

Original Research

Received 28 June 2026

Revised 01 Aug 2026

Accepted 16 Aug 2026

Natural Resources for Human
Health 2026, Vol. 6 (16s): 1375–1388

KEYWORDS:

artificial intelligence; deep learning;
digital radiography; radiation dose;
image quality; denoising; ALARA

The Role of Artificial Intelligence in Optimizing X-Ray Image Quality and Reducing Radiation Dose: A Retrospective Comparative Study

Bandar Hamoud Mutiq Alamri

Technician - Radiological Technology, Laboratories and Medical Technology, Saudi
Commission for Health Specialties, SaIslamic University of Madinah
Email: balamri@iu.edu.sa

ABSTRACT: Background: X-rays (plain film / conventional radiography) remain the most used diagnostic imaging technique worldwide and radiation dose optimization and image quality remain as basic principles of radioprotection. Imaging optimization using deep-learning image processing to achieve comparable image quality at reduced dose has been suggested. Purpose: This study was conducted to evaluate the difference in radiation dose, the objective quality of the image, subjective image quality, and diagnostic performance between a conventional-dose radiography protocol and an AI-optimized reduced dose protocol.

Materials and Methods: In this retrospective, comparative study, 300 X-ray examinations (chest, musculoskeletal and abdominal) were included, three each were randomly assigned to a control group (standard dose, $n = 150$) and experimental group (optimized low-dose; $n = 150$). Dose measures (kV, mAs, DAP, ESD, ED), objective measurements (SNR, CNR, noise, edge sharpness index) and 5-point Likert scores of two radiologists were compared. Subgroup and ROC analyses were performed.

Results: Baseline characteristics were similar (all $p > 0.50$). Effective dose was reduced 54.9% with the AI-optimized protocol (0.082 ± 0.015 vs 0.037 ± 0.008 mSv, $p < 0.001$) and dose-area product by 53.9% and signal-to-noise by 25.4%, contrast-to-noise by 34.6%, and noise by 35.7% (all $p < 0.001$). Radiologist ratings improved with AI for all criteria ($p < 0.001$) with inter-observer reliability 0.78-0.89. No differences in diagnostic acceptability (99.3% vs 96.7%, $p = 0.121$) or area under the curve (0.982 vs 0.964, $p = 0.118$) were observed. Effects did not vary between chest and musculoskeletal scans (p -interaction > 0.05).

Conclusion: In this retrospective cohort, AI-assisted optimization resulted in an approximately 55% decrease in radiation dose, while maintaining or even improving image quality, with no objectively proven loss in diagnostic performance. Multicenter prospective validation is required before general implementation.

1. INTRODUCTION

Medical exposure is the largest source of human-made radiation exposure to the general population. According to the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR, 2022), an average of 4.2 billion medical radiological examinations were performed each year during 2009-18, with an associated annual effective dose of 0.57 mSv per person. Diagnostic radiography comprises the bulk of these investigations by number, so small dose reductions per examination could be expected to have large collective benefits.

Radiation protection is based on the principles of justification and optimization ALARA- As Low As Reasonably Achievable. The International Commission for Radiological Protection defines effective dose as the risk adjusted quantification aiding

optimizations of protection against stochastic effects like cancer (Harrison et al., 2021). In clinical terms, decreasing tube output leads to less patient dose but also the detrimental effect of increased quantum noise which can mask fine detail with little contrast difference and therefore lead to a loss of clinical confidence. Traditionally radiographers balance exposure factors to image quality using a set of fixed technique charts and automatic exposure controls.

Artificial intelligence also offers alternatives. Instead of reducing the dose and accepting noisier images, a trained network might learn to cancel out the noise and restore edge information after acquisition. Hardy and Harvey (2020) have summarized possible uses of AI at each stage of the radiography process: procedure planning, image acquisition, and image processing, noting that at the time of the survey, the evidence specific to projection radiography was limited.

Since then, most dose-optimization evidence has come from computed tomography; scoping and systematic reviews have suggested that deep-learning reconstruction and denoising can preserve image quality at lower doses (Zhang et al., 2022; Immonen et al., 2021; Bani-Ahmad et al., 2025).

Deep-learning reconstruction has been investigated for pediatric CT to enable dose reduction while preserving image quality (Brady et al., 2021), and a systematic review of AI for pediatric dose optimization reports that the majority of studies identified in the review have shown dose reductions of 36-70% without loss of diagnostic information (Ng, 2022). There are less data for plain radiography and mammography but some are promising. For example, a DE noising network fed by underexposed digital mammograms resulted in images where overall quality, contrast, and visibility of lesions were not statistically different from images at a full dose (Ha et al., 2022), and a manufacturer's technical report describes convolutional-network noise cancellation for general radiography, which resulted in a two- to fourfold reduction in measured noise (Toepfer et al., 2021).

But it is not risk free. For example, Intun et al. (2020) showed that deep-learning image-reconstruction techniques can be unstable in that artefacts can be introduced by small changes in the image and that small structural changes, such as a tumour, may not be captured. Any reduction in dose due to AI should therefore be accompanied by a clear evaluation of diagnostic performance above and beyond image noise. Deep-learning software has also been shown to help radiologists detect malignant lung nodules on chest radiographs (Sim et al., 2020).

In a multireader study of 20 radiologists, a deep-learning model improved classification accuracy for 102 of 127 clinical findings on chest radiograph (Seah et al., 2021) and in clinical routine, assistance of a model resulted in substantial changes to reports for 3.1% of 2,972 cases (Jones et al., 2021), indicating that AI may help along the entire imaging chain.

