Clinical Validation of a Deep Learning Model for Segmenting and Quantifying Intracranial and Ventricular Volumes on Computed Tomography

Bruna Garbes Gonçalves Pinto[1*†], Tayran Milá Mendes Olegário[1†], Pedro Vinicius Alves Silva[1], Gabriel Monteiro Ferracioli[1], Artur José Marques Paulo[1], Klaus Schumacher[1], Mateus Trinconi Cunha[1, 2], Henrique Min Ho Lee[1], Mariana Athaniel Silva Rodrigues[1], Felipe Campos Kitamura[3], Joselisa Peres Queiroz de Paiva[1], Rafael Maffei Loureiro[1]

[1*]Image Department, Hospital Israelita Albert Einstein, São Paulo, Brazil.
[2]Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Canada.
[3*]Departament of Diagnostic Imaging, Universidade Federal de São Paulo, São Paulo, Brazil.

*Corresponding author(s). E-mail(s): bruna.goncalves@einstein.br;
†These authors contributed equally to this work.

1 Supplementary Material

The stages of generating the pre-segmentation masks of the intracranial volume include:

1. **Image size verification:** If the image has less than 500 slices, it is resized to a new image with 500 slices.
2. **Application of Gaussian filter:** A Gaussian filter with a parameter $\sigma = 1$ is applied to the image to reduce noise, uniformizing some image regions and thus facilitating the segmentation process.
3. **Thresholding application:** A mask is generated by applying a threshold value to the image. The mask includes all voxels with values between -15 and 50 Hounsfield units.

4. **Erosion operation application:** An erosion operation is applied to the resulting mask from the previous step to ”erode” the sides of the mask and thus disconnect the various structures present in the mask.
5. **Selection of the largest connected region:** then the largest connected region of the mask is selected. This largest connected region is expected to correspond to part of the intracranial volume region.
6. **Dilation operation application:** A dilation operation is then applied to the largest connected region to ”compensate” for previously applied erosion operation. It is expected that after applying this operation, only the intracranial volume region will be obtained. This dilation will only affect voxels with values between -50HU and 150HU.
7. **Cutting the bottom region of the mask:** Finally, a heuristic is applied that traverses the image from bottom to top and calculates the number of connected regions in the axial slices. If only one connected region appears in the slice, it is excluded; the heuristic is terminated if more than one connected region appears.

Table 1 Kruskal-Wallis and Effect Sizes for Volumes by Patient Sex

	P-Value	η^2	Magnitude
Intracranial Volume - Ground truth	<0.001	0.38	Large
Intracranial Volume - DeepCTE3D	<0.001	0.36	Large
Ventricular Volume - Ground truth	<0.001	0.02	Small
Ventricular Volume - DeepCTE3D	0.02	0.01	Small

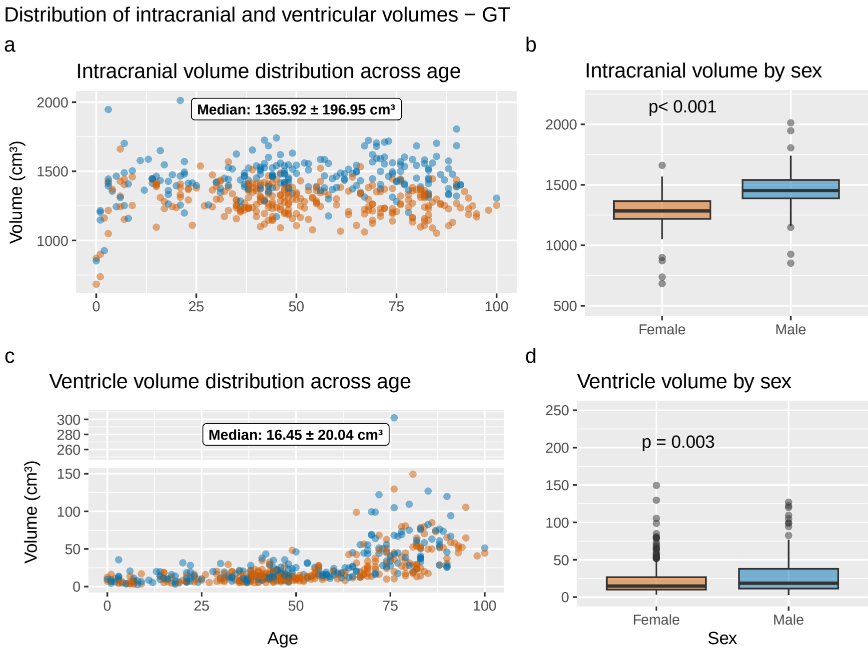

Fig. 1 Volume distributions of ground-truth segmentations. a) Intracranial volume by patient age. b) Intracranial volume by patient sex. c) Ventricular volume by patient age. d) Ventricular volumes by patient sex.

