Automated Anatomical Identification and Standardization for Medical Images

Qingrui Wang, B.S.[1], Teri Sippel Schmidt, M.S.[1], Paul Nagy, Ph.D.[1], Blake E. Dewey, Ph.D.[1,2]
[1]Biomedical Informatics and Data Science, Johns Hopkins University, Baltimore, MD, USA.
[2]Department of Neurology, Johns Hopkins University, Baltimore, MD, USA.

Background

Accurate antomical labels in medical imaging metadata is critical for research. However, prior studies have identified substantial inconsistencies; notably, the DICOM tag "Body Part Examined" (0018,0015) is incorrectly annotated in approximately 15% of cases.[1] Additionally, the free-text nature of the "Series Description" (0008,103E) DICOM tagmakes standardization challenging.[2] Improving metadata accuracy by validating anatomical labels can significantly enhance research data utilization, especially for large-scale clinically acquired datasets. The medical imaging extension to the OMOP CDM allows linkage to low-level DICOM image data empowering users to build cohorts and analyses at the imaging level.[3] This project introduces a workflow for automatically validating anatomical labels and mapping them to SNOMED CT concepts for standardization, facilitating accurate querying and integration into the OMOP CDM.

Methods

To automate and standardized anatomies, we employed TotalSegmentator, an open-source, pre-trained deep learning tool, to automatically identify anatomical structures in CT and MR images.[4,5] Segmentation outputs were compared against original DICOM metadata ("Body Part Examined"). If original labels matched segmentation results, we marked labels "correct" and mismatches or absent labels were flagged for manual review. After validation, we mapped verified labels to corresponding SNOMED CT codes and terms (**Figure 1**). Additional anatomies identified by TotalSegmentator were also stored for future reference.

We analyzed publicly available TCIA datasets, comprising 13 CT and 40 MR scans from various scanners and institutions across 16 anatomical regions. Data varied in positioning (prone, supine), imaging planes (coronal, sagittal, axial), structural granularity (e.g., spine, lumbar spine), and series (e.g., T1, T2, FRFSE, TSE, HASTE, DWI, ADC, MRCP). Of the 53 scans, 51 had explicit anatomical labels, one was labeled "NOT SPECIFIED," and another "UNKNOWN." All images were in DICOM format.

