TPM Vol. 32, No. S8, 2025
ISSN: 1972-6325
https://www.tpmap.org/

Open Access

EVALUATION OF INTER-RATER AGREEMENT IN 1.5T MAGNETIC RESONANCE BRAIN IMAGING ARTIFACTS REDUCTION BEFORE AND AFTER IMAGE QUALITY OPTIMIZATION"AND PERCEPTION OF RADIOLOGY PROFESSIONALS ON IT. A RELIABILITY STUDY

JYOTI[1]*, DR. AVINASH MUNSHI[2], DR. DEEPAK SINGH[2], MS. MOHINI GUPTA[3], MS. RASHI SARAWAT[4], MR. GOPI KISHAN[5], MR. PIYUSH PANDEY[6], MR. MOHIT KUMAR DAHIYA[7]

[1]PH.D. SCHOLAR, DEPARTMENT OF RADIATION & IMAGING TECHNOLOGY, NIMS COLLEGE OF ALLIED & HEALTH CARE SCIENCES, NIMS UNIVERSITY RAJASTHAN, JAIPUR.
[2]PROFESSOR, DEPARTMENT OF RADIATION & IMAGING TECHNOLOGY, NIMS HOSPITAL, NIMS UNIVERSITY RAJASTHAN, JAIPUR.
[2]ASSISTANT PROFESSOR&RADIOLOGIST, DEPARTMENT OF RADIODIAGNOSIS, SGT MEDICAL COLLAGE, HOSPITAL & RESEARCH INSTITUTE, BUDHERA, GURUGRAM
[6]RADIOGRAPHER, DEPARTMENT OF RADIODIAGNOSIS, GTB HOSPITAL, DILSHAD [3]ASSISTANT PROFESSOR, DEPARTMENT OF RADIOLOGY, SCHOOL OF ALLIED HEALTH SCIENCES, GALGOTIAS UNIVERSITY, GREATER NOIDA, INDIA ,203201
[4]ASSISTANT PROFESSOR, DEPARTMENT OF RADIOLOGY AND IMAGING TECHNOLOGY, NOIDA INTERNATIONAL UNIVERSITY, NOIDA.
[5]ASSISTANT PROFESSOR, DEPARTMENT OF RADIO IMAGING TECHNOLOGY, MITHILA INSTITUTE OF TECHNOLOGY & MEDICAL SCIENCE (MITMS), BIBIPAKAR, DARBHANGA
[6]ASSISTANT PROFESSOR, DEPARTMENT OF RADIOLOGY AND IMAGING TECHNOLOGY,ERA UNIVERSITY, LUCKNOW.
[7]RADIOGRAPHER, DEPARTMENT OF RADIODIAGNOSIS, GTB HOSPITAL, DILSHAD GARDEN, DELHI.

ABSTRACT

Background: Magnetic Resonance Imaging (MRI) at 1.5 Tesla is widely used in clinical practice for brain imaging because of its accessibility and reliable diagnostic performance. However, image artifacts such as motion, susceptibility, and flow-related distortions often compromise diagnostic accuracy and interpretation[6] . These artifacts can mask pathology or create misleading appearances, thereby reducing confidence in clinical decision-making.

Several strategies have been proposed to optimize MRI image quality, including adjustments in acquisition protocols, use of artifact suppression techniques, and application of advanced post-processing tools[7] . Despite these efforts, assessment of image quality remains largely subjective, with potential variability among different raters[8] .

Methods: A comparative study was conducted on 41 patients undergoing routine brain MRI at 1.5T. Conventional images (pre-intervention) were compared with optimized images (post-intervention), incorporating protocol adjustments. Two independent skilled and senior radiologists, blinded to acquisition status, assessed image overall image quality using a 5-point Likert scale. Inter-rater agreement was measured with Cohen's kappa test[9]

Results: A total of 53 patients underwent brain MRI at 1.5T during the study period. Of these, 12 cases were excluded according to the predefined exclusion criteria. The final analysis was therefore conducted on 41 patients, which met the sample size requirement for this exploratory study. Where male candidates are 35 and females were 18 only (randomly selected)

In the analysis of 41 brain MRI cases at 1.5T, artifacts were initially identified in 17% of scans by Rater 1 (R1) and in 10% by Rater 2 (R2). Following the implementation of image quality optimization strategies, a substantial reduction in artifact prevalence was observed. Specifically, 86% of the artifact-affected cases demonstrated improvement according to R1, while 75% showed improvement according to R2. These findings suggest that the applied strategies were effective in mitigating artifacts and enhancing overall image quality.

