Evaluation of Synthetically Generated CT for use in Transcranial Focused Ultrasound Procedures

Han Liu[a,†], **Michelle K. Sigona**[b,c,†], **Thomas J. Manuel**[b,c], **Li Min Chen**[c,d], **Benoit M. Dawant**[e], **Charles F. Caskey**[b,c,d,*]

[a]Dept. of Computer Science, Vanderbilt University, Nashville, TN 37235, USA
[b]Dept. of Biomedical Engineering, Vanderbilt University, Nashville, TN 37235, USA
[c]Institute of Imaging Science, Vanderbilt University, Nashville, TN 37235, USA
[d]Dept. of Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN 37235, USA
[e]Dept. of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN 37235, USA
[†]These authors contributed equally to this work

1 SUPPLEMENTARY MATERIAL

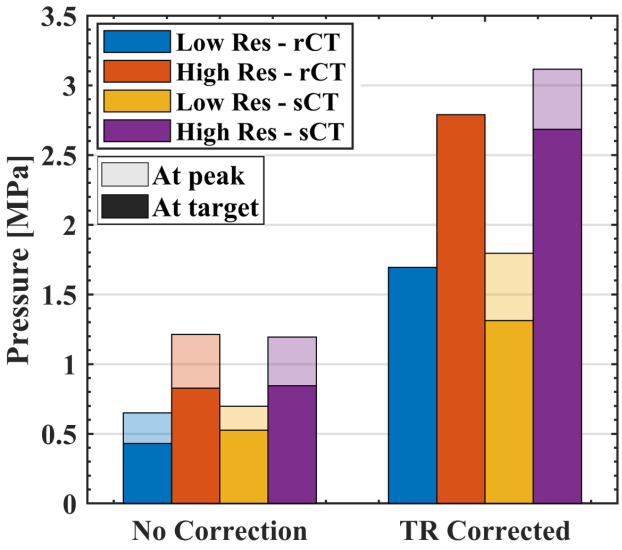

Fig S1 Comparison of intracranial pressure at different resolutions for one test case. A single case was simulated with no phase correction and phases from time-reversal at a resolution greater than 7 PPW (High Res) and compared with simulations performed at 4.3 PPW maintained in our study (Low Res). Averaging all simulations in this test case (N=4), peak intracranial pressure (transparent bars) was $42.4 \pm 2.5\%$ greater for high resolution simulations than low resolution simulations, and $43.9 \pm 5.9\%$ for the same comparison with pressure at the intended target (solid bars). rCT: real CT, sCT: synthetic CT, TR: time-reversal

Table S1 Peak intracranial pressure is summarized from high and low resolution simulations. The percent difference is calculated for different comparison groups (rCT vs sCT, high res vs low res).

	No Correction			TR Corrected		
	High Resolution [MPa]	**Low Resolution [MPa]**	**% Difference**	**High Resolution [MPa]**	**Low Resolution [MPa]**	**% Difference**
rCT	1.21	0.65	46.3%	2.79	1.69	39.4%
sCT	1.19	0.70	41.2%	3.12	1.79	42.6%
% Difference	1.6%	7.1%		11.7%	6.0%	

Table S2 Differences in focal position of the peak intracranial pressure and focal volume between rCT and sCT is presented for high and low resolution simulations.

	No Correction		TR Corrected	
	Focal Position [mm]	**Focal Volume [%]**	**Focal Position [mm]**	**Focal Volume [%]**
High Resolution	2.9	13.2%	0.3	1.2%
Low Resolution	0	12.0%	0.5	1.6%

1.1 5-Fold Cross Validation

To further estimate the skill of our sCT generator with the data available, we performed a 5-fold cross validation by splitting the entire dataset into training/validation/testing of 66, 10, and 10 samples in five configurations. The 5-fold cross validation analyses are performed to evaluate the image similarity between synthetic and real CT skulls and compare skull metrics from Kranion. Acoustic simulations are not included due to the constraints of the computational resources and simulation time.