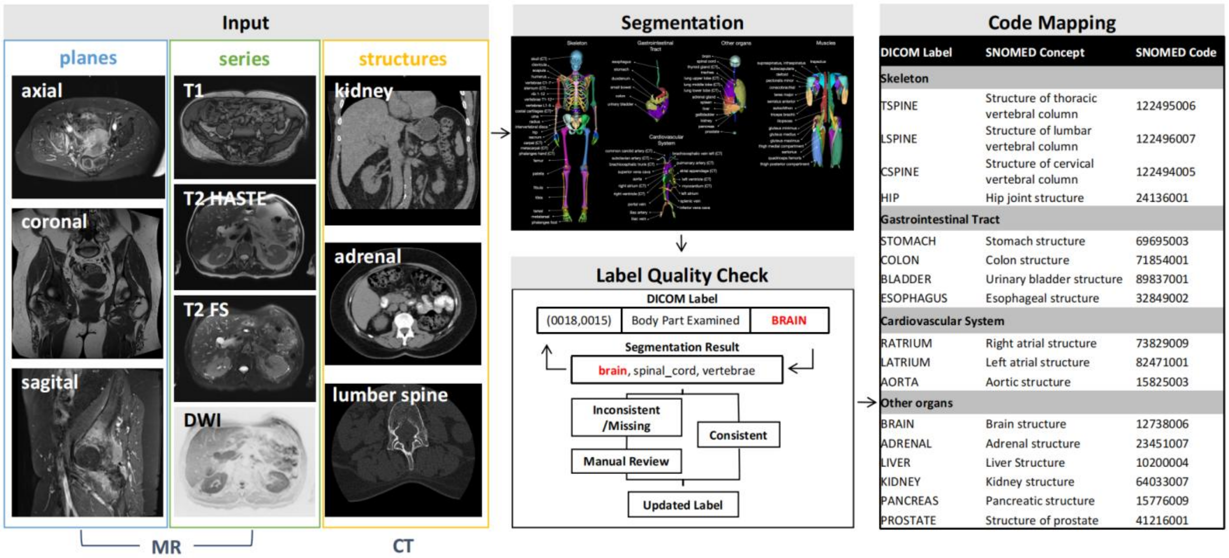

Figure 1. Project framework

Results

Out of 53 scans, 50 were consistent between original labels and segmentation results. For the two scans without original labels, reasonable labels were provided automatically by TotalSegmentator. One labeled scan (Pancreas, T2 MRCP coronal) was flagged because the segmentator failed detection, due to poor image quality. In addition to the primary label, an average of 70.00 additional anatomical classes were detected for CT and 15.38 for MR (due to different grouping of structures). See **Table 1** for example classifications.

Table 1. Automated anatomical identification and standardization examples

(red indicates a match with Body Part Examined)

Modality	Series Description	Study Description	Body Part Examined	Detected Structures by TotalSegmentator	Needs Review	SNOMED concept	SNOMED value
MR	COR T2 MRCP NAV	MR ABDOMEN NO CONTRASTAB	PANCREAS	colon, duodenum, femur_left, femur_right, gallbladder, gluteus_medius_left, gluteus_medius_right, gluteus_minimus_left, gluteus_minimus_right, hip_left, hip_right, iliac_artery_left, iliac_artery_right, iliac_vena_left, iliac_vena_right, iliopsoas_left, iliopsoas_right, intervertebral_discs, liver, lung_left, lung_right, sacrum, small_bowel, spleen, stomach, urinary_bladder, vertebrae	Yes	Pancreatic structure	15776009
MR	t1fl3dtragkv1	Avanto RoutineHead	NOT SPECIFIED	brain, intervertebral_discs, spinal_cord, vertebrae	Yes	Brain structure	12738006
MR	AXIAL T2 FLAIR ProcessedCaPTk	MR HEAD W AND WO IV CONTRAST TUMOR NAVIGATION	UNKNOWN	brain, spinal_cord	Yes	Brain structure	12738006
MR	TSE45 split	NCI PDMR Tumor Characterization	ABDOMEN	**gallbladder**	No	Structure of abdominal cross-sectional segment of trunk	818981001