Table 2 Kruskal-Wallis and Effect Sizes for Volumes by Equipment

	P-Value	η^2	Magnitude
Intracranial Volume - GT	0.20	< 0.01	Small
Intracranial Volume - DeepCTE3D	0.25	< 0.01	Small
Ventricular Volume - GT	< 0.001	0.03	Small
Ventricular Volume - DeepCTE3D	< 0.001	0.05	Small

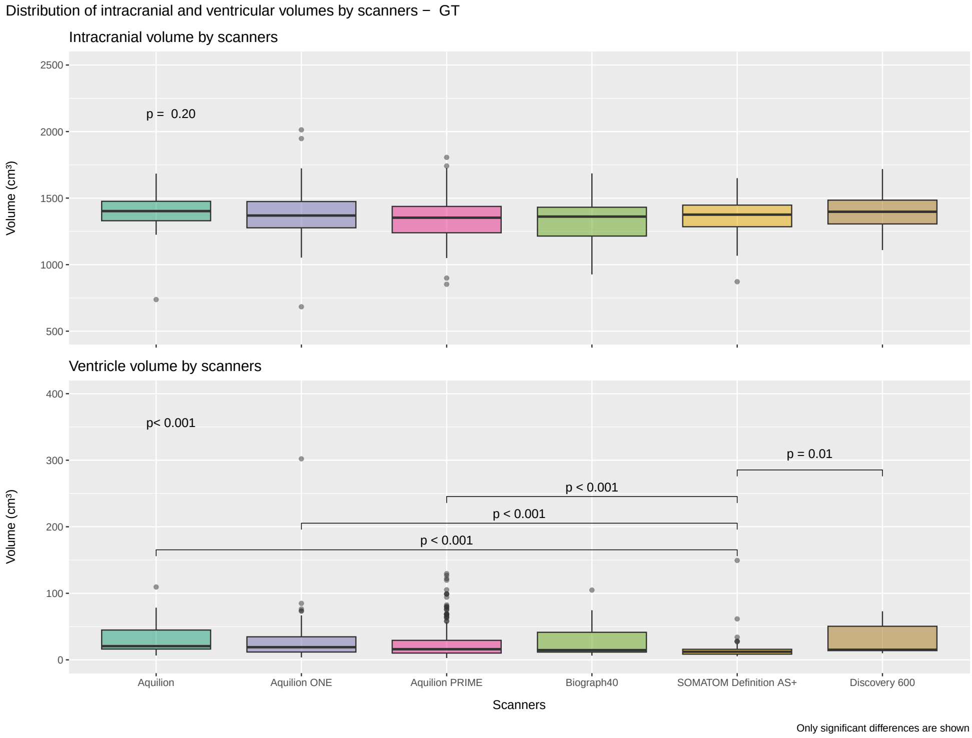

Fig. 2 Intracranial and ventricular volumes by scanner model - Ground Truth.

