CONCORDANCE AND DISCORDANCE BETWEEN RADIOLOGY RESIDENTS AND CONSULTANT RADIOLOGIST INTERPRETATION OF CT BRAIN

Madiha Pervaiz[1], Ummara Siddique Umer[2], Muhammad Abdullah[3], Ghulam Ghaus[4], Muhammad Kamran Khan[5], Muhammad Sohail[6], Hammad Ur Rehman[7]

How to cite this article

Pervaiz M, Umer US, Abdullah M, Ghaus G, Khan MK, Sohail M, et al. Concordance and Discordance Between Radiology Residents and Consultant Radiologist Interpretation Of CT Brain. J Gandhara Med Dent Sci. 2024;11(4):7-11

Date of Submission: 23-10-2024
Date Revised: 30-08-2024
Date Acceptance: 04-09-2024

[1]Resident, Department of Radiology, Rehman Medical Institute, Peshawar
[3]Research Fellow, Department of Radiology, Rehman Medical Institute, Peshawar
[4]Consultant, Radiologist, Head of Department, Rehman Medical Institute, Peshawar
[5]Fellow, Department of Radiology, Rehman Medical Institute, Peshawar
[6]Assistant Professor, Department of Neurology, Rehman Medical Institute, Peshawar
[7]Resident, Department of Radiology, Rehman Medical Institute, Peshawar

Correspondence

[2]Ummara Siddique Umer, Associate Professor Radiology, Rehman Medical Institute, Peshawar
✆: +92-321-5220449
✉: ummara.umer@rmi.edu.pk

ABSTRACT

OBJECTIVES

The primary objective of this study is to assess the degree of concordance and discordance between the interpretations of computed tomography (CT) brain images by resident and consultant radiologists while emphasizing the critical significance of accurate image interpretation for informed clinical decision-making.

METHODOLOGY

The evaluation of radiology reports for CT Brain interpretation through a prospective analysis at the Radiology Department of Rehman Medical Institute over two years, from 1st October 2020 to 31st October 2022. A total of 198 patients who underwent cranial CT scans were interpreted by residents (R1, R2, R3, R4). Following this, the consultant radiologists reviewed the images and completed their reports. The reports of the residents and the consultant radiologists were then compared, and concordance was achieved when the residents' reports were consistent with the final radiologist's reports. The data collected were recorded in Microsoft Excel. The statistical analysis was performed using SPSS version 22 (IBM Corp., Armonk, NY), and the kappa coefficient was used to determine the level of agreement between residents and consultants.

RESULTS

Among the 198 CT Head reports evaluated, 186 of them were in agreement with the final report of the consultant radiologist. Of the correctly diagnosed cases, R1 correctly diagnosed 46 cases, R2 correctly diagnosed 80 cases, R3 correctly diagnosed 54 cases, and R4 correctly diagnosed 6 cases. Our study achieved a percentage agreement of 93.93, with a Cohen's kappa coefficient of 0.8.

CONCLUSION

The overall concordance rate between residents and consultant radiologists was 93.93%, with a kappa coefficient 0.8. This high kappa coefficient indicates strong agreement between the two groups.

KEYWORDS: *Concordance, Radiology, Diagnosis, Consultant*

INTRODUCTION

Computed tomography (CT) is a widely utilized initial imaging modality in the diagnosis and management of neurological disorders and cranio-cerebral trauma.[1] The accurate and timely interpretation of CT images is essential for precise diagnosis, appropriate clinical decisions, and optimal patient outcomes. Radiologists are trained during their residency programs to interpret radiological images, and it has been demonstrated that brief educational efforts and structured training programs can significantly improve the proficiency of radiology residents in interpreting CT images.[2] During residency, residents are initially trained in the basic imaging of the brain using CT. By the end of their first year, residents are provided with hands-on exposure to the reporting of CT brain images, which is then assessed in detail by a qualified radiologist. The level of agreement between residents and consultant radiologists in interpreting CT brain images is a crucial consideration, as residents' interpretations can significantly impact patient management. Inter-observer variability in the interpretation of radiological images, including head and chest CT scans, is a common issue impacting patient management and treatment. Inter-observer variability in the interpretation of radiological images, including head and chest CT scans, is a common issue impacting patient management and