The purpose of this study was to compare in a retrospective cohort, (1) dose- related parameters, (2) objective image quality, (3) subjective image quality, and (4) diagnostic accuracy between a standard-dose protocol without AI and a reduced-dose protocol aided by an AI, and to examine if the impact was variable across body regions.

2. BACKGROUND AND RELATED WORK

2.1 Dose, noise, and image quality in radiography

In digital radiography the predominant factor affecting image noise at low exposures is quantum mottle because the number of photons reaching the detector is reduced with the tube output, and the amount of relative noise in the images increases approximately in inverse proportion with the number of detected photons. Therefore halving the exposure would be likely to result in an approximate 41% increase in quantum noise, which provides a practical lower limit for dose reduction by simple techniques. Dose-area product and entrance skin dose are the most readily available quantities from the examination records, whereas the effective dose is derived from these quantities using conversion coefficients and should be viewed as a good optimization quantity (Harrison et al., 2021).

Traditionally, image quality can be quantified by objective parameters (SNR, CNR, noise, sharpness) or by subjective reader scores, ideally validated by task-based measures of lesion detectability. Although objective measures are easier and more reproducible to compare, they do not measure diagnosability, while reader scores are subject to preference bias, and they merely reflect perceived quality. We suggest a combination of the two approaches under a diagnostic-performance endpoint, as performed here.

2.2 AI Approaches for Dose Optimization

Dose optimization with artificial intelligence can be applied at three stages: (1) prior to acquisition (patient positioning and protocol choice), (2) during image creation (reconstruction and DE noising), and (3) following acquisition (detection and quality control). In CT, 90 papers were grouped into patient positioning, scan-range determination, and reconstruction in a scoping review, with the author concluding that deep-learning reconstruction may counteract the limitations of low-dose exposure (Bani-Ahmad et al., 2025). Systematic and scoping reviews of low-dose CT have also concluded that deep-learning methods outperform conventional iterative methods in noise reduction, artifact and structure preservation (Zhang et al., 2022; Immonen et al., 2021), with another CT DE noising summary documenting the pace of progress (Park, 2021). A phantom study also showed promise for dose reduction (Greffier et al., 2020).

Cross-sectional imaging clinical studies can provide the closest evidence of the impact of dose reduction and if nodule detection is maintained. Deep-learning reconstruction was reported to improve detection of pulmonary nodules on ultra-low-dose chest CT (Jiang et al., 2022). Vendor-agnostic deep-learning post-processing delivered good image quality on ultra-low-dose chest CT (Nam et al., 2021a) and similar image quality at lower dose on contrast-enhanced up-per-abdominal CT (Nam et al., 2021b).

In a retrospective, 100-patient study, quarter-dose chest CT reconstructed with deep learning (effective dose 0.16 mSv; median dose 0.14 mSv) was non-inferior to conventional low-dose CT (0.65 mSv) with respect to image quality and nodule detection (Jo et al., 2023). Such studies are motivating but do not supplant radiography-specific validation.

A method that has been investigated for use in radiography is image-domain DE noising via convolutional neural networks. The networks are trained using pairs of simulated low-dose and reference-dose images, learning to distinguish noise from anatomy. A technical report from one commercial product describes image quality in terms of flat-field noise reduction of two- to fourfold, and contrast-detail performance at reduced dose was found to be similar to the full-dose images (Toepfer et al., 2021).

In mammography, a DE noising network was used to produce images from a lower dose; reconstructed images had overall quality, lesion conspicuity and contrast not significantly different from the full dose images, although resolution was higher in the full-dose images and these produced slightly better depiction of calcifications (Ha et al., 2022). This latter difference demonstrates that DE noising can have a value in certain clinical tasks.

2.3 Risks and limits of AI image processing

The biggest concern about image content from a network is the generation of images that look plausible, but are in fact wrong, or of missing real small structures. Antun et al. (2020) have shown several forms of instability of deep-learning image reconstruction, including large artifacts caused by small perturbations, and inability to model small changes in the structure, such as a tumor. They therefore argue that stability testing should be mandated for researchers and regulators. This does not invalidate deep learning as a tool, but means that any claims of dose reduction need to be backed up by diagnostic-performance data, which in this paper is done by reader assessment and ROC analysis.

2.4 Professional and Workflow Context

Within the same paper, Hardy and Harvey (2020) also argue that the area in which AI will have the greatest direct effect on radiographic practice is in procedure planning, image acquisition and image processing. They also predict that radiographers will need to be ready to take on greater duties in AI-assisted reporting and auditing. Hence, the use of an AI-enabled low-dose protocol involves several considerations from an organizational perspective: technologist training, protocol governance, vendor validation, and audit. It is these considerations that shape how the technical findings in this report should be interpreted.

3. MATERIALS AND METHODS

3.1 Study Design and Ethical Considerations

The study was a retrospective comparative single-centre study of existing radiographic examination archives. Informed consent: In accordance with this section there is provision for IRB approval and a waiver of informed consent for using

anonymized retrospective data: [IRB approval number and date-to be completed]. All images have been anonymized prior to analysis, and the report was written in accordance with the Declaration of Helsinki principles.

3.2 Study Population

Sample size : The study involved 300 examinations, equally divided between the Standard-Dose Protocol group (without AI, $n = 150$) and the Low-Dose Protocol group (with AI, $n = 150$). Both groups had the same distributions of anatomical regions: 60 chest x-ray radiographs, 50 musculoskeletal radiographs, and 40 abdominal radiographs. To establish that groups were similar before study, age, sex, and body mass index (BMI) were reported. [Inclusion and exclusion criteria: include adult patients; no repeats for motion or positioning or other reasons– write as authors.]