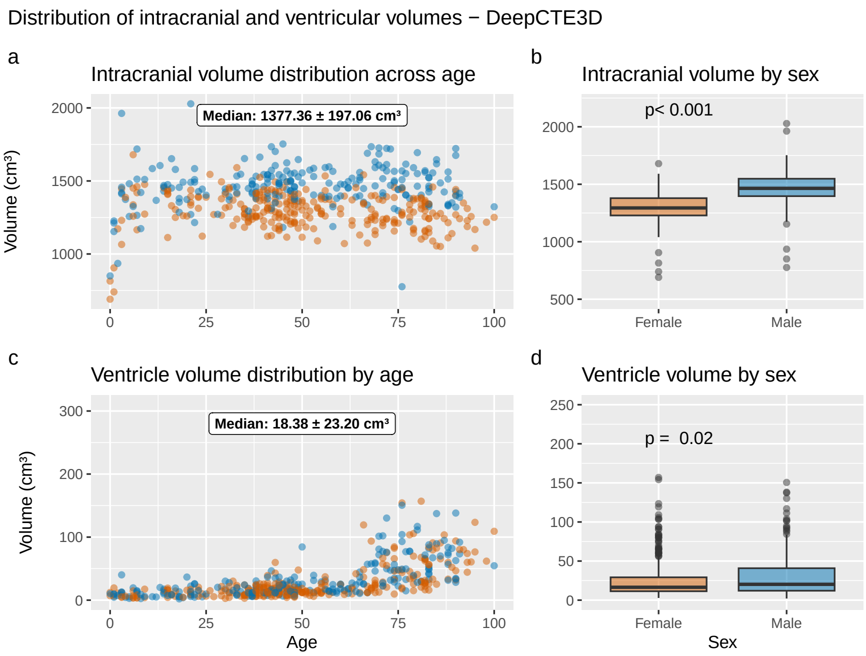

Fig. 3 Volume distributions of DeepCTE3D segmentations. a) Intracranial volume by patient age. b) Intracranial volume by patient sex. c) Ventricular volume by patient age. d) Ventricular volume by patient sex.

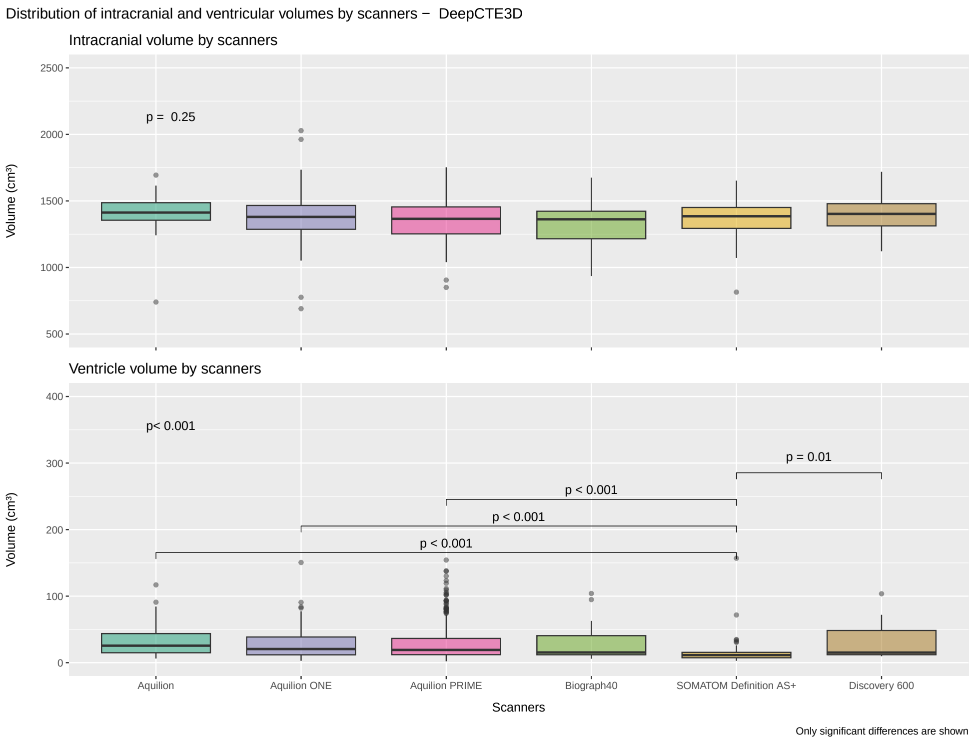

Fig. 4 Volumes by scanner model - DeepCTE3D.

Table 3 Kruskal-Wallis and Effect Sizes for Relative Differences by Patient Sex and Scanners

Structure	Group	P-Value	η^2	Magnitude
Intracranial volumes	Patient Sex	<0.001	0.03	Small
Intracranial volumes	Scanner	<0.001	0.12	Moderate
Lateral ventricles volumes	Patient Sex	0.03	0.02	Small
Lateral ventricle volumes	Scanner	<0.001	0.23	Large

Table 4 Effect Sizes for Intracranial Relative Differences Between Scanner Type Pairwise Comparisons

Scanner A	Scanner B	P-Value	r	Magnitude
Aquilion	Biograph40	<0.001	0.49	Moderate
Aquilion	SOMATOM Definition AS+	<0.001	0.35	Moderate
Aquilion	Discovery 600	<0.001	0.55	Large
Aquilion ONE	Biograph40	<0.001	0.35	Moderate
Aquilion ONE	SOMATOM Definition AS+	<0.001	0.37	Moderate
Aquilion ONE	Discovery 600	<0.001	0.33	Moderate
Aquilion PRIME	Biograph40	<0.001	0.28	Small
Aquilion PRIME	SOMATOM Definition AS+	<0.001	0.31	Moderate
Aquilion PRIME	Discovery 600	<0.001	0.27	Small
SOMATOM Definition AS+	Discovery 600	0.049	0.24	Small

Table 5 Effect Sizes for Lateral Ventricle Relative Differences Between Scanner Type Pairwise Comparisons.

