

Concept Annotation Guidelines

The concept extraction task builds toward the assertion and relation tasks of the challenge. Therefore, from the wealth of concept and concept types that are found in clinical records, only a small subset that directly contribute to these tasks have been identified.

This guideline describes the specific types of information that should be annotated for the concept extraction task and provides examples similar to those that may be found in the clinical texts. The instances that should be marked along with examples of the surrounding text that should be included in the annotation are described. Instances in this guideline marked in **BLUE** are correctly annotated concepts. Instances marked in **RED** are terms that should not be marked.

General Guidelines for Concepts

Only complete noun phrases (NPs) and adjective phrases (APs) should be marked. Terms that fit concept semantic rules, but that are only used as modifiers in a noun phrase should not be marked.

- *The man was **obese** . The **obese** man came to the clinic .*
- *She developed **diabetes** . She takes **diabetes** medication .*
 - ***diabetes medication** is one concept and the full NP should be marked.*
- *Patient underwent **catheterization** . **Catheterization** report showed ...*
- *Patient arrived in the **surgery** suite . **Surgery** was performed .*

Include all modifiers with concepts when they appear in the same phrase except for assertion modifiers. (Please see assertion annotation guidelines for a description).

- ***bilateral DVT***
- ***some recurrent angina***
- ***high grade LAD lesion***
- ***chronic hepatitis***
- ***cataract surgery***
- ***diabetes medication***
- ***head CT***
- ***chest X-ray***
- ***no fever***
- ***possible tamponade***

You can include up to one prepositional phrase (PP) following a markable concept if the PP does not contain a markable concept and either indicates an organ/body part or can be rearranged to eliminate the PP (we later call this the PP test).

- *pain in chest* -> indicates body part and can be rearranged to *chest pain*
- *changes in mental status* -> can be rearranged to *mental status changes*
- *defect in lobe of liver* -> only mark the first PP, can be rearranged to *lobe defect*
- *removal of mass* -> PP contains another concept, so mark as two separate concepts: *removal* and *mass*

Include articles and possessives

- *his cancer*
- *her medications*
- *the surgery*
- *an appendectomy*
- *the patient 's tumor*

Conjunctions and other syntax that denote lists should be included if they occur within the modifiers or are connected by a common set of modifiers. If the portions of the list are otherwise independent, they should not be included. Similarly, when concepts are mentioned in more than one way in the same noun phrase (such as the definition of an acronym or where a generic and a brand name of a drug are used together), the concepts should be marked together.

- *metastases in the liver and pancreas*
- *behavioral / mental changes*
- *diarrhea , nausea , and vomiting* -> three separate concepts *diarrhea*, *nausea* and *vomiting*
- *Zocor (Simvastatin) 20 mg p.o. once daily .*
- *DM (diabetes mellitus)*

Concepts should be mentioned in relation to the patient or someone else in the note. Section headers that provide formatting, but that are not specific to a person are not marked.

- *HISTORY OF PRESENT ILLNESS :*
- *The patient 's illness continued to worsen .*
- *DISCHARGE MEDICATIONS :*
- *The patient should remain on his home medications .*

Categories of Concepts

Concepts are defined in three general categories that are each annotated separately.

- 1) *Medical Problems*: phrases that contain observations made by patients or clinicians about the patient's body or mind that are thought to be abnormal or caused by a disease. They are loosely based on the UMLS semantic types of pathologic functions, disease or syndrome, mental or behavioral dysfunction, cell or molecular dysfunction, congenital abnormality, acquired abnormality, injury or poisoning, anatomic abnormality, neoplastic process, virus/bacterium, sign or symptom, but are not limited by UMLS coverage.

What should be annotated?

- Medical problems need to be things that are wrong with the patient and can be treated.
- They need to be complete noun phrases (NPs) or adjective phrases (APs).
- They need to be able to be modified with a negation modifier.
- They need to belong to one of the above semantic types, but do not need to be UMLS concepts.

All four of these conditions have to hold.

