



HAL
open science

Knowledge-based semantic enrichment of medical imaging data for automatic phenotyping and pattern discovery in metastatic lung cancer

Francesco Cremonesi, Lucie Chambon, Nelson MokkaDEM, Huyen Thi Trang Nguyen, Oliver Humbert, Marco Lorenzi

► To cite this version:

Francesco Cremonesi, Lucie Chambon, Nelson MokkaDEM, Huyen Thi Trang Nguyen, Oliver Humbert, et al.. Knowledge-based semantic enrichment of medical imaging data for automatic phenotyping and pattern discovery in metastatic lung cancer. 2025. hal-04879690

HAL Id: hal-04879690

<https://hal.science/hal-04879690v1>

Preprint submitted on 7 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Knowledge-based semantic enrichment of medical imaging data for automatic phenotyping and pattern discovery in metastatic lung cancer

Francesco Cremonesi (1), Lucie Chambon (1) , Nelson MokkaDem (2) , Huyen Thi Trang Nguyen (1) , Oliver Humbert (2) , Marco Lorenzi (3, 4, 1)

1 EPIONE - E-Patient : Images, données & mOdèles pour la médeciNe numériquE

2 UNICANCER/CAL - Centre de Lutte contre le Cancer Antoine Lacassagne [Nice]

3 Inria - Institut National de Recherche en Informatique et en Automatique

4 ASCLEPIOS - Analysis and Simulation of Biomedical Images

Abstract

Introduction

The emergence of AI models for the analysis of 18FDG PET/CT images has opened the way for automatic quantification of clinically-relevant parameters in metastatic cancer patients such as e.g. the number, metabolic volume, and dispersion of lesions. However, current approaches lack specificity as they do not provide information about the localisation of lesions, which is a major clinical parameter for patient management. These methods also lack generality, as they fail to integrate the medical community's body of anatomical knowledge, which can nowadays be found in machine-readable ontology formats. Such integration may improve analysis depth as well as reusability for secondary purposes of the data, by enabling reasoning at different anatomical scales and linking with other sources of information, such as clinical records. To fill this gap, we present in this work a novel framework for the enrichment of 18FDG PET-based lesion annotations through the automatic segmentation of anatomical structures and inclusion of ontological knowledge.

Methods

Starting from aligned whole body 18FDG PET/CT images from a cohort of metastatic non-small cell lung cancer (NSCLC) patients undergoing immunotherapy treatment (n=101), we obtain automatic organ segmentations for over 100 anatomical structures using the publicly available TotalSegmentator (TS) tool applied on CT scans, and combine them with manual lesion annotation (including both segmentation and gold-standard anatomic labels of lesions) by nuclear physicians on the PET images. We map the TotalSegmentator structures to the Foundational Model of Anatomy (FMA) ontology, which in turn allows us to associate each structure with the gold standard labels by querying the ontology for relationships such as *subClassOf*, *part_of*, and others.

We then compute the intersection between the lesion volumes contoured by human experts on PET and the anatomical structure segmentations on CT, and combine intersections belonging to the same class as defined by the gold standard label. We localise lesions by finding the class with the maximal intersection, allowing us to directly compare to the expert annotations.

We insert the lesion location information into a Knowledge Graph (KG) integrating the FMA ontology, restricting our analysis to locations present in the set of structures segmented by TotalSegmentator. From our KG we extract the co-occurrence of lesions for each patient at two semantic scales: the raw TotalSegmentator structures and the gold standard anatomic labels. In both cases, we perform KMeans clustering to identify patterns of lesion co-occurrence.

Results

We found a good overlap between gold-standard anatomic localisations of lesions and automatically segmented anatomical structures by TS: compared to gold-standard labels, our method successfully localized 1568 out of 1926 lesions (79.9 %).

The co-occurrence of lesions across selected anatomical structures is shown in Figure 1. Unsupervised clustering of the co-occurrence patterns shows distinct groups of lesion distributions associated with different survival likelihood. The most pathological clusters mapped to liver and bones (OS 24M rate resp. 0.12 and 0.29), in accordance with published literature (PMID: 25130083). On the contrary, clusters mapping to lesions in muscles and adrenal glands were associated with better prognosis (OS 24M rate resp. 1.00 and 0.73). When applying the clustering on the segmented structures directly, without ontologies, no clinically-relevant subgroups were found (OS 24M rates from 0.33 to 0.46).

Conclusion

This study introduces the methodology of augmenting annotations of tumor lesions on 18FDG PET images with automatic organ segmentation from CT and anatomical knowledge from ontologies. Our approach improved disease phenotyping by automatically characterizing the distribution of lesions across clinically relevant anatomical groupings and identifying patterns of lesion dispersion correlating with survival parameters.

Methods

Starting from aligned whole body 18FDG PET/CT images from a cohort of metastatic non-small cell lung cancer (NSCLC) patients undergoing immunotherapy treatment (n=101), we obtain automatic organ segmentations for over 100 anatomical structures using the publicly available TotalSegmentator (TS) tool applied on CT scans, and combine them with manual lesion annotation (including both segmentation and gold-standard anatomic labels of lesions) by nuclear physicians on the PET images (see Table 1).

