Table of Contents

ANNOTATION GUIDELINES FOR CLINICAL FINDINGS IN RADIOLOGY REPORTS

The goal of this annotation is to extract two different clinical findings from radiology reports, namely Medical problem finding, and Lesion finding. Each finding is represented by an event consisting of a trigger and multiple arguments. The annotation process involves identifying text spans within the notes that directly associate with the different clinical information as well as the relationships among them. Each piece of information (entity) is related to an event trigger, which link all information together cohesively. Figure 1 shows the different clinical entities with the event triggers marked in red. The entities with * are categorical with pre-defined values. The following sections describe how each piece of information will be annotated. Not all entities are present in a radiology report. However, when annotating an entity within a clinical finding, the corresponding trigger (red) should always be identified and annotated first.

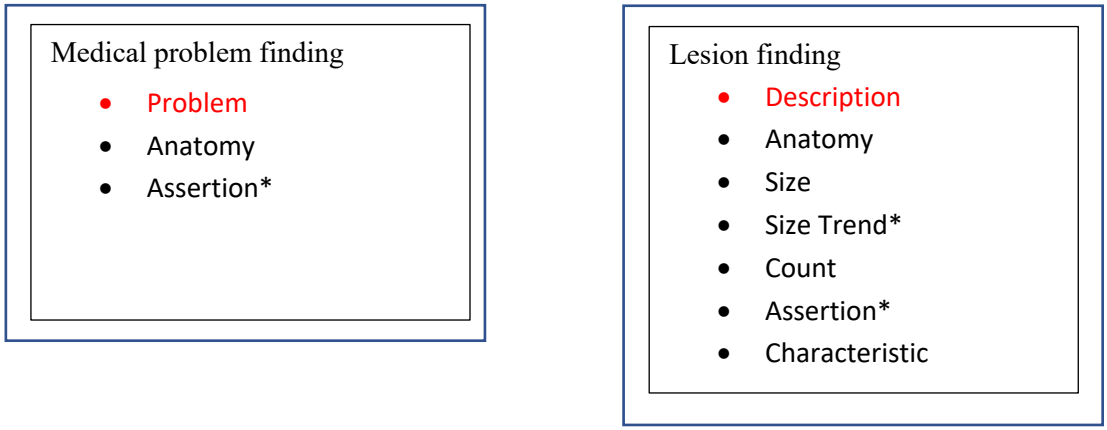


Figure 1. Clinical finding events and their associated arguments

Annotation is done on the BRAT tool. When highlighting a text span, a window will pop up showing the selections.