treatment.[3] Factors contributing to this variability include differences in experience, knowledge, and interpretation skills among radiologists. However, training programs and structured reporting can help improve inter-observer variability, ensure comprehensive reporting findings, avoid missing minor details, and reduce errors in the interpretation of radiological images .[4] However, training programs and structured reporting can help improve inter-observer variability, ensure full reporting findings, avoid missing minor details, and reduce errors in the interpretation of radiological images.[5] Prior studies have examined the agreement between radiology residents and consultants in various imaging modalities, including CT brain. A survey conducted by Guérin G and fellows showed CT head findings by residents were concordant in 88.1% of cases with a kappa value of 0.86.[6] Similarly, another study by William K.Erly had an agreement rate of 91%, with a disagreement rate of 7% and a significant disagreement rate of 2%. The level of training was substantial (P = .032), which means upper-level residents had higher rates of agreement than junior residents in the evaluation of head CT in the emergency department. However, other studies have reported lesser levels of concordance between residents and consultants.[7] Improving agreement between residents and consultants is essential for ensuring accurate and timely diagnosis and appropriate patient care and management. Discordance between residents and consultants can lead to delays in treatment, unnecessary testing, and potentially harmful interventions. This study aims to determine the level of concordance and discordance between these two groups, identify common areas of disagreement, shortcomings and assess the impact of additional training, explore potential strategies for improving agreement, and highlight the importance of accurate interpretation for patient-appropriate patient care, management, and patient outcomes.

METHODOLOGY

This study was carried out through a cross-sectional, prospectively analyzed dataset at the Radiology Department of Rehman Medical Institute in Peshawar, Pakistan, spanning two years, from 1st October 2020 to 31st October 2022. Residents and consultant radiologists evaluated one hundred ninety-eight patients who underwent cranial computed tomography (CT) scans. The Rehman Medical Institute-Research Ethics Committee approved the study, with an Institutional Review Board (IRB) number of RMI/RMI-REC/Article Approval/92. A Microsoft Excel-based chart was created to assess the level of agreement and disagreement between the two groups in interpreting the CT head. The analysis included two groups: residents, consisting of first-year (R1), second-year (R2), third-year (R3), and final-year (R4) residents, with a total of 14 residents, who were further divided into four subgroups. The first subgroup (R1) comprised four residents from the first year, four from the second year, four from the third year, and two from the final year. The second group (B) comprised four radiology consultants. Computed tomography (CT) Head examination was conducted using a 128-slice CT system (Aquilion Toshiba, Japan) with unenhanced acquisitions of the head from 2cm below the base of the skull to the vertex in a plane parallel to the base of the skull. The following parameters were used: kV, 300 mA, rotation time of 0.75 seconds, and scan thickness of 0.5-1 mm. A resident performed the initial interpretation of the 128 MDCT Head images within the first hour of the scan, and the images were reviewed on the RMI PACS (picture archiving and communication system). The consultant radiologists, with a minimum of 5 years of clinical experience, were blinded to the resident's report findings and reviewed the images. All relevant information, such as age, gender, medical history, and cranial CT findings by the resident and consultant, was recorded on a datasheet in MS Excel. The analysis of the abnormal conclusions and evaluation of the concordance and discordance rate between both groups were conducted. CT scans with incomplete or missing reports or that were not interpretable due to technical issues were excluded from the analysis. The reviewers classified each report as concordant or discordant based on agreement with the final diagnosis. Discordant reports were further categorized based on the type and severity of the discrepancy. Statistical analysis was performed using SPSS version 22 (IBM Corp., Armonk, NY) and the kappa coefficient to determine the level of agreement between residents and consultants. A high kappa coefficient indicated strong agreement, while a low coefficient suggested poor agreement.

