SIIM CAIMI26 Conference on Artificial Intelligence in Medical Imaging
OCT 26-27 | University of Pennsylvania, Philadelphia, PA

Automated Sagittal MRI Morphometry for Signs of Intracranial Pressure Dysregulation

Isaac Anthony, University of Utah; Troy Hutchins, MD; Miriam Peckham, MD; Kazem Hasemizadeh, PhD; Tyler Richards, MD

Briefly describe what attendees will see during your demo.
The sagittal T1-weighted sequence of a brain MRI is uploaded, and then the tool runs end-to-end. The tool automatically finds the midsagittal slice, segments the sellar and brainstem structures, and computes the Bern score components and related sagittal measurements. Attendees will see an interactive display: the segmentations and measurement overlays drawn on the image, a click-to-enlarge feature for any image, each measurement shown against its threshold alongside the non-contrast Bern score, and a validated-versus-in-development split so it's clear which measurements are tested against a board-certified neuroradiologist.

Problem & Motivation: What clinical or operational problem does your MVP address?
Signs of intracranial pressure dysregulation are often missed on routine brain MRI. Many of the features that flag them are quantitative (cistern heights and widths, distances, and angles), have low inter-reader reliability, and are rarely measured by hand. The radiologist often receives insufficient patient history, or the patient does not disclose symptoms suggestive of pressure dysregulation, so clinical suspicion is low, and these measurements are not typically performed; the classic morphologic imaging features of SIH, IIH, and NPH can be overlooked. Having these measurements automatically performed and reported when they are abnormal creates the potential for additional patients to get the care they need for treatable but potentially debilitating conditions. The motivating case is the Bern SIH score (Dobrocky et al., 2019), whose non-contrast components have validated cutoffs that are seldom measured during routine reads.

What You Built: Describe your MVP, tool, or workflow in plain language
An automated sagittal MRI morphometry tool for SIH markers, with a web app. The tool accepts a sagittal T1-weighted sequence in NIfTI format. The tool automatically localizes the mid-sagittal image, segments the relevant structures, places the calipers, and computes the set of measurements described above. Next, the tool displays the segmentations, measurement overlays, the non-contrast Bern score, and each measurement against its threshold.

Tech Stack & Development Approach: What tools, frameworks, or methods did you use to build it?
The tool was developed using Python. The model is a 2.5D U-Net in MONAI/PyTorch (three adjacent sagittal slices as input channels), with a self-supervised encoder pretrained by masked autoencoding on 1,222 unlabeled brain MRIs before fine-tuning on hand-labeled brains from 100 healthy adults in public datasets spanning three scanner vendors (GE, Philips, Siemens) and two field strengths (1.5T, 3T): 65 for model development (57 train, 8 validation), 15 for held-out testing, and 20 for rater agreement. Preprocessing and I/O use SimpleITK. The measurements are custom geometry on the segmentation masks. Test-time

augmentation is used at inference. The web app uses Streamlit with a custom CSS heads-up display and matplotlib for the overlays. Training was run locally on a single RTX 4070, and labeling was performed in 3D Slicer. A frozen train/validation split and a regression test keep the reported numbers reproducible. Much of the app and supporting tools were built iteratively with an AI coding assistant. The model design, labeling, and initial validation were done by the medical student developer, with input and expert validation by a board-certified neuroradiologist.

Validation & Evidence: Have you done any informal testing or validation? If so, what did you find?
On 20 held-out healthy brains, the automated measurements were compared to manual blinded measurements performed by a medical student and a board-certified neuroradiologist with 7 years of experience. Agreement with the neuroradiologist and model was good (ICC 0.85 mamillopontine distance, 0.90 suprasellar height, 0.76 prepontine width), with Bland-Altman biases under 0.5 mm. For binary Bern classification, agreement with the neuroradiologist was substantial, including a kappa of 0.80 for mamillopontine distance and prepontine width (both outperformed published expert inter-reader agreement; Callen et al., 2023) and a kappa of 0.69 for suprasellar cistern height. The non-contrast Bern score kappa was similar between the tool and expert (0.74) compared to the two human readers (0.79). The other sagittal measurements (pituitary, pontomesencephalic angle, tonsillar descent, callosal height and elevation) are computed but not yet validated.

Feedback You're Seeking: What specific feedback, help, or collaboration are you hoping to get from the CAIMI26 community?
First, feedback on how to optimize the interface and experience so radiologists could seamlessly incorporate this tool into their workflow, adding value without becoming overly intrusive. Second, what are some potential methods for integration into PACS, reporting software, or other software interfaces, and which would be the best fit? More broadly, it'd be helpful to hear which of these measurements practicing radiologists would find worth having on screen, and which are noise. Pointers, collaborators, and "don't bother with X" are all welcome.

Known Limitations & Honest Failures: What isn't working yet, or what have you already tried that failed?
The tool has only been validated in healthy adults and, in practice, would be used in patients with concern for intracranial pressure dysregulation, so this reflects measurement agreement rather than diagnostic accuracy. All readers were given only the sagittal slice on which the model performed its measurements, which may positively bias agreement by removing the slice-selection variability that occurs in real-world practice. The tool currently covers the non-contrast components of the Bern score; the contrast-dependent components require additional sequences, and now that IRB approval has been granted, training a contrast-specific module is planned.

Potential for Impact: If this MVP were fully developed and deployed, who would benefit and how?
Opportunistic screening to detect abnormal measurements related to intracranial pressure dysregulation that could then be mentioned in the report and correlated with patient symptoms. This would also increase measurement consistency across patients, which could lead to more accurate measurement cutoffs for these conditions and increase clinical efficiency by automating measurements. This framework is already built to extend to other planes (axial T2 for IIH markers, coronal T1 for NPH) and other intracranial measurements, further increasing clinical efficiency once the models are trained.