Scanner A	Scanner B	P-Value	r	Magnitude
Aquilion	Aquilion PRIME	0.02	0.15	Small
Aquilion	SOMATOM Definition AS+	<0.001	0.44	Moderate
Aquilion ONE	Aquilion PRIME	<0.001	0.41	Moderate
Aquilion ONE	SOMATOM Definition AS+	<0.001	0.29	Small
Aquilion PRIME	Biograph40	<0.001	0.23	Small
Aquilion PRIME	SOMATOM Definition AS+	<0.001	0.50	Large
Aquilion PRIME	Discovery 600	= 0.002	0.19	Small
SOMATOM Definition AS+	Discovery 600	0.04	0.26	Small

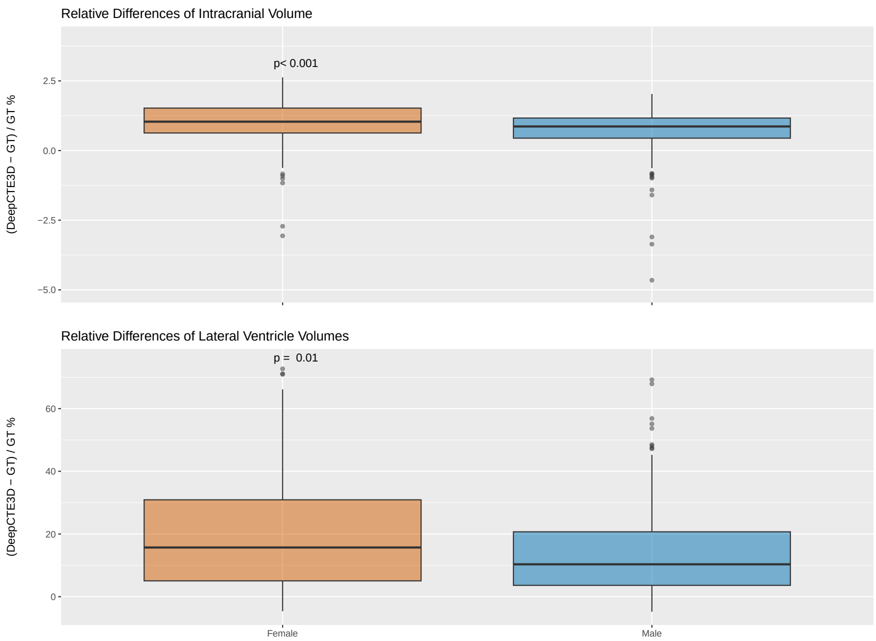

Fig. 5 Relative differences by patient sex a) Relative differences for intracranial volume b) Relative Difference for lateral ventricle volumes.

