Technical Brief Series

Smart Noise Cancellation Processing: Providing a New Level of Clarity in Digital Radiography and a Foundation to Reduce Dose

Authors: *Lori Barski (MS), Karin Toepfer (PhD), Jim Sehnert (PhD) and Levon Vogelsang (PhD)*

Carestream Health, Inc., has developed a convolutional neural network (CNN)-based noise-suppression approach called Smart Noise Cancellation (SNC) that significantly reduces image noise while retaining fine spatial detail.[1] SNC is an optional feature for Eclipse, the intelligent platform that serves as the backbone of Carestream's image processing. SNC works in harmony with EVP Plus[2] image processing, which renders projection radiographs with superior image quality. The synergy of the two – SNC and EVP Plus – results in remarkable image quality improvements.

This white paper provides an overview of the SNC technology and includes clinical reader study results that demonstrate a significant improvement in dose efficiency. This means that SNC provides users with the ability to improve image quality at nominal-dose levels or preserve image quality at a reduced-dose level. A summary of the approach used – artificial intelligence (AI) by means of a CNN – is discussed in Section 2. Objective image-quality measurements are presented in Section 3, including measurements of noise in uniform areas, sharpness and contrast-detail performance that are characterized from phantom captures. Subjective image-quality results, which were measured during two separate reader studies, are presented in Sections 4a and 4b, followed by a brief discussion of the configurability of SNC in Section 5.

1. Introduction

Best practice in medical X-ray imaging employs the principle of ALARA – "as low as reasonably achievable" – for dose management. A consequence of this principle is that imaging is performed with a dose just high enough to confidently achieve diagnosis.[3] As a result, images tend to contain noise that reduces clarity and masks anatomical structures, resulting in degraded image quality. Medical image processing utilizes traditional noise suppression approaches that can lead to some loss of fine image detail. In recent years, noise reduction with deep convolutional neural networks has been shown to better preserve image detail.[4] The benefits of CNN-based noise reduction are improved image quality, reduced-dose, increased contrast-to-noise and radiographs that are easier to read.

Figure 1 on the next page provides a visual illustration of SNC. (Due to the very fine detail of the images, which are enhanced by the removal of noise, please zoom into the images to better visualize the SNC effect if you are reading a soft copy of this paper.) The left shows the noisy ankle joint as originally captured, the center shows the ankle joint after SNC, and the right shows the difference between the two images that corresponds to the noise field. Note the absence of spatial detail in the noise field.

This advanced noise-suppression method may offer benefits in:

- Dose reduction, e.g. from 400 speed to 800 speed for CsI and from 320 speed to 500 speed for GOS panels.
- Gridless imaging (i.e., SmartGrid), where scatter removal typically leads to increased noise appearance.
- Neonatal and pediatric imaging, where imaging at the lowest possible dose is critical.
- Portable imaging, where conditions challenge acquisition of optimal images.
- General radiography, to improve the clarity of anatomical features in the processed images.

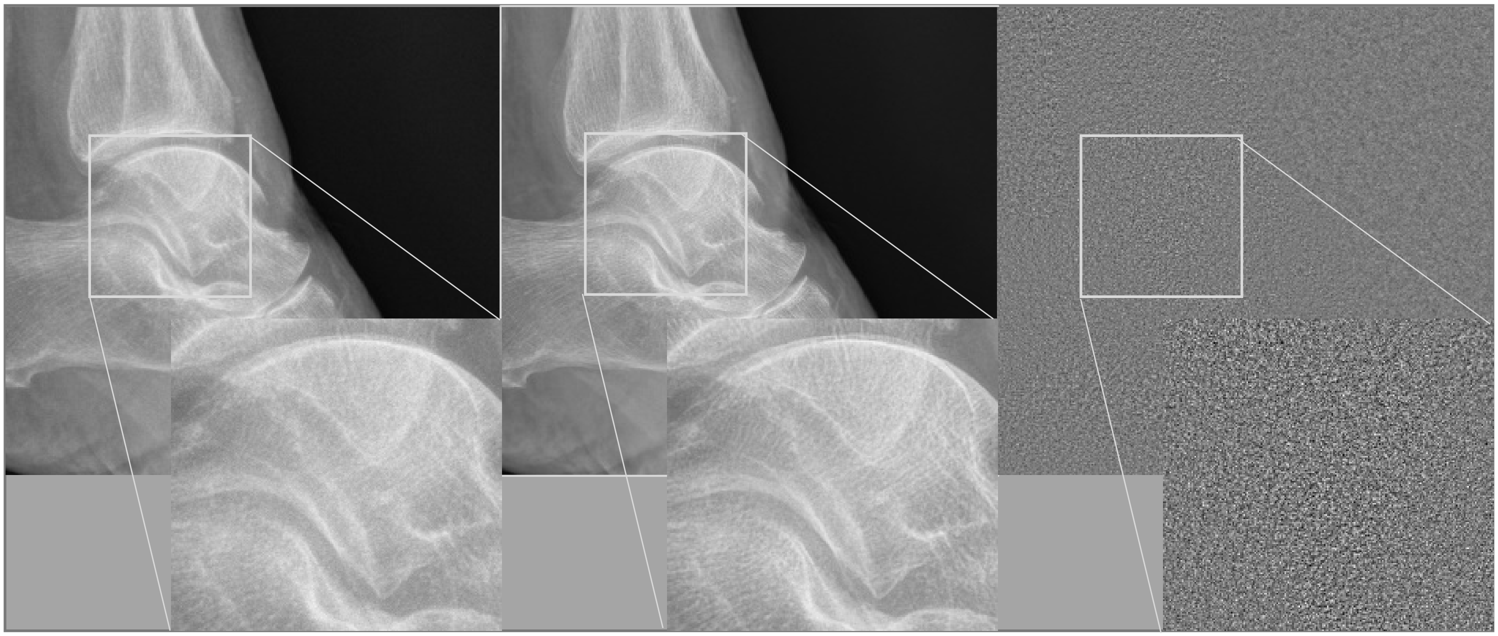


Figure 1. Ankle joint of lateral foot (60 kVp, 0.4 mAs, 40″ SID, IEC EI 115); Left – Original image with noise; Center – Image after SNC; Right – Predicted noise field shown with a window of -13 to + 13 code values.

2. The SNC Algorithm

SNC uses a deep convolutional neural network[5] trained to predict a noise field from an input image (Figure 2). The network, a U-Net architecture[6], was trained using low-noise/high-noise image pairs of clinical patient, cadaver and anthropomorphic phantom images representing a wide variety of general radiographic exams. The high-noise images were produced by using image simulations[7] to create a lower-dose equivalent of the input original (low-noise) images. The simulated high-noise images were equivalent to 40% of the original dose of the input original (low-noise) image. These noise simulations were based on a validated physical noise model of a-Si-based flat-panel detectors incorporating exposure-dependent X-ray quanta and detector-panel structure noise, exposure-independent electronic noise and the spatial texture of the noise. The original input images were used as the higher dose aims for training. Using simulated noise eliminates misregistration (misregistration would cause an artificial loss of sharpness) and many examples of noisy images are readily available without repeated patient exposures.

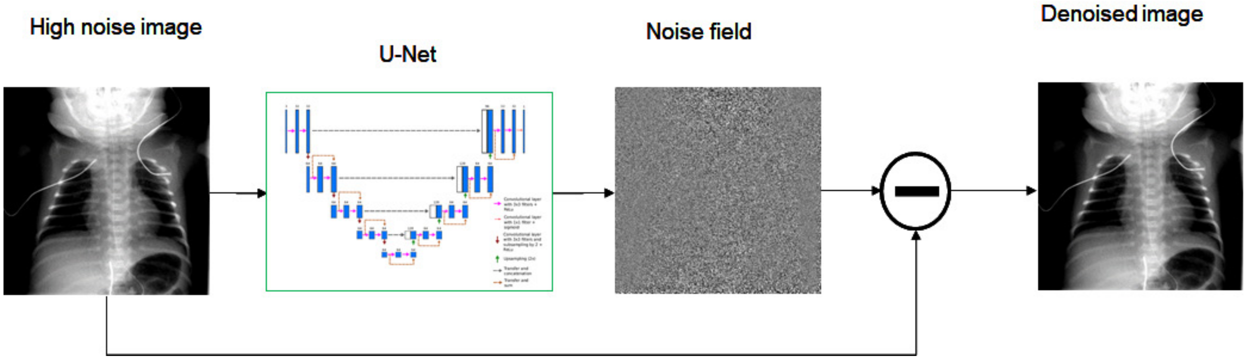


Figure 2. Diagram illustrating the model pipeline. A noisy image is added to a network that predicts a noise field. This noise field is then subtracted from the input image to produce the noise-reduced image.

During the training process, 480 small patches of 128 × 128 pixels were randomly sampled from each input image with a 2560 x 3072 pixel matrix. Patch selection was randomized for each batch in the optimization, resulting in more than 25 million patches used during training. The weights of the U-Net were optimized based on the mean absolute error-loss function between the predicted and the aim noise field.

Different CNN noise models were trained for each type of flat-panel detector in the Carestream portfolio. Detectors were grouped by pixel spacing, flat-panel detector type and scintillator technology (cesium iodide, CSI, and gadolinium oxysulfide, GOS) as shown in Table 1.

Detector Type	Scintillator	Pixel Pitch
DRX Plus 3543/4343	GOS	0.139
DRX Plus 3543C/4343C	CsI	0.139
DRX Plus 2530C	CsI	0.098
DRX-1	GOS	0.139
DRX1-C, DRX 2530C	CsI	0.139

Table 1. Carestream detector types for which CNN noise models were created.

3. Objective Performance on CARESTREAM DRX Plus Detectors

Background

While CNN-based noise suppression is a nonlinear process, it is nevertheless useful to characterize its performance, in terms of image quality, by using traditional analysis methods with nonclinical data. And because SNC is the first step in the image processing chain after the detector receives the raw images, analysis performed with and without SNC is suitable for comparing the image quality differences.