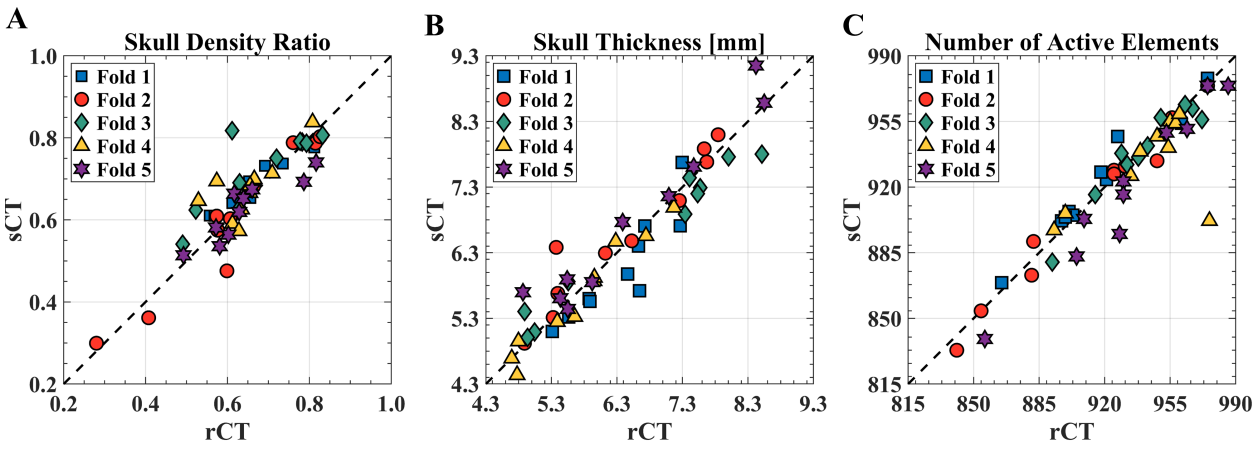

Fig S2 Strong correlation is observed between rCT and sCT for all three skull metrics when plotting skulls generated from each fold. Fold 1 is the original 10 testing skulls used in the paper. The line of unity is indicated by the black dashed line. The skull density ratio (A) and skull thickness (B) demonstrated a larger range of skull values were captured in certain folds such as skulls in fold 2. A test case in Fold 4 demonstrated low correlation for the number of active elements (C)

Table S3 Mean ± standard deviation for skull density ratio, skull thickness and number of active elements are reported for the real and synthetic skulls of each fold. The relationship between rCT and sCT is indicated by Pearson's correlation coefficient, *r*.

Fold #	Skull Density Ratio			Skull Thickness [mm]			Number of Active Elements		
	rCT	sCT	*r*	rCT	sCT	*r*	rCT	sCT	*r*
1	0.65 ± 0.08	0.67 ± 0.06	0.96	6.4 ± 0.7	6.1 ± 0.8	0.90	917 ± 32	922 ± 31	0.98
2	0.61 ± 0.17	0.60 ± 0.18	0.96	6.4 ± 1.1	6.6 ± 1.1	0.96	912 ± 45	910 ± 46	0.99
3	0.68 ± 0.12	0.72 ± 0.10	0.82	6.7 ± 1.4	6.6 ± 1.1	0.98	940 ± 25	938 ± 25	0.95
4	0.65 ± 0.08	0.67 ± 0.07	0.72	5.7 ± 0.8	5.7 ± 0.9	0.98	942 ± 27	933 ± 23	0.57
5	0.64 ± 0.10	0.62 ± 0.07	0.89	6.5 ± 1.3	6.8 ± 1.3	0.97	934 ± 38	921 ± 43	0.98

The skull metrics from 10 testing CTs targeting the left ventral intermediate nucleus of the thalamus for each fold are plotted in Figure S2A-C. With the addition of 40 new skulls (10 from each fold) we observed even with a larger range of skull metrics (skull density ratio, rCT = 0.28-0.83 and skull thickness rCT = 4.7-8.5) we found strong positive correlation between rCT and sCT. Table S3 contains the mean skull density ratio, skull thickness, and number of active elements for each fold, demonstrating good agreement between rCT and sCT of each metric, further indicated by the Pearson's correlation coefficient greater than 0.5 for all paired comparisons between rCT

and sCT. However, because some folds performed worse than our original testing cases (fold 1), this indicates larger number of testing cases are required or larger number of training pairs should be used to train the network.