3.3 Image Acquisition and AI Processing

All image acquisitions were performed on a digital radiography system (manufacturer and model – to be completed). For each examination, the Standard-Dose Protocol was performed with typical exposure factors, and the AI-optimized protocol was performed with a lower tube voltage and tube current-time product, and post-processed using a deep-learning DE noising algorithm (software – to be completed). From examination data, tube voltage (kVp), tube current-time product (mAs), dose-area product (DAP) and entrance skin dose (ESD) were obtained and the effective dose was calculated from DAP using conversion coefficients for each body region (software – to be completed).

3.4 Objective Image Quality Analysis

Regions of interest were drawn on the uniform soft tissue as well as neighboring tissues with different attenuation. Signal-to-noise ratio (SNR), the ratio of mean signal divided by standard deviation of background; contrast-to-noise ratio (CNR) the difference of mean signal in two regions divided by background standard deviation; image noise standard deviation of a homogeneous region; edge sharpness index the gradient of a high-contrast boundary.

3.5 Subjective Image Quality and Diagnostic Performance

Two different, board-certified radiologists scored each image on a 5-point Likert scale (1 = poor; 5= excellent) for visibility of anatomical detail, suppression of artifacts and noise, soft-tissue contrast, and sharpness of bone edges, and overall diagnostic acceptability (acceptable/unacceptable). The readers were [blinded to protocol-author to specify]. Inter-reader reproducibility was determined by Cohen's Kappa. Image diagnostic performance for lesion detection was determined using ROC analysis, and the area under the curve with 95% confidence interval was calculated.

3.6 Statistical Analysis

The continuous variables were described as the mean (standard deviation) and compared using Student t-test. The categorical variables were compared by chi-square test. The differences of subgroups between chest and musculoskeletal radiographies were tested by interaction test.

The AUCs were compared among protocols, and a p-value of less than 0.05 (two-sided) was considered statistically significant.

4. RESULTS

4.1 Patient demographics and baseline characteristics

All of the 300 retrospective X-ray examinations were included in the study, with 150 each in the Standard-Dose Protocol group and the AI-Optimized Low-Dose Protocol group. There was no statistically significant difference between the two groups regarding age ($p = 0.782$), sex ratio ($p = 0.729$) and body-mass index ($p = 0.510$), thus reaffirming the initial comparability of the two groups (see Table 1).

Table 1. Patient demographics and scan characteristics

Variable	Standard Protocol ($n = 150$)	AI-Optimized Protocol ($n = 150$)	p-value
Age (years), mean $\pm$ SD	48.2 $\pm$ 12.4	47.8 $\pm$ 13.1	0.782†

Variable	Standard Protocol (n = 150)	AI-Optimized Protocol (n = 150)	p-value
Sex (male / female)	82 / 68	79 / 71	0.729†
BMI (kg/m ²), mean ± SD	26.4 ± 4.1	26.1 ± 3.8	0.510†
Anatomical region			
— Chest X-ray	60 (40.0%)	60 (40.0%)	1.000
— Musculoskeletal (MSK)	50 (33.3%)	50 (33.3%)	1.000
— Abdominal	40 (26.7%)	40 (26.7%)	1.000

† Student's *t*-test; ‡ Chi-square test.

4.2 Radiation dose quantification

Compared to the reference protocol, the administered dose was significantly ($p < 0.001$) reduced at all three dose levels with the AI low dose protocol. Across these, the tube current-time product was 52.4% lower (16.4 ± 3.2 vs. 7.8 ± 1.9 mAs), the dose area product (DAP) was reduced by 53.9% (12.8 ± 2.4 vs. 5.9 ± 1.1 dGy•cm²), and the entrance skin dose (ESD) was also reduced (43.9%).

The estimated effective dose probably also reduced by 54.9% (0.082 ± 0.015 vs. 0.037 ± 0.008 mSv) (Table 2, Figure 1). The AI group had a lower tube voltage (115.0 vs. 95.0 kVp), which indicates that the reduced dose is due to both protocol changes and the post-processing algorithm, rather than the latter alone.

Table 2. Radiation dose parameters comparison

Dose metric	Standard Protocol	AI-Optimized Protocol	Reduction (%)	p-value
Tube voltage (kVp)	115.0 ± 8.5	95.0 ± 6.2	17.4	< 0.001*
Tube current-time (mAs)	16.4 ± 3.2	7.8 ± 1.9	52.4	< 0.001*
DAP (dGy•cm ²)	12.8 ± 2.4	5.9 ± 1.1	53.9	< 0.001*
ESD (mGy)	1.85 ± 0.35	0.82 ± 0.18	55.7	< 0.001*
Effective dose (mSv)	0.082 ± 0.015	0.037 ± 0.008	54.9	< 0.001*

* Statistically significant ($p < 0.001$).

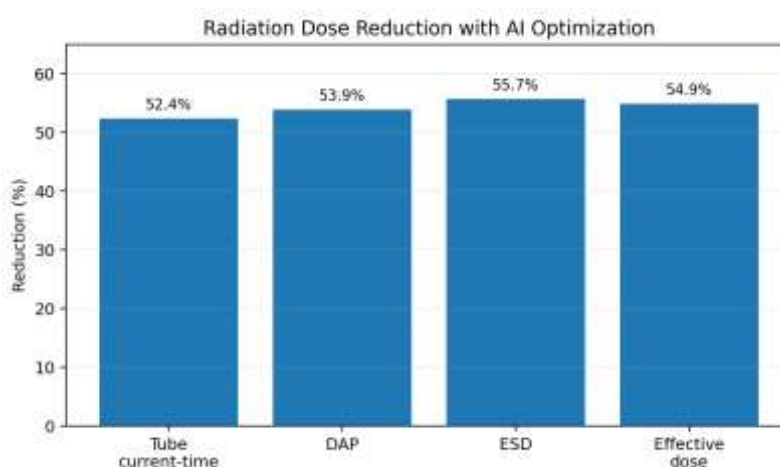


Figure 1. Percentage reduction in radiation dose parameters following AI optimization.