Examples of Medical Problems to annotate:

- a) Phrases that name a disease, syndrome, sign, or symptom
 - *Ativan 0.25 to 0.5 mg IV q 4 to 6 hours prn anxiety*
 - What was treated? *anxiety*. (can also be negated: "no anxiety")
 - *the wound was noted to be clean with mild serous drainage*
 - Question test: What was treated? *The wound*.
 - Question test: What was treated? *mild serous drainage*.
 - Question test: What was treated? *clean?* -> does not fit the definition of something being wrong with the patient.
 - *an echocardiogram revealed a pericardial effusion and possible tamponade clinically.*
 - *the drop in hematocrit was secondary to ...*
 - PP Test: can be rearranged to *the hematocrit drop*
 - *the patient has had increasing dyspnea on exertion*
 - *he underwent a resection of a skull base chordoma*
- b) Mental or behavioral status observations
 - *His mental status changes remained stable*
 - *the patient developed a sudden change in her mental status*
 - PP test: *a sudden mental status change*
 - *she did well except for some episodes of confusion*
 - *the patient seemed subdued*
 - *the patient was unresponsive*

- c) Virus and bacterium
 - *Blood cultures were positive for **S. Veridans***
 - *procured sample to rule out **MRSA***
- d) Injury
 - *patient arrived with **a broken arm***
 - *examined **the deep gash in her head***
- e) Abnormalities
 - ***The defects** were found*
 - *chest x-ray revealed **an abnormality***
- f) Test results explicitly stated to be abnormal
 - ***low blood pressure***
 - ***moderately decreased ejection fraction***

Examples of concepts that should **not be marked** as Medical Problems:

- a) Statements about normal states of health
 - *she had been **comfortable** on the Morphine drip*
 - ***wound** was **clean***
 - *He was **dynamically stable** on his Dobutamine*
 - *converted back to **normal sinus rhythm***
 - *serum uric acid was **normal***
- b) Physiologic measurements, vital signs, or the tests that measure them and statements of bodily function (even if the value of the test could be inferred to be outside normal range)
 - ***EF 35%***
 - ***Blood pressure 145/95***
 - *will monitor **BUN / Cr** after giving Lasix*
 - ***the patient 's respiratory status** continued to decline*
 - What was treated? *Respiratory status?* → wrong
 - ***hematocrit** dropped*
 - What was treated? *Hematocrit?* -> wrong
- c) Verbs that describe the outcome of an event
 - *the patient **defervesced***
 - ***the tumor** was **growing***
 - *the patient 's respiratory status continued to **decline***
 - *the patient 's LDH **climbed***
- d) Alcohol, illicit drug, and tobacco use
 - *2 packs **cigarettes** per day*
 - *moderate **alcohol** use*
 - *he uses **recreational drugs***

- 2) *Treatments*: phrases that describe procedures, interventions, and substances given to a patient in an effort to resolve a medical problem. They are loosely based on the UMLS semantic types therapeutic or preventive procedure, medical device, steroid, pharmacologic substance, biomedical or dental material, antibiotic, clinical drug, and drug delivery device. Other concepts that are treatments but that may not be found in UMLS are also included. Treatments that a patient had, will have, may have in the future, or are explicitly mentioned that the patient will not have are all marked as treatments.

What should be annotated?

- In general, the annotated concepts need to be things that can treat a problem. They need to sensibly answer the question “what fixed the medical problem?”
- They need to be complete noun phrases (NPs) or adjective phrases (APs).
- They need to belong to one of the above semantic types, but do not need to be UMLS concepts.

All three of these conditions have to hold.