Table 1 Gold-standard labels for anatomical locations of lesions as identified by nuclear physicians.

Gold-standard anatomical location label	Description	FMA label
Mediastinum		
Cervical		
Lung		
Bone		
Adrenal gland		
SD	Sub-Diaphragmatic	
Liver		
Brain		
Pancreas		
CP	Peritoneal carcinomatosis	
Thyroid		
Muscle		
Axilla		
Spleen		
C1	Obsolete – Bone	
C2	Obsolete - Mediastinum	

Mapping automatic organ segmentation structures to gold-standard labels for lesion localization

This section describes the process of mapping the Totalsegmentator structure labels to the gold-standard lesion localization labels. The approach uses the Foundational Model of Anatomy (FMA) ontology as a bridge between the two concept sets.

Baseline mapping

We mapped the Totalsegmentator structures to the FMA ontology using a semi-manual approach: a first heuristic based on matching the Totalsegmentator name to the FMA label allowed to identify roughly 70% of mappings, while the rest was mapped manually. A manual revision based on human oversight was also required, as several subtleties needed to be fixed in order to improve the quality of the mapping, as shown in Table 2. The final results are shown in Table 3.

Table 2 Challenges encountered while mapping Totalsegmentator structures to FMA concepts.

Structure	Mapped FMA label	Comments
-----------	------------------	----------

Autochthon	Set of erector spinae	The mapped structure is not a subclassOf "Muscle organ", because it is a set of muscles.
Sacrum	Sacral Vertebra	The alternative mapping (i.e. "Sacrum") was discarded because it is not a SubClassOf "Bone Organ".
Hip	Hip Bone	The alternative mapping (i.e. "Hip") was discarded because Totalsegmentator only segments the hip bone, not the whole hip region.

Table 3 Mapping of Totalsegmentator structures to FMA ontology

Structure	fma_id	fma_label
adrenal_gland_left	http://purl.org/sig/ont/fma/fma15630	Left adrenal gland
adrenal_gland_right	http://purl.org/sig/ont/fma/fma15629	Right adrenal gland
aorta	http://purl.org/sig/ont/fma/fma3734	Aorta
autochthon_left	http://purl.org/sig/ont/fma/fma22732	Set of left erector spinae
autochthon_right	http://purl.org/sig/ont/fma/fma22731	Set of right erector spinae
brain	http://purl.org/sig/ont/fma/fma50801	Brain
clavicula_left	http://purl.org/sig/ont/fma/fma13323	Left clavicle
clavicula_right	http://purl.org/sig/ont/fma/fma13322	Right clavicle
colon	http://purl.org/sig/ont/fma/fma14543	Colon
duodenum	http://purl.org/sig/ont/fma/fma7206	Duodenum
esophagus	http://purl.org/sig/ont/fma/fma7131	Esophagus
face	http://purl.org/sig/ont/fma/fma24728	Face
femur_left	http://purl.org/sig/ont/fma/fma24475	Left femur
femur_right	http://purl.org/sig/ont/fma/fma24474	Right femur
gallbladder	http://purl.org/sig/ont/fma/fma7202	Gallbladder
gluteus_maximus_left	http://purl.org/sig/ont/fma/fma22329	Left gluteus maximus
gluteus_maximus_right	http://purl.org/sig/ont/fma/fma22328	Right gluteus maximus
gluteus_medius_left	http://purl.org/sig/ont/fma/fma22331	Left gluteus medius
gluteus_medius_right	http://purl.org/sig/ont/fma/fma22330	Right gluteus medius
gluteus_minimus_left	http://purl.org/sig/ont/fma/fma22333	Left gluteus minimus
gluteus_minimus_right	http://purl.org/sig/ont/fma/fma22332	Right gluteus minimus
heart_atrium_left	http://purl.org/sig/ont/fma/fma7097	Left atrium
heart_atrium_right	http://purl.org/sig/ont/fma/fma7096	Right atrium
heart_myocardium	http://purl.org/sig/ont/fma/fma9462	Myocardium
heart_ventricle_left	http://purl.org/sig/ont/fma/fma7101	Left ventricle
heart_ventricle_right	http://purl.org/sig/ont/fma/fma7098	Right ventricle
hip_left	http://purl.org/sig/ont/fma/fma16587	Left hip bone
hip_right	http://purl.org/sig/ont/fma/fma16586	Right hip bone
humerus_left	http://purl.org/sig/ont/fma/fma23131	Left humerus
humerus_right	http://purl.org/sig/ont/fma/fma23130	Right humerus
iliac_artery_left	http://purl.org/sig/ont/fma/fma14766	Left common iliac artery
iliac_artery_right	http://purl.org/sig/ont/fma/fma14765	Right common iliac artery
iliac_vena_left	http://purl.org/sig/ont/fma/fma21388	Left common iliac vein