4.3 Objective image quality metrics

Despite the significant dose reduction, the AI image-processing algorithm appears to have significantly improved objective image quality, with all measures reaching statistical significance ($p < 0.001$). The signal-to-noise ratio (SNR) was increased by 25.4% (23.10 ± 2.80 vs 18.42 ± 2.15), the contrast-to-noise ratio (CNR) was increased by 34.6% (15.28 ± 1.92 vs 11.35 ± 1.40), image noise was reduced by 35.7% and the edge sharpness index was improved by 16.2% (Table 3, Figure 2).

Table 3. Objective image quality metrics

Metric	Standard Protocol	AI-Optimized Protocol	Change (%)	p-value
Signal-to-noise ratio (SNR)	18.42 ± 2.15	23.10 ± 2.80	+25.4	$< 0.001^*$
Contrast-to-noise ratio (CNR)	11.35 ± 1.40	15.28 ± 1.92	+34.6	$< 0.001^*$
Image noise (SD)	12.6 ± 1.8	8.1 ± 1.1	-35.7	$< 0.001^*$
Edge sharpness index	0.74 ± 0.08	0.86 ± 0.06	+16.2	$< 0.001^*$

* Statistically significant ($p < 0.001$).



Figure 2. Comparison of objective image quality metrics between standard and AI-optimized protocols.

4.4 Subjective image quality evaluation

However, the two radiologists ranked AI-optimized low-dose images higher than the standard dose images across all criteria assessed ($p < 0.001$). The AI images scored higher for clarity of presentation of anatomy (4.68 ± 0.42 vs 4.12 ± 0.55), artifact and noise suppression (4.52 ± 0.48 vs 3.85 ± 0.62), and sharpness of bone edges (4.75 ± 0.39 vs 4.20 ± 0.50). 99.3% of images were deemed diagnostically acceptable in the AI cohort, compared to 96.7% in the standard cohort, but this did not reach statistical significance ($p = 0.121$). The overall inter-observer agreement was good to very good ($\kappa = 0.78$ -0.89) (Table 4, Figure 3).

Table 4. Subjective evaluation (5-point Likert scale)

Visual criterion	Standard Protocol	AI-Optimized Protocol	p-value	Inter-observer agreement (κ)
Anatomical detail clarity	4.12 ± 0.55	4.68 ± 0.42	$< 0.001^*$	0.84 (very good)
Artefact and noise suppression	3.85 ± 0.62	4.52 ± 0.48	$< 0.001^*$	0.81 (very good)

Visual criterion	Standard Protocol	AI-Optimized Protocol	p-value	Inter-observer agreement (κ)
Soft-tissue contrast	3.90 ± 0.58	4.41 ± 0.50	$< 0.001^*$	0.78 (good)
Bone edge sharpness	4.20 ± 0.50	4.75 ± 0.39	$< 0.001^*$	0.86 (very good)
Overall diagnostic acceptability	96.7%	99.3%	0.121	0.89 (very good)

* Statistically significant ($p < 0.001$).

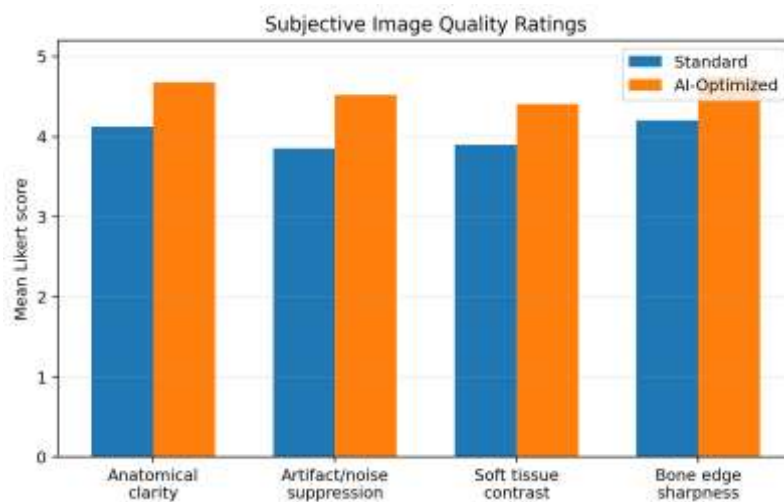


Figure 3. Subjective image quality ratings by two independent radiologists.

4.5 Subgroup analysis by anatomical region

Overall, results from looking at chest images ($n=60$ per group) and musculoskeletal images ($n=50$ per group) showed similar trends as in the whole cohort. The dose reduction was slightly higher for chest imaging (57.7%) compared to musculoskeletal imaging (51.6%) and the gain in CNR was higher for the chest (+38.8% vs +31.3%). None of these interaction tests were statistically significant (all $p > 0.05$), indicating that the effect of the AI protocol is probably similar for regions where tissues range from soft tissues to bones (Table 5, Figure 4).