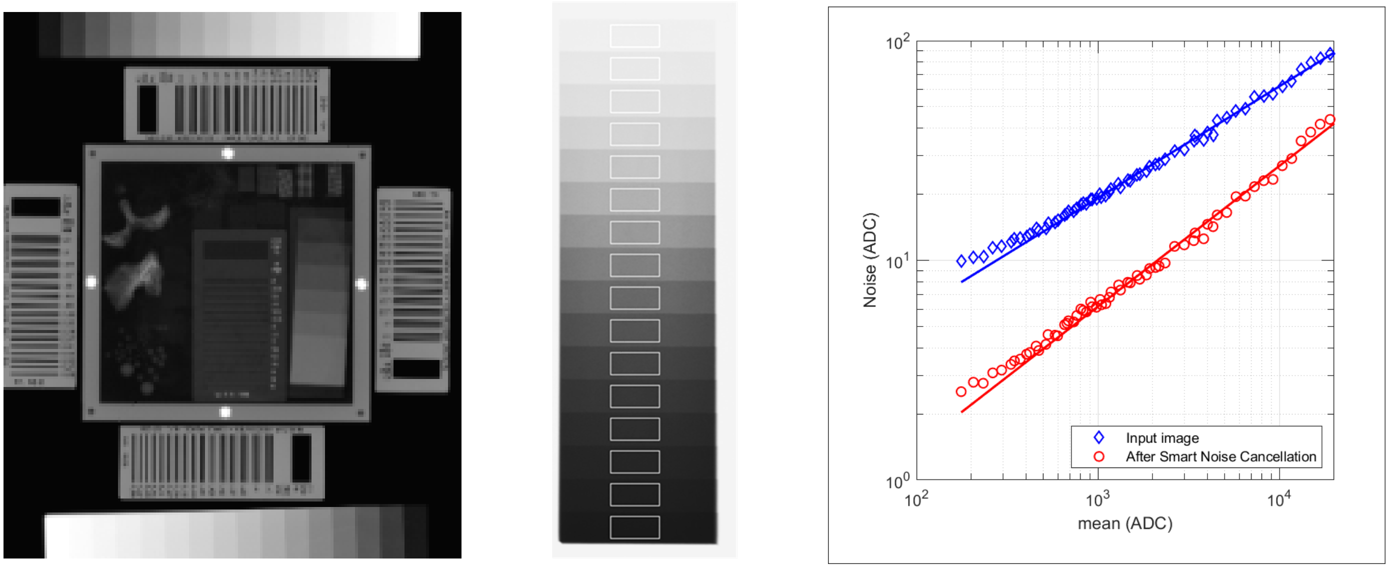
Several objective features were chosen for analysis. Noise in uniform areas, sharpness, and the rendering of low-contrast objects and objects with fine detail were all characterized from test-phantom captures. A second form of objective testing was done based upon disease feature simulation. Disease features consisted of 10 mm lung nodules[8] and a 0.5 mm high-contrast feature at two contrast levels. A mathematical observer, specifically a channelized Hotelling observer, was employed to demonstrate increased detectability of disease features with SNC. Details of this analysis are provided in Reference 9[9].

Noise reduction in uniform image areas

A special test phantom shown in Figure 3a (next page) was used that contained aluminum step tablets, resolution targets, small acrylic beads, wire mesh, bone chips and other features for the qualitative and quantitative evaluation of image quality. This phantom was imaged at 80 kVp, 0.5 mm Cu / 1 mm Al filtration, 180 cm source-to-image distance (SID), and at four exposure levels: 0.5 mAs, 1.0 mAs, 2.0 mAs, and 10 mAs. Figure 3b shows an image of a step tablet from the phantom together with the regions of interest that were used for analysis of the mean and standard deviation, the latter serving as a measure of noise from uniform areas. Noise measurements are shown in Figure 3c for Carestream's DRX Plus 3543C detector. The solid blue line indicates quantum-limited behavior (Noise $\propto$ mean$^{0.5}$) and matches well to the noise measurements made prior to application of SNC. Noise reduction ranged between 4x at low exposures and 2x at higher exposures. In terms of quantum noise, a 2x noise reduction corresponds to the image appearance of a 4x higher exposure.

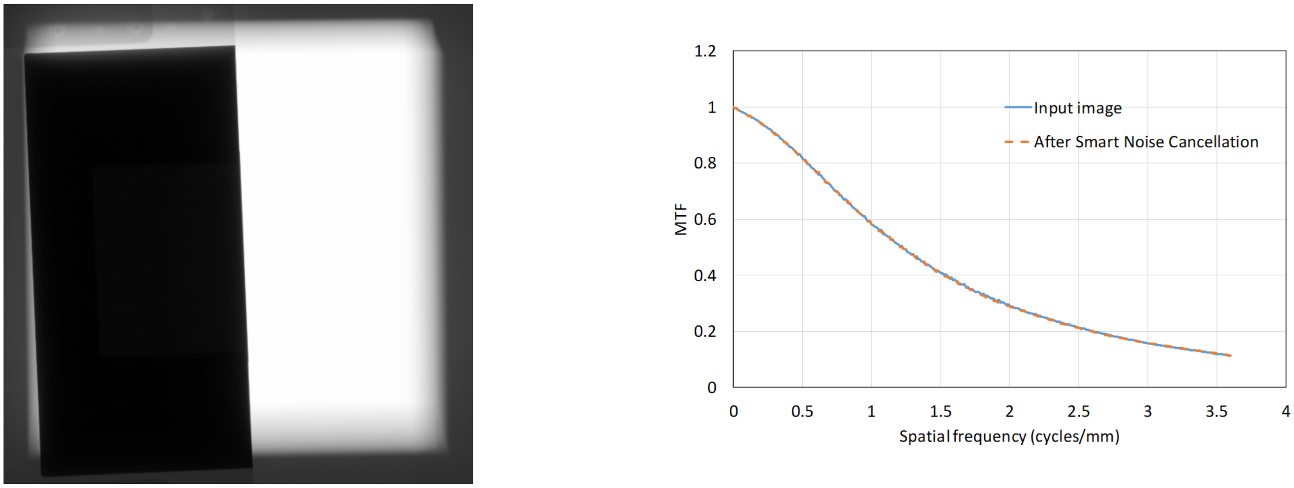
Preservation of high-contrast sharpness

The modulation-transfer function (MTF) was calculated from the acquisition of an edge target conforming to the IEC 62220-1-1 standard for DQE measurement under RQA-5 beam conditions. The exposure level was chosen at approximately 3.2x the normal exposure level for each detector. The normal exposure level corresponds to 2.5 μGy for detectors with a CsI(Tl) scintillator (ISO 400 speed) and 3.1 μGy for detectors with a GOS scintillator (ISO 320 speed). An image of the edge target is shown on the next page in Figure 4a. Preservation of high-contrast sharpness is demonstrated by the MTF in Figure 4b for the DRX Plus 3543C detector – there was no MTF loss after SNC was performed.



a *b* *c*

Figure 3. Uniform area noise reduction.



a *b*

Figure 4. Preservation of high-contrast sharpness.

Preservation of low-contrast and high-frequency detail

Contrast-detail analysis is a common procedure used to characterize the detectability of low-contrast and fine (high-frequency) detail in an X-ray imaging system including the X-ray detector, medical image display and human visual system. The Artinis CDRAD Phantom 2.0[10], used for this purpose, is a 265 x 265 x 10 mm^3 PMMA tablet with a matrix of 15 rows and columns containing cylindrical holes of variable diameter and depth. The layout of the phantom is shown in Figure 5a on the next page. Using this phantom, a contrast-detail curve was generated to represent a plot of minimum visible feature size (hole diameter) as a function of contrast (hole depth). Images of the CDRAD 2.0 phantom were acquired at 70 kVp with a 12:1 203 lp/cm grid to represent general radiography images. The phantom was sandwiched between two 5 cm-thick sheets of polymethyl methacrylate (PMMA) to simulate a thicker anatomy. All images were acquired on a CPI Indico 100 X-Ray generator, without additional filtration, at SID = 183 cm and with a small focal spot (0.6 mm). Images were acquired at five exposure levels corresponding to the detector entrance air kerma under the phantom of 1 µGy, 1.25 µGy, 2.5 µGy, 5 µGy and 10 µGy.

All images were scored with Artinis CDRAD Analyzer 2.1.15 software to produce contrast-detail curves and the inverse of Image Quality Score, IQF_{inv}, before and after SNC. Eight replicated images were included in each score and the confidence level was set to 5.e-5 in the software. IQF_{inv} was calculated according to the following equation:

$$IQF_{inv} = \frac{100}{\sum_{i=1}^{15} C_i \times D_{i,th}},$$

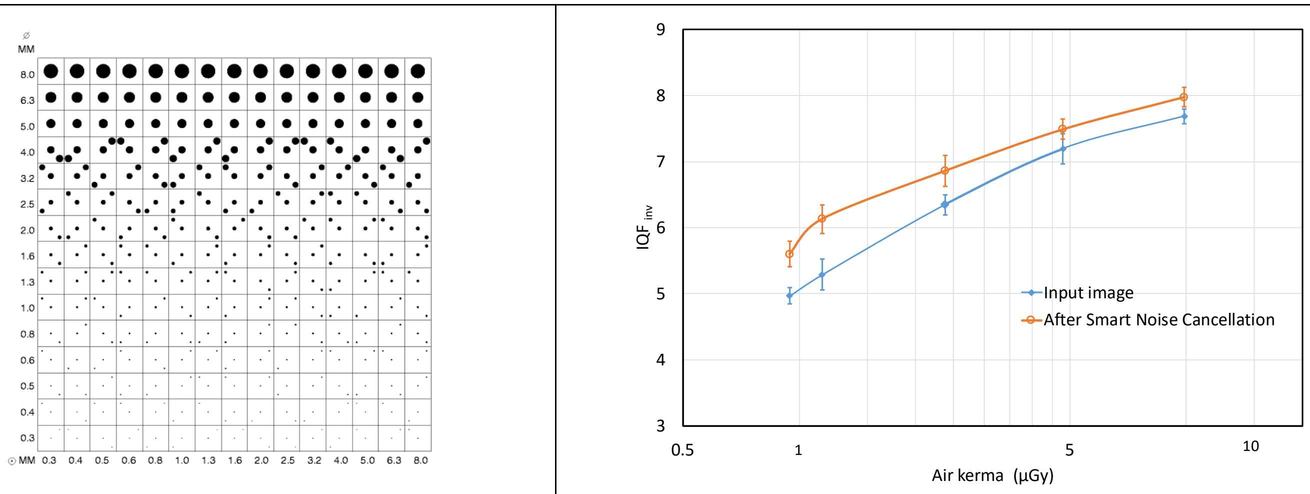
C_i represents the depth value (contrast) of the object (visible hole) in column *i* and $D_{i,th}$ is the threshold (smallest visible) diameter in contrast column *i*. High detectability of a system corresponds to a high IQF_{inv} score.

