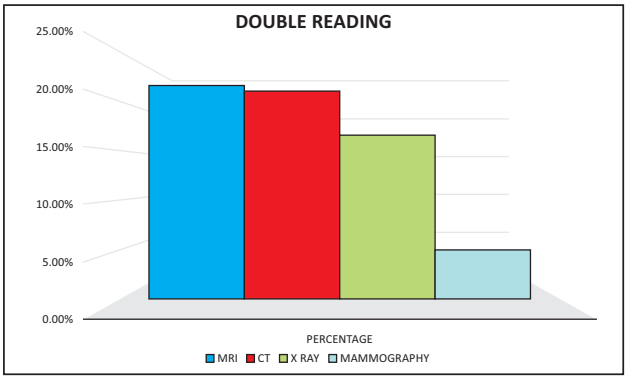

Figure 3: We found double reading rate of 21% for CT and MRI, 17% for X ray images, and 5% for mammograms.

2. Working Hours Consumed:

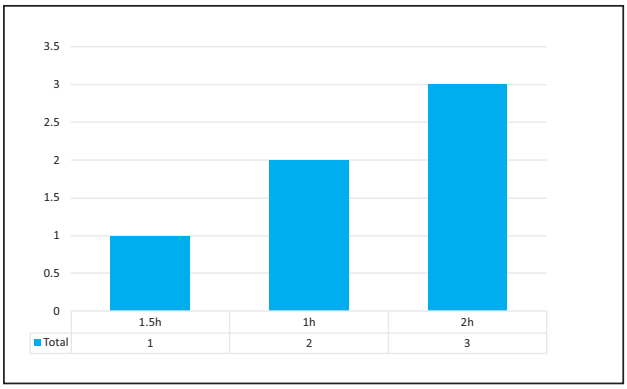

Figure 4: Bar chart showing double reading caused upto 25% increase in time consumed for reporting.

changes than other second readers. The severity of the radiological findings was increased in 30% of the clinically important changes. Double reading caused upto 25% increase in time consumed for reporting.

3. Clinically important changes:
Chest radiologists (Radiologists reporting the chest scans) made more clinically important changes than other second readers. Clinically important changes were made less frequently with CT abdomen. The severity of clinically important findings in double read radiological findings was increased in 30%.

Discussion

Misinterpreting the radiological examinations has been an important contributing factor to diagnostic errors, especially among less experienced yet qualified radiologists. Qualifying for standard radiology reporting is different from clinical approach which one learns with experience. Misinterpretation has been seen even among highly qualified young radiologists with little experience in the field. In such cases, double reading reduces interpretation errors and increases sensitivity. Diagnostic errors remain a relentless and inevitable occurrence in radiology. As proposed by Melvin et al, discrepancy meetings are an important and effective forum for review of errors in radiology.[11,12] It is postulated that 4% of radiologist s daily work will contain errors.[13] Review of literature reveals that an autopsy study of a group of patients showed that radiological misinterpretation caused 8%, and contributed to another 33% of diagnostic errors in patients with relevant imaging.[14] Most radiologists hold a very firm view on the concept of double reading; their argument is either for or against i.e. it reduces error and improves quality but it is time consuming with wastage of time and resources. In a radiology department, however it is crucial for improvement in quality services, The main goal should be to keep double reading among peers with similar sub-specialization interests. It causes significant decrease in rate of misses and overcalls with the aim of establishing the added value of double reading by human observers. Royal college of Radiologists has recommended peer feedback for all radiology department s regular learning from Discrepancies meetings.[15] In the United States, similarly 5% of cases require continuous peer review for on-going credentialing by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) for some time.