RESULTS

The study involved a total of 198 patients. Most of the patients were female, comprising 102 (54.8%) of the sample, while male patients accounted for 96 (48.48%). The mean age of the patients was 99.5 ± 57.3, and out of the 198 cases, only 12 were missed by the residents, while the remaining 186 cases agreed with the consultant radiologist. The consultant radiologist accurately reported all cases and did not miss any abnormal findings. The residents were divided into four subgroups based on their year of residency: R1, R2, R3,

and R4. Among the misinterpreted cases, first-year and second-year residents missed two cases each, while R3 and R4, who performed the best, missed no abnormal findings. In the correctly diagnosed cases, R1 diagnosed 46 cases correctly, R2 diagnosed 80 cases, R3 diagnosed 54 cases, and R4 diagnosed six cases correctly. The overall concordance between the resident and the consultant radiologist was 93.9% (kappa = 0.87; p = 0.0001), indicating substantial agreement between the two groups. The p-value of <0.05 was considered significant.

Table 1: Stratification for the Level of Radiology Residents

	Year of Training				Total
	R1	R2	R3	R4	
Confirmed	46	80	54	06	186
Missed	02	02	08	00	12

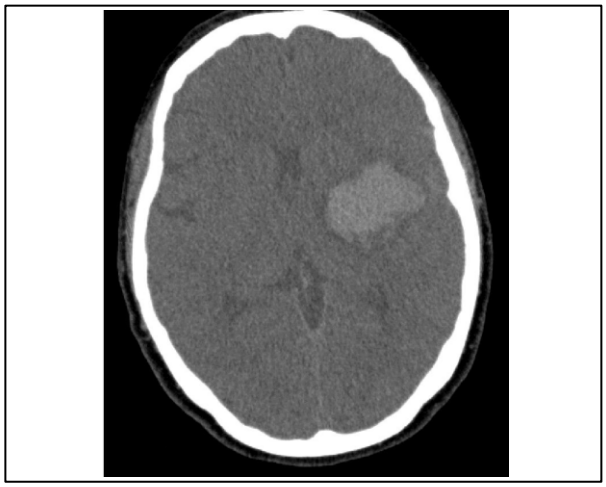

Figure 1: Axial Non-Contrast CT Head Reveals Left Basal Ganglia Bleed with Mild Vasogenic Edema

Figure 2: Left MCA Territory Infarct

Table 2: Shows Various Pathologies and Their Prevalence When Diagnosed on CT Head

Diagnosis	Frequency	%age
Normal Study	106	53.5
Hematoma	18	9.1
Infarct	36	18.2
Infarct with hemorrhagic conversion	08	4.0
Fracture	16	8.1
Farr's disease	02	1.0
SAH	04	2.0
Cerebral edema	02	1.0
Tonsillar herniation	02	1.0
Normal pressure hydrocephalus.	02	1.0
Basilar artery thrombus	02	1.0
Total	198	100.0

Table 3: Concordance and Discordance between Residents and Consultant Radiologists

Significant Abnormality by Radiology Resident	Significant abnormality by Consultant Radiologist		Kappa Value	P-Value
	Yes	No		
Yes	80	Yes	0.87	<0.0001
No	12	No		

*(kappa = 0.87, p = 0.0001, Concordance = 93.93)

DISCUSSION

Medical imaging plays a crucial role in diagnosing and managing various diseases. Radiology residents interpret medical images, including head CT and chest, abdomen, and pelvis CT scans, as part of their training. However, studies have shown high inter-observer variability in the interpretation of these images by radiology residents and junior attending radiologists.[8,9,10] Several factors may contribute to variability in an agreement between residents and consultants, including differences in training, level of exposure, experience, confidence, level of professionalism, and dedication toward a specialty. Some studies have suggested that additional training may improve agreement between residents and consultants. For example, the agreement rate between residents and radiologists increased with the senior residency year level for CTPA interpretations. There was a higher kappa for final-year residents (0.9) as compared to 2nd residents (0.7).[11] Several studies have shown that the level of agreement between residents and consultants in various imaging modalities may be related to the level of interest or training, but there is still limited knowledge on the specific issues that contribute to discordance and how to address them. A study by Khorasani et al. found only fair agreement between residents and attending radiologists in interpreting CT abdomen and pelvis images, with a kappa coefficient of 0.34.[12] Understanding the challenges and shortcomings and looking for solutions to improve agreement between residents and consultants may lead to more accurate interpretations of