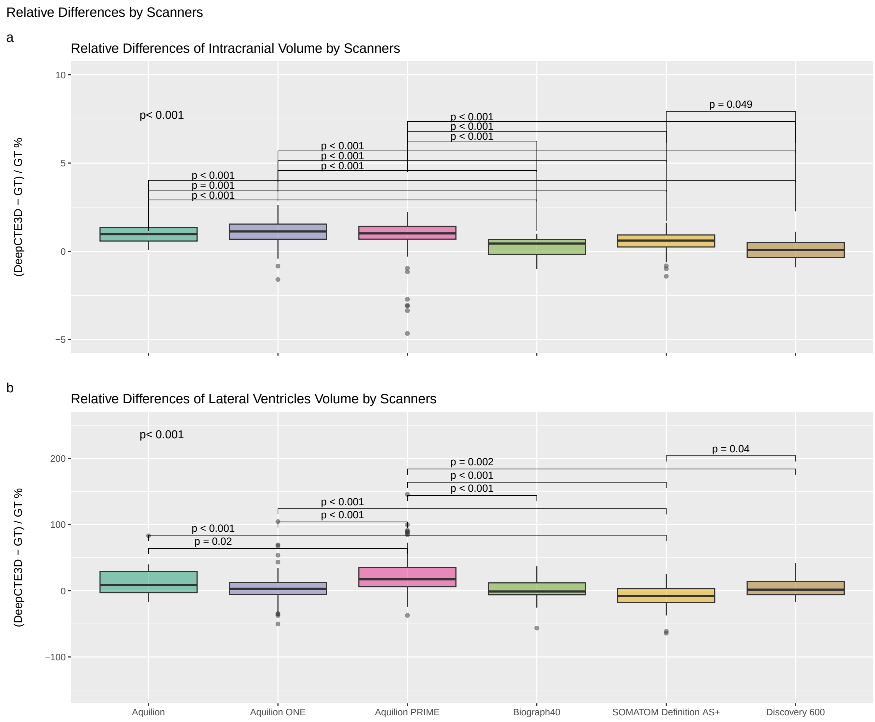

Fig. 6 Relative differences by scanners a) Relative differences for intracranial volume b) Relative Difference for lateral ventricle volumes.

Table 6 Kruskal-Wallis and Effect Sizes for Metrics by Patient Sex

	P-Value	η^2	Magnitude
Intracranial Dice	0.03	0.01	Small
Intracranial Hausdorff	0.15	<0.01	Small
Ventricular Dice	<0.001	0.02	Small
Ventricular Hausdorff	0.82	<0.01	Small

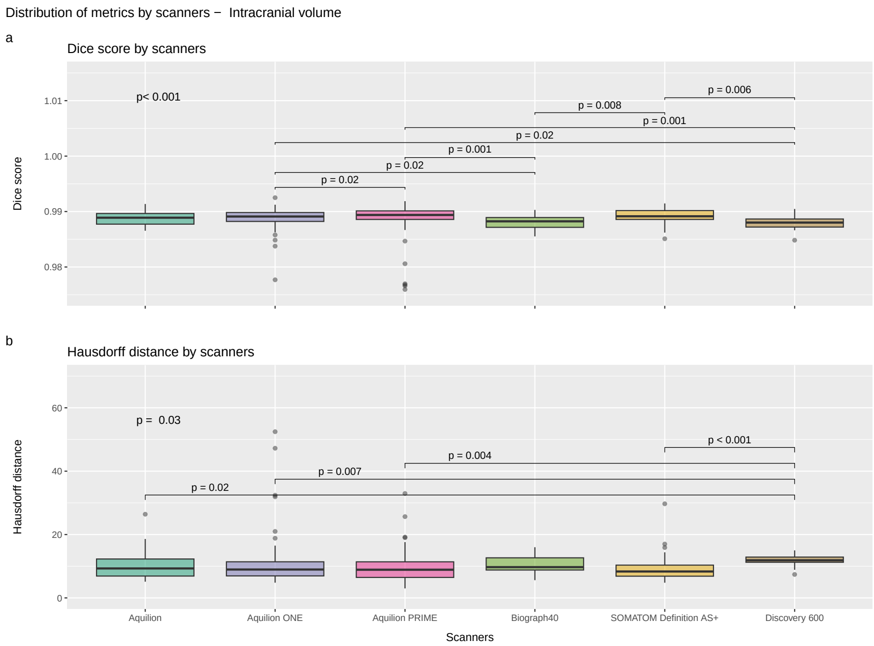

Fig. 7 Dice and Hausdorff distributions of intracranial segmentation across scanner models. a) Dice Score across models. b) Hausdorff Distance across models

Table 7 Kruskal-Wallis and Effect Sizes for Metrics by Equipment

	P-Value	η^2	Magnitude
Intracranial Dice	< 0.001	0.04	Small
Intracranial Hausdorff	0.03	0.02	Small
Ventricular Dice	<0.001	0.04	Small
Ventricular Hausdorff	0.51	<0.01	Small