Figure 5b shows the quality score IQF_{inv} plotted as a function of air kerma at the detector. The plot demonstrates increased detection scores (IQF_{inv}) after SNC is applied. Figure 5c illustrates the noise reduction for one of the hole pairs of the phantom matrix.

To elucidate the source of the image quality improvements after SNC, Figure 5d shows an example of the contrast-detail curves for a 70 kVp capture of the CDRAD phantom on the DRX Plus 3543C detector at 1.25 µGy air kerma, corresponding to an ISO 800-speed exposure. The improvements were mainly seen for low-contrast objects (shallow hole depth), where objects of smaller diameter could be better detected after SNC.

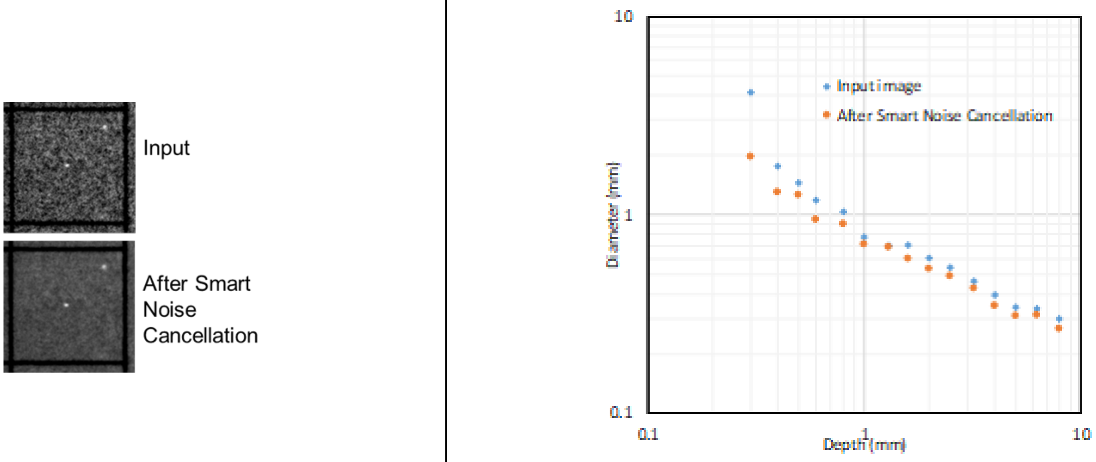
In summary, objective measures used to assess image quality with SNC demonstrated that:

- A 2x to 4x noise reduction in uniform areas was attained.
- High-contrast sharpness was preserved.
- A 10% to 20% improvement in contrast-detail scores on the CDRAD 2.0 phantom was attained.



a. Artinis CDRAD Phantom

b. IQF_{inv} scores vs detector air kerma



c. Hole pair before and after SNC

d. Contrast-detail curve before and after SNC, 1.25 µGy

Figure 5. Contrast-detail results for Carestream DRX Plus 3543C Detector.

Comparison of Scintillator Technology

Carestream's detector portfolio offers a choice of two scintillators – gadolinium oxysulfide (GOS) and cesium Iodide (CsI). GOS scintillators provide a cost-effective offering with good image quality and reduced-dose compared with computed radiography. The CsI scintillator is a premium offering, delivering the highest image quality at the lowest dose based on its higher X-ray absorption and improved light management, due to the columnar structure of the scintillator material compared with GOS. As a result, images acquired on a CsI detector have a higher signal-to-noise ratio (SNR) than images acquired on GOS at the same input exposure (dose).

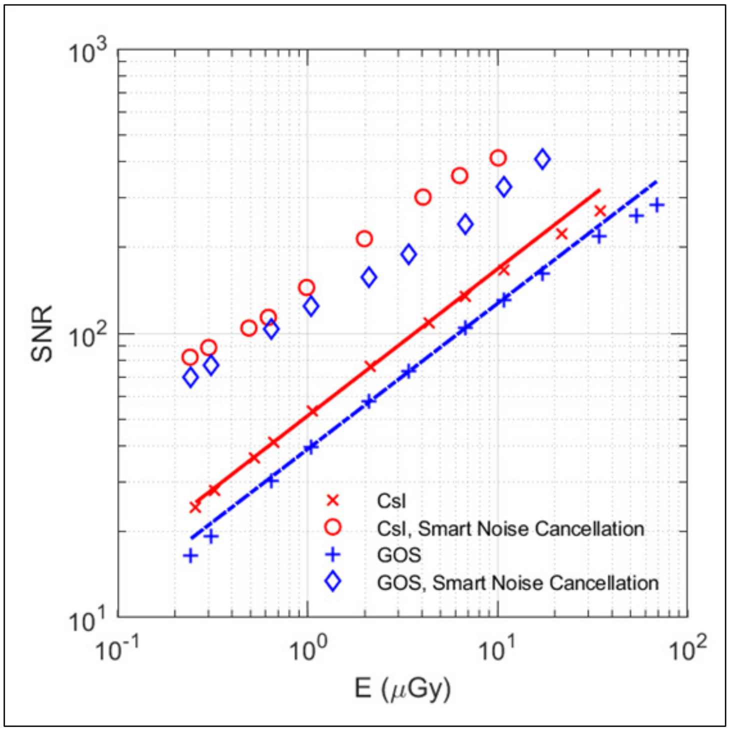


Figure 6. Signal-to-noise ratio (SNR) of flat fields vs air kerma under RQA-5 beam conditions; CsI (DRX Plus 3543C) and GOS (DRX Plus 3543) before and after SNC.

This is illustrated by the fitted red and blue lines in Figure 6, where SNR was measured for both scintillators at RQA-5 beam quality. Figure 6 also shows that, after application of SNC, the SNR for both scintillator technologies is significantly improved. Moreover, the SNR with the GOS scintillator after SNC has been applied is higher than that of the CsI scintillator without applying SNC. This observation extends to more complex anatomical images, as illustrated in Figure 7. In this example, SNC enables the noise associated with the GOS scintillator to be reduced to a level that is comparable to a CsI scintillator. The acquisitions were performed at 500 speed, using 75 kVp, 6.3 mAs, 40 ln/cm 6:1 grid, IEC EI 129 (CsI), IEC EI 126 (GOS).

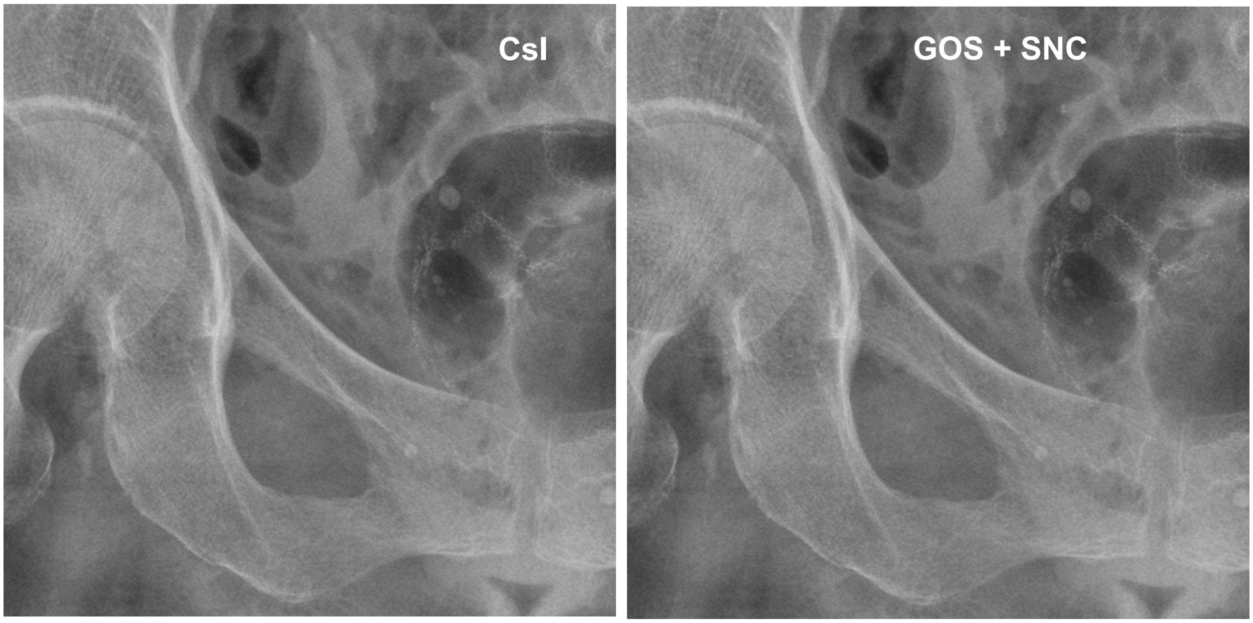


Figure 7. Comparison of an ISO 500-speed pelvis exam acquired on CsI vs. ISO 500-speed pelvis exam acquired on GOS with SNC.