CT brain images and better patient care, management, and improved outcomes. Several studies have investigated inter-observer variability in radiology interpretation, particularly in the context of emergency. One study found high inter-observer variability in interpreting cranial CT scans among residents and junior attending radiologists. Another study showed that the overall rate of misinterpretations was low for nighttime CT scans. However, no death occurred. Junior residents missed more correct findings.[13] Several factors contribute to inter-observer variability in radiology interpretation. These include differences in aptitude, experience, knowledge, and interpretation skills among radiologists. In addition, terminology used in radiology reports may not effectively convey diagnostic certainty, leading to misinterpretations.[14] In our study, we compared radiology residents and attending radiologists in interpreting CT scans. We found good agreement overall, with the best agreement seen with second-year radiology residents. Additional training led to a significant improvement in agreement between residents and senior radiologists in interpreting chest CT scans. One factor affecting agreement between residents and consultants is the difference in experience, knowledge, and terminology. Additional training, such as structured lectures, case-based discussions, hands-on experiences, and extra shifts, may improve residents' knowledge and confidence in interpreting CT scans. A study by Khorasani et al. found that A study found that radiology reports often lacked precise terminology, leading to miscommunication and errors. Standardization of terminology and implementation of quality control measures may be necessary to improve diagnostic accuracy and reduce discrepancies in interpretation.[15] Overall, understanding the factors contributing to discordance/ disagreement between residents and consultants in interpreting CT brain images is crucial for improving patient outcomes. By identifying common areas of disagreement and implementing strategies to enhance agreement, such as additional training, standardization of terminology, and intensified training programs, structure-based lectures in radiology departments can ensure accurate interpretation of CT brain images and optimal patient management.[16] There is a need to assess the level of agreement between radiology residents and consultant radiologists in interpreting CT brain images. To address the issue of inter-observer variability in radiology interpretation, several training programs have been developed that evaluated the effectiveness of a structured curriculum, intensified training programs for radiology residents in reducing errors in CT image interpretation and found that the curriculum led to a significant reduction in errors. Another approach to improving inter-observer variability is using structured reporting. This approach provides a standardized framework and templates for report generation that enhances the clarity and consistency of radiology reports. It also makes it easier for radiologists and residents to identify and communicate diagnostic issues related to image interpretation. Johnson et al. conducted a cohort study comparing structured reporting with conventional dictation and found that structured reporting significantly reduced errors and improved the clarity of radiology reports.[17,18] The studies highlight the issue of inter-observer variability in radiology interpretation, especially among radiology residents and junior attending radiologists. Factors contributing to this variability include differences in experience, knowledge, and interpretation skills. Comprehensive training programs for radiology residents are crucial to developing their interpretation skills and understanding of diagnostic terminology. AI systems can reduce variability by providing additional information to assist radiologists in image analysis, leading to more accurate and consistent interpretation. This can improve patient outcomes and impact clinical trials and research studies. Healthcare providers need to minimize variability and ensure timely and accurate reports.

LIMITATIONS

The study was conducted at a single centre, which may limit generalizability to other settings. In conclusion, inter-observer variability remains a significant challenge in radiology interpretation. Comprehensive training programs, structured reporting, templates, and additional technology can help reduce this variability and improve accuracy and consistency. It's worth noting that the studies cited in this discussion encompass both cranial and chest CT scans, suggesting that inter-observer variability is an issue across different types of medical images.

CONCLUSIONS

The issue of inter-observer variability in interpreting radiological images can impact patient management. Factors contributing to this include differences in experience, knowledge, and interpretation skills among radiologists. Training programs and structured reporting can help mitigate this issue. The study found that while most patients had average results, a substantial number were diagnosed with medical and surgical neurological conditions. It also highlighted the diagnostic performance of different resident doctors, with R2 performing the best and R3 performing the worst. Missed cases were also identified, presenting an area

for improvement.