iliac_vena_right	http://purl.org/sig/ont/fma/fma21387	Right common iliac vein
iliopsoas_left	http://purl.org/sig/ont/fma/fma64920	Left iliopsoas
iliopsoas_right	http://purl.org/sig/ont/fma/fma64919	Right iliopsoas
inferior_vena_cava	http://purl.org/sig/ont/fma/fma10951	Inferior vena cava
kidney_left	http://purl.org/sig/ont/fma/fma7205	Left kidney
kidney_right	http://purl.org/sig/ont/fma/fma7204	Right kidney
liver	http://purl.org/sig/ont/fma/fma7197	Liver
lung_lower_lobe_left	http://purl.org/sig/ont/fma/fma7371	Lower lobe of left lung
lung_lower_lobe_right	http://purl.org/sig/ont/fma/fma7337	Lower lobe of right lung
lung_middle_lobe_right	http://purl.org/sig/ont/fma/fma7383	Middle lobe of right lung
lung_upper_lobe_left	http://purl.org/sig/ont/fma/fma7370	Upper lobe of left lung
lung_upper_lobe_right	http://purl.org/sig/ont/fma/fma7333	Upper lobe of right lung
pancreas	http://purl.org/sig/ont/fma/fma7198	Pancreas
pulmonary_artery	http://purl.org/sig/ont/fma/fma66326	Pulmonary artery
rib_left_1	http://purl.org/sig/ont/fma/fma7987	Left first rib
rib_left_2	http://purl.org/sig/ont/fma/fma8012	Left second rib
rib_left_3	http://purl.org/sig/ont/fma/fma8039	Left third rib
rib_left_4	http://purl.org/sig/ont/fma/fma8148	Left fourth rib
rib_left_5	http://purl.org/sig/ont/fma/fma8093	Left fifth rib
rib_left_6	http://purl.org/sig/ont/fma/fma8202	Left sixth rib
rib_left_7	http://purl.org/sig/ont/fma/fma8256	Left seventh rib
rib_left_8	http://purl.org/sig/ont/fma/fma8310	Left eighth rib
rib_left_9	http://purl.org/sig/ont/fma/fma8391	Left ninth rib
rib_left_10	http://purl.org/sig/ont/fma/fma8472	Left tenth rib
rib_left_11	http://purl.org/sig/ont/fma/fma8532	Left eleventh rib
rib_left_12	http://purl.org/sig/ont/fma/fma8534	Left twelfth rib
rib_right_1	http://purl.org/sig/ont/fma/fma7857	Right first rib
rib_right_2	http://purl.org/sig/ont/fma/fma7882	Right second rib
rib_right_3	http://purl.org/sig/ont/fma/fma7909	Right third rib
rib_right_4	http://purl.org/sig/ont/fma/fma7957	Right fourth rib
rib_right_5	http://purl.org/sig/ont/fma/fma8066	Right fifth rib
rib_right_6	http://purl.org/sig/ont/fma/fma8175	Right sixth rib
rib_right_7	http://purl.org/sig/ont/fma/fma8229	Right seventh rib
rib_right_8	http://purl.org/sig/ont/fma/fma8283	Right eighth rib
rib_right_9	http://purl.org/sig/ont/fma/fma8364	Right ninth rib
rib_right_10	http://purl.org/sig/ont/fma/fma8445	Right tenth rib
rib_right_11	http://purl.org/sig/ont/fma/fma8531	Right eleventh rib
rib_right_12	http://purl.org/sig/ont/fma/fma8533	Right twelfth rib
sacrum	http://purl.org/sig/ont/fma/fma12526	Sacral vertebra
scapula_left	http://purl.org/sig/ont/fma/fma13396	Left scapula
scapula_right	http://purl.org/sig/ont/fma/fma13395	Right scapula
small_bowel	http://purl.org/sig/ont/fma/fma7200	Small intestine
spleen	http://purl.org/sig/ont/fma/fma7196	Spleen