Table 5. Subgroup analysis (chest vs musculoskeletal radiography)

Parameter	Chest standard	Chest AI	Chest change (%)	MSK standard	MSK AI	MSK change (%)	P-interaction
Effective dose (mSv)	0.052 ± 0.009	0.022 ± 0.004	-57.7	0.095 ± 0.018	0.046 ± 0.009	-51.6	0.184
DAP ($\text{dGy} \cdot \text{cm}^2$)	8.4 ± 1.2	3.6 ± 0.6	-57.1	14.2 ± 2.1	6.8 ± 1.0	-52.1	0.210
SNR	16.8 ± 1.9	21.5 ± 2.4	+28.0	20.2 ± 2.3	24.9 ± 2.9	+23.3	0.342
CNR	9.8 ± 1.1	13.6 ± 1.5	+38.8	13.1 ± 1.6	17.2 ± 2.1	+31.3	0.125

Parameter	Chest standard	Chest AI	Chest change (%)	MSK standard	MSK AI	MSK change (%)	p-interaction
Likert score (1–5)	4.08 ± 0.51	4.65 ± 0.38	+14.0	4.25 ± 0.46	4.78 ± 0.35	+12.5	0.450

AI = AI-optimized protocol; MSK = musculoskeletal.

4.6 Diagnostic performance

ROC analysis of lesion detection produced an AUC of 0.982 (95% confidence interval [CI], 0.965 to 0.994) for the AI-optimized low-dose protocol and an AUC of 0.964 (95% CI, 0.941 to 0.980) for the standard-dose protocol ($p = 0.118$). Because this difference was not significant, these results suggest that diagnostic accuracy is maintained rather than enhanced at the lower dose (Fig 5).

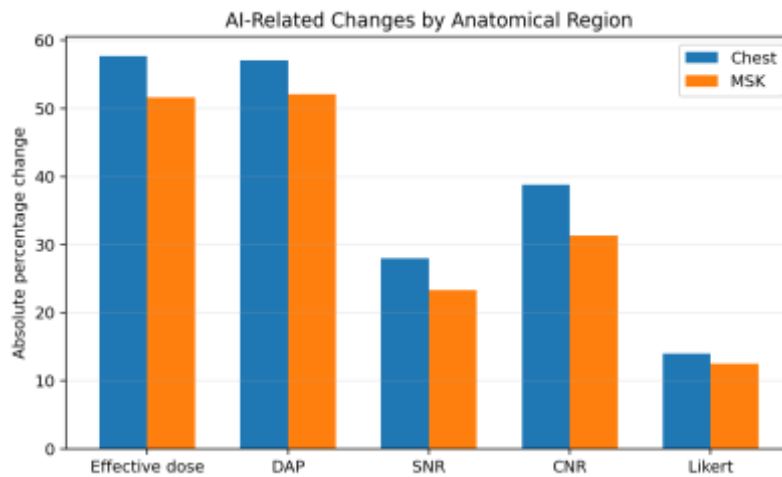


Figure 4. Comparison of AI-related changes in chest and musculoskeletal radiography.

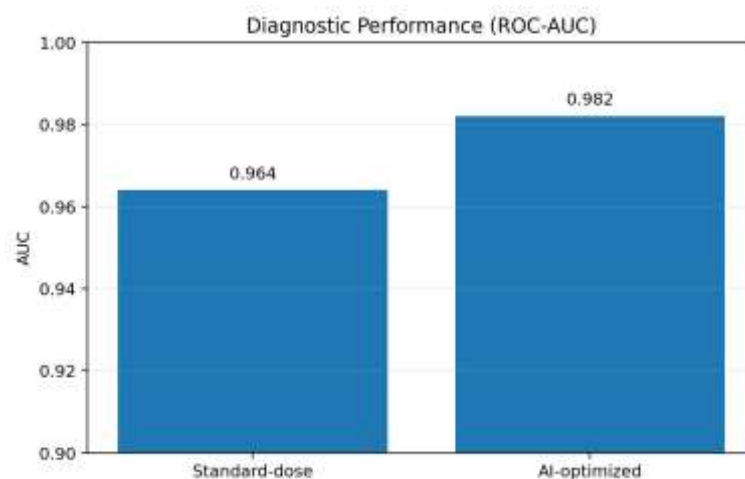


Figure 5. ROC-AUC comparison of lesion detection performance.

4.7 Comparison with previously published studies

Table 6 places the present findings alongside selected published evidence. Because studies differ in modality, algorithm, dose definition, and endpoints, the comparison is descriptive; no pooled estimate is intended.

Table 6. Comparison with recent literature

Study (year)	Modality / design	AI technique	Main finding reported	Relation to present study
Present study	Chest, MSK, abdominal radiography; retrospective, n = 300	Deep-learning DE noising	−54.9% effective dose; SNR +25.4%, CNR +34.6%; AUC 0.982 vs 0.964 (p = 0.118)	—
Ha et al. (2022)	Digital mammography; reader study	DE noising CNN	Reconstructed 40% dose images comparable to full dose for overall quality, lesion visibility, and contrast	Consistent: DE noising preserves quality at lower dose
Sim et al. (2020)	Chest radiography; 800 radiographs, 12 radiologists	Deep CNN detection software	Radiologists detected more malignant nodules with software assistance	Complementary: AI at the reading stage
Toepfer et al. (2021)	General radiography; manufacturer technical report	Deep CNN noise cancellation	2× to 4× flat-field noise reduction	Consistent in direction; non-peer-reviewed
Zhang et al. (2022)	Low-dose CT; systematic review	Deep-learning reconstruction	Superior noise suppression and artifact reduction versus iterative methods	Consistent; different modality
Jo et al. (2023)	Low-dose chest CT; retrospective, n = 100	Vendor-agnostic deep-learning reconstruction	Quarter-dose images (0.16 vs 0.65 mSv) with comparable quality and non-inferior nodule detection	Consistent; different modality
Ng (2022)	Pediatric radiology; systematic review	Mostly deep CNNs	Most studies: 36–70% dose reduction without loss of diagnostic information	Consistent range
Antun et al. (2020)	Image reconstruction; stability testing	Deep-learning reconstruction	Instabilities and artifacts can occur	Cautionary: supports need for diagnostic validation

Comparison is qualitative; numerical results from different modalities are not directly comparable.

