

## Purpose

This document provides details of the post processing step in the sites of mets prediction pipeline, which uses mapping files for term normalisation. Example scenario is given to aid explanation.

**Created by:** Tay See Boon / Low Guat Hwa

**Last Update:** 11 Jul 2023

## List of Mapping Files

These are used to normalize the body part, sequentially.

s/n	File Name	Purpose	Usage scenario
1	01_map_anatomical_descriptor_to_body_part.csv	to map anatomical descriptor (that is not linked to a body part in the predicted relation) to a target body part	1
2	02_map_cancer_body_part_anat.csv	to map cancer imaging finding + body part to a target body part (nodal diseases) Eg adenopathy + abdominal > abdominopelvic lymph node	2
3	03_map_normalized_body_part.csv	to normalize a body part value Eg hepatic > liver	3
4	04_map_merge_body_part.csv	to merge 2 body part values to one value Eg breast + nipple > breast	4
5	05_map_drop_body_part.csv	to remove unwanted body part values from final model output Eg surgical_bed	5

## How to use the mapping file

To ensure consistency and accuracy, the extracted chunks (anatomical\_descriptor and/or body\_part) are normalized using rule-based logic for the final model predictions.

### Scenario 1

In the case where only the anatomical descriptor is extracted without being linked to any specific body part, you can use mapping file 1 to map it to a target body part.

Example: Stable left [upper lobe](#) (*anatomical\_descriptor*) metastases (*cancer\_imaging\_finding*).

Processing logic:

Using mapping file 1: we search for “upper\_lobe” and retrieve the value “lung”.

<b>anatomical_descriptor → map_to</b> (*refer as output1_body_part)
upper lobe → lung

### Scenario 2

In the case where the report includes the presence of nodal diseases, you can use mapping file 2 to map cancer imaging finding + body part to a target body part.

Example: Increased size of the [paraaortic](#) (*anatomical\_descriptor*) [adenopathy](#) (*cancer\_imaging\_finding*).

Processing logic:

Using mapping file 1: we search for “paraaortic” and retrieve the value “abdominal”.

<b>anatomical_descriptor → map_to</b> (*refer as output1_body_part)
paraaortic → abdominal

Using mapping file 2: we search for “adenopathy” and “abdominal” to get “abdominopelvic lymph node”

<b>cancer_imaging_finding + body_part_anat</b> *linked body_part/ output1_body_part	<b>body_part</b> (*refer as output2_body_part)
adenopathy + abdominal	abdominopelvic lymph node

### **Scenario 3**

Use mapping file 3 to normalise all body parts found to target body parts.

Example: Largely stable scattered hepatic (*body\_part*) metastases (*cancer imaging finding*)

Processing logic:

Using mapping file 3: we search for “hepatic” and retrieve the value “liver”.

<b>body_part</b> → <b>normalized_body_part</b> (*refer as output3_body_part)
hepatic → liver

### **Scenario 4**

In the case where some body\_parts obtained indicate the same anatomical region or organ, you can make use of mapping file 4 to combine them. The purpose of this step is to enhance precision by refining and eliminating the duplicated similar body areas or body parts.

Example: Interval increase in size and number of multiple enhancing foci throughout the right breast (*body\_part*) and right nipple (*body\_part*), suspicious for tumour deposits.

Processing logic:

Using mapping file 4: we search for “breast” + “nipple” to obtain the value “breast”.

<b>body_part1</b>	<b>body_part2</b>	<b>output_body_part</b> (*refer as output4_body_part)
breast	nipple	breast

### **Scenario 5**

In the case where the body\_part mentioned does not correspond to an organ, you can use mapping file to remove them from final model output. The purpose of this step is to enhance and obtain a clean predicted output.

Example: surgical bed, anastomotic site

Processing logic:

Using mapping file 5: we search for the values “surgical bed”, “anastomotic site” and dropped them from final model outputs.

## Post processing flow chart

# Post Processing (term normalization) flow chart