4. Subjective Performance of SNC

The greatest risk of noise suppression is the possibility of inadvertently removing important information. The assessment of this risk is best accomplished by performing a controlled observer study that evaluates image quality based upon human observers with appropriate domain knowledge. Two of these studies were performed, first to assess the image-quality impact of SNC, and the second to assess the dose-reduction impact of SNC.

4a. Image-Quality Impact of SNC – Isodose Study

The objectives of the image-quality study were to demonstrate that images processed with SNC deliver a quality level that is equivalent to or better than images processed with EVP Plus alone, and that SNC processing is safe and effective.

Two U.S. board-certified radiologists (specializing in diagnostic radiology) evaluated 67 pairs of human clinical and cadaveric subjects acquired on five detector types (Table 1). Exposures ranged from 200-1000 speed (image selection biased towards low-exposure cases) with an IEC EI distribution shown in Figure 8. Varied exam types and patient sizes were evaluated. The evaluations were performed on a PACS workstation configured with two diagnostic monitors calibrated to the DICOM grayscale standard-display function.

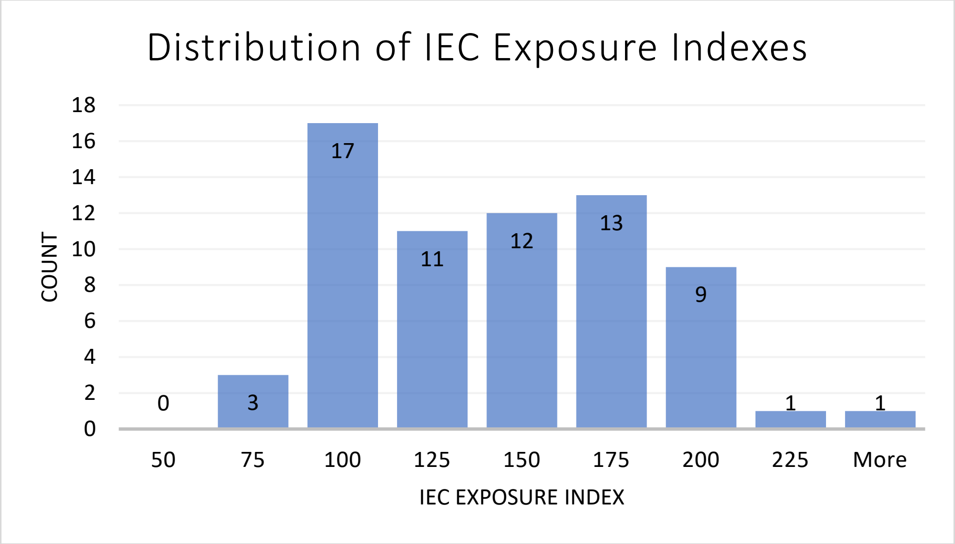


Figure 8. Distribution of IEC Exposure indices in the isodose reader study.

Pairs consisted of the same image processed with default EVP Plus image-processing software, which includes traditional noise suppression, and EVP Plus with SNC, where 100% of the predicted noise field was removed. Prior to the study, EVP Plus with SNC was tuned to have default processing that turns off the traditional noise suppression (a separate capability in EVP Plus) and takes advantage of SNC's noise reduction by optimizing sharpness.

The images from each pair were randomly placed left/right on the PACS workstation monitor and the pairs were randomly distributed into five reading worklists. The worklists were shuffled for each reader so that no reader could read images in the same order. Readers were blinded to the image treatments (i.e. which image was on the left vs. right).

The images were evaluated pairwise using a five-point visual-difference preference scale tied to diagnostic confidence, as described in Table 2. The readers were instructed to use the preference scale such that slightly preferable ratings would likely not impact the diagnosis and strongly preferable ratings would likely impact the diagnosis. In addition, the overall diagnostic capability of each image in the pair was rated using the RadLex[11] scale, as described on the next page in Table 3. Reader comments were also recorded.

Rating Score	Score Description
-2	Left image strongly preferable, probable diagnostic impact
-1	Left image slightly preferable, no diagnostic impact
0	No preference
+1	Right image slightly preferable, no diagnostic impact
+2	Right image strongly preferable, probable diagnostic impact

Table 2. Five-point Visual Difference Preference scale.

Score	Term	Definition
1	Non-Diagnostic	Unacceptable for diagnostic purposes. Little or no clinically usable diagnostic information (e.g., gross underexposure, system failure or extensive motion artifact). Almost all such imaging should be repeated. Similar to International Labor Office (ILO) classification #4: "Unacceptable."
2	Limited	Acceptable, with some technical defect (motion artifact, body habitus/poor X-ray penetration, or patient positioning may limit visualization of some body regions but is still adequate for diagnostic purposes). Not as much diagnostic information as is typical for an examination of this type, but likely sufficient. Similar to ILO classification #3: "Poor, with some technical defect but still acceptable."
3	Diagnostic	Image quality that would routinely be expected when imaging cooperative patients. Similar to ILO classification #2: "Acceptable, with no technical defect likely to impair classification of the radiograph."
4	Exemplary	Good, most adequate for diagnostic purposes. Image quality that can serve as an example that should be emulated. Similar to (ILO) classification #1: "Good."

Table 3. RadLex Scale for Diagnostic Capability rating.

The readers were trained on the PACS workstation. Each observer started the evaluation at a different point in the reading worklist to guard against learning bias. Observers could adjust window width/window level, pan, magnify and synchronized pan/magnify on the PACS workstation to evaluate image quality.

After the observers finished rating the images, left/right preference ratings were decoded so that positive values indicated favor for SNC. Similarly, RadLex ratings were decoded to map left/right ratings to their corresponding treatment: EVP Plus or EVP Plus with SNC.

Isodose Reader Study Results

Table 4 presents descriptive statistics of all ratings and the distribution of RadLex ratings is presented in Figure 9. The median RadLex rating of EVP Plus with SNC was 4 (Exemplary). An average RadLex rating difference of 0.5 (half of a rating level) is a meaningful difference that indicates a substantial difference in image quality. Inference testing (paired t-test) of the RadLex rating differences, testing if the mean difference was greater than 0.5, is summarized in Table 5. It demonstrates that EVP Plus processing with SNC yields diagnostic quality ratings that are substantially higher than the EVP Plus processing alone, with a 95% confidence level (*indicates a significant p-value result).

	EVP Plus RadLex	**EVP Plus w/ SNC RadLex**	**Pair Preference (+ favor for SNC)**
Mean	2.9	3.6	1.3
Std. Error Mean	0.03	0.04	0.06
Median	3	4	1.0
Std. Deviation	0.34	0.50	0.66
95% Conf. Interval	(2.83, 2.95)	(3.48,3.65)	(1.23, 1.46)
Count	133	133	133

Table 4. Descriptive statistics for all RadLex and Preference ratings.

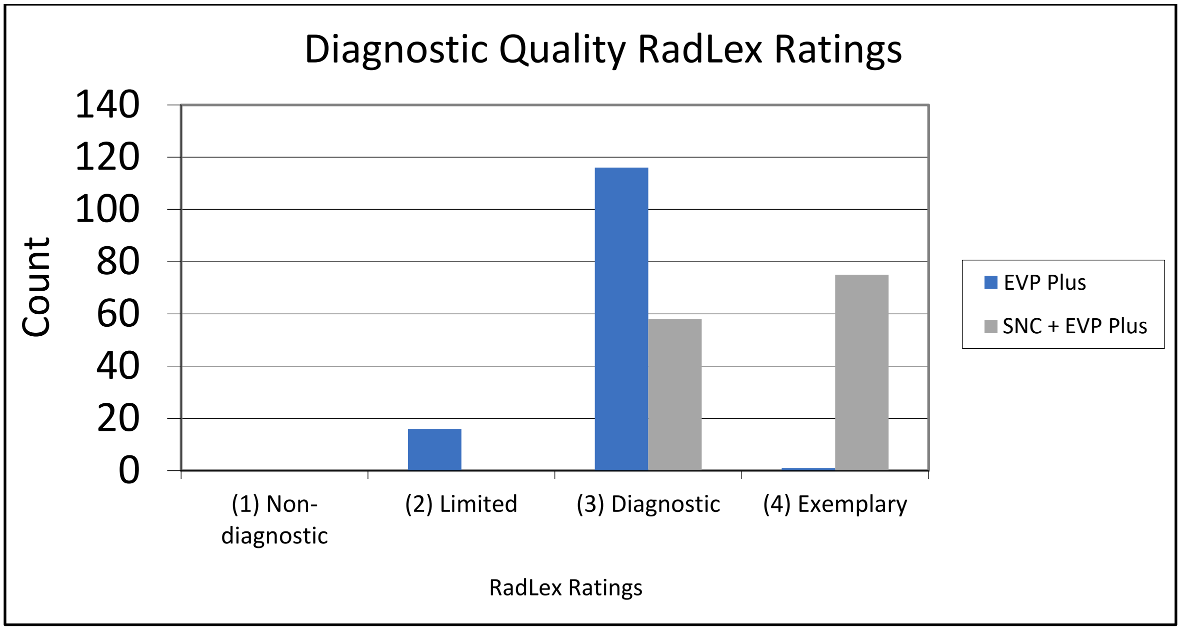


Figure 9. Distribution of RadLex ratings for all readers.