Also**,** A cross-sectional survey was conducted among radiologists, radiographers, and radiology technologists working in various healthcare settings. The survey included questions assessing demographic information, professional background, knowledge of image quality optimization applications in radiology, attitudes towards image quality optimization integration, perceived benefits and challenges, and overall awareness of image quality optimization advancements in the field.

The preliminary analysis indicated that while a majority of radiology professionals recognize the potential benefits of image quality optimization in enhancing diagnostic accuracy, there is a significant variation in the level of knowledge and understanding of image quality optimization applications. The perception of image quality optimization was generally positive, with respondents acknowledging its role in improving patient outcomes radiological procedures. However, there was a notable demand for more comprehensive educational programs to bridge the knowledge gap.
Conclusion: Image quality optimization strategies effectively reduced artifacts in 1.5T brain MRI, with inter-rater agreement fairly observed. While improvements were consistent, subjective variability highlights the need for integrating objective measures in future evaluations.
Keywords: Brain MRI, 1.5 Tesla, image quality, artifacts,Inter-Rater Agreement

INTRODUCTION

Magnetic Resonance Imaging (MRI) is a non-invasive diagnostic tool widely used in neuroimaging due to its superior soft tissue contrast and multiplanar capability. However, the presence of image artifacts remains a significant limitation, often compromising diagnostic accuracy and clinical decision-making. Artifacts may arise from patient-related factors, hardware limitations, or sequence-dependent parameters, and their impact is particularly pronounced in 1.5 Tesla (T) MRI systems, which continue to be the most commonly used scanners worldwide.[1]
Improving image quality in routine brain MRI has been the focus of several optimization strategies, including adjustments in acquisition protocols, artifact suppression techniques, and post-processing enhancements. These interventions aim to minimize motion artifacts, improve contrast-to-noise ratio, and enhance overall image clarity.[2] Despite such advances, the evaluation of image quality often depends on subjective assessments by radiologists or trained raters, which may lead to variability. Hence, assessing inter-rater agreement becomes crucial in determining the reliability of these evaluations.[3]
Cohen’s kappa (κ) statistic is a widely accepted method for measuring agreement between observers beyond chance. It provides an objective estimate of reliability in studies where categorical ratings are employed, such as the evaluation of artifact reduction and image quality.[4] According to the commonly accepted interpretation, κ values between 0.41–0.60 indicate moderate agreement, 0.61–0.80 substantial agreement, and values above 0.81 almost perfect agreement.[5]
Given the importance of artifact reduction in improving diagnostic outcomes, this study aims to evaluate inter-rater agreement in the assessment of brain MRI images acquired at 1.5T before and after the application of image quality optimization strategies. By employing kappa statistics, the study seeks to quantify the reliability of rater assessments and provide insight into the effectiveness of optimization interventions.
The study also to evaluate the attitude and perception, among radiology professionals towards image quality optimisation measures. By examining these factors, we can better understand the challenges and opportunities that image quality optimisation presents in the context of medical radiology and imaging technology. The findings will inform the development of targeted educational initiatives and support systems to enhance the proficiency and confidence of radiology professionals in leveraging image quality optimisation for improved clinical outcomes.

METHODOLOGY:

Data collection: This research employs a Comparative study design to compare the image quality in routine brain Magnetic Resonance Imaging (MRI) scans at 1.5T before and after implementing image quality improvement strategies. The study aims to assess the inter-rater agreement using Cohen's Kappa statistics.
The target population for this study comprises routine brain MRI scans performed at a 1.5T MRI system. A convenience sampling method will be employed to select the MRI scans from patients who underwent routine brain imaging during a specified period. The data collection process will span a defined timeframe to ensure an adequate sample size.
Specific image quality improvement strategies for routine brain MRI at 1.5T will be identified and implemented during the study. These strategies may include protocol optimization, hardware adjustments,. The details of each strategy and the rationale behind their selection will be documented.
Two trained radiologists will independently and blindly rate the image quality of each MRI scan before and after implementing the image quality improvement strategies.. Prior to the rating process, the raters will undergo a comprehensive training session to ensure consistent and reliable assessments.
Patient data and confidentiality will be handled in compliance with ethical guidelines.
TECHNIQUE:
To estimate the sample size needed for an exploratory assessment of inter-rater agreement using Cohen’s Kappa.
Formula for Sample Size for Estimating Cohen’s Kappa
n=Z2·P·(1−P)E2n=E2Z2·P·(1−P)
Where:

- n: required sample size (number of MRI scans)

- ZZ: Z-score corresponding to the desired confidence level
(for 95% confidence, Z=1.96Z=1.96)
- PP: expected proportion of agreement (i.e., expected Kappa value)
- EE: desired precision (margin of error around the estimate)