stomach	http://purl.org/sig/ont/fma/fma7148	Stomach
trachea	http://purl.org/sig/ont/fma/fma7394	Trachea
urinary_bladder	http://purl.org/sig/ont/fma/fma15900	Urinary bladder
vertebrae_S1	http://purl.org/sig/ont/fma/fma13077	First sacral vertebra
vertebrae_C1	http://purl.org/sig/ont/fma/fma12519	Atlas
vertebrae_C2	http://purl.org/sig/ont/fma/fma12520	Axis
vertebrae_C3	http://purl.org/sig/ont/fma/fma12521	Third cervical vertebra
vertebrae_C4	http://purl.org/sig/ont/fma/fma12522	Fourth cervical vertebra
vertebrae_C5	http://purl.org/sig/ont/fma/fma12523	Fifth cervical vertebra
vertebrae_C6	http://purl.org/sig/ont/fma/fma12524	Sixth cervical vertebra
vertebrae_C7	http://purl.org/sig/ont/fma/fma12525	Seventh cervical vertebra
vertebrae_L1	http://purl.org/sig/ont/fma/fma13072	First lumbar vertebra
vertebrae_L2	http://purl.org/sig/ont/fma/fma13073	Second lumbar vertebra
vertebrae_L3	http://purl.org/sig/ont/fma/fma13074	Third lumbar vertebra
vertebrae_L4	http://purl.org/sig/ont/fma/fma13075	Fourth lumbar vertebra
vertebrae_L5	http://purl.org/sig/ont/fma/fma13076	Fifth lumbar vertebra
vertebrae_T1	http://purl.org/sig/ont/fma/fma9165	First thoracic vertebra
vertebrae_T2	http://purl.org/sig/ont/fma/fma9187	Second thoracic vertebra
vertebrae_T3	http://purl.org/sig/ont/fma/fma9209	Third thoracic vertebra
vertebrae_T4	http://purl.org/sig/ont/fma/fma9248	Fourth thoracic vertebra
vertebrae_T5	http://purl.org/sig/ont/fma/fma9922	Fifth thoracic vertebra
vertebrae_T6	http://purl.org/sig/ont/fma/fma9945	Sixth thoracic vertebra
vertebrae_T7	http://purl.org/sig/ont/fma/fma9968	Seventh thoracic vertebra
vertebrae_T8	http://purl.org/sig/ont/fma/fma9991	Eighth thoracic vertebra
vertebrae_T9	http://purl.org/sig/ont/fma/fma10014	Ninth thoracic vertebra
vertebrae_T10	http://purl.org/sig/ont/fma/fma10037	Tenth thoracic vertebra
vertebrae_T11	http://purl.org/sig/ont/fma/fma10059	Eleventh thoracic vertebra
vertebrae_T12	http://purl.org/sig/ont/fma/fma10081	Twelfth thoracic vertebra
heart	http://purl.org/sig/ont/fma/fma7088	Heart
pulmonary_vein	http://purl.org/sig/ont/fma/fma66643	Pulmonary vein
brachiocephalic_trunk	http://purl.org/sig/ont/fma/fma3932	Brachiocephalic artery
subclavian_artery_right	http://purl.org/sig/ont/fma/fma3953	Right subclavian artery
subclavian_artery_left	http://purl.org/sig/ont/fma/fma4694	Left subclavian artery
common_carotid_artery_right	http://purl.org/sig/ont/fma/fma3941	Right common carotid artery
common_carotid_artery_left	http://purl.org/sig/ont/fma/fma4058	Left common carotid artery
brachiocephalic_vein_left	http://purl.org/sig/ont/fma/fma4761	Left brachiocephalic vein
brachiocephalic_vein_right	http://purl.org/sig/ont/fma/fma4751	Right brachiocephalic vein
superior_vena_cava	http://purl.org/sig/ont/fma/fma4720	Superior vena cava
prostate	http://purl.org/sig/ont/fma/fma9600	Prostate
femur	http://purl.org/sig/ont/fma/fma9611	Femur
patella	http://purl.org/sig/ont/fma/fma24485	Patella
tibia	http://purl.org/sig/ont/fma/fma24476	Tibia
fibula	http://purl.org/sig/ont/fma/fma24479	Fibula

humerus	http://purl.org/sig/ont/fma/fma13303	Humerus
ulna	http://purl.org/sig/ont/fma/fma23466	Ulna
radius	http://purl.org/sig/ont/fma/fma23463	Radius
sternum	http://purl.org/sig/ont/fma/fma7485	Sternum
skull	http://purl.org/sig/ont/fma/fma0326476	Cranial bone
subcutaneous_fat	http://purl.org/sig/ont/fma/fma74315	Subcutaneous adipose tissue
skeletal_muscle	http://purl.org/sig/ont/fma/fma0329248	Skeletal muscle
spinal_cord	http://purl.org/sig/ont/fma/fma7647	Spinal cord
kidney	http://purl.org/sig/ont/fma/fma7203	Kidney
adrenal_gland	http://purl.org/sig/ont/fma/fma9604	Adrenal gland
thyroid_gland	http://purl.org/sig/ont/fma/fma9603	Thyroid gland
vertebrae_lumbar	http://purl.org/sig/ont/fma/fma9921	Lumbar vertebra
vertebrae_thoracic	http://purl.org/sig/ont/fma/fma9139	Thoracic vertebra
vertebrae_cervical	http://purl.org/sig/ont/fma/fma9915	Cervical vertebra
ribs	http://purl.org/sig/ont/fma/fma7574	Rib
scapula	http://purl.org/sig/ont/fma/fma13394	Scapula
hip	http://purl.org/sig/ont/fma/fma16585	Hip bone
gluteus_maximus	http://purl.org/sig/ont/fma/fma22314	Gluteus maximus
gluteus_medius	http://purl.org/sig/ont/fma/fma22315	Gluteus medius
gluteus_minimus	http://purl.org/sig/ont/fma/fma22317	Gluteus minimus
iliopsoas	http://purl.org/sig/ont/fma/fma64918	Iliopsoas

We also manually mapped the gold-standard anatomical location labels to the FMA label, as shown in Table 1. Note that the mapped concepts for the gold standard labels are not necessarily subclasses of the “Organ” concept. For example, the Mediastinum is not an organ, but it is part of the “Anatomical cluster” hierarchy (see Figure 1).