In our study most of the time consumed in double reading was due to perceptual misses, accounting for up to 75% cases. In this regards our study is comparable to the study of JJ Donald et al.[16] Previously published data has established multiple psychophysiological factors as possible culprits for these errors. A few noteworthy mentionable include observer alertness level, workload, distraction factors, fatigue, reporting speed and not ideal viewing conditions; such errors tend to be sporadic in nature.[17,18]

The double reading rate for CXR in our study was 24% where errors mostly occurred related to a lung nodule (<3 cm) or a small pneumothorax being missed. In this regards our study was comparable to that of Quekel et al 19 who found 19% missing rate for lung nodules missed having mean diameter of 16mm. Previous literature have reported an even higher incidence of such misses. As a possible measure for reducing such errors it was emphasized to include the comparison with previous imaging, including previous CXR if available for reducing any possible diagnostic error. Results of application of this practice will be published in another study.

Double reading was highest for MR (41%) and second highest for the CT reports (33%). Most of the time consumed in CT double reporting was due to the false negative errors in reporting like incidental subsegmental pulmonary artery emboli, anomaly of vascular structures, bone lesions, omental nodularity and incidentelomas with few MR reports having false negative errors related to secondary ramifications in perianal fistulae, discrepancies in grading of morbidly adherent placentas, under-call of posterior serosal plaque for possibility of deep pelvic endometriosis. In much of these aspects our study coincided with McCreadie[20] and Oliverstudies.[21] For reducing such errors, use of different windows in CT reporting was emphasized as proposed by Horton et al.[22]

Segregation of time consumed in double reporting due to errors have been postulated to be a more frequent occurrence in pediatric and geriatric age groups,[23] however our study was unable to properly extrapolate the significance of age group in collection of this data due to smaller sample size of pediatric

population and lack of pediatric specialist approach of the reporting radiologists. In this regards our study was again comparable to that of JJ Donald et al.

This study aimed to apply principles of quality assurance in departmental working after using world renowned guidelines. In the process of data collection it was observed that time consumed in double reading was adding to the workload of consultants. Therefore, benefit of double reading must be balanced by the considerable number of working hours a systematic double-reading scheme requires. Waiting time for report should be at least 48 hours in non-emergency cases, thus giving adequate time for double read. However, the study lacked some integral components. The sample size fell short for consolidated search of errors and double reading in different modalities. The study has the potential of being refined if sample selection is specified for individual modality rather than combined multimodality approach for data collection. Another possible room for improvement lies in the segregation of systems for reviewing the errors in reporting as well as the time consumed in double reading. This is backed by suggestions by some radiologists that it might be more efficient to strive for sub-specialized readers than to implement double reading, considering the wastage of time and resources it is associated with.

There is also a general consensus among the authors to go regional after implementing these changes in approach to quality assurance. Inclusion of other hospitals in data collection process with help not only in reviewing and reshaping the model of quality assurance in the region but also improve patient care by reducing the number of hospital visits.

Qualifying for standard radiology reporting is different from clinical approach which one learns with experience. However, the main goal of a double read report should be to keep double reading among peers with similar sub-specialization interests as it causes significant decrease in rate of misses and overcalls with the aim of establishing the added value of double reading by human observers.

Conclusion

We concluded from our results that Double reading has a major impact on workflow and output directly by consuming 20-25% more of the working hours. The rates of clinically important changes to radiology reports following double reading indicate that quality assurance of radiological interpretation is warranted.

Conflict of Interest: Declared none by authors.

References

1. Sabih A, Sabih Q, Khan AN. Image perception and interpretation of abnormalities; can we believe our eyes? Can we do something about it?. Insights into imaging. Feb 2011; **2(1):** 47-55.

2. Garland LH. Studies on accuracy of diagnostic procedures. AJR. 1959; **82:** 25-38.

3. Brady AP. Error and discrepancy in radiology: inevitable or avoidable?. Insights into imaging. Feb 2017; **8(1):** 171-82.

4. Kim YW, Mansfield LT. Fool me twice: delayed diagnoses in radiology with emphasis on perpetuated errors. American journal of roentgenology. Mar 2014; **202(3):** 465-70.

5. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive error in radiology. American Journal of Roentgenology. Apr 2017; **208(4):** 739-49.