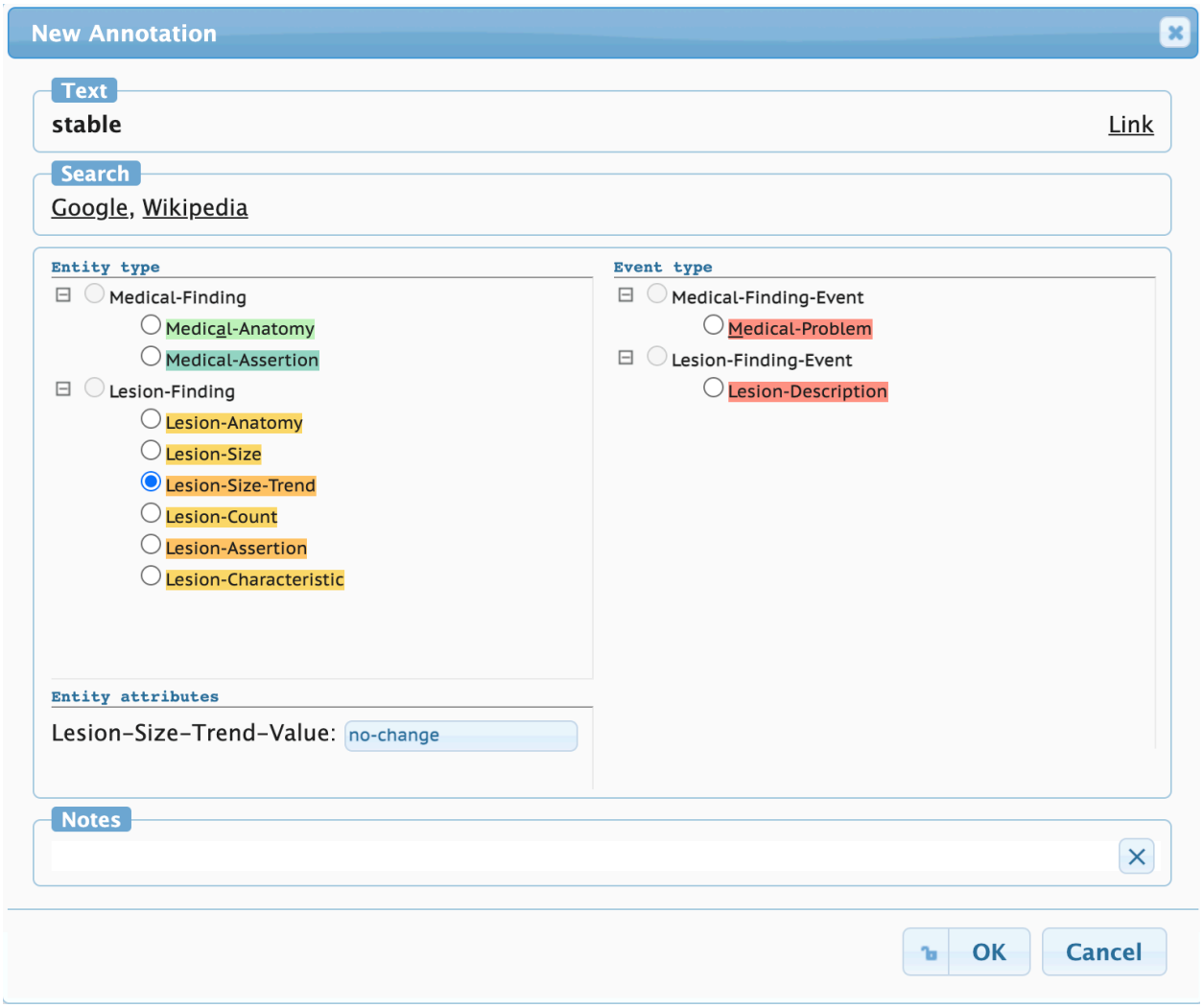


Figure 2. BRAT tool entity selections for new annotation

The right panel indicates the event triggers for the clinical findings. The left panel shows the entity types that are associated with each event trigger. The medical finding entities are highlighted in green while the lesion finding entities are in yellow. The ones in dark green and dark yellow are categorical which require selecting one of the possible values in the drop-down box down below. More details will be provided in the following sections. In general, entities highlighted in red are the event triggers and should be annotated first before others.

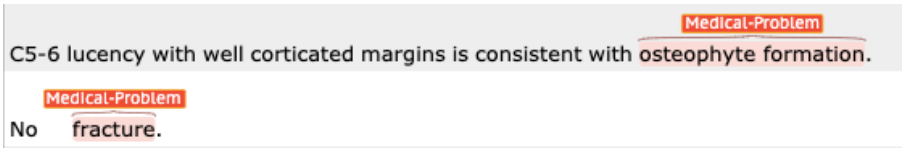
Annotating a text span involves selecting the entire text span with a mouse, and then choose one of the entities on Figure 2. If the entity is categorical, a drop-down box will be presented on the window. Choose the appropriate value from the drop-down box.

(1) Medical problem finding

Medical problem findings are abnormal pathological process uncovered by the radiology imaging test, such as cirrhosis, air-trapping, fracture, and effusion. A medical finding includes problem description, affected anatomy, and assertion.

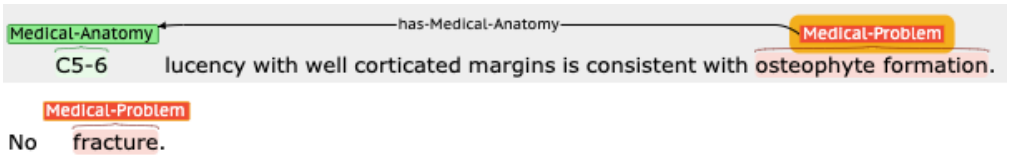
1.1 Problem Description (required)

The description of medical problem serves as the event trigger. The text span can be a multi-word phase that identifies the actual medical problem, such as "**osteophyte formation**" and "**fracture**".