Calculation
- moderate agreement: P=0.6P=0.6
- 95% confidence level: Z=1.96Z=1.96
- margin of error of ±0.15: E=0.15E=0.15

formula:
n=(1.96)2·0.6·(1−0.6)(0.15)2n=(0.15)2(1.96)2·0.6·(1−0.6)
n=3.8416·0.6·0.40.0225=3.8416·0.240.0225=0.9219840.0225n=0.02253.8416·0.6·0.4=0.02253.8416·0.24 =0.02250.921984n≈41n≈41
Hence, required exploratory sample size = 41 MRI scans
This will give you a 95% confidence interval for Cohen's Kappa with a ±0.15 margin of error around an expected Kappa of 0.6.
SELECTION CRITERIA OF PATIENTS
INCLUSION CRITERIA:
- Patient those who are coming for routine brain scan.
- Subjects of either sex will be recruited with the age from 18 to 60 years.
- Only patients who are willing to participate in the study would be selected.

EXCLUSION CRITERIA:
- Patients with acute trauma.
- Pregnant patients
- Patient below the age of 18 and elderly patients above the age of 60.
- Patients who are not willing to participate in the study would be not selected.
- Patients with absolute contraindications.
- Incomplete Imaging Sequences: MRI scans with incomplete imaging sequences, missing essential sequences, or significant motion artifacts will be excluded.

Cohen's Kappa Analysis:
Cohen's Kappa statistics will be calculated to evaluate the inter-rater agreement between the two radiologists. This analysis will measure the level of agreement beyond what is expected by chance, and it will provide insights into the consistency and reliability of image quality assessment before and after implementing the improvement strategies[9].
Data analysis according to patients age group:
The current exploratorystudy a total 41 patient's data has been assessed (the actual total 53 patient's data has been assessed where, 12 were excluded (as per exclusion criteria) and total 41 as per sample size has been selected).

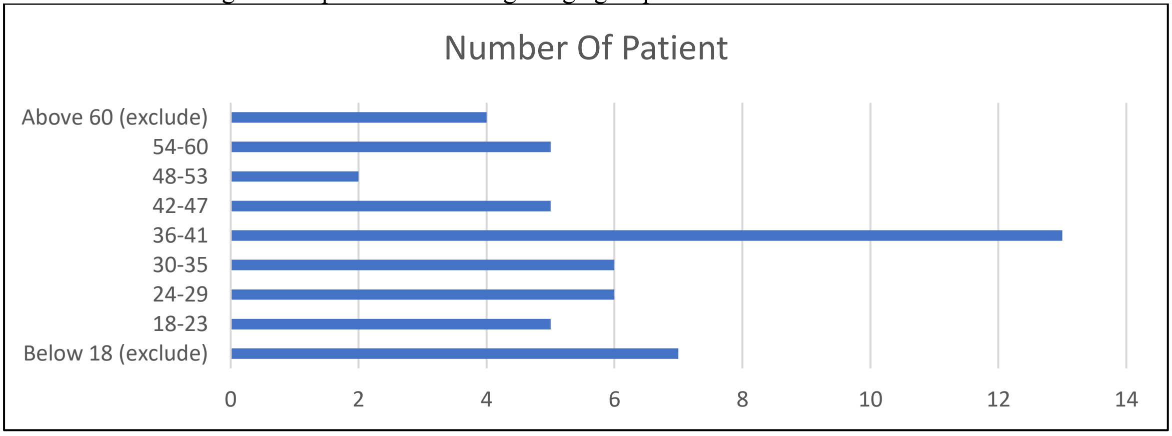
Total Age Range
- Minimum age = 18
- Maximum age = 60
- Total range = 60 − 18 = 42 years

Divide into 7 Equal Intervals
- 42 years ÷ 7 groups = 6 years per group

Age Group	Number Of Patient	No of Excluded patients
Below 18 (exclude)	7	7
18-23	5	-
24-29	6	-
30-35	6	-
36-41	13	1
42-47	5	-
48-53	2	-
54-60	5	-

Above 60 (exclude)	4	4
7 age groups	53	12(11age factor + 1artifact)