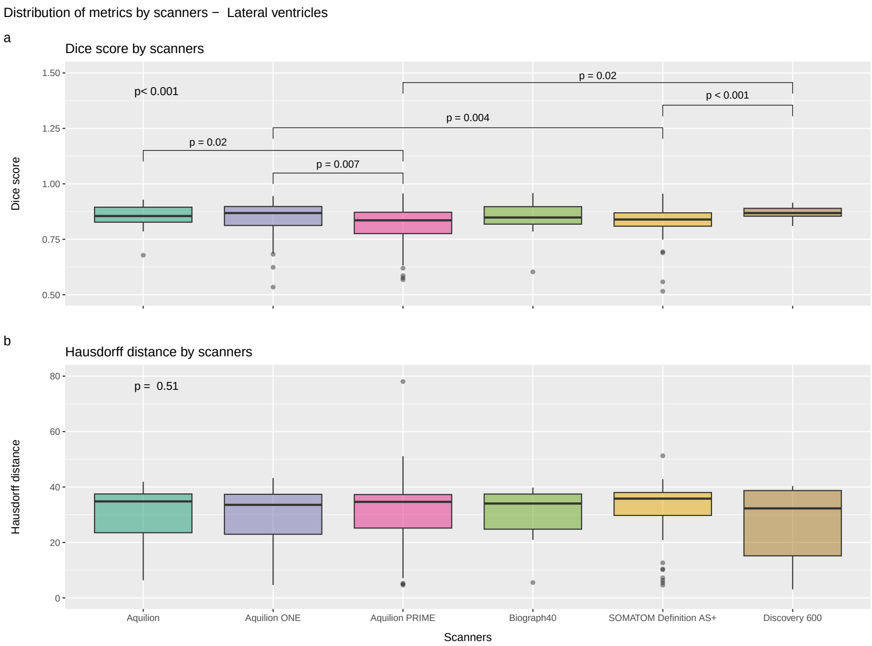

Fig. 8 Dice and Hausdorff distributions of ventricular segmentation across scanner models. a) Dice Score across equipment. b) Hausdorff Distance across scanner models

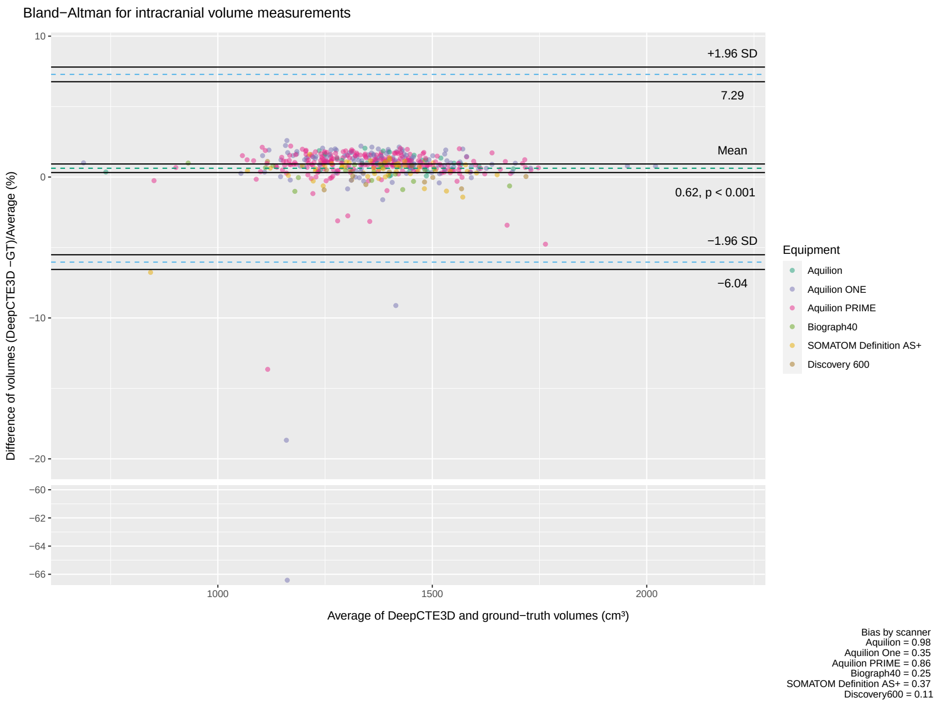

Fig. 9 Bland-Altman plot assessing agreement between DeepCTE3D-generated and ground-truth intracranial volumes colored by scanner type.