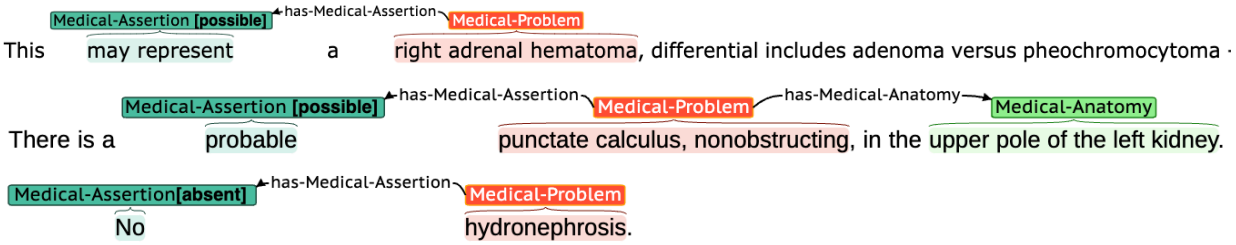
1.2 Anatomy

Medical finding anatomy entity is a text span which captures one or more body parts associated with the medical problem, such as "**C5-6**" in the following example. Notice that a link "**has-Medical-Anatomy**" needs to be created from the trigger "**osteophyte formation**" to "**C5-6**" indicating their relationship. The link can be created on the UI by simply dragging an arrow from the event trigger to the anatomy entity.



1.3 Assertion

Assertion is a categorical value (*possible, absent*) indicating the likelihood of the medical problem. The following shows some assertions highlighted in dark green. Assertion has a default value of *present*. If no other explicit assertion value is annotated for the medical problem, it is implied that the medical problem has a *present* assertion.



A link needs to be created from the event trigger to each assertion entity indicating their relationship. The following table presents some examples of each category. The underlined text spans are the medical problem event triggers.

value	**Examples**
possible	There is a ***possible*** nondisplaced L5 spinous process fracture. Liver: There is a mildly nodular contour of the liver as before, ***possibly*** representing cirrhosis.
Present (default)	Calcified atherosclerosis of the LAD. C5-6 lucency with well corticated margins is consistent with osteophyte formation.
absent	***No evidence*** of radiopaque nephrolith. Visualized osseous structures show ***no*** acute osseous abnormality.

When annotating a text span with this entity type or any categorical types, make sure a corresponding entity attribute is selected from the bottom drop-down on the BRAT tool. Select the values from the corresponding drop-down for each type as shown in figure 3.