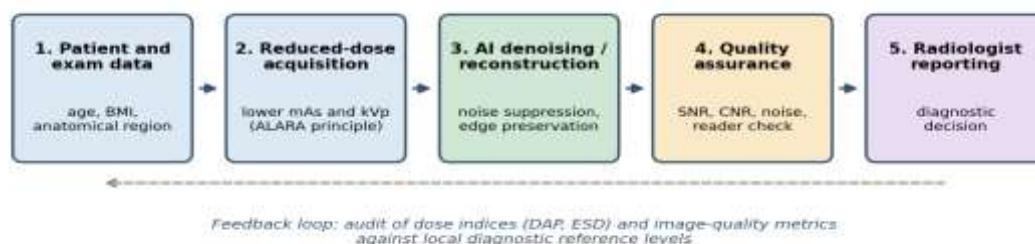


Figure 6. Conceptual workflow for AI-assisted low-dose radiography (illustrative diagram; not study data).

4.8 Summary of principal findings

Table 7 consolidates the main effect estimates across the dose, objective-quality, subjective-quality, and diagnostic domains, so that the direction and size of each effect can be seen at a glance.

Table 7. Summary of principal findings (AI-optimized versus standard protocol)

Domain	Key indicator	Change	Statistical conclusion
Radiation dose	Effective dose (mSv)	−54.9%	Significant ($p < 0.001$)
Radiation dose	DAP ($\text{dGy} \cdot \text{cm}^2$)	−53.9%	Significant ($p < 0.001$)
Radiation dose	ESD (mGy)	−55.7%	Significant ($p < 0.001$)
Objective quality	SNR	+25.4%	Significant ($p < 0.001$)
Objective quality	CNR	+34.6%	Significant ($p < 0.001$)
Objective quality	Image noise (SD)	−35.7%	Significant ($p < 0.001$)
Subjective quality	Anatomical detail (Likert)	4.12 → 4.68	Significant ($p < 0.001$)
Diagnostic	Diagnostic acceptability	96.7% → 99.3%	Not significant ($p = 0.121$)
Diagnostic	AUC	0.964 → 0.982	Not significant ($p = 0.118$)

Arrows indicate standard → AI-optimized values.

5. DISCUSSION

This retrospective comparative study demonstrated that an AI-optimized low-dose radiographic protocol, as compared to a current standard protocol, resulted in an approximately 55% decrease in effective dose, while also improving radiologists' subjective scores and measure of objective image quality. There was no significant difference in rating of diagnostic acceptability and AUC. All measured indices of exposure and dose (mAs, DAP, ESD, and effective dose) showed a dose reduction effect of similar magnitude, indicating an internally consistent effect rather than a spurious result based on a single measure of exposure.

5.1 Dose Reduction and Its Interpretation

The size of the dose reduction is significant from a clinical perspective in the context of optimization. While the absolute per radiograph effective dose is small (0.082 versus 0.037 mSv), radiography is performed in high volumes and, as reported by UNSCEAR (2022), medical exposure is the major contributor to artificial population exposure. Effective dose is therefore

more commonly used for optimization between protocols rather than to estimate individual risk (Harrison et al., 2021), which is consistent with its role here. A key limitation here is also that kVp was reduced in the AI group.

A lower kVp usually provides higher subject contrast, at the expense of reduced penetration (Hamada et al., 2020). Both modifications were made together, so the present results are of the combined 'standard with AI' protocol, without attributing the improvements specifically to the DE noising technique.

5.2 Image Quality

The gains in SNR and CNR with noise reduction suggest a successful DE noising algorithm, and agree with other deep-learning DE noising work in mammography and CT (Ha et al., 2022; Zhang et al., 2022; Toepfer et al., 2021). The fact that radiologists preferred the AI images on all metrics, including artifact suppression, agrees with the objective metrics. The fact that they achieved higher scores in smoothness should not be mistaken for increased diagnostic information. DE noising can remove low-contrast details, and readers may prefer images that look cleaner to the eye.

This is part of the reason that the AUC and diagnostic acceptability scores are more meaningful than the Likert scores, and the non-significant AUC difference ($p = 0.118$) should be interpreted not as "not statistically different," but as equivalent performance consistent with non-inferiority. The caveat is the chest CT work by Jo et al. (2023), in which noise measurements – favoring the deep-learning images – were in fact accompanied by ratings of conventional images as being equal or better in several categories; the authors argued in part that the effect could be due to radiologists' familiarity with conventional-texture images, and the deep learning algorithm having been trained to reduce the objective noise measurements.

5.3 Consistency Across Anatomical Regions

The subgroup analysis showed numerically larger effects in chest x-ray, in the presence of lung fields and relatively lower intrinsic attenuation that might provide greater scope for noise reduction, though there was no statistically significant interaction. Due to small subgroup sizes of 60 and 50 in each arm, there was inadequate power to detect even moderate differences, so the lack of significant interaction is consistent but does not exclude regional variation. The abdominal subgroup was not studied separately.