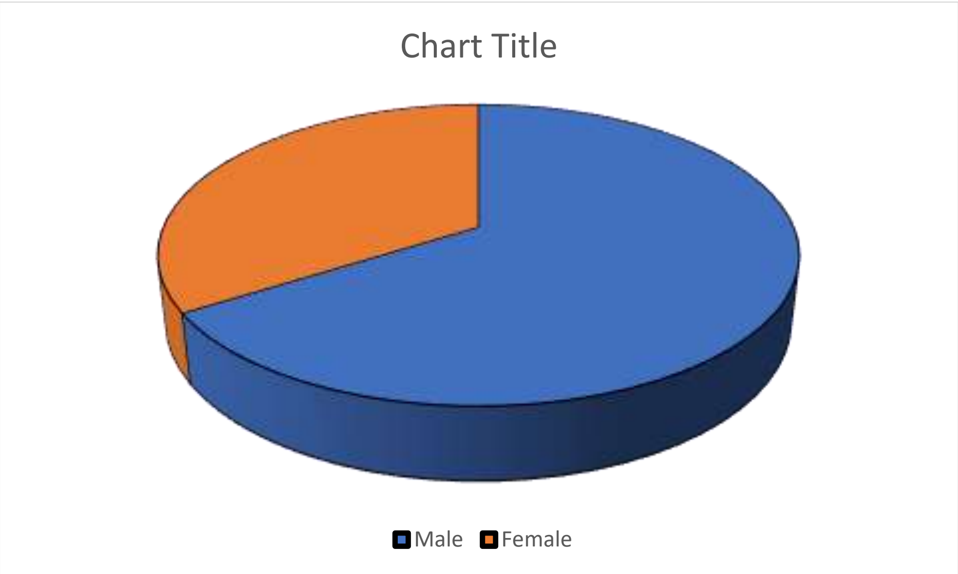
Table no 4.1 showing data of patients according to age group



Graph 4.1 shows the ratio of age group to the number of patients who underwent for MRI scan

Gender	Number of patient
Male	35
Female	18
TOTAL	53

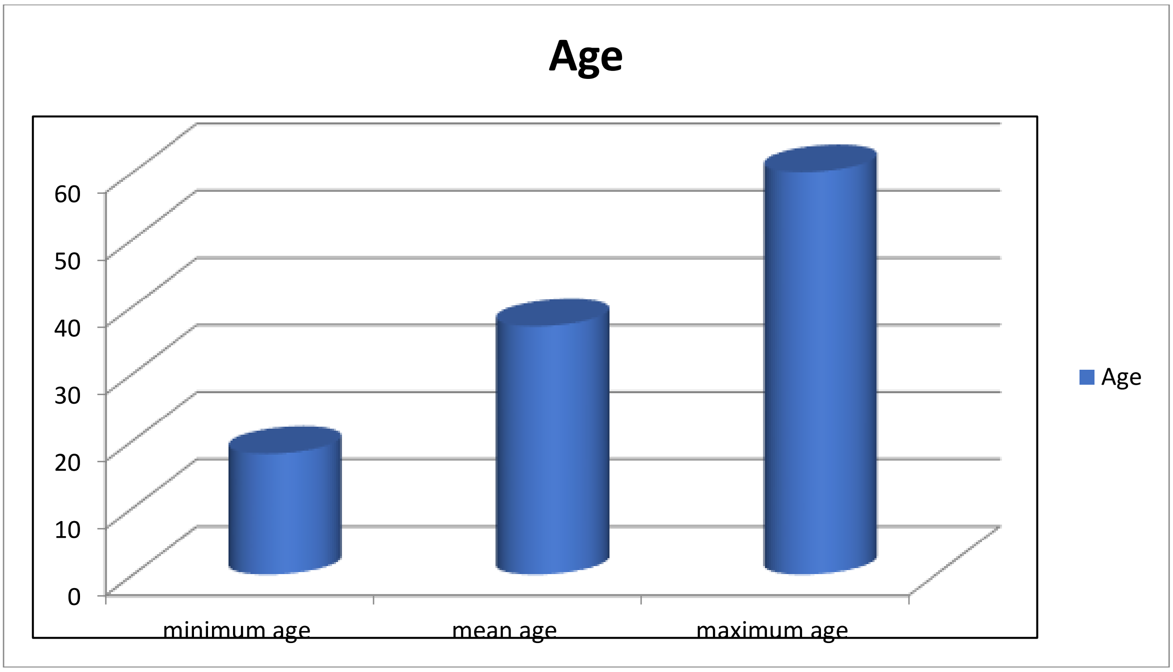
Table 4.2 shows the ratio of gender to the number of patients who underwent for MRI scan.



Age Group	AGE
Minimum age	18
Mean age	39
Maximum age	60

Graph 4.2 shows the ratio of gender to the number of patients who underwent for MRI scan

Table 4.3 shows the ratio of age range to the number of patients who underwent for MRI scan



Graph 4.3 shows the ratio of gender to the number of patients who underwent for MRI scan

Kappa Statics

Reliability is an important part of any research study. The static of kappa coherences assessment is the inter-rater reliability of 2 raters in a particular sample.

