Image generation procedure

1) Noise generation procedure

To create images with varying levels of dosage, realistic noise was added to the 2016 Low-dose CT Grand Challenge (1) and National Cancer Center (NCC), South Korea datasets. Note that N_0 represents the average number of photons emitted during a time interval. The average number of photons transmitted is a non-linear function of the line integral of the attenuation coefficient ($N_d = N_0 \exp(-\int \mu ds)$), where μ and s are the linear attenuation coefficient and length element along the photon path, respectively (2). To compute the line integral in detector bin *i*, the following formula is used, where $\kappa = Poisson(N_d)$ represents the number of photons transmitted at bin *i*, and it follows a Poisson distribution with an average value of N_d.

$$p_i = \ln \frac{N_0}{\kappa} = lnN_0 - ln\kappa$$

We used a precise physics-based method for generating noise (3, 4) and inserted Poisson noise into the projection data. This allowed us to obtain two additional noise levels equivalent to 50%, and 10% of the full-dose for the 2016 Low-dose CT Grand Challenge dataset, as well as three additional noise levels equivalent to 50%, 25%, and 10% of the full-dose for the NCC dataset. To acquire the projection data corresponding to the line integral, we forward-projected full-dose CT images, resulting in 360˚ of projection data with the gantry rotated over 2π.

$$\tilde{n}_{2D} = Poisson(N_d)$$

Afterwards, the filtered back-projection algorithm was utilized to create the 3D image with noise, denoted as $\tilde{n}_{3D}$ from the 2D noisy image, $\tilde{n}_{2D}$. This was accomplished using the back-projector, denoted as A^T, and the reconstruction kernel, denoted as h.

$$\tilde{n}_{3D} = A^T h(\tilde{n}_{2D})$$

In the end, the reconstructed $\tilde{n}_{3D}$ images were transformed into Hounsfield units (HU) through linear scaling, with the mean pixel values of the fat and air regions used as reference points. Then, the values were normalized to a range of 0 to 1. Upon adding the supplementary data to the initial dataset, we obtained a new dataset consisting of four distinct dose levels. These dose levels correspond to 100%, 50%, 25%, and 10% of the full-dose.

2) Streak generation procedure

Streak artifacts are generated using a similar pipeline to noise insertion, but by reconstructing $\tilde{n}_{3D}$ with different parameters of the number of projections in use (projection stack size) and angular increment per rotation, rather than by inserting noise into the projection data.

Streak level	Projection stack size	Angular increment
1	720	0.5
2	360	1
3	180	2

References

(1) Acknowledge: NIH grants EB017095 and EB017185 (Cynthia McCollough, PI) from the National Institute of Biomedical Imaging and Bioengineering, and American Association of Physicists in Medicine, Low Dose CT Grand Challenge Dataset, https://app.box.com/s/eaw4jddb53keg1bptavvvd1sf4x3pe9h, (related paper: Moen, T. R., Chen, B., Holmes, D. R., III, Duan, X., Yu, Z., Yu, L., Leng, S., Fletcher, J. G., & McCollough, C. H. (2020). Low dose CT image and projection dataset. Medical Physics. https://doi.org/10.1002/mp.14594)

(2) Macovski A 1983 Medical Imaging Systems (Hoboken, NJ: Prentice–Hall)

(3) Gholizadeh-Ansari M, Alirezaie J and Babyn P 2020 J. Digit. imaging 33 504–15

(4) Kim B, Shim H and Baek J 2021 Med. Image Anal. 71 102065