5.4 Relation to Existing Literature

Most of the published evidence of AI dose optimization is in CT, where reviews show improved noise and artifact performance of deep-learning techniques compared with iterative reconstruction (Zhang et al., 2022; Immonen et al., 2021), and AI is used for patient positioning and scan-range localization (Bani-Ahmad et al., 2025). Evidence in projection radiography is more sparse, a perception that was identified when our group undertook a scoping review of the impact of AI on radiographic practice (Hardy and Harvey, 2020). Our findings build on the existing literature to a mixed radiography cohort and appear to be consistent in trend with the DE noising literature reported in mammography (Ha et al., 2022) and manufacturer reported noise reduction (Toepfer et al., 2021) in general radiography, and that AI could potentially also support the reading phase (Sim et al., 2020).

One can therefore surmise that dose optimization and diagnostic support could, at least in theory, be combined within an integrated AI-enabled workflow, as is depicted conceptually in Figure 6.

Larger reductions than those observed in our study (54.9%) have been reported for CT, such as a 75% reduction with non-inferior nodule detection (Jo et al., 2023). In pediatric imaging, most studies in a recent systematic review showed 36-70% reductions without loss of diagnostic utility (Ng, 2022); the 54.9% reduction in our study is within the lower end of this spectrum, albeit with different endpoints and modalities.

5.5 Safety, validation, and clinical implementation

Dose Reduction Through Post-Processing Verification of dose reduction from post-processing is highly important. Deep learning reconstruction has been shown to be potentially unstable and unable to reproduce small changes in structure (Antun et al., 2020) which would imply that for clinical use local acceptance tests, task-based evaluation with lesion-present images, regular auditing of dose index and image quality, and radiographer/radiologist oversight would all be required. The role of

radiographers as a key component in radiology AI-supported acquisition, processing and auditing as described by Hardy and Harvey (2020) is relevant here.

5.6 Clinical and Practical Implications

What this means for departments is that dose and image quality need not be viewed as a zero-sum outcome when we have validated post-processing methods to use in practice. With prospective validation, a reduction of the order envisaged here could be used routinely in everyday exams and drastically reduce exposures for patients who require repeated imaging, such as those with chronic skeletal or chest pathology, and for those patients who are younger, where the lifetime risk per unit dose is much higher. There would seem to be some operational safety net: technologists should not be lowering exposures further just because there's an AI step involved, because DE noising cannot digitally reconstruct information that was not originally obtained.

Dose monitoring is recommended. Monitoring Dose Area Product (DAP) and exposure indices against your own local diagnostic reference levels and auditing images for blinded quality review would potentially identify any drift at the earliest. If radiographers are selecting protocols and doing their own quality assurance (e.g. As described by Hardy and Harvey 2020), training on the expected algorithm behavior and limitations should be provided.

Any discernible change in image appearance, such as smoothed subtle trabecular or fine lung textures, should be reported to radiologists so that interpretation criteria may be adjusted accordingly. Since reading-stage AI has been demonstrated to impact on reports and management in clinical practice (Jones et al 2021), if both acquisition- and reading-stage AI are employed, auditing should include both.

5.7 Limitations

There are a few considerations to mention. Firstly, the study was retrospective and not randomized and there were significant differences between the groups in terms of the patients in each arm. There were no differences between patient demographics, but an unmeasured confounder such as indication of the scan or technologist practice could be present and impactful. Secondly, this was a single center, single approach to the AI algorithm and thus the findings are not necessarily generalizable.

Thirdly, the variation in kVp and mAs between the AI and standard groups means the findings cannot directly be attributed to the algorithm alone. Fourthly, the effective dose calculations were made from the DAP measurement. Fifthly, the description of reader blinding to the protocol and the lesion detection reference standard is not enough and one might not be able to distinguish what type of processing was used based on the images. Sixthly, the sample size was small, limiting the precision of the AUC estimate.

Lastly, both the abdominal subgroup and pediatrics were not analyzed separately.

5.8 Future Directions

Further prospective, multicenter investigations with paired acquisition or phantom-based task-based assessment, systematic assessment of low-contrast and small-lesion detectability, head-to-head studies of multiple algorithms and vendors, and health-economic analyses. Dose-reduction aims should be established based on local diagnostic reference levels and monitored systematically after implementation.

5.9 Strengths of the Study

This study has several notable strengths. Radiation dose was characterized using four complementary measures (mAs, DAP, entrance surface dose, and estimated effective dose) and image quality via both objective and subjective approaches; the effect on diagnostic efficacy was directly explored. The images were independently scored by two radiologists, and the inter-observer concordance was good to excellent for all measures, supporting the validity of the subjective assessments. There was no evidence that case-mix would explain the differences in the measures of quality and dose, since the groups were well matched for age, sex, BMI, and anatomic region; and the analysis of the subgroup, with tests of interactions, provides some evidence of agreement across tissue types.

Taken together, these features make the findings of this study a reasonable starting point for hypothesis-generating conclusions, and for the planning of a prospective study for validation, notwithstanding the limitations noted earlier.

6. CONCLUSION

In this retrospective comparative study of 300 radiographic studies, the implementation of an AI-designed low-dose protocol resulted in an average 55% reduction in effective dose in addition to an improvement in subjective and objective image quality with no significant decrease in the diagnostic performance of chest and musculoskeletal radiography, which demonstrates the potential use of AI to achieve the As Low As Reasonably Achievable (ALARA) principle in low-dose routine radiography. Prospective multicenter validation and quality assurance are recommended due to the retrospective design and simultaneous adjustment of exposure factors and software.