CONFLICT OF INTEREST: None

FUNDING SOURCES: None

REFERENCES

1. Kim JJ, Gean AD: Imaging for the diagnosis and management of traumatic brain injury. Neurotherapeutics. 2011, 8:39-53.
2. Ganeshan D, Duong PA, Probyn L, et al.: Structured reporting in radiology. Academic radiology. 20181, 25:66-73.
3. Carney E, Kempf J, DeCarvalho V, Yudd A, Nosher J: Preliminary interpretations of after-hours CT and sonography by radiology residents versus final interpretations by body imaging radiologists at a level 1 trauma center. American Journal of Roentgenology. 2003, 181:367-73.
4. Guérin G, Jamali S, Soto CA, Guilbert F, Raymond J: Interobserver agreement in the interpretation of outpatient head CT scans in an academic neuroradiology practice. American Journal of Neuroradiology. 20151, 36:24-9.
5. Erly WK, Berger WG, Krupinski E, Seeger JF, Guisto JA: Radiology resident evaluation of head CT scan orders in the emergency department. American journal of neuroradiology. 2002, 23:103-7.
6. Wu MZ, McInnes MD, Blair Macdonald D, Kielar AZ, Duigenan S: CT in adults: systematic review and meta-analysis of interpretation discrepancy rates. Radiology. 2014, 270:717-35.
7. Babiarz LS, Yousem DM: Quality control in neuroradiology: discrepancies in image interpretation among academic neuroradiologists. American journal of neuroradiology. 20121, 33:37-42.
8. McCoubrie P, FitzGerald R: Commentary on discrepancies in discrepancy meetings. Clinical Radiology. 20141, 69:11-2.
9. Tamjeedi B, Correa J, Semionov A, Mesurolle B: Interobserver agreement between on-call radiology resident and general radiologist interpretations of CT pulmonary angiograms and CT venograms. PLoS One. 2015, 4:0126116.
10. Khorasani R, Bates DW, Teeger S, et al.: Is terminology used effectively to convey diagnostic certainty in radiology reports?. Acad Radiol. 2000, 7:1091-8.
11. Vaattovaara E, Nikki M, Nevalainen M, Ilmarinen M, Tervonen O: Discrepancies in interpretation of nighttime emergency computed tomography scans by radiology residents. Acta Radiologica Open. 2018, 7:2058460118807234.
12. Waite S, Grigorian A, Alexander RG, Macknik SL, Carrasco M, Heeger DJ, Martinez-Conde S: Analysis of perceptual expertise in radiology-Current knowledge and a new perspective. Frontiers in human neuroscience. 2019, 25:213.
13. Geijer H, Geijer M: Added value of double reading in diagnostic radiology, a systematic review. Insights into imaging. 2018, 9:287-301.
14. Zwaan L, Kok EM, van der Gijp A: Radiology education: a radiology curriculum for all medical students?. Diagnosis. 20171, 4:185-9.
15. Johnson AJ, Chen MY, Swan JS, Applegate KE, Littenberg B: Cohort study of structured reporting compared with conventional dictation. Radiology. 2009, 253:74-80.
16. Goldberg-Stein S, Chernyak V: Adding value in radiology reporting. Journal of the American College of Radiology. 20191, 16:1292-8.
17. Richardson ML, Garwood ER, Lee Y, Li MD, Lo HS, Nagaraju A, Nguyen XV: Probyn L, Rajiah P, Sin J, Wasnik AP. Noninterpretive uses of artificial intelligence in radiology. Academic Radiology. 2021, 1:1225-35.
18. Mutasa S, Sun S, Ha R. Understanding artificial intelligence based radiology studies: What is overfitting?. Clinical imaging. 2020 Sep 1;65:96-9.

CONTRIBUTORS

1. ***Madiha Pervaiz*** *- Concept & Design; Data Acquisition; Data Analysis/Interpretation; Drafting Manuscript; Critical Revision; Supervision*
2. ***Ummara Siddique Umer*** *- Data Analysis/Interpretation; Drafting Manuscript; Critical Revision; Supervision; Final Approval*
3. ***Muhammad Abdullah*** *- Data Acquisition; Data Analysis/Interpretation; Drafting Manuscript; Ciritical Revision*
4. ***Ghulam Ghaus*** *- Data Analysis/Interpretation; Drafting Manuscript; Critical Revision; Supervision; Final Approval*
5. ***Muhammad Kamran Khan*** *- Concept & Design; Data Acquisition; Data Analysis/Interpretation; Drafting Manuscript; Critical Revision*
6. ***Muhammad Sohail*** *- Data Acquisition; Data Analysis/Interpretation; Drafting Manuscript; Critical Revision*
7. ***Hammad ur Rehman*** *- Concept & Design; Data Acquisition; Data Analysis/Interpretation; Critical Revision; Supervision*