Figure 1 Class hierarchy for Mediastinum in the FMA ontology

Finally, we are left with the challenge of describing the mapping between the Totalsegmentator structures and the gold-standard lesion localization labels. The level of granularity of the gold-standard labels varies widely, ranging from single organs (e.g. Liver) to aggregate entities such as organ sets (e.g. Bones is the set of all bones) and regions in the anatomical space (e.g. the Mediastinum).

Our approach consists of defining a set of SparQL queries searching for paths from a gold-standard label to a Totalsegmentator structure. The queries must be adapted to the level of granularity and type of gold-standard label. In cases where the gold-standard label is the organ itself, or a superclass of a Totalsegmentator structure, a simple query structured around the code shown in Figure 2 is sufficient.

```
{ #catch the same individuals
  ?lesion_organ a ?lesion_organ_class .
  ?totalseg_organ a ?totalsegclass .
  ?totalsegclass rdfs:subClassOf+ ?lesion_organ_class
}
UNION
{ # catch the same classes (that are not punned)
  ?lesion_organ rdfs:subClassOf? ?totalseg_organ .
  ?totalseg_organ rdfs:subClassOf? ?lesion_organ .
}
UNION
{ # unpunned classes that are subclasses of lesions
  ?totalseg_organ rdfs:subClassOf+ ?lesion_organ .
}
```

Figure 2 Sample query to identify cases where the Totalsegmentator structure and the gold-standard lesions are the same concepts, or have a subClassOf relationship.

In cases where the gold-standard label represents an aggregate of organs, we exploit the FMA relationships *constitutional_part_of*, *regional_part_of*, *member_of*, *contained_in* to identify a path from a Totalsegmentator structure to a gold-standard label. Additional care must be taken to account for subclasses: we link a Totalsegmentator structure to a gold-standard label also in cases where a parent class of the structure is contained in the gold-standard label, as well as cases where the structure is contained in a child class of the gold-standard label. For example, the FMA concept fma:7371 “Lower lobe of left lung” is asserted as *regional_part_of* fma:7310 “Left lung”, and fma:7310 “Left lung” is a *subClassOf* the gold-standard label concept fma:7195 “Lung”. The sample query shown in Figure 3 covers these cases. The full query is shown in Figure 4.

```
{ ?totalseg_organ rdfs:subClassOf* ?y .
  ?y (fmapred:constitutional_part_of | fmapred:regional_part_of |
fmapred:member_of | fmapred:contained_in)+ ?x .
  ?x rdfs:subClassOf* ?lesion_organ }
```

Figure 3 Sample query to identify cases where the gold-standard label is an aggregate of structures.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX fma: <http://purl.org/sig/ont/fma/fma>
PREFIX fmapred: <http://purl.org/sig/ont/fma/>
```