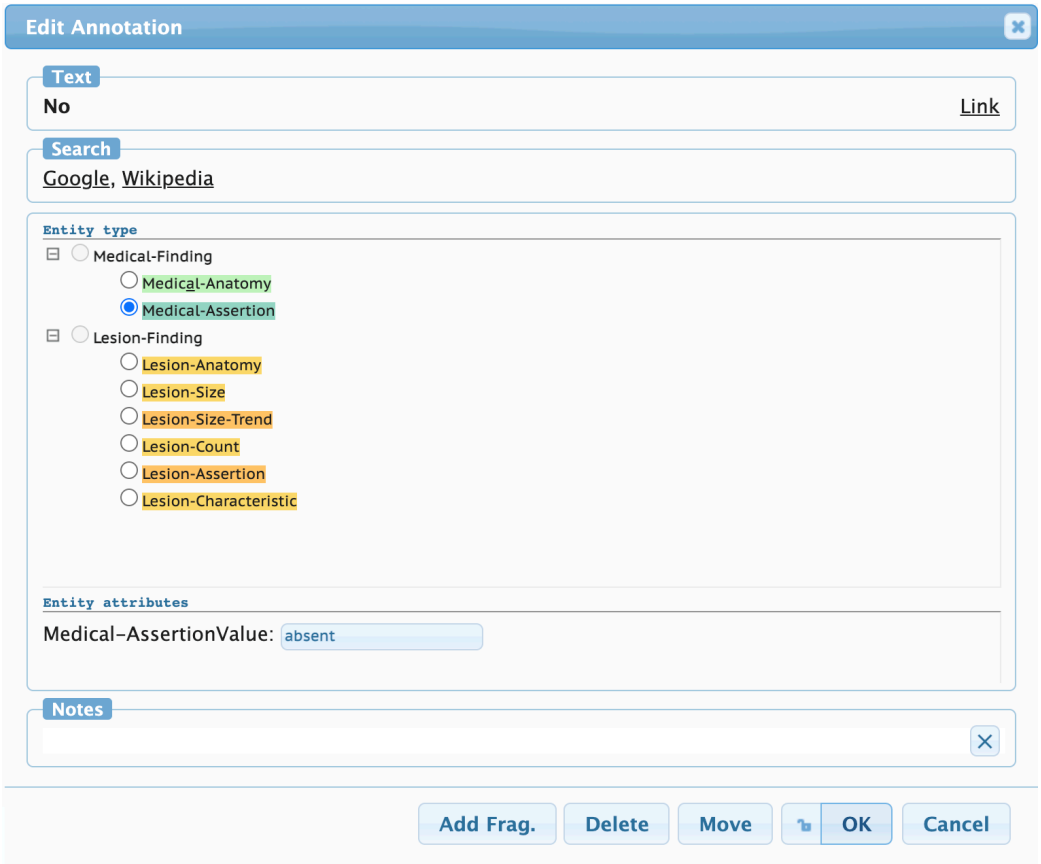


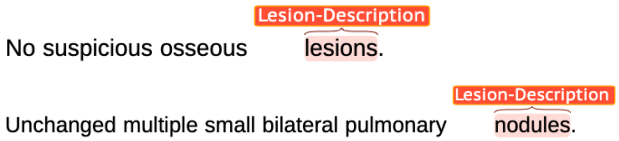
Figure 3. Brat annotation screen showing categorical drop-down at the bottom.

(2) Lesion finding

Lesion finding describes the extent of lesion development that can be observed on the imaging, which includes description, anatomy, lesion size, size trend, count and assertions. Noun phrases containing anatomical location as part of a lesion description, (e.g. *brain lesion* or *pulmonary nodules*) should be annotated as two separate entities, i.e. lesion-anatomy (*brain, pulmonary*) and lesion-description (*lesion, nodules*).

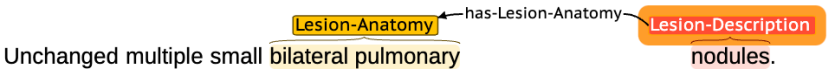
2.1 Lesion description (required)

The description of lesion finding serves as the event trigger and is mandatory. Common text spans are ("mass", "node", "nodule", "nodular opacity", "lesion"). "Opacity" on its own is considered a medical problem.



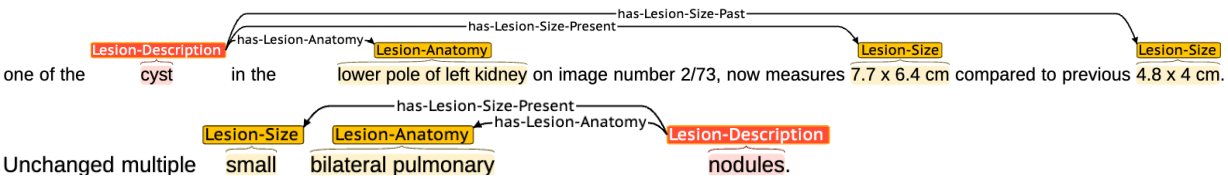
2.2 Anatomy

Lesion anatomy entity is a text span capturing one or more body parts where the lesion is located, such as "**bilateral pulmonary**" in this example. The links labelled "**has-Lesion-Anatomy**" indicate the relations between the **Lesion-Anatomy** entities and the corresponding lesion descriptions.



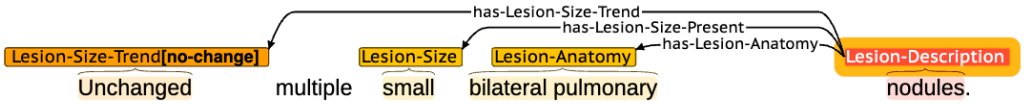
2.3 Size

Some lesion descriptions contain size, such as "**7.7 x 6.4 cm**" and "**4.8 x 4 cm**" in the example. The links labelled "**has-Lesion-Size-Present**" indicate the relations between the **Lesion-Size** entities and the corresponding lesion descriptions. A separate relation "**has-Lesion-Size-Past**" should be used to link to lesion sizes in the past exams.



2.4 Size trend

Some lesion descriptions contain size trend which is a categorical value (new, increasing, deceasing, no-change), such as the word "**Unchanged**" in the following example. The links labelled "**has-Lesion-Size-Trend**" indicate the relations between the **Lesion-Size-Trend** entities and the corresponding lesion descriptions.