Kappa concurrence is a degree of calculation of accuracy and reliability, agreements. The agreement is measurement is measured between 2 raters (judges). Both the 2 raters separately and blindly judge the MRI Image data both pre and post image improvement strategies and rate the in various categories like image contrast, clarity, artifact and overall quality.

$$k = \frac{p_o - p_e}{1 - p_e}$$

Where P_o is the relative observed agreement among raters and P_e is the hypothetical probability of chance agreement. It can be measured in two ways. First is Inter – rater reliability: it is to evaluate the degree of agreement between the choices made by two (or more independent judges). On the other hand, second is Intra rater reliability: it is to evaluate the degree of agreement presented by the same person at a distance of time.

Interpret the kappa statics

Kappa should always less than or equal to 1. It can be negative as well that happens when both observers agreed less than that would be expected by the chance.

The following point that are necessary for kappa calculation for 2 raters are: -

- Both judges agree to include
- Both judges agree to exclude
- Only the first judge wants to include
- Only the second judge wants to include[9]

Also, A cross-sectional survey was conducted among radiologists, radiographers, and radiology technologists working in various healthcare settings. The survey included questions assessing demographic information, professional background, knowledge of image quality optimization applications in radiology, attitudes towards image quality optimization integration, perceived benefits and challenges, and overall awareness of image quality optimization advancements in the field.

RESULT:

Calculating Kappa coherence statistics: -

Both the 2 raters separately and blindly judge the MRI Image data, both pre and post image improvement strategies and rate the image in various categories as follow: -

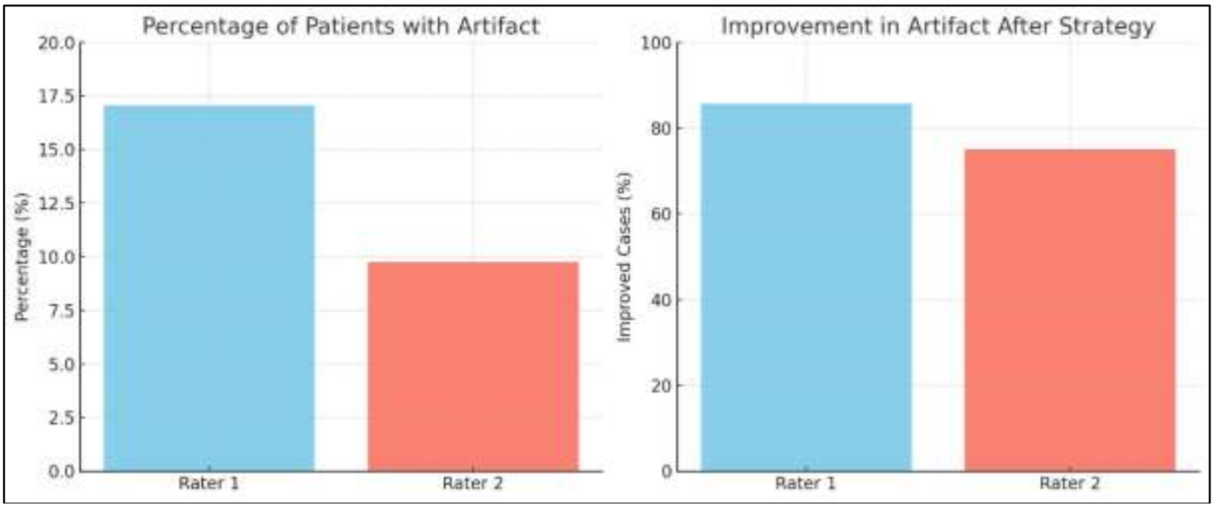
1. ARTIFACT: - MRI artifacts are unwanted distortions or anomalies that appear on MRI images and can degrade image quality, mislead interpretation, or obscure pathology.

In this study both the 2 raters separately and blindly judge the MR Image data for both the pre and the post image improvement strategies and rated the image on the basis of its Artifact for both before and after application of image quality improvement strategies. The rating criteria includes YES and NO

YES: The "after image quality improvement strategies" image is recorded as improved (and denoted as "YES") only if the rater has not rates the image with artifact in compare to the before application of image quality improvement strategies.

NO: The "after image quality improvement strategies" image is recorded as not improved (anddenoted as "NO") only if the rater has rated to the image with artifact in compare to the "before application of image quality improvement strategies" Image.

Out of 41 patients 7 patient's image shows artifact according to rater 1 and out of these 7 patients 6 have reduced or minimal artifact after implementation of image quality improvement strategies and out of 41 patients 4 patient's image shows artifact according to rater 2 and out of these 4 patients 3 have reduced or minimal artifact after implementation of image quality improvement strategies"



- Left chart: Shows how many patients had artifacts, as detected by each rater.
  - Rater 1: ~17%
  - Rater 2: ~10%
- Right chart: Shows how many of those cases improved after quality strategies.
  - Rater 1: ~86% of artifact cases improved
  - Rater 2: 75% of artifact cases improved

This indicates that most artifact cases improved after implementing image quality strategies, with both raters reporting good effectiveness.