```
select distinct ?lesion_organ ?totalseg_organ ?organlabel ?lesionorganlabel
#?organclasslabel ?lesionorganclasslabel
where {
  values ?lesion_organ {fma:14543 fma:18252 fma:24864 fma:5018 fma:5022
fma:50801 fma:7163 fma:7195 fma:7196 fma:7197 fma:7198 fma:7203 fma:79646
fma:9578 fma:9583 fma:9584 fma:9601 fma:9603 fma:9604 fma:9608 fma:9826
fma:12266 fma:9738 } .
  values ?totalseg_organ {fma:329248 fma:10014 fma:10037 fma:10059 fma:10081
fma:10951 fma:12519 fma:12520 fma:12521 fma:12522 fma:12523 fma:12524
fma:12525 fma:13072 fma:13073 fma:13074 fma:13075 fma:13076 fma:13077
fma:13303 fma:13322 fma:13323 fma:13394 fma:13395 fma:13396 fma:14543
fma:14765 fma:14766 fma:15629 fma:15630 fma:15900 fma:12526 fma:21387
fma:21388 fma:22314 fma:22315 fma:22317 fma:22328 fma:22329 fma:22330
fma:22331 fma:22332 fma:22333 fma:22731 fma:22732 fma:23130 fma:23131
fma:23463 fma:23466 fma:24474 fma:24475 fma:24476 fma:24479 fma:24485
fma:24728 fma:3734 fma:3932 fma:3941 fma:3953 fma:4058 fma:4694 fma:4720
fma:4751 fma:4761 fma:50801 fma:64918 fma:64919 fma:64920 fma:66326 fma:66643
fma:7088 fma:7096 fma:7097 fma:7098 fma:7101 fma:7131 fma:7148 fma:7196
fma:7197 fma:7198 fma:7200 fma:7202 fma:7203 fma:7204 fma:7205 fma:7206
fma:7333 fma:7337 fma:7370 fma:7371 fma:7383 fma:7394 fma:74315 fma:7485
fma:7574 fma:7647 fma:7857 fma:7882 fma:7909 fma:7957 fma:7987 fma:8012
fma:8039 fma:8066 fma:8093 fma:8148 fma:8175 fma:8202 fma:8229 fma:8256
fma:8283 fma:8310 fma:8364 fma:8391 fma:8445 fma:8472 fma:8531 fma:8532
fma:8533 fma:8534 fma:9139 fma:9165 fma:9187 fma:9209 fma:9248 fma:9462
fma:9600 fma:9603 fma:9604 fma:9611 fma:9915 fma:9921 fma:9922 fma:9945
fma:9968 fma:9991 fma:16586 fma:16587 fma:0326476} .
  { ?totalseg_organ rdfs:subClassOf* ?y .
    ?y (fmapred:constitutional_part_of | fmapred:regional_part_of |
fmapred:member_of | fmapred:contained_in)+ ?x .
    ?x rdfs:subClassOf* ?lesion_organ
  }
  UNION
  { ?lesion_organ (fmapred:constitutional_part | fmapred:regional_part |
fmapred:member)* ?intermediate .
    ?intermediate (fmapred:constitutional_part_of |
fmapred:regional_part_of | fmapred:member_of)* ?totalseg_organ . }
  UNION
  { #catch the same individuals
    ?lesion_organ a ?lesion_organ_class .
    ?totalseg_organ a ?totalsegclass .
    ?totalsegclass rdfs:subClassOf+ ?lesion_organ_class
  }
  UNION
  { # catch the same classes (that are not punned)
    ?lesion_organ rdfs:subClassOf? ?totalseg_organ .
    ?totalseg_organ rdfs:subClassOf? ?lesion_organ .
  }
  UNION
  { # unpunned classes that are subclasses of lesions
    ?totalseg_organ rdfs:subClassOf+ ?lesion_organ .
  }
  ?totalseg_organ rdfs:label ?organlabel .
```

Figure 4 Full SparQL query to link Totalsegmentator structures to gold-standard labels. This query is executed on the rdf representation of the FMA ontology.

Mapping improvement through heuristics

The process described thus far is helpful to establish a good baseline, but some imperfections remain, as shown in Table 4.

Table 4 Remaining issues after establishing a baseline link between Totalsegmentator structures and gold-standard labels.

No.	Description	Example