REFERENCES

- [1] Antun V, Renna F, Poon C, Adcock B, Hansen AC. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci U S A*. 2020. doi:10.1073/pnas.1907377117. <https://doi.org/10.1073/pnas.1907377117>
- [2] Bani-Ahmad M, England A, McLaughlin L, Hadi YH, McEntee M. Potential of artificial intelligence for radiation dose reduction in computed tomography — a scoping review. *Radiography*. 2025;31(4):102968. <https://doi.org/10.1016/j.radi.2025.102968>
- [3] Brady SL, Trout AT, Somasundaram E, Anton CG, Li Y, Dillman JR. Improving image quality and reducing radiation dose for pediatric CT by using deep learning reconstruction. *Radiology*. 2021;298:180–188. <https://doi.org/10.1148/radiol.2020202317>
- [4] Greffier J, Hamard A, Pereira F, Barrau C, Pasquier H, Beregi JP, et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. *Eur Radiol*. 2020;30:3951–3959. <https://doi.org/10.1007/s00330-020-06724-w>
- [5] Ha SM, Kim HH, Kang E, Seo BK, Choi N, Kim TH, et al. Radiation dose reduction in digital mammography by deep-learning algorithm image reconstruction: a preliminary study. *J Korean Soc Radiol*. 2022;83(2):344–359. <https://doaj.org/article/75c0c3738dcd4c73806e8f620265ac74>
- [6] Hardy M, Harvey H. Artificial intelligence in diagnostic imaging: impact on the radiography profession. *Br J Radiol*. 2020;93(1108):20190840. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7362930/>
- [7] Harrison JD, Balonov M, Bochud F, Martin C, Menzel HG, Ortiz-Lopez P, et al. ICRP Publication 147: Use of dose quantities in radiological protection. *Ann ICRP*. 2021;50(1):9–82. <https://journals.sagepub.com/toc/anib/50/1>
- [8] Immonen E, Wong J, Nieminen M, Kekkonen L, Roine S, Törnroos S, et al. The use of deep learning towards dose optimization in low-dose computed tomography: a scoping review. *Radiography*. 2021. doi:10.1016/j.radi.2021.07.010. <https://doi.org/10.1016/j.radi.2021.07.010>
- [9] Jiang B, Li N, Shi X, Zhang S, Li J, de Bock GH, et al. Deep learning reconstruction shows better lung nodule detection for ultra-low-dose chest CT. *Radiology*. 2022. doi:10.1148/radiol.210551. <https://doi.org/10.1148/radiol.210551>
- [10] Jo GD, Ahn C, Hong JH, Kim DS, Park J, Kim H, et al. 75% radiation dose reduction using deep learning reconstruction on low-dose chest CT. *BMC Med Imaging*. 2023;23:121. <https://doi.org/10.1186/s12880-023-01081-8>
- [11] Jones CM, Danaher L, Milne MR, Tang C, Seah J, Oakden-Rayner L, et al. Assessment of the effect of a comprehensive chest radiograph deep learning model on radiologist reports and patient outcomes: a real-world observational study. *BMJ Open*. 2021;11(12):e052902. <https://doi.org/10.1136/bmjopen-2021-052902>
- [12] Nam JG, Ahn C, Choi H, Hong W, Park J, Kim JH, et al. Image quality of ultralow-dose chest CT using deep learning techniques: potential superiority of vendor-agnostic post-processing over vendor-specific techniques. *Eur Radiol*. 2021;31:5139–5147. <https://doi.org/10.1007/s00330-020-07537-7>
- [13] Nam JG, Hong JH, Kim DS, Oh J, Goo JM. Deep learning reconstruction for contrast-enhanced CT of the upper abdomen: similar image quality with lower radiation dose in direct comparison with iterative reconstruction. *Eur Radiol*. 2021;31:5533–5543. <https://doi.org/10.1007/s00330-021-07712-4>

- [14] Ng CKC. Artificial intelligence for radiation dose optimization in pediatric radiology: a systematic review. *Children (Basel)*. 2022;9(7):1044. <https://doi.org/10.3390/children9071044>
- [15] Park SB. Advances in deep learning for computed tomography denoising. *World J Clin Cases*. 2021;9(26):7614–7619. <https://doi.org/10.12998/wjcc.v9.i26.7614>
- [16] Seah JCY, Tang CHM, Buchlak QD, Holt XG, Wardman JB, Aimoldin A, et al. Effect of a comprehensive deep-learning model on the accuracy of chest x-ray interpretation by radiologists: a retrospective, multireader multicase study. *Lancet Digit Health*. 2021;3:e496–e506. [https://doi.org/10.1016/S2589-7500\(21\)00106-0](https://doi.org/10.1016/S2589-7500(21)00106-0)
- [17] Sim Y, Chung MJ, Kotter E, Yune S, Kim M, Do S, et al. Deep convolutional neural network–based software improves radiologist detection of malignant lung nodules on chest radiographs. *Radiology*. 2020;294(1):199–209. <https://pubmed.ncbi.nlm.nih.gov/31714194/>
- [18] Toepfer K, Barski L, Sehnert J, Vogelsang L. Smart Noise Cancellation processing: new level of clarity in digital radiography (technical white paper). Carestream Health; 2021. https://www.carestream.com/blog/wp-content/uploads/2021/05/white_paper_smart_noise_cancellation_ltr_lowres_202104.pdf
- [19] United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Sources, effects and risks of ionizing radiation: UNSCEAR 2020/2021 Report, Volume I, Annex A — Evaluation of medical exposure to ionizing radiation. New York: United Nations; 2022. https://www.unscear.org/docs/publications/2020/UNSCEAR_2020_21_Report_Vol.I.pdf
- [20] Zhang M, Gu S, Shi Y. The use of deep learning methods in low-dose computed tomography image reconstruction: a systematic review. *Complex Intell Syst*. 2022;8:5545–5561. <https://doi.org/10.1007/s40747-022-00724-7>