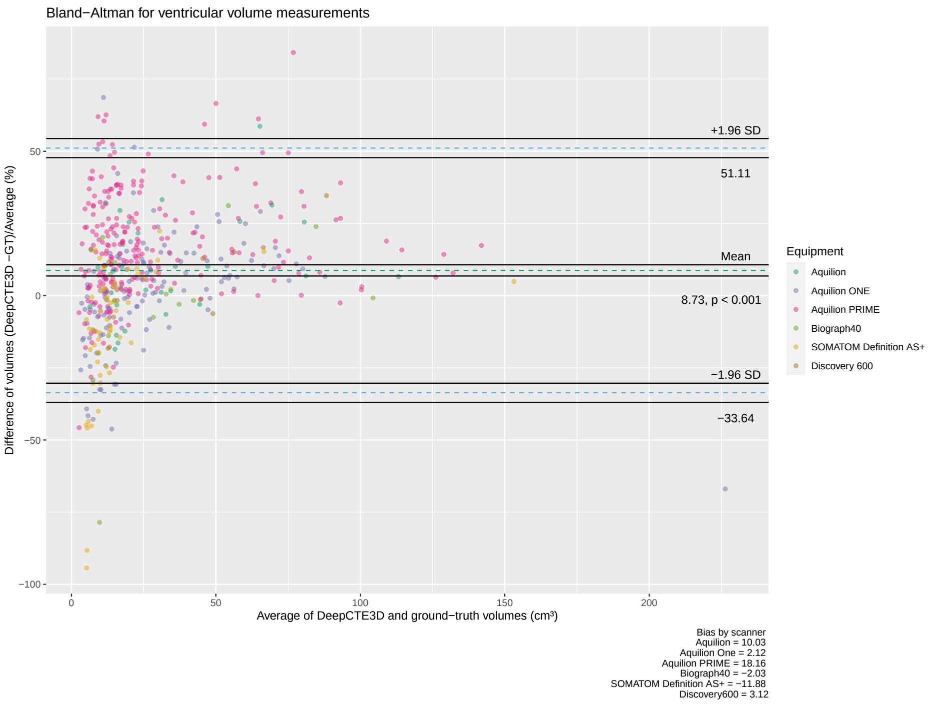

Fig. 10 Bland-Altman plot assessing agreement between DeepCTE3D-generated and ground-truth ventricular volumes colored by scanner type.

Table 8 Bias and Effect Size by Equipment for Intracranial Volumes.

Scanner A	Scanner B	P-Value	r	Magnitude
Aquilion	Biograph40	< 0.001	0.49	Moderate
Aquilion	SOMATOM Definition AS+	0.001	0.35	Moderate
Aquilion	Discovery 600	< 0.001	0.55	Large
Aquilion ONE	Biograph40	< 0.001	0.35	Moderate
Aquilion ONE	SOMATOM Definition AS+	< 0.001	0.37	Moderate
Aquilion ONE	Discovery 600	< 0.001	0.33	Moderate
Aquilion PRIME	Biograph40	< 0.001	0.28	Small
Aquilion PRIME	SOMATOM Definition AS+	< 0.001	0.31	Moderate
Aquilion PRIME	Discovery 600	< 0.001	0.27	Small
SOMATOM Definition AS+	Discovery 600	0.049	0.24	Small

Table 9 Bias and Effect Size by Equipment for Lateral Ventricles Volumes.

Scanner A	Scanner B	P-Value	r	Magnitude
Aquilion	Aquilion PRIME	0.02	0.15	Small
Aquilion	SOMATOM Definition AS+	< 0.001	0.44	Moderate
Aquilion ONE	Aquilion PRIME	< 0.001	0.41	Moderate
Aquilion ONE	SOMATOM Definition AS+	< 0.001	0.29	Small
Aquilion PRIME	Biograph40	< 0.001	0.23	Small
Aquilion PRIME	SOMATOM Definition AS+	< 0.001	0.50	Large
Aquilion PRIME	Discovery 600	0.002	0.19	Small
SOMATOM Definition AS+	Discovery 600	0.04	0.26	Small

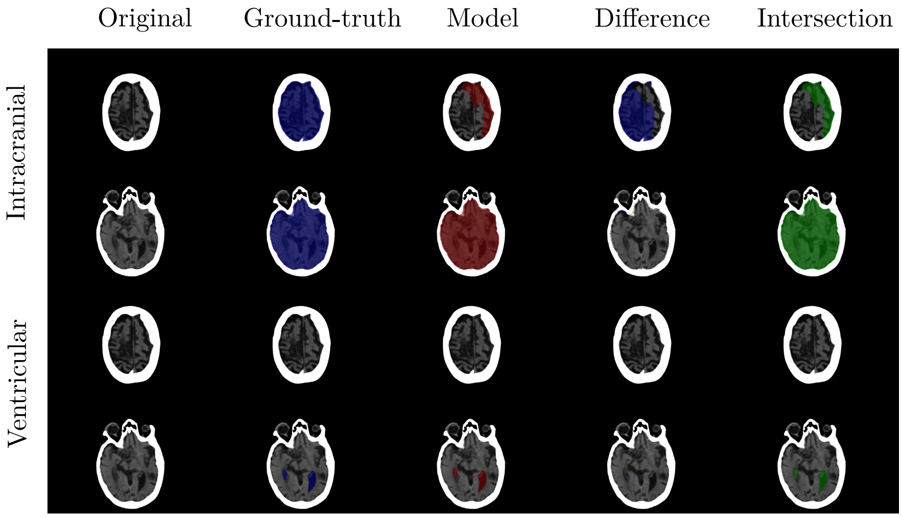

Fig. 11 Example 1 of poor model performance: This case exhibits an angulated visualization of the head, a large frontoparietal encephalomalacia on the right side, and asymmetric ventricular ectasia. Segmentation errors occur predominantly in the high convexity on the same side as the encephalomalacia but also extend into regions with normal brain parenchyma. In the ventricular areas, the model overflowed the boundaries of the lateral ventricles.