1	The same Totalsegmentator structure is linked with multiple gold-standard labels	Adrenal glands are matched with both the "Adrenal gland" gold-standard label and the "Sub-diaphragmatic" label. The latter match happens because of the following chain of relationships "Adrenal gland" member_of "Set of viscera of abdomen" contained_in "Abdominal cavity".
2	No valid path can be found between the Totalsegmentator structure and the correct gold-standard label	The "Set of left erector spinae" is not linked to "Muscle organ" by any valid relationship.

Issue number 1 is solved by establishing a priority among the different ways in which a Totalsegmentator structure may be linked to the gold-standard label:

1. The Structure is the same concept as the gold-standard label;
2. The Structure is a sub-class of the gold-standard label (in case of multiple matches, choose the one that minimizes the number of intermediate subclasses);
3. The Structure is linked to the gold-standard label through one or more of the *constitutional_part_of*, *regional_part_of*, *member_of*, *contained_in* relationships.

Issue number 2 was solved manually, due to the small number of structures falling in this category: the erector spinae and iliopsoas muscles.

The final result is shown in Table 5.

Table 5 Mapping of Totalsegmentator structures to gold-standard labels, both represented through their corresponding FMA concepts.

totalsegmentator	gold-standard label
Stomach	Abdominal cavity
Small intestine	Abdominal cavity
Adrenal gland	Adrenal gland
Left adrenal gland	Adrenal gland
Right adrenal gland	Adrenal gland
Gallbladder	Biliary system
Sternum	Bone organ
Cranial bone	Bone organ
Right scapula	Bone organ
Left scapula	Bone organ
Scapula	Bone organ

Sacral vertebra	Bone organ
Rib	Bone organ
Right ninth rib	Bone organ
Right eighth rib	Bone organ
Right sixth rib	Bone organ
Tibia	Bone organ
Right fifth rib	Bone organ
Right fourth rib	Bone organ
Right third rib	Bone organ
Right second rib	Bone organ
Right twelfth rib	Bone organ
Right eleventh rib	Bone organ
Right tenth rib	Bone organ
Right first rib	Bone organ
Left ninth rib	Bone organ
Left eighth rib	Bone organ
Left seventh rib	Bone organ
Right seventh rib	Bone organ
Ulna	Bone organ
Axis	Bone organ
Left sixth rib	Bone organ
Lumbar vertebra	Bone organ
Cervical vertebra	Bone organ
Ninth thoracic vertebra	Bone organ
Eighth thoracic vertebra	Bone organ
Seventh thoracic vertebra	Bone organ
Sixth thoracic vertebra	Bone organ
Fifth thoracic vertebra	Bone organ
Fourth thoracic vertebra	Bone organ
Third thoracic vertebra	Bone organ
Second thoracic vertebra	Bone organ
Twelfth thoracic vertebra	Bone organ
Eleventh thoracic vertebra	Bone organ
Tenth thoracic vertebra	Bone organ
First thoracic vertebra	Bone organ
First sacral vertebra	Bone organ
Fifth lumbar vertebra	Bone organ
Fourth lumbar vertebra	Bone organ
Third lumbar vertebra	Bone organ
Second lumbar vertebra	Bone organ
First lumbar vertebra	Bone organ
Seventh cervical vertebra	Bone organ
Sixth cervical vertebra	Bone organ
Fifth cervical vertebra	Bone organ
Fourth cervical vertebra	Bone organ
Third cervical vertebra	Bone organ
Atlas	Bone organ
Left fifth rib	Bone organ

Left fourth rib	Bone organ
Left third rib	Bone organ
Left femur	Bone organ
Femur	Bone organ
Thoracic vertebra	Bone organ
Right clavicle	Bone organ
Left hip bone	Bone organ
Right hip bone	Bone organ
Left clavicle	Bone organ
Humerus	Bone organ
Left humerus	Bone organ
Right humerus	Bone organ