Examples of Treatments to annotate:

- a) Medication names, brand names and generics as well as collective names for groups of medications. Include amount, dose, frequency, and mode only when used as modifiers - do not mark when part of *signatura*
- *Lasix 20 mg b.i.d. by mouth*
 - *800 mg ibuprofen*
 - *She was started on IV Ciprofloxacin*
 - *regular sliding scale insulin*
 - *Vitamin K therapy , low-dose ASA , statin*
 - *Patient was placed on a morphine drip*
- b) Biological substances
- *creatinine improved with hydration and packed red blood cells*
 - *the patient remained on IV hydration therapy*
 - *lipids*
 - *he did not require a transfusion*
- c) Drug and treatment delivery devices
- *ICD shocks*
 - *he was responsive to having an NG feeding tube placed with Jevity feeds started*
 - *the patient uses her respirator at night .*
 - *the patient remained hemodynamically stable on the ventilator*
- d) Procedures and devices or hardware involved in those procedures
- *the patient had a bronchoalveolar lavage performed*
 - *After tube removal , she was started on...*
 - *Significant for radiation therapy after his surgery*
 - *the staples were removed*
 - *status post myocardial infarction with stent placement*

- *Total abdominal hysterectomy, bilateral salpingoophorectomy, bilateral lymph node dissection, omentectomy, lysis of adhesions, removal of mass from surface of transverse colon*
- *after months of physical therapy, the patient gained strength*
- *the patient required a cane*

e) General terms referring to the patient's treatments

- *the patient's asthma medication*
- *this drug seemed to reduce ...*
- *her treatment*

Examples of concepts that should not be marked as Treatments:

a) Verbs that indicate application of a treatment

- *Her dressing was changed*
- *...had PPM and AICD placed*
- *Tube was removed*
- *The patient was intubated at that time*
- *he had the right knee aspirated*
- *The sheaths were removed and heparin was restarted*

b) Locations and persons providing treatment

- *need consult with rheumatology*
- *patient discharged to physical therapy*
- *patient admitted to ICU*

3) *Tests*: phrases that describe procedures, panels, and measures that are done to a patient or a body fluid or sample in order to discover, rule out, or find more information about a medical problem. They are loosely based on the UMLS semantic types laboratory procedure, diagnostic procedure, but also include instances not covered by UMLS.

What should be annotated?

- In general, the annotated concepts need to be things that help determine whether the patient has a problem. Tests are analogous to treatments, but instead of *fixing* the problem, they are used to find more information out about a problem. They need to sensibly answer the question "what was done to find out if there was a medical problem?"
- They need to be complete NPs or APs.
- They need to belong to one of the above semantic types, but do not need to be UMLS concepts.

All three of these conditions have to hold.

Examples of Tests to annotate:

- a) Procedures performed on the patient
 - *An abdominal ultrasound* was performed showing no stones .
 - *Chest x-ray* revealed clear lungs
 - *Cardiac catheterization* revealed coronary artery lesions
 - *An echocardiogram* revealed *an ejection fraction* of 25%
 - *A lung biopsy* was performed , revealing chorio carcinoma pathologically
 - *EKG* revealed normal sinus rhythm with no ST changes .
 - *A mediastinoscopy* was also performed at that time .

- b) Panels and tests run on patient body fluids
 - *His urinalysis* showed 10-20 granular casts
 - *CBC* was unremarkable .
 - *Metabolic 20 panel* was unremarkable .
 - *Blood cultures* were positive for *S. Veridans*
 - *WBC* : 12.8
 - *CRE* : 1.2

- c) Physiologic measures and vital signs
 - *Blood pressure* 120/80
 - *Pulse* 40

- d) Examinations and evaluations of the patient
 - *Cardiac exam* revealed an irregular rate
 - *Rectal exam* was heme negative
 - rash apparent upon *evaluation*
 - *workup* for liver cancer

Examples of concepts that **should not** be marked as Tests:

- a) Verbs that indicate application of a treatment
 - *The patient was **evaluated*** for repair of false femoral aneurysm
 - *A lung biopsy* was **performed** , revealing chorio carcinoma pathologically
 - *EKG **revealed*** normal sinus rhythm with no ST changes .
 - *A mediastinoscopy* was also **performed** at that time .

- b) Test values and measurements
 - *Blood pressure* **120/80**
 - *Pulse* : **40**
 - *HCT* : **30.1**

- c) Mentions of tests that are stated as problems (these are marked as problems) or that are mentioned as normal values
 - *The patient has **high blood pressure***
 - ***elevated enzymes** were noted*
 - *normal white blood cell count*