MR	AXI T2 3mm HR	RESSONANCIA MAGNETICA DE PELVE	BLADDER	colon, femur_left, femur_right, hip_left, hip_right, iliac_artery_left, iliac_artery_right, iliac_vena_left, iliac_vena_right, iliopsoas_left, iliopsoas_right, prostate, sacrum, **urinary_bladder**	No	Urinary bladder structure	89837001
MR	SAG T2 sSH	RESSONANCIA MAGNETICA DE PELVE	BLADDER	aorta, autochthon_left, autochthon_right, colon, gluteus_maximus_left, gluteus_maximus_right, hip_left, hip_right, iliac_artery_left, iliac_artery_right, iliac_vena_left, iliac_vena_right, iliopsoas_left, iliopsoas_right, inferior_vena_cava, intervertebral_discs, prostate, sacrum, small_bowel, spinal_cord, **urinary_bladder**, vertebrae	No	Urinary bladder structure	89837001
MR	SAG T1 TFE POST MPR	MR BRAIN WOW CONTRAST	BRAIN	**brain**, spinal_cord, vertebrae	No	Brain structure	12738006
CT	Coloprone 1.0 B30f	Abdomen24ACR INColoIRB2415-04 Adult	COLON	adrenal_gland_left, adrenal_gland_right, aorta, autochthon_left, autochthon_right, **colon**, costal_cartilages, duodenum, esophagus, femur_left, femur_right, gallbladder, gluteus_maximus_left, gluteus_maximus_right, gluteus_medius_left, gluteus_medius_right, gluteus_minimus_left, gluteus_minimus_right, heart, hip_left, hip_right, iliac_artery_left, iliac_artery_right, iliac_vena_left, iliac_vena_right, iliopsoas_left, iliopsoas_right, inferior_vena_cava, kidney_left, kidney_right, liver, lung_lower_lobe_left, lung_lower_lobe_right, lung_middle_lobe_right, lung_upper_lobe_left, pancreas, portal_vein_and_splenic_vein, rib_left_10, rib_left_11, rib_left_12, rib_left_6, rib_left_7, rib_left_8, rib_left_9, rib_right_10, rib_right_11, rib_right_12, rib_right_5, rib_right_6, rib_right_7, rib_right_8, rib_right_9, sacrum, small_bowel, spinal_cord, spleen, sternum, stomach, urinary_bladder, vertebrae_L1, vertebrae_L2, vertebrae_L3, vertebrae_L4, vertebrae_L5, vertebrae_S1, vertebrae_T10, vertebrae_T11, vertebrae_T12, vertebrae_T7, vertebrae_T8, vertebrae_T9	No	Colon structure	71854001
CT	Colosupine 1.0 B30f	Abdomen24ACR INColoIRB2415-04 Adult	COLON	adrenal_gland_left, adrenal_gland_right, aorta, autochthon_left, autochthon_right, **colon**, costal_cartilages, duodenum, esophagus, femur_left, femur_right, gallbladder, gluteus_maximus_left, gluteus_maximus_right, gluteus_medius_left, gluteus_medius_right, gluteus_minimus_left, gluteus_minimus_right, heart, hip_left, hip_right, iliac_artery_left, iliac_artery_right, iliac_vena_left, iliac_vena_right, iliopsoas_left, iliopsoas_right, inferior_vena_cava, kidney_left, kidney_right, liver, lung_lower_lobe_left, lung_lower_lobe_right, lung_middle_lobe_right, lung_upper_lobe_left, pancreas, portal_vein_and_splenic_vein, rib_left_10, rib_left_11, rib_left_12, rib_left_5, rib_left_6, rib_left_7, rib_left_8, rib_left_9, rib_right_10, rib_right_11, rib_right_12, rib_right_5, rib_right_6, rib_right_7, rib_right_8, rib_right_9, sacrum, scapula_right, small_bowel, spinal_cord, spleen, sternum, stomach, urinary_bladder, vertebrae_L1, vertebrae_L2, vertebrae_L3, vertebrae_L4, vertebrae_L5, vertebrae_S1, vertebrae_T10, vertebrae_T11, vertebrae_T12, vertebrae_T8, vertebrae_T9	No	Colon structure	71854001

Conclusion

This project introduces an automated workflow for verifying and standardizing anatomical metadata in medical imaging, significantly enhancing data quality and interoperability within the OHDSI community. Validated and standardized anatomy metadata beyond the Body Part Examined are critical for fine-tuned cohorts of specific structures. Future work includes expanding validation to larger datasets and integrating this workflow directly into OMOP CDM imaging extensions to power real-world imaging network studies.

References

1. Gueld, M. O., Kohnen, M., Keysers, D., Schubert, H., Wein, B. B., Bredno, J., & Lehmann, T. M. (2002, May). Quality of DICOM header information for image categorization. In Medical imaging 2002: PACS and integrated medical information systems: design and evaluation (Vol. 4685, pp. 280-287). SPIE.
2. Gauriau R, Bridge C, Chen L, et al. Using DICOM Metadata for Radiological Image Series Categorization: a Feasibility Study on Large Clinical Brain MRI Datasets. J Digit Imaging. 2020;33(3):747-762. doi:10.1007/s10278-019-00308-x
3. Park WY, Jeon K, Schmidt TS, et al. Development of Medical Imaging Data Standardization for Imaging-Based Observational Research: OMOP Common Data Model Extension. J Imaging Inform Med. 2024;37(2):899-908. doi:10.1007/s10278-024-00982-6
4. Wasserthal, J., Breit, H. C., Meyer, M. T., Pradella, M., Hinck, D., Sauter, A. W., ... & Segeroth, M. (2023). TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. Radiology: Artificial Intelligence, 5(5), e230024.
5. Akinci D’Antonoli, T., Berger, L. K., Indrakanti, A. K., Vishwanathan, N., Weiss, J., Jung, M., ... & Wasserthal, J. (2025). TotalSegmentator MRI: Robust sequence-independent segmentation of multiple anatomic structures in MRI. Radiology, 314(2), e241613.