Metric	Hypothesis Statement	Test	Estimation of Paired Difference					
			Mean	Std. Dev.	SE Mean	95% Lower Bound for mean diff	t-statistic	p-value
RadLex Difference	Ho: mean diff = 0.5 Ha: mean diff > 0.5	Paired t-test	0.68	0.53	0.05	0.60	3.84	0.000*

*Significant

Table 5. Paired t-test results of RadLex rating differences.

An average Preference rating greater than 0.5 is considered a threshold level that indicates a substantial reader preference. The one-sample t-test was used to determine if the mean preference was greater than 0.5, and the result is summarized in Table 6. The mean preference (positive values indicate favor for EVP Plus with SNC) is greater than 0.5, supporting the conclusion that EVP Plus with SNC is substantially more preferred over EVP Plus alone, with 95% confidence.

Metric	Hypothesis Statement	Test	Estimation of Preference					
			Mean	Std. Dev.	SE Mean	95% Lower Bound for μ	t-statistic	p-value
Preference	Ho: mean ≤ 0.5 Ha: mean > 0.5	1-sample t-test	1.35	0.66	0.06	1.25	14.70	0.000*

*Significant

Table 6. One-sample t-test of Preference ratings.

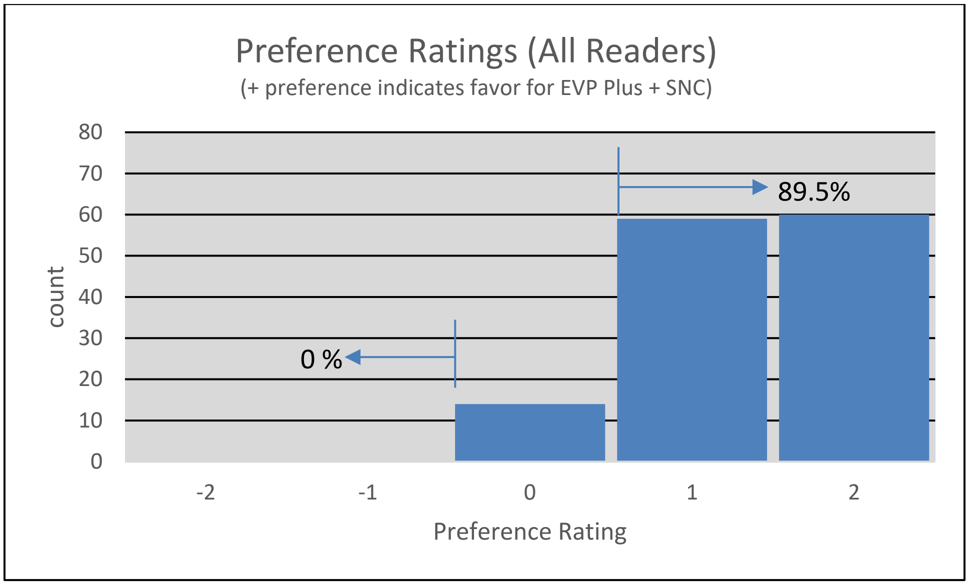


Figure 10. Distribution of Preference ratings for all readers.

Figure 10 shows the distribution of all Preference ratings and that 89.5% of all ratings showed slight to strong preference for the SNC processing.

Table 7 is a paired-comparison contingency table of RadLex ratings. Zero counts at the top of the shaded diagonal areas indicate zero instances of the EVP Plus processing as being rated more highly than EVP Plus processing with SNC. There were four instances of images processed with default EVP Plus processing as being rated Limited (RadLex rating = 2), but with SNC were rated Exemplary (RadLex rating = 4). Likewise, there were 12 instances of images processed with default EVP Plus processing as being rated Limited, but with SNC they were rated as Diagnostic (RadLex rating = 3). And finally, there were 70 instances of images processed with default EVP Plus processing rated Diagnostic, but with SNC were rated Exemplary.

		EVP Plus RadLex Ratings				
	Counts % of Row	**1 Non-Diagnostic**	**2 Limited**	**3 Diagnostic**	**4 Exemplary**	Total
EVP Plus + SNC RadLex Ratings	**1 Non-Diagnostic**	0 0.00 %	0 0.00 %	0 0.00 %	0 0.00 %	0 0.00 %
	2 Limited	0 0.00 %	0 0.00 %	0 0.00 %	0 0.00 %	0 0.00 %
	3 Diagnostic	0 0.00 %	12 20.69 %	46 79.31 %	0 0.00 %	58 100.00 %
	4 Exemplary	0 0.00 %	4 5.33 %	70 93.33 %	1 1.33 %	75 100.00 %
	Total	0 0.00 %	16 12.03 %	116 87.22 %	1 0.75 %	133 100.00 %

Table 7. RadLex paired-comparison contingency table.

These increases in diagnostic capability clearly indicate that EVP Plus with SNC provided significant improvements in image quality.

Reader variability was not a significant source of variation in the study (ANOVA with H_o: Readers are equal; H_a: Readers are not equal; p = 0.572). Likewise, detector type was not a significant source of variation (p = 0.264) and exposure speed was not a significant source of variation (p = 0.518).

In conclusion, the image-quality assessment by board-certified radiologists yielded a strong signal that EVP Plus with SNC significantly improves image quality and is strongly preferred.

Examples of IQ impact with SNC Processing

Figures 11 thru 14 are additional examples of SNC. Figure 11 demonstrates its benefit when combined with SmartGrid – a software-based scatter-reduction feature. Figure 12 demonstrates SNC's benefit on an adult elbow and includes a rendering of the noise field. Notice the lack of structure and edges in the noise field. Figure 13 demonstrates SNC on a low-exposure pediatric arm along with a rendering of the noise field. Figure 14 demonstrates SNC processing on a low-exposure pediatric babygram acquired on the DRX Plus 2530C detector.

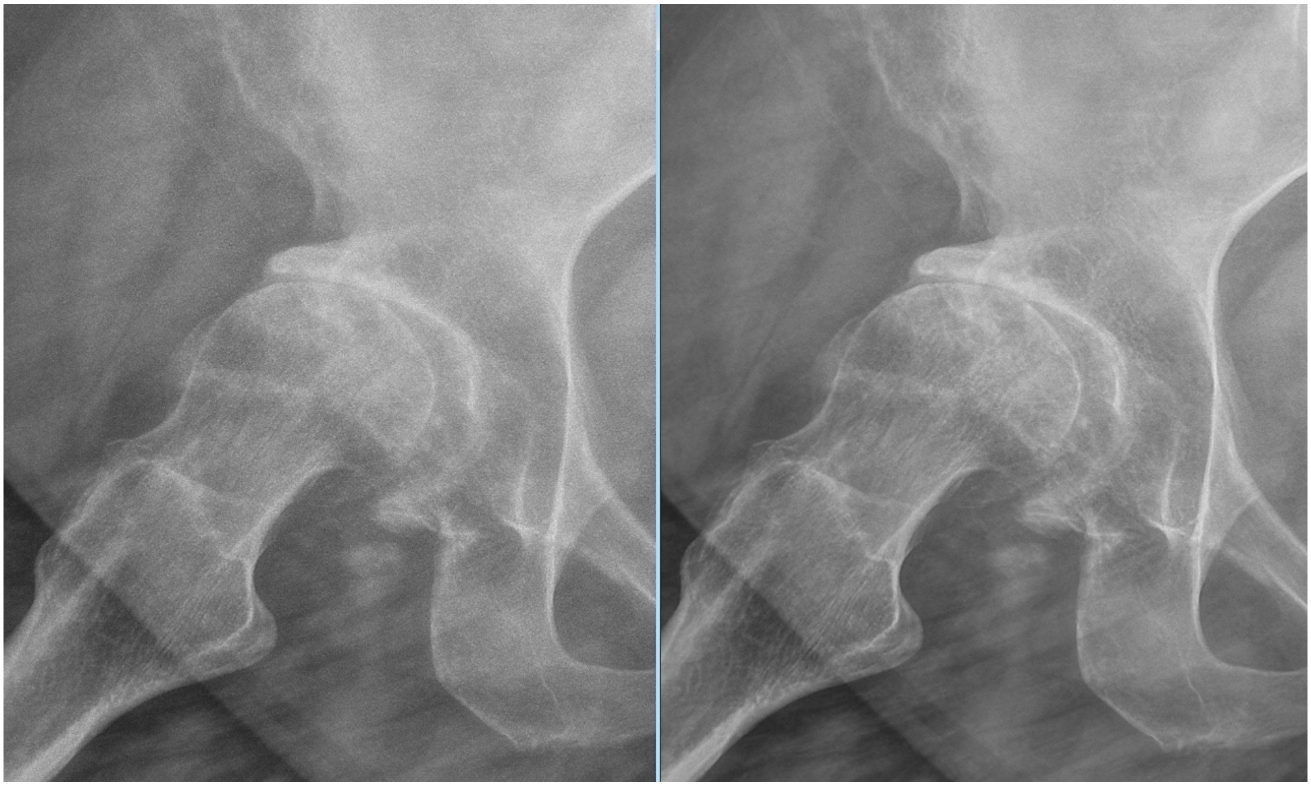


Figure 11. Left – EVP Plus with scatter suppression (SmartGrid), EI 158; Right – Same image using EVP Plus with scatter suppression (SmartGrid) and SNC resulting in improved clarity.

Figure 12. Adult elbow on DRX Plus 2530C: 55 kVp, 0.36 mAs, IEC EI 69. Left Image: EVP Plus default processing; Middle: EVP Plus with SNC; Right: Noise field.