2. Overall Quality:-

In this study both the 2 raters separately and blindly judge the MR Image data for both the pre and the post image improvement strategies and rated the image on the basis of its Overall Quality of image for both before and after application of image quality improvement strategies. The rating criteria includes rating scale where 1= Poor; 2= Fair; 3= Good; 4= Very Good and 5=Excellent. Which is later classified in to two categories i.e. "YES" and "NO"

Where, YES: The "after image quality improvement strategies" image is recorded as improved (and denoted as "YES") only if the rater has given more rates to the image in compare to the before application of image quality improvement strategies.

NO: The "after image quality improvement strategies" image is recorded as not improved (and denoted as "NO") only if the rater has given equal or less rates to the image in compare to the "before application of image quality improvement strategies" Image.

Calculating Kappa coherence statistics for MR image Overall Image Quality.

1. "Rater 1" finds that 37 out of 41 patients image data have YES or improved image quality.
2. "Raters 1" finds that 04 out of 41 patients image data have NO or same or not improved image quality.
3. "Rater 2" finds that 38 out of 41 patients image data have YES or improved image quality.
4. "Raters 2" finds that 03 out of 41 patients image data have NO or same or not improved image quality.
5. Both the radiologist (Rater 1 and Rater 2) agreed that 36 out of the 41 patients image data have YES or improved image quality
6. (leaving 01 patient where the doctors disagreed from each other in a peaceful manner).
7. Both the radiologist (Rater 1 and Rater 2) agreed that 02 out of the 41 patients image data have NO or same or not improved image quality.

8. (leaving 02 patients where the doctors disagreed from each other in a peaceful manner).
The Kappa statistic is calculated using the following formula:

Observed agreement - chance agreement 1-chance agreement

1. First step: -
filling 2 X 2 table as follows:

	R1			
R2		Yes	No	total
	Yes	36	02	38
	No	01	02	03
	total	37	04	41

The observed agreement is: (X + Y) / N
Where, "X" = both the raters (radiologist) agreed to include the patients as a positive find.
And, "Y" = both the raters (radiologist) disagreed to include the patients as a positive find or agreed to exclude the patient as negative finding.
N = total no of observation (Patients)
= (36 + 02) / 41

The observed agreement is = 0.92
The observed agreement percentage is: [(a + d) / N] x 100
= 0.78 x 100 = 92.68%
2. Second step: -
To calculate the chance agreement: -
note that "R1" found 37/41 patients to have improved image quality and 04/41 to not have improved image quality
And "R2" found 38/41 patients to have improved image quality and 03/41 to not have improved image quality.
formula for "chance of agreement": - Pe=[(a+b)/N×(a+c)N] + [(c+d)/N×(b+d)/N]
Where,

	R1			
R2		Yes	No	total
	Yes	a	c	a+c
	No	b	d	b+d
	total	a+b	c+d	N

- First term = expected Yes agreement
- Second term = expected No agreement

i.e Pe = [(a+b)/N×(a+c)N] + [(c+d)/N×(b+d)/N]
Pe = [37/41×38/41] + [04/41×03/41]
Pe= 0.84
3. Third step: - To find the value of Cohen's Kappa and to calculate the formula is as follow:

Observed agreement Po - chance agreement Pe 1-chance agreement Pe

i.e. The observed agreement is Po = 0.92
and the chance of agreement is Pe = 0.84