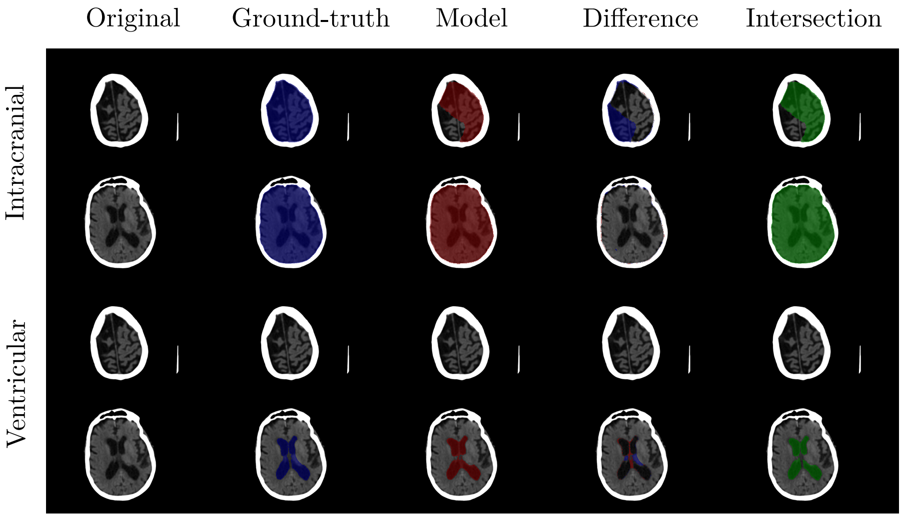

Fig. 12 Example 2 of poor model performance: This case exhibits a focal frontoparietal subarachnoid enlargement on the right side, moderate ventricular ectasia, and a mild degree of parenchymal atrophy. Segmentation errors occur predominantly in the high convexity on the same side as the enlargement but also extend into regions with normal brain parenchyma.

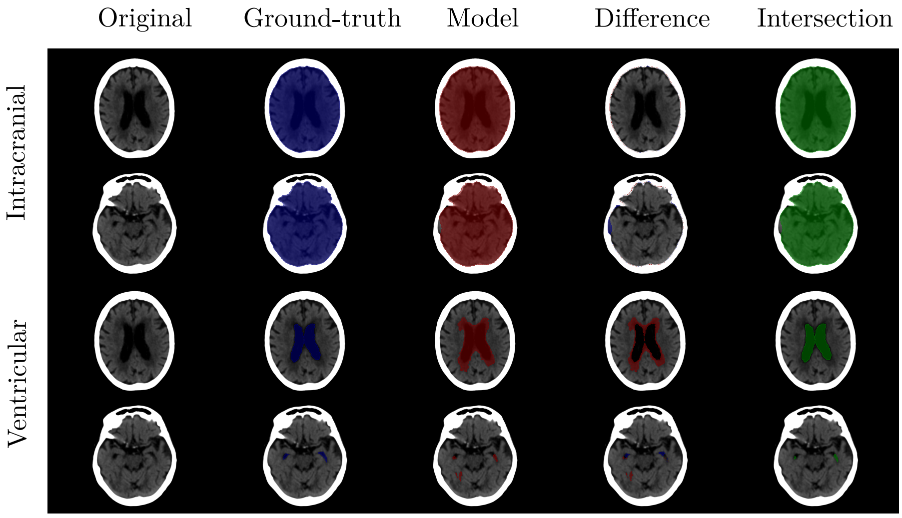

Fig. 13 Example 3 of poor model performance: This case exhibits a greater degree of white matter disease, a moderate degree of parenchymal atrophy with proportional ventricular ectasia, and a small thalamic lacune (not shown). Segmentation errors predominantly occur in the ventricular spaces, overflowing the boundaries, with some subarachnoid spaces also being misclassified (on the right side). The lacune was erroneously segmented as part of the ventricular space.