6. Geijer H, Geijer M. Added value of double reading in diagnostic radiology, a systematic review. Insights into imaging. 2018 Jun; **9(3):** 287-301.

7. Goddard P, Leslie A, Jones A, Wakeley C, Kabala J. Error in radiology. The British journal of radiology. Oct 2001; **74(886):** 949-51.

8. Birkelo CC, Chamberlain WE, Phelps PS, Schools PE, Zacks D, Yerushalmy J. Tuberculosis case finding: A comparison of the effectiveness of various roentgenographic and photofluorographic methods. Journal of the American Medical Association. Feb 1947; **133(6):** 359-66.

9. Garland LH. On the scientific evaluation of diagnostic procedures: Presidential address thirty-fourth annual meeting of the Radiological Society of North America. Radiology. Mar 1949; **52(3):** 309-28.

10. Lauritzen PM, Andersen JG, Stokke MV, Tennstrand AL, Aamodt R, Heggelund T, Dahl FA, Sandb k G, Hurlen P, Gulbrandsen P. Radiologist-initiated double reading of abdominal CT: retrospective analysis of the clinical importance of changes to radiology reports. BMJ quality & safety. Aug 2016; **25(8):** 595-603.

11. Melvin C, Bodley R, Booth A, Meagher T, Record C, Savage P. Managing errors in radiology: a working model. Clinical radiology. Sep 2004; **59(9):** 841-5.

12. Royal Colege of Radiologists Standards for radiological discrepancy meetings. London, UK : RCR ; 2007

13. Jackson VP, Cushing T, Abujudeh HH, Borgstede JP, Chin KW, Grimes CK, Larson DB, Larson PA, Pyatt Jr RS, Thorwarth Jr WT. RADPEER scoring white paper. Journal of the American College of Radiology. Jan 2009; **6(1):** 21-5.

14. G.S. Heriot GS, P. McKelvie P, A.G. Pitman AG. Diagnostic errors in patients dying in hospital: radiology s contribution. J. Med. Imaging Radiat. Oncol. 2009; **53(2):** 188-93.

15. The Royal College of Radiologists. Quality assurance in radiology reporting: peer feedback. London, UK, 2014. 1-18.

16. Donald JJ, Barnard SA. Common patterns in 558 diagnostic radiology errors. Journal of medical imaging and radiation oncology. Apr 2012; **56(2):** 173-8.

17. Kundel HL. Perception errors in chest radiography. InSeminars in Respiratory Medicine Jul 1989; **10(3):** pp. 203-210). Copyright' 1989 by Thieme Medical Publishers, Inc.

18. Renfrew DL, Franken Jr EA, Berbaum KS, Weigelt FH, Abu-Yousef MM. Error in radiology: classification and lessons in 182 cases presented at a problem case conference. Radiology. Apr 1992; **183(1):** 145-50.

19. Quekel LG, Kessels AG, Goei R, van Engelshoven JM. Miss rate of lung cancer on the chest radiograph in clinical practice. Chest. Mar 1999; **115(3):** 720-4.

20. Berland LL, Silverman SG, Gore RM, Mayo-Smith WW, Megibow AJ, Yee J, Brink JA, Baker ME, Federle MP, Foley WD, Francis IR. Managing incidental findings on abdominal CT: white paper of the ACR incidental findings committee. Journal of the American College of Radiology. Oct 2010; **7(10):** 754-73.

21. McCreadie G, Oliver TB. Eight CT lessons that we learned the hard way: an analysis of current patterns of radiological error and discrepancy with particular emphasis on CT. Clinical radiology. May 2009; **64(5):** 491-9.

22. Horton KM, Johnson PT, Fishman EK. MDCT of the abdomen: common misdiagnoses at a busy academic center. American Journal of Roentgenology. Mar 2010; **194(3):** 660-7.

23. Diaz S, Ekberg O. The frequency of diagnostic errors in radiologic reports depends on the patient's age. ActaRadiologica. Oct 2010; **51(8):** 934-8.