Hence,

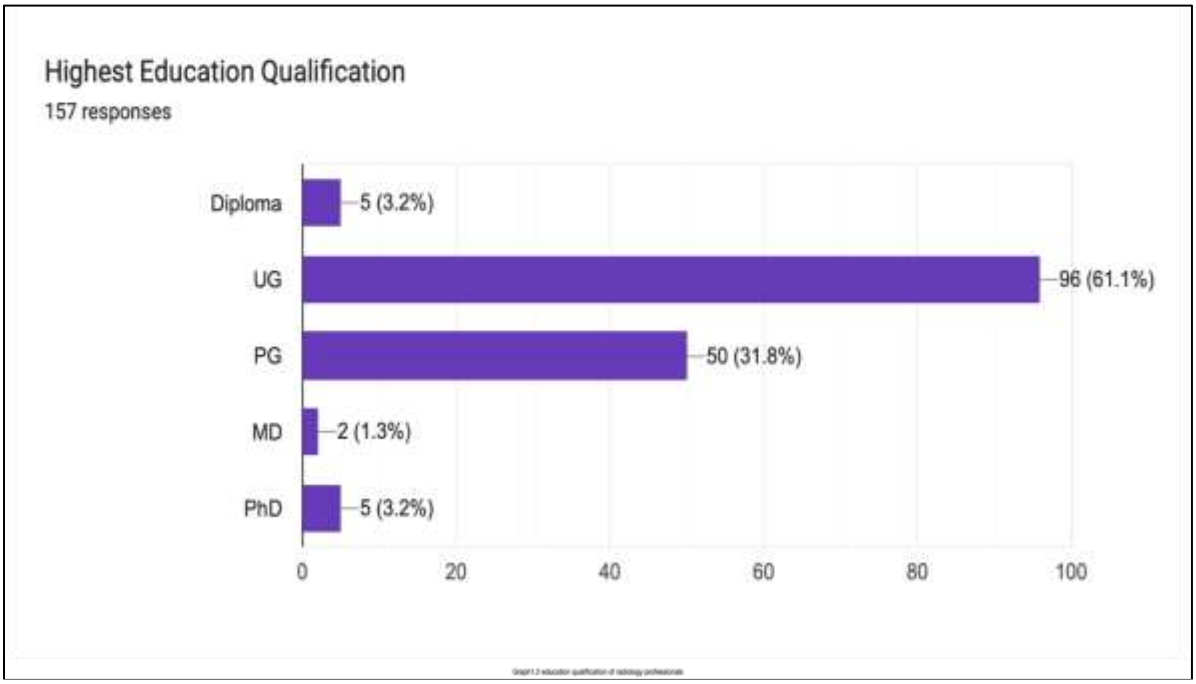
$$\text{Kappa} = \frac{0.92 - 0.84}{1-0.84}$$

$$\text{Kappa} = 0.53$$

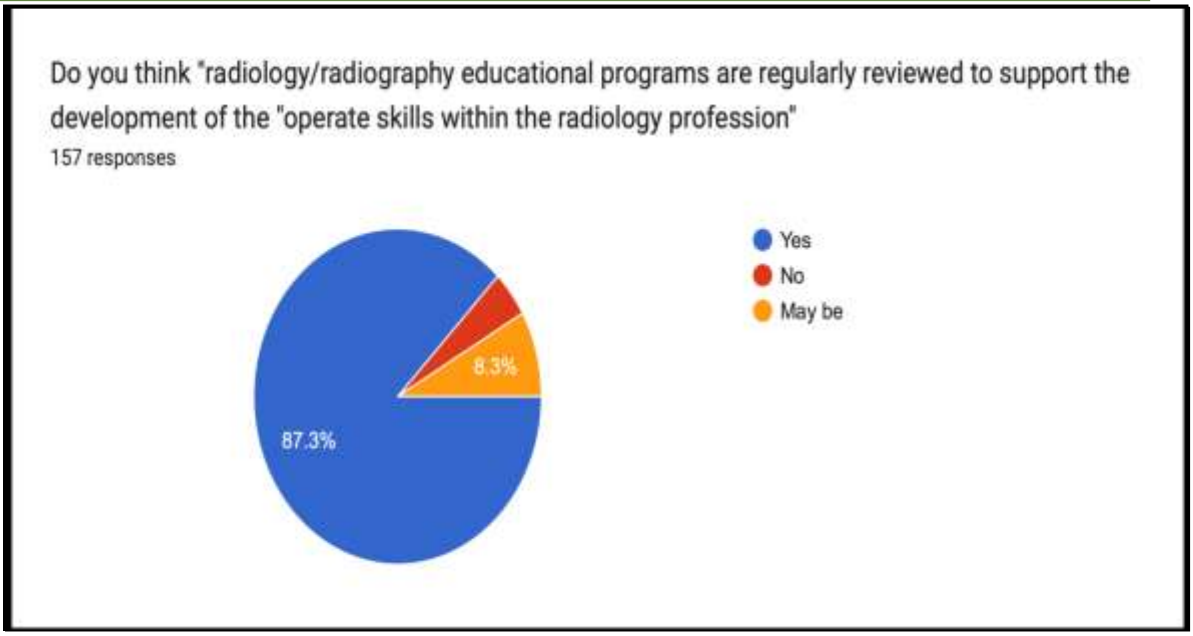
Kappa= 0.53
95% confidence interval: From 0.068 to 0.997
A kappa value of 0.53 indicates moderate agreement between observers.