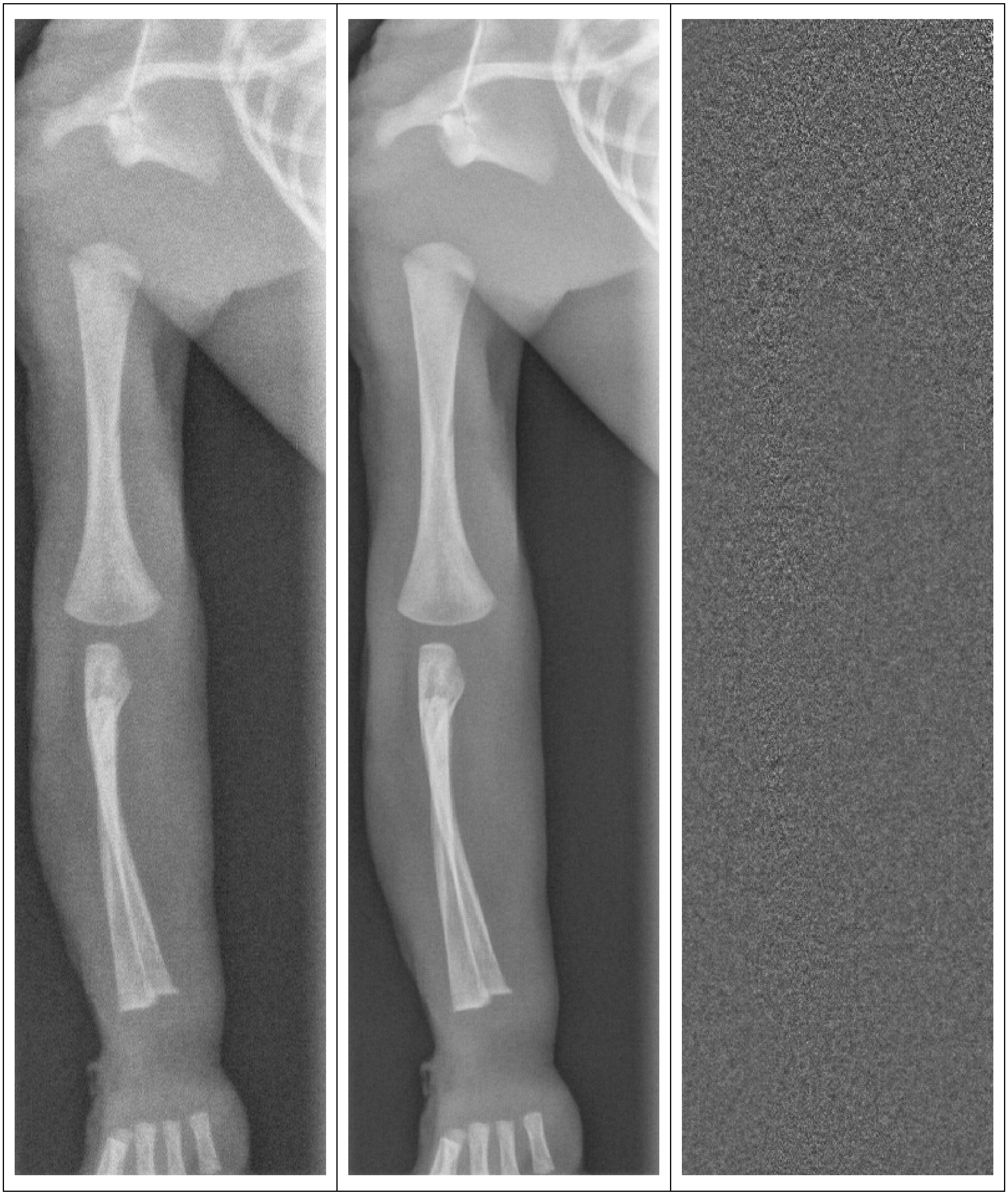


Figure 13. Infant arm on DRX1 (GOS): 43 kVp, 46" SID, 1 mAs, IEC EI 154. Left image: EVP Plus default processing; Middle: EVP Plus with SNC; Right: Noise field.

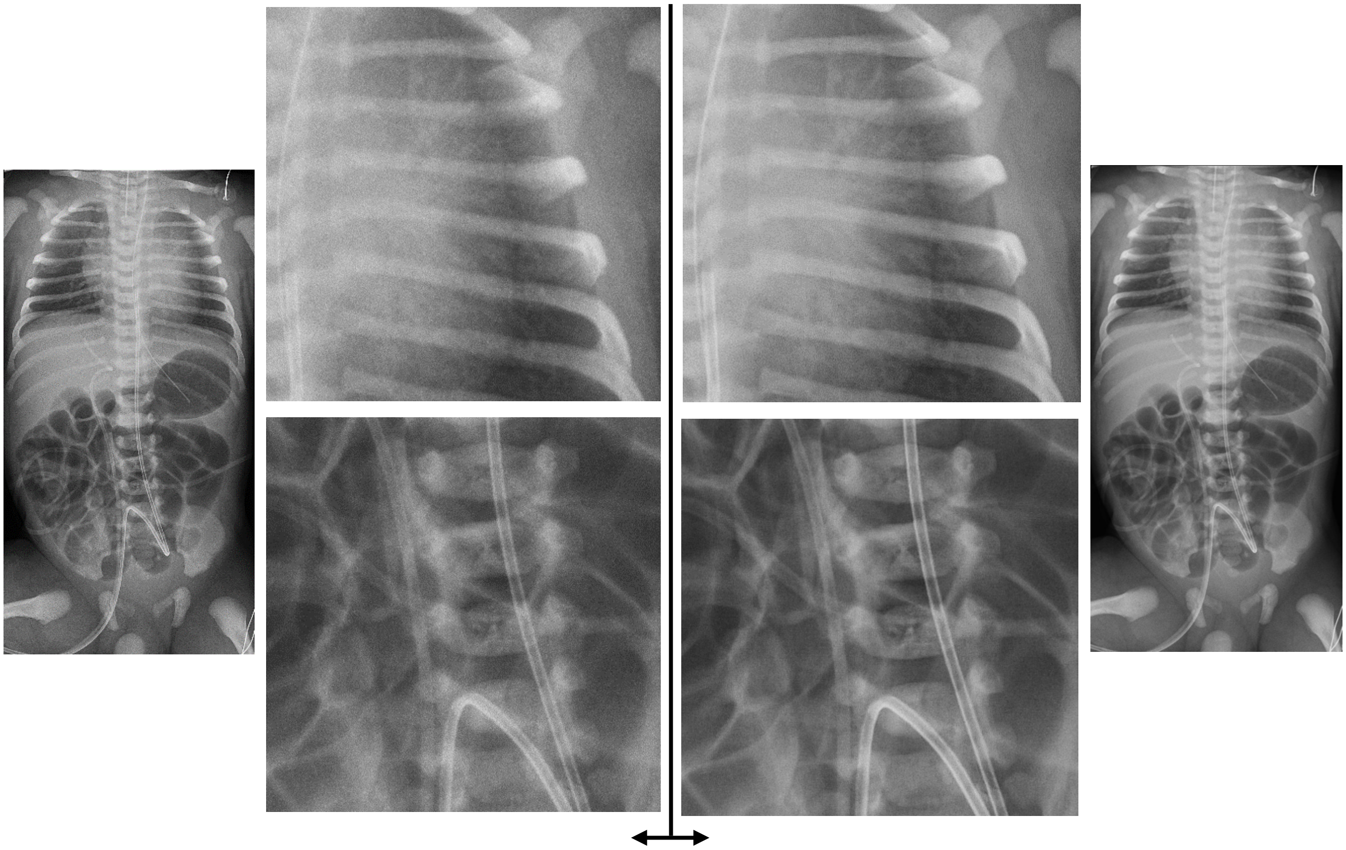


Figure 14. Pediatric babygram on DRX Plus 2530 detector: 55 kVp, 1 mAs, IEC EI 80. Left half: Default processing; Right half – SNC processing.

4b. Image-Quality Impact of SNC – Dose-Reduction Study

The primary objective of the reduced-dose study was to demonstrate that reduced-dose images processed with SNC and EVP Plus deliver radiographic image quality that is as good as or better than corresponding images acquired at nominal dose processed with EVP Plus alone. A secondary objective was to evaluate reduced-dose image quality as it relates to detector type, image type and exposure level. The design of this study was very similar to the image-quality isodose study previously described.

Three U.S. board-certified radiologists specializing in diagnostic radiology evaluated 60 pairs of human clinical and cadaveric subjects captured on five detector types (GOS and CsI panels from Carestream's DRX1 and DRX Plus family of detectors; see Table 1). Pairs consisted of a nominally exposed image with EVP Plus processing (without SNC) and a reduced-dose image, ranging from 35%-60% dose reduction, processed with EVP Plus and SNC. Various exams, dose-reduction levels and patient sizes were used in the study.

Two reduced-dose image types were included in the study. Independent cadaver acquisitions across a range of dose reductions were used for the DRX Plus family of detectors. Clinical (live patient) reduced-dose images were simulated at prescribed dose-reduction levels, which were dependent on scintillator type for each detector (40% reduction for GOS and 50% reduction for CsI panels). Noise simulation was used to eliminate the need to acquire multiple exposures of live subjects. The same noise-simulation methodology successfully used to develop the SNC algorithm was applied to a random assortment of clinical images collected under a trade trial agreement. Image type (low-dose cadaver or simulated low-dose clinical) was a factor in the study to determine if these different types of low-dose images were a significant source of variability in the study results.

The images from each pair were randomly placed left/right on the PACS workstation monitors and the pairs were randomly distributed into four reading worklists. The worklists were shuffled for each reader so that no reader could read images in the same order. Readers were blinded to the image treatments (i.e. which image was on the left vs. right). Each radiologist had a different worklist order. The readers were encouraged to pan/zoom and adjust window width/window level to fully explore and compare the images.

Like the Isodose Study, the readers first rated the pair for preference using a five-point relative scale tied to diagnostic confidence (Table 2) and then rated the left and right images for diagnostic quality using the RadLex scale (Table 3). Comments were also recorded.

After the ratings were completed, they were decoded using the same method described in the first study, such that positive preference values indicate favor for the reduced-dose images processed with SNC and EVP Plus.

