Appendix 1 Additional experiment to investigate the performance of dural tail sign detection model for other tumors

Our model was also tested on other tumor groups to verify its ability to distinguish meningiomas from other tumors. The hypothesis of the additional experiment was that most other types of brain tumors do not exhibit the dural tail sign, unlike meningiomas, and therefore the trained model may not be able to predict the dural tail sign in other tumor dataset.

The other tumor group dataset consists of 41 patients (45 scan; 15,260 images) who were pathologically confirmed as having non-meningioma tumors (e.g. metastatic carcinoma, glioblastoma, and adenoma), with brain magnetic resonance imaging (MRI) taken from single institution between March 2016 and June 2022. The MRI acquisition for the other tumor dataset was the same as the dural tail sign dataset and normal dataset in the manuscript, and the preprocessing such as applying the brain mask for false positive reduction were performed the same as the test and normal dataset.

For evaluation, to investigate whether that the trained model detects dural tail signs around other tumor lesions, we manually evaluated the predicted boxes where the dural tail sign was detected around the lesion as false positives, instead of evaluating entire false positives for each scan.

As a result, false positives were observed around the tumor lesion in 26 patients, and some false positives around other tumor lesion were due to peritumoral enhancing vessels in clearly parenchymal locations, not near the dura.

Appendix 2 External validation

The total 96 scans in meningioma dataset from The Cancer Imaging Archive, we excluded 20 scans without the observation of dural tail sign through the labeling process, and selected 76 scans for the external validation set. The external validation dataset consisted mostly of axial contrast-enhanced 3 dimensional T1 weighted image (CE 3D T1WI) with an image matrix size of 256×256. Each image was resliced into the sagittal plane for evaluation using the trained model. We performed the evaluation process for the external validation set using the same procedure as that used for the internal test dataset, excluding the brain masking process because the external validation dataset does not include 3D T1WI. As a results for the external validation, a sensitivity and false positive average were measured as 36.84% and 9.21, respectively.

Table S1 Manufacture and scan parameters for internal and external validation datasets

Manufacturer	Internal validation dataset (testset; 45 scans)	External validation dataset (76 scans)				
	Philips Medical Systems	GE Medical Systems	GE Medical Systems	GE Medical Systems	Siemens	Toshiba
Manufacturer model	Achieva	DISCOVERY MR750	SIGNA EXCITE	Signa HDxt	Symphony	Titan
Study description	T1 3D TFE SAG-Prohence	Ax FSPGR BRAVO post	AX 3D FSPGR, AX 3D stealth	AX 3D stealth	t1_mpr_ns_ax stealth	3D T1 AX STEALTH
Number of scans	45	1	23	50	1	1
Field strength (T)	3.0	3.0	3.0	3.0	1.5	1.5
Repetition time (ms)	9.43	8.13	7.94	8.00	2.14	6.20
Echo time (ms)	4.60	3.16	3.06	3.06	3.93	3.20
Inversion time (ms)	–	450.00	450.00	450.00	1,100.00	600.00
Resolution	512×512	256×256	256×256	256×256	256×256	256×256
Pixel spacing	0.49	1.02	1.02	1.01	1.02	1.00
Intensity range	1,933.04	16,091.00	18,513.52	19,620.82	352.00	4,295.00

https://dx.doi.org/10.21037/qims-23-114

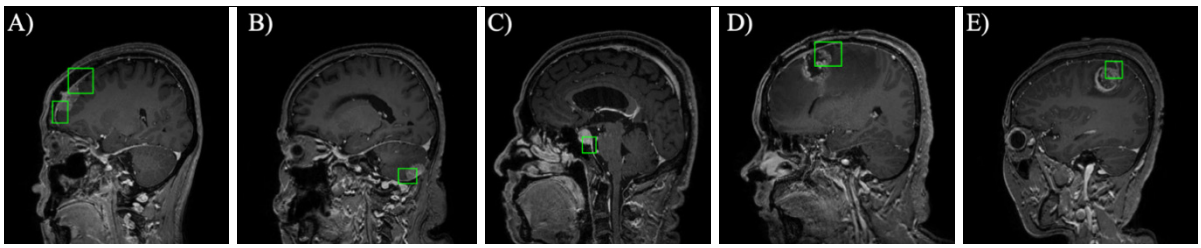

Figure S1 Examples of false-positive cases from the normal dataset. False positive around the (A) transverse sinus, (B) superior sagittal sinus, and (C) enhanced meningeal vessel are represented by the green-boundary boxes; (D-F) other examples observed at low frequencies.

https://dx.doi.org/10.21037/qims-23-114