As, the kappa test analyses value can be classified as: -
- 0.01 – 0.20 slight agreement
- 0.21 – 0.40 good agreement
- 0.41 – 0.60 moderate agreement
- 0.61 – 0.80 substantial agreement
- 0.81 – 1.00 almost perfect or perfect agreement

kappa is always less than or equal to 1. A value of 1 implies perfect agreement and values less than 1 imply less than perfect agreement.
It’s possible that kappa is negative. This means that the two observers agreed less than would be expected just by chance.
The result for the above study shows a good kappa value, which is as follow:
The artifact shows helps in improving image quality as in the total 41 case data 17% data shoes artifact according to R1 and 10% according to R2.however, many of these cases improved after quality strategies shows 86% of artifact cases improved according to R1 75% of artifact cases improved as per R2.
Also**,** A cross-sectional survey was conducted among radiologists, radiographers, and radiology technologists working in various healthcare settings. The survey included questions assessing demographic information, professional background, knowledge of image quality optimization applications in radiology, attitudes towards image quality optimization integration, perceived benefits and challenges, and overall awareness of image quality optimization advancements in the field.
The preliminary analysis indicated that while a majority of radiology professionals recognize the potential benefits of image quality optimization in enhancing diagnostic accuracy, there is a significant variation in the level of knowledge and understanding of image quality optimization applications. The perception of image quality optimization was generally positive, with respondents acknowledging its role in improving patient outcomes radiological procedures. However, there was a notable demand for more comprehensive educational programs to bridge the knowledge gap.
The survey results from 157 radiology professionals regarding their knowledge, attitude, perception, and awareness of artificial intelligence (AI) in medical radiology and imaging technology are presented below. Various radiology professionals have taken part in the online survey from various institutions, healthcare organisations as shown in the graph below:



A great no percentage of professionals are be interested in taking a course based on Al training within the Radiology/Radio-technology sector also many thinks think "radiology/radiography educational programs are regularly reviewed to support the development of the "operate skills within the radiology profession". As shown in the pie chart below:



The attitude of radiology professionals is very positively concerned and elevated towards introduction of AI in radiology and healthcare sector. Many are concerned about the might be some ethical issues associated with the use of Al in Radiology but they are positive about the Al as is might going to improve the routine workload of Radiology professionals.

DISCUSSION:

The present study evaluated the effectiveness of image quality optimization strategies in reducing artifacts in 1.5T brain MRI and assessed the reliability of rater agreement using Cohen's kappa statistic. Artifacts were initially observed in 17% of cases by R1 and in 10% of cases by R2, highlighting a modest level of variability between raters in the baseline assessment of artifact prevalence. Following optimization strategies, the majority of artifact-affected cases showed improvement, with 86% improvement reported by R1 and 75% by R2. These results indicate that the implemented interventions were effective in mitigating artifacts and improving image quality.

The inter-rater agreement, reflected by a kappa value of 0.49, suggests a moderate level of agreement between raters. According to established interpretative guidelines, κ values between 0.41 and 0.60 are considered to indicate moderate agreement, values between 0.61 and 0.80 substantial agreement, and values above 0.81 almost perfect agreement.[6] This moderate agreement highlights that, although the raters independently recognized improvements in image quality, some subjectivity persists in artifact assessment

CONCLUSION:

This study demonstrated that image quality optimization strategies were effective in reducing artifacts in 1.5T brain MRI, with significant improvements observed by both raters. The inter-rater agreement, reflected by a kappa value of 0.49, indicated moderate reliability in artifact assessment. These findings highlight that while optimization techniques enhance image quality, subjective variability in evaluation persists. Incorporating objective image quality measures alongside structured rater assessments may further strengthen the reliability and reproducibility of future studies.

REFERENCES

1. Kanal E, Barkovich AJ, Bell C, Borgstede JP, Bradley WG Jr, Froelich JW, et al. ACR guidance document on MR safe practices: 2013. *J MagnReson Imaging*. 2013;37(3):501–30.
2. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters. *J MagnReson Imaging*. 2007;26(2):375–85.
3. Kidoh M, Shinoda K, Kitajima M, Isogawa K, Morita K, Uetani H, et al. Evaluation of image quality: subjective and objective approaches. *Radiographics*. 2020;40(6):1682–96.
4. Wood ML, Henkelman RM. MR image artifacts from periodic motion. *Med Phys*. 1985;12(2):143–51.
5. Dietrich O, Reiser MF, Schoenberg SO. Artifacts in MRI. *J MagnReson Imaging*. 2008;27(3):671–81.
6. Kidoh M, Shinoda K, Kitajima M, Isogawa K, Morita K, Uetani H, et al. Evaluation of image quality: subjective and objective approaches. *Radiographics*. 2020;40(6):1682–96.
7. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005;85(3):257-68.