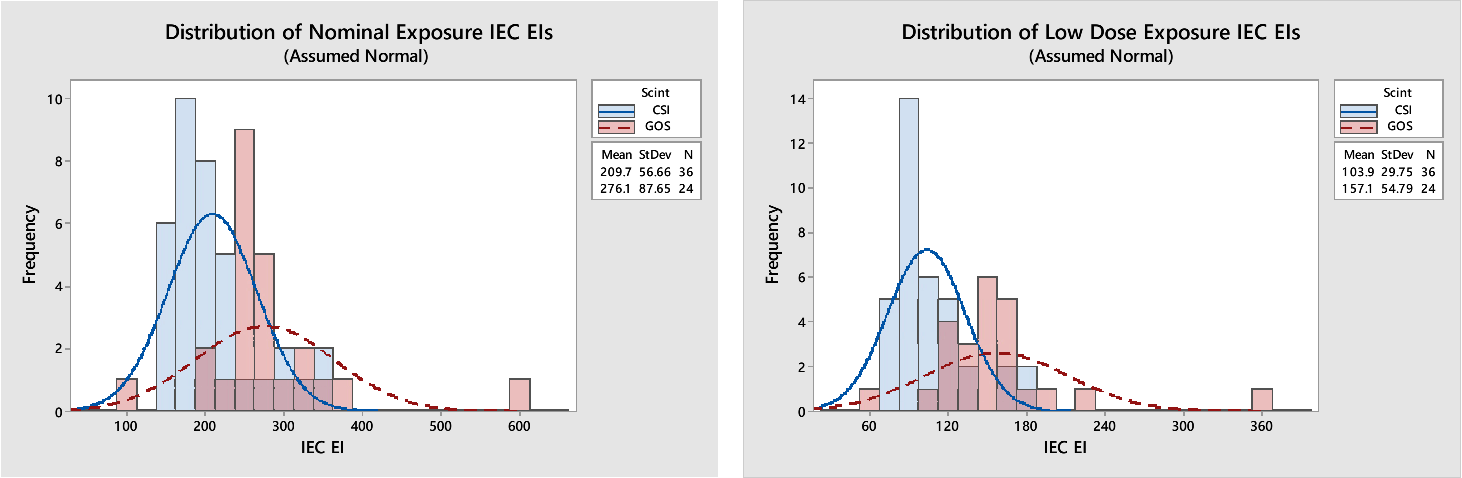
Dose-Reduction Reader Study Results

Descriptive statistics of all ratings are shown in Table 8. The median RadLex rating of reduced-dose images processed with EVP Plus and SNC was "Exemplary" (4), regardless of image type (i.e. cadaver pairs or clinical noise-simulated pairs) as compared to the nominal-dose images, which had a median RadLex rating of 3. An average Preference rating greater than 0.5 (half of a rating level) is a meaningful difference that indicates a substantial reader preference. The mean combined preference of 0.67 (with a 95% lower bound of 0.56) suggests significant preference for low-dose images processed with SNC. The difference between image types for both Preference and RadLex ratings is inconsequential. The non-overlapping confidence interval of the mean RadLex ratings suggests significant differences in favor of the reduced-dose images processed with SNC.

	Pair Preference (+ for Reduced-Dose w/SNC)			**Nominal EVP Plus RadLex Rating**			**Reduced-Dose w/ SNC RadLex Rating**		
Image Type	**Cad**	**Clin**	**Combined**	**Cad**	**Clin**	**Combined**	**Cad**	**Clin**	**Combined**
Mean	0.57	0.78	0.67	3.27	3.20	3.24	3.50	3.57	3.53
Std. Error	0.09	0.08	0.06	0.05	0.06	0.04	0.06	0.05	0.04
Median	1	1	1	3	3	3	4	4	4
Std. Dev.	0.84	0.73	0.79	0.52	0.56	0.54	0.52	0.52	0.52
95% Conf. Int.	All: (0.56, 0.79)			All: (3.16, 3.32)			All: (3.46, 3.61)		
Count	90	90	180	90	90	180	90	90	180

Table 8. Low-dose study descriptive statistics for all RadLex and Preference ratings.

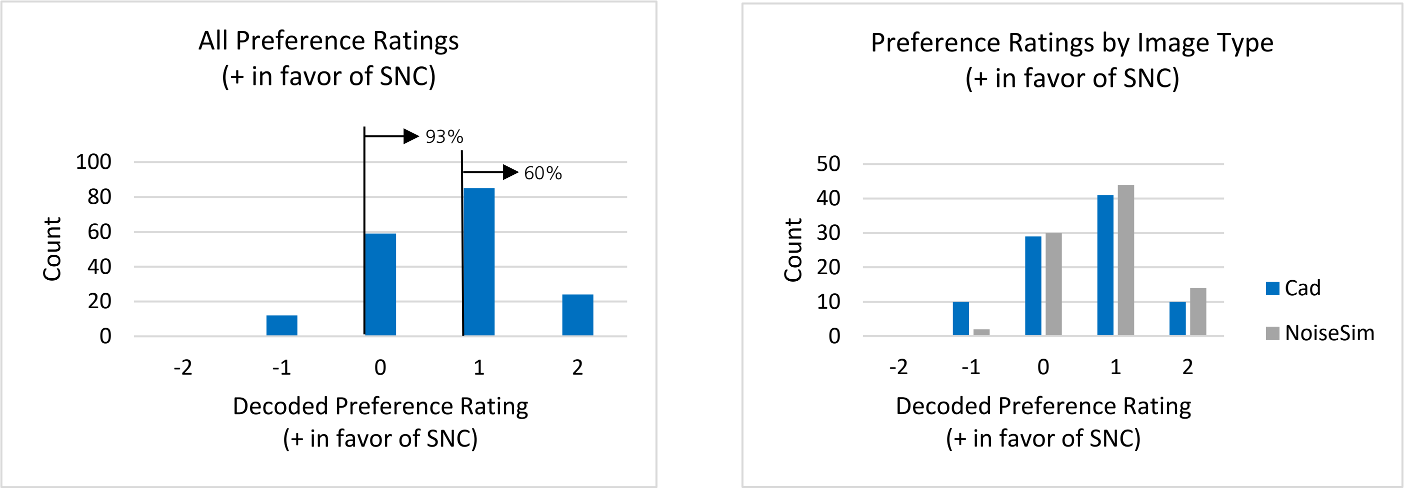
The distribution of both the nominal and reduced-dose IEC-exposure indices is shown in Figure 15 on the next page. Figure 15a demonstrates a bimodal distribution of the nominal exposures, where the two populations correspond to the two types of scintillators used in the study, GOS and CsI. The lower (left) peak corresponds to the exposure indices of CsI acquisitions and the upper (right) peak corresponds to GOS acquisitions. Similarly, Figure 15b shows the distribution of the reduced-dose IEC-exposure index values used in the study.



a

b

Figure 15. Distribution of nominal and low-dose IEC-exposure indices.



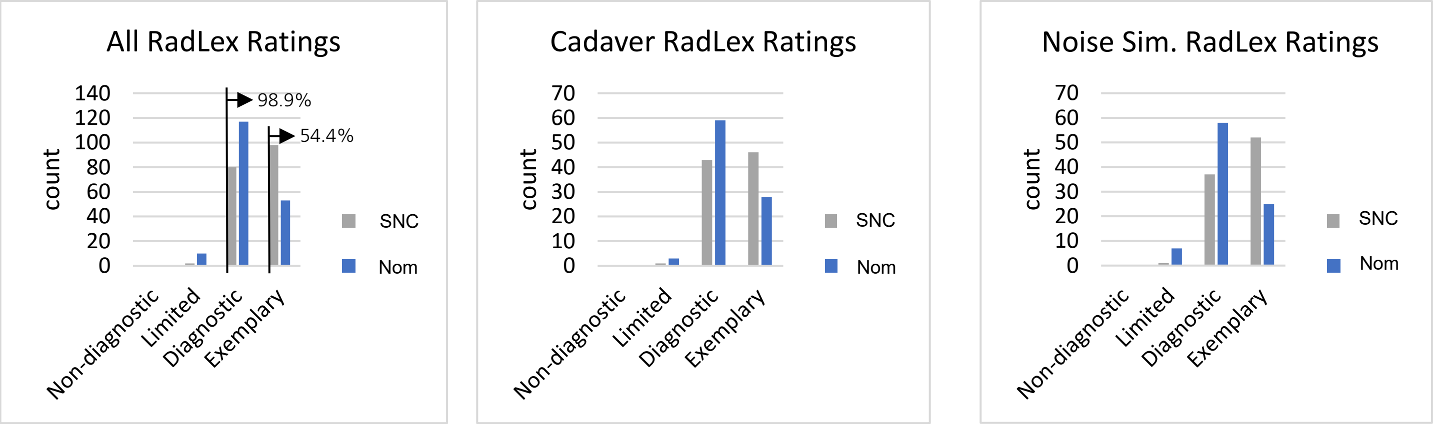
a

b

Figure 16. Distribution of Preference ratings: a. combined image types; b. by image type.

The distribution of Preference ratings is shown in Figure 16. Figure 16a demonstrates the overall distribution of the Preference ratings, whereas Figure 16b shows the Preference ratings broken down by image type (cadaveric or live subjects, the latter of which adds simulated noise to produce the reduced-dose image). Of the Preference ratings, 93% reflected no preference or a preference in favor of the reduced-dose images processed with SNC. Slight to strong preference in favor of the reduced-dose images processed with SNC and EVP Plus was observed in 60% of the image pairs. This evidence suggests an overall preference for the reduced-dose images processed with SNC and EVP Plus (see Table 11). Preference ratings by image type (Figure 16b) support the observation that the preference response is similar between cadaver image types and clinical images with simulated noise added.