Right femur	Bone organ
Fibula	Bone organ
Patella	Bone organ
Left second rib	Bone organ
Left twelfth rib	Bone organ
Left eleventh rib	Bone organ
Left tenth rib	Bone organ
Left first rib	Bone organ
Radius	Bone organ
Brain	Brain
Colon	Colon
Left kidney	Kidney
Kidney	Kidney
Right kidney	Kidney
Liver	Liver
Middle lobe of right lung	Lung
Upper lobe of right lung	Lung
Upper lobe of left lung	Lung
Lower lobe of left lung	Lung
Lower lobe of right lung	Lung
Aorta	Mediastinum
Right brachiocephalic vein	Mediastinum
Esophagus	Mediastinum
Brachiocephalic artery	Mediastinum
Trachea	Mediastinum
Myocardium	Mediastinum
Right ventricle	Mediastinum
Left ventricle	Mediastinum
Heart	Mediastinum
Left atrium	Mediastinum
Right atrium	Mediastinum
Set of left erector spinae	Muscle organ
Set of right erector spinae	Muscle organ
Right iliopsoas	Muscle organ
Left iliopsoas	Muscle organ
Right gluteus minimus	Muscle organ

Left gluteus minimus	Muscle organ
Left gluteus maximus	Muscle organ
Right gluteus maximus	Muscle organ
Gluteus medius	Muscle organ
Left gluteus medius	Muscle organ
Right gluteus medius	Muscle organ
Gluteus minimus	Muscle organ
Gluteus maximus	Muscle organ
Pancreas	Pancreas
Prostate	Prostate
Spleen	Spleen
Thyroid gland	Thyroid gland

Lesion localization

We preprocessed the lesion annotation masks provided by nuclear physicians by computing the connected components and considering each connected component as a separate lesion. We then computed the intersection between each Totalsegmentator structure and each lesion's connected component separately. For each component, we aggregated the intersections for structures that are mapped to the same gold-standard label, and finally assigned as lesion localization the label with the largest intersection.

For example, the red lesion in Figure 5 intersects the aorta on 50 voxels and the esophagus on 70 voxels. Since both aorta and esophagus are mapped by the ontology to the mediastinum, we sum these contributions for a total intersection of 120 voxels, and compare it with the total intersection for other lesion locations. Since the mediastinum is the label with the maximal intersection, we predict the lesion location as mediastinum.

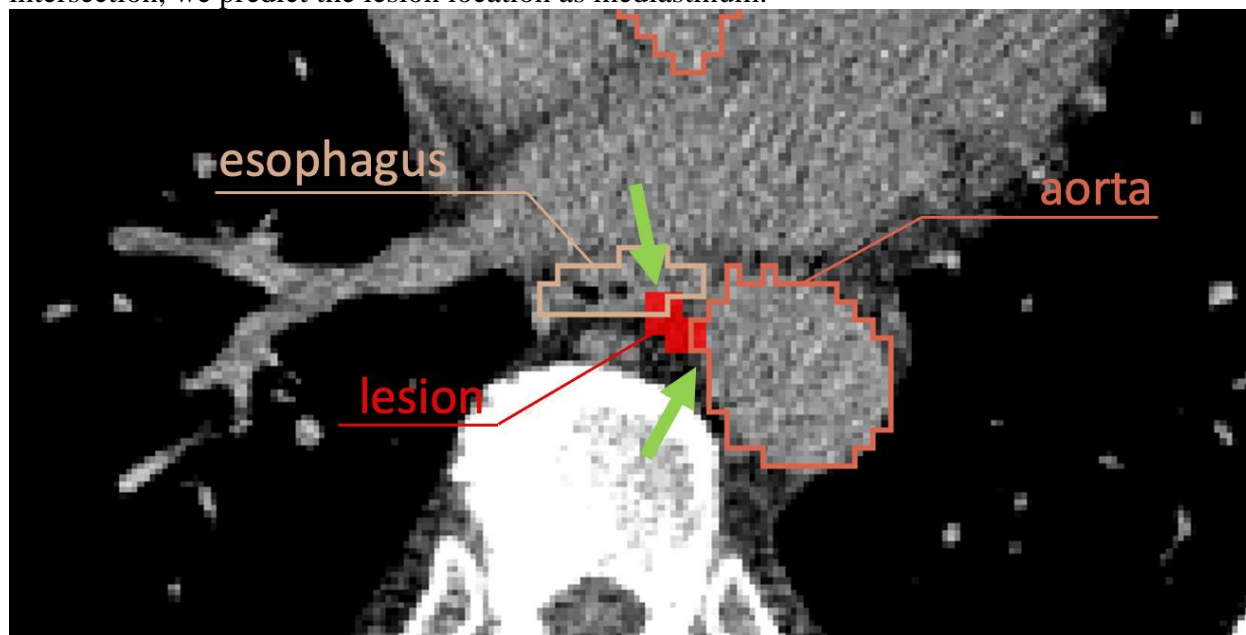


Figure 5 Intersection between lesion masks annotated by nuclear physicians and Totalsegmentator segmentations.

Knowledge graph

Patient stratification

We represent each patient by a set of features defined as the number of lesions with

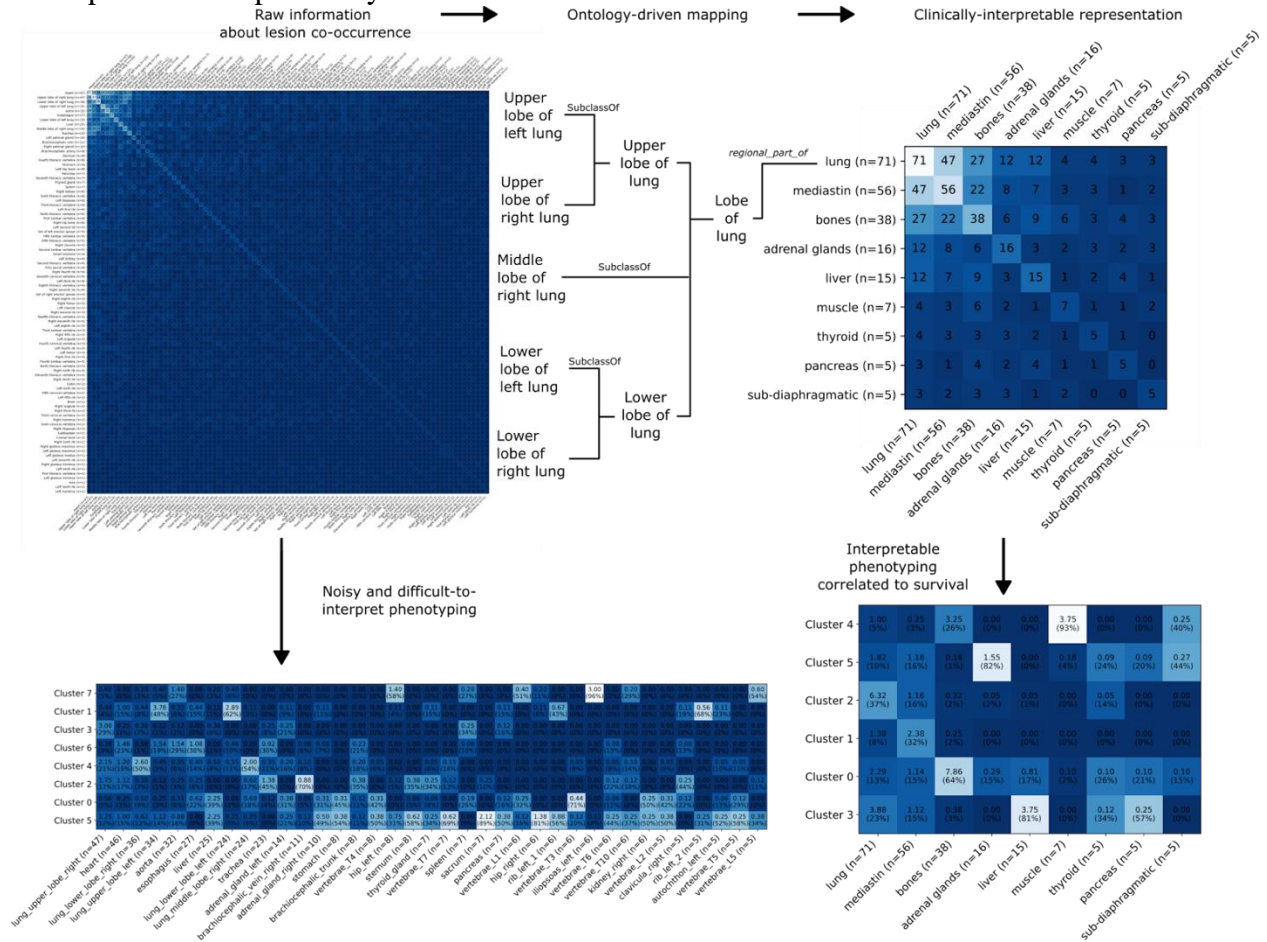


Figure 1: *Top*: Co-occurrences of lesion locations. On the left, raw lesion locations as identified by the automatic segmentation tool. In the center, an example of the mapping to gold standard labels grounded on the ontological relationships between anatomical structures. On the right, clinically-interpretable representation of the co-occurrence information, leading to improved patient phenotyping.

Bottom: Phenotyping through patient clustering. On the left, groups arising from clustering on the raw lesion locations as identified by the automatic segmentation tool. On the right, groups arising from clustering on the lesion locations after the ontology-driven mapping to gold standard labels, leading to more interpretable clusters with a stronger correlation to overall survival.