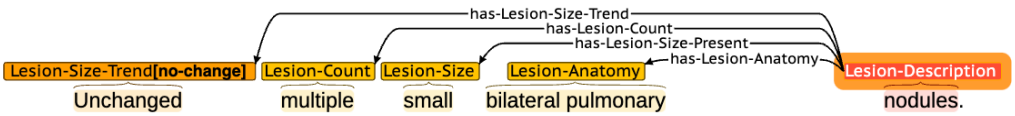


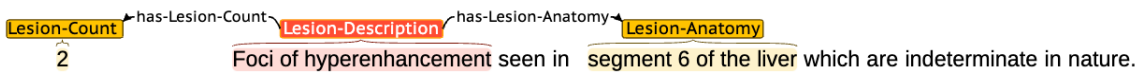
The following table presents some examples of each category. The underlined text spans are the lesion finding event triggers.

value	**examples**
new	On the current exam, there is a ***newly*** identified hypodense well-delineated mass
increasing	One nonhypermetabolic lymph node which is **increased in size**. More peripherally, there is 10 x 9 mm nodule adjacent to the suture line (4/63) which is ***gradually increasing in size*** since 2/17/2017, too small to characterize on PET .
decreasing	***Decreasing size*** of the hypodense lesion within the inferior aspect of the right hepatic lobe now measuring 0.6 cm compared to 1.5 cm on 06/04/2014 The mass in the proximal ureter has ***decreased significantly in size***, currently measuring 3 mm (4/104), ***decreased*** from 7 x 8 mm.
no-change	There is a hypoattenuating left adrenal nodule that has been increasing in size since 2009 though it is ***unchanged*** since May. Enlarged inferior mediastinal and right hilar lymph nodes are ***unchanged*** since January.

2.5 Count

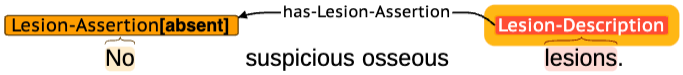
Some lesion findings include the number of nodules or lesions, such as "**multiple**" in the following example. The link "**has-Lesion-Count**" indicates the relation between the Lesion-Count entity and the corresponding lesion description.





2.6 Assertion

Assertion is a categorical value (possible, absent) indicating the likelihood of the lesion finding, such as the word "**no**" in this example. The link "**has-Lesion-Assertion**" indicate the relation between the assertion entity and the corresponding lesion description. Like Medical Assertion, Lesion Assertion also has a default value of *present*. If no other explicit assertion value is annotated for the lesion finding, it is implied that the lesion finding has a *present* assertion.

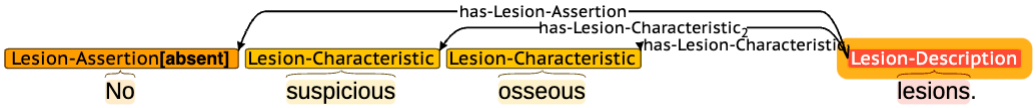


The following table presents some examples of each category. The underlined text spans are the lesion finding event triggers.

value	**Examples**
possible	***Cannot completely exclude*** mass. Focal lesion seen in segment five shows delayed phase contrast washout indeterminate in nature, ***possibly*** dysplastic nodule vs. low grade HCC.
present (default)	Stable segment 7 metastasis status post radiation therapy. Intense FDG uptake (max SUV 17.1) is noted within 27 x 21 mm nodule in left lower lobe (4/76), consistent with biopsy-proven invasive adenocarcinoma.
absent	Findings: ***No*** suspicious enhancing nodule is seen. ***No*** obvious intracystic septations or mural nodularity are seen.

2.7 Characteristic

Characteristic attribute indicates the lesion characteristics such as the word "**osseous**" in this example. The link "**has-Lesion-Characteristc**" indicates the relation between the Characteristic entity and the corresponding lesion description.



Notes:

- Avoid annotate articles (e.g. a, an, the), and unnecessary adjectives.
- Avoid annotate overlapping text spans. i.e. text spans overlapped with more than one annotation.
- Typical lesion description noun phrases ("mass", "node", "nodule", "nodular opacities", "lesion")
- "opacity" itself is considered a medical problem.
- Noun phrases containing anatomical location as part of a lesion description, (e.g. *brain lesion* or *pulmonary nodules*) should be annotated as two separate entities, i.e. lesion-anatomy (*brain, pulmonary*) and lesion-description (*lesion, nodules*).
- Avoid annotating assertion modifiers ('likely', 'possible') for non-triggers. E.g. in the span, *much less likely a metastatic lesion*, "much less likely" is not Lesion-Assertion (possible) for "lesion", as it is describing the extent of metastasis.