Figure 17 on the next page illustrates the frequency distribution of RadLex ratings, overall (17a) and broken apart by image type (17b and 17c). No instance of images rated as Non-Diagnostic occurred, while greater than 98% of the reduced-dose SNC ratings were “Diagnostic” (3) or higher, and greater than 54% were rated “Exemplary” (4). Most nominal-dose images with EVP Plus were rated “Diagnostic” (3), whereas most reduced-dose SNC ratings were rated “Exemplary” (4). When the distributions are examined by image type, similar trends are demonstrated when comparing Figures 17b and 17c.



a *b* *c*

Figure 17. Distribution of RadLex ratings: a. combined image types; b. by cadaver image type; c. by noise simulation image type.

		Nominal-Dose RadLex Ratings				
	Counts % of Row	**1 Non-Diagnostic**	**2 Limited**	**3 Diagnostic**	**4 Exemplary**	Total
SNC Reduced-Dose RadLex Ratings	**1 Non-Diagnostic**	0 0.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%
	2 Limited	0 0.00%	2 100.00%	0 0.00%	0 0.00%	2 100.00%
	3 Diagnostic	0 0.00%	7 8.75%	71 88.75%	2 2.5%	80 100.00%
	4 Exemplary	0 0.00%	1 1.02%	46 46.94%	51 52.04%	98 100.00%
	Total	0 0.00%	10 5.56%	117 65.00%	53 29.44%	180 100.00%

Table 9. RadLex paired-comparison contingency table (both image types combined).

Table 9 is a paired-comparison contingency table of RadLex ratings, both image types combined. The contingency table facilitates pairwise comparison of reduced-dose images with SNC RadLex ratings versus nominal-dose images with EVP Plus alone.

For 99% (178/180) of the paired ratings for both image types combined, the reduced-dose images with SNC were rated the same or higher than the nominal-dose images with EVP Plus alone, and 30% of the ratings (54/180, green shaded boxes) were instances of the reduced-dose images with SNC rated more highly than the nominal-dose images, suggesting that the reduced-dose images with SNC are superior to the nominal-dose images.

Preference-rating analysis of variance (single-factor ANOVA testing) was performed to identify significant factors in the study data and is summarized in Table 10. Only one factor, the readers of the study, was found to be significant (p = 0.000). Image type (p = 0.073), detector type (p = 0.256) and dose-reduction level (p = 0.075) were determined to be insignificant.

Table 11 summarizes the findings from two hypothesis tests designed to determine whether favor was demonstrated for the reduced-dose images with SNC compared to nominal-dose images without SNC. The paired t-test was used to assess the RadLex ratings, whereas the one-sample t-test was used to assess the Preference ratings.

Factor	Number of Variables	Hypothesis Statement Regarding Preference	F-test Statistic	p-value	Statistical Significance
Readers	3	Ho: All means are equal. Ha: Not all means are equal.	9.39	0.000*	p-value < alpha 0.05; **Reject Ho**. Conclude that not all population means are equal.
Image Type	2	Ho: All means are equal. Ha: Not all means are equal.	3.26	0.073	p-value > alpha 0.05; **Fail to reject Ho**. Not enough evidence to conclude that not all population means are equal.
Detector Type	5	Ho: All means are equal. Ha: Not all means are equal.	1.34	0.256	p-value > alpha 0.05; **Fail to reject Ho**. Not enough evidence to conclude that not all population means are equal.
Dose Reduction	6	Ho: All means are equal. Ha: Not all means are equal.	2.04	0.075	p-value > alpha 0.05; **Fail to reject Ho**. Not enough evidence to conclude that not all population means are equal.

*Significant

Table 10. Preference-rating single-factor ANOVA results.

Rating Factor	Hypothesis Statement	Test	mean	t-statistic	p-value	Statistical Significance
RadLex Difference (SNC Reduced-Dose – Nom. Dose)	Ho: mean diff = 0 Ha: mean diff > 0	Paired t-test	0.29	8.02	0.000*	p-value < alpha 0.05: **Reject Ho**…mean difference is greater than 0.
Preference	Ho: mean <= 0.5 Ha: mean > 0.5	One-sample t-test	0.67	2.93	0.002*	p-value < alpha 0.05; **Reject Ho**…mean is greater than 0.5

*Significant

Table 11. Hypothesis testing using t-statistic.

The paired t-test result indicates that the mean RadLex difference is greater than 0. The 95% confidence intervals for this difference are (0.22, 0.37), so we conclude that the reduced-dose images with SNC yield RadLex ratings greater than the nominal-dose image processed with EVP Plus alone. The one-sample t-test result for the mean Preference ratings indicates that the mean Preference is greater than 0.5 (a substantial difference). The 95% confidence interval for the Preference rating is (0.56, 0.79), so we conclude that the reduced-dose images with SNC are substantially more preferred over the nominal dose images with EVP Plus alone.

Safety and effectiveness of using SNC at reduced-dose was demonstrated with a one-sample t-test, to determine whether the diagnostic quality of the average reduced-dose ratings with SNC is greater than Limited (2). Table 12 summarizes this test.

Rating Factor	Hypothesis Statement	Test	mean	t-statistic	p-value	Statistical Significance
SNC Reduced-Dose RadLex	Ho: mean <= 2.0 Ha: mean > 2.0	One-sample t-test	3.53	39.40	0.000*	p-value < alpha 0.05; **Reject Ho**…mean is greater than 2.0

*Significant

Table 12. One-sample t-test for safety and effectiveness of using SNC at reduced-dose.

The mean diagnostic quality rating of 3.53 is significantly greater than the test level of 2 (“Limited”) and supports the conclusion that using SNC on reduced-dose images is both safe and effective.

Table 13 on the next page summarizes the dose-reduction levels realized in this reader study. CsI-based acquisitions at up to 800 ISO speed dose reduction on all exam types (all body parts, projections and patient sizes) demonstrated superior image quality with SNC when compared to their corresponding nominal ISO 400 speed acquisitions without SNC. Likewise, GOS-based acquisitions at up to 500 ISO speed dose reduction on all exam types demonstrated superior image quality with SNC when compared to their corresponding nominal ISO 320 speed acquisitions.

Scintillator Type	Exam	Nominal ISO Speed	Reduced-Dose ISO Speed
Cesium Iodide, CsI	All	400	800
Gadolinium Oxysulfide, GOS	All	320	500

Table 13. Summary of dose-reduction levels.

Reduced-dose image type (cadaver vs. live subjects with simulated noise) was not a significant factor in the study. Reduced-dose images with SNC were clearly preferred (mean preference of 0.67), and 60% of the Preference ratings were slightly to strongly in favor of the reduced-dose images with SNC. The median RadLex rating for reduced-dose images processed with SNC was "Exemplary" (4) and more than 98% of these images were rated as "Diagnostic" (3) or "Exemplary" (4).

5. Customized Noise Reduction

Objective measurements and subjective ratings demonstrate that SNC processing can reduce noise while simultaneously retaining fine spatial detail. The objective measurements and reduced-dose study present reasonable evidence that meaningful dose reduction is possible. But because the desired level of noise is subjective (e.g. some radiologists expect to see a certain degree of noise in images, which assures them that the patient was not over-exposed) and its impact can be substantial, Carestream has enabled users to select their preferred level of noise reduction. The "Noise-Adjustment Level" parameter is available to the user via the Image Processing Preference Editor and enables the key operator to set the amount of noise that is removed, from 100% (the full noise field) to 50% (half magnitude of the noise field). SNC processing is available with Carestream's ImageView software.

6. Conclusion

Images processed with EVP Plus and SNC demonstrate significant improvements in image quality and provide a level of clarity never before achieved in projection radiography. Objective testing demonstrates that SNC processing enables a 2x to 4x noise reduction in uniform areas, preserves high-frequency sharpness and improves contrast detail. Subjective evaluation of images from five detector types, a wide range of exams and a wide range of exposure levels corroborate these results. Furthermore, the dose-reduction study demonstrates that SNC used at reduced-dose (e.g. up to 800 ISO speed for CsI and up to 500 ISO speed for GOS panels) yields image quality as good or better than images exposed at nominal-dose levels (e.g. 400 speed for CsI and 320 speed for GOS panels) without SNC. With the preservation of fine detail and the removal of noise, image clarity has reached an all-new level in digital radiography – all to the benefit of care providers and their patients.

[1] Smart Noise Cancellation Technical Paper

[2] EVP (Enhanced Visualization Processing) Plus is the Image-Processing Rendering arm of Carestream's Eclipse Image Processing Engine.

[3] https://www.carestream.com/blog/2020/04/21/understanding-and-managing-noise-sources-in-x-ray-imaging/

[4] Hu Chen, Yi Zhang, "Low-Dose CT with a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN)", 2017, https://arxiv.org/ftp/arxiv/papers/1702/1702.00288.pdf

[5] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, Kaori Togashi, "Convolutional Neural Networks: An Overview and Application in Radiology," 2018, https://doi.org/10.1007/s13244-018-0639-9

[6] Olaf Ronneberger, Philipp Fischer, Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015, https://arxiv.org/pdf/1505.04597.pdf

[7] US Patent 7480,365 B1 (K. Topfer, J. Ellinwood, "Dose Reduced Digital Medical Image Simulations")

[8] E. Samei, M. J. Flynn, and W. R. Eyler, "Simulation of Subtle Lung Nodules in Projection Chest Radiography," Radiology, 202, 117-124 (1997)

[9] Smart Noise Cancellation Technical Paper

[10] M.A.O. Thijssen, K.R. Bijkerk, R.J.M. van der Burght, "Manual Contrast-Detail Phantom CDRAD type 2.0," University Hospital Nijmegen, St. Radboud, 1998

[11] Radiological Society of North America (2010) RadLex: A Lexicon for Uniform Indexing and Retrieval of Radiology Information Resources. Available via http://www.rsna.org/radlex/. Accessed 22 Oct 2010.

Carestream

"Rx only"

CAT 2000 289 03/22