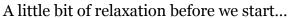# HebUMLS Guidelines for the Annotation of Disorders and Chemical and Drugs in OHC Posts

Latest draft – 05 Mar, 2020 Original
draft – January 2011
Primary Author – Noémie Elhadad
Participants – Yonatan Biton, Michael Elhadad

Derived from Share Guidelines 2011 by Noemie Elhadad

A little bit of relaxation before we start...

Contents

Overview

The goal of this annotation task is to go through posts from Online Health Communities (OHC), find mentions of disorders and chemicals and drugs, and annotate them. In a second step, we will verify whether a specific UMLS CUI (Concept Unique Identifier) corresponds to the mentions annotated in the first step.

1. Identifying a Mention of a Disorder (span)

1.1. A disorder mention is a concept in UMLS part of the Disorder Semantic Group

We define a **disorder mention** if as any **span of text** that can be **mapped to a concept** in the **UMLS** terminology, which belongs to the **Disorder semantic group**. A concept is in the Disorder semantic group if it belongs to one of the following UMLS semantic types:

- Congenital Abnormality
- Acquired Abnormality
- Injury or Poisoning
- Pathologic Function
- Disease or Syndrome
- Mental or Behavioral Dysfunction
- Cell or Molecular Dysfunction
- Experimental Model of Disease
- Finding
- Anatomical Abnormality
- Neoplastic Process
- Signs and Symptoms

Example 1.1.a

This is a gentleman who presents with colon cancer .
מדובר בגבר עם סימנים של סרטן מעי הגס

The span "colon cancer" is a disorder mention. It corresponds to CUI C0007102 (Preferred term: Malignant tumor of colon) in UMLS, which is under the semantic type "Neoplastic Process."

Example 1.1.b

No chest pain .
בלי כאבים בחזה

The span "chest pain" is a disorder mention. It corresponds to CUI C0008031 (Preferred term: Chest Pain) in UMLS, which is under the semantic type "Sign or Symptom."

Example 1.1.c

The patient's glucose is 107.

There is no disorder mention in this sentence.

Example 1.1.d

The patient was continued on nebulizers while she was admitted.

There is no disorder mention in this sentence.

1.2 A disorder mention is a reasonable synonym of the lexical form of a UMLS Disorder Group CUI

A disorder mention must be annotated, even if the text span is not one of the lexical forms in the terminology, but is a **reasonable synonym**. **Example 1.2.a**

He felt left-sided   `facial droop`

The span "facial droop" is a disorder mention. The lexical form is not part of UMLS, but can be mapped to CUI C0427055 (Preferred term: Facial Paresis) in UMLS.

Example 1.2.b

His `face` was `weak`.

The disjoint span "face … weak" is a disorder mention. Even though it can be mapped to CUI C0427055 (Preferred term: Facial Paresis), we **do not annotate disjoint terms.**

1.3 A disorder mention is explicit

A concept should not be annotated unless it is specifically mentioned by name. That is **descriptions of disorders should not be annotated**. Similarly, there should **not be any inference** made to decide that a disorder is mentioned.

Example 1.3.a

The patient's chem profile reveals a 138 sodium, 4.0 potassium, 106 chloride, 19 bicarb, 43 and 2.2 are the BUN and creatinine, indicating probable blood in the GI tract.

There is no disorder mention in this sentence, even though it describes the disorder GI bleed.

Example 1.3.b

EF 10-15%.

There is no disorder mention in this sentence, even though an ejection fraction of 10-15% strongly indicates the disorder congestive heart failure.

Example 1.3.d

She had a   `Crohn`   flare with symptoms of rehydration and   `bowel obstruction`   Her current symptoms is reminiscent of this.

The first sentence contains two disorder mentions ("Crohn" and "bowel obstruction"). There is no disorder mention in the second sentence: "this," while referring to a disorder, is not a disorder mention.

1.4 A disorder mention is the most specific disorder conveyed in the text

Only the **most specific named disorder** may be annotated.

Example 1.4.a

The patient has a <mark>small bowel obstruction</mark>.

The sentence contains only one disorder mention, "small bowel obstruction." It corresponds to CUI C0235329 (preferred term: Small bowel obstruction). Only one mention is annotated, even though there are sub-spans of text corresponding to other disorders: the span "obstruction" can be mapped to CUI C0028778 (preferred term: Obstruction), and the span "bowel obstruction" can be mapped to CUI C0021843 (preferred term: Intestinal Obstruction). Because "small bowel obstruction" is more specific than "bowel obstruction" and "obstruction," it is the disorder chosen for this annotation.

Example 1.4.b

The patient was found to have left <mark>lower extremity DVT</mark>.

The sentence contains only one disorder mention, "lower extremity DVT." It corresponds to CUI C0340708 (preferred term: Deep vein thrombosis of lower limb).
The span "DVT" can be mapped to CUI C0149871 (preferred term: Deep Vein Thrombosis) is present, but is not annotated because it is part of a more specific disorder in the sentence.

Example 1.4.c

The patient has <mark>severe pre-eclampsia</mark>.

The sentence contains only one disorder mention, "severe pre-eclempsia." It corresponds to CUI C0341950 (preferred term: Severe pre-eclampsia). The sub-span "pre-eclampsia" can be mapped to CUI C0032914 (preferred term: Pre-eclampsia) but is not annotated as it is more general.

Example 1.4.d

The patient has <mark>chronic gingivitis</mark>.

The sentence contains only one disorder mention, "chronic gingivitis." It corresponds to CUI C0008684 (preferred term: chronic gingivitis). The sub-span "gingivitis" can be mapped to CUI C0017574 (preferred term: gingivitis) but is not annotated as it is more general.

1.5 A disorder mention does not contain any negation in its span

When looking for a concept in the terminology, do not map to concepts that have negations in them.

Example 1.5.a

No pain.

The span "pain" is a disorder mention (C0030193, preferred term: Pain), even though there is a CUI for the span "no pain" (C0234225, preferred term: Absence of pain).

1.6 A disorder mention does not contain any mention of temporality in its span

When looking for a concept in the terminology, do not map to concepts that have information about past or history.

Example 1.6.a

Hx of stroke.

The span "stroke" is a disorder mention (C0038454, preferred term: Cerebrovascular accident), even though there is a SNOMED-CT Disorder CUI for the span "Hx of stroke" (C0559159, preferred term: "H/O: CVA").

1.7 A disorder mention is not linked to any syntactic construct

Most of the time, a disorder mention is either a whole or a subset of a noun phrase, including any prepositional phrase attached to a noun phrase. But in some cases, the disorder mention is not conveyed through a noun phrase only.

Example 1.7.a

I had a tumor of the skin removed.

The span "tumor of the skin" is a disorder mention. It corresponds to CUI C0037286 (preferred term: Skin Neoplasms). The determiner at the edge of the span "a" is not part of the annotation, but the preposition "of" is part of the disorder mention.

Example 1.7.b

A tumor was found in the left ovary.

The disjoint span "tumor … ovary" is a disorder mention. It corresponds to C0919267 (preferred term: ovarian neoplasm). We do NOT annotate it because it is a disjoint mention.

1.8 The span for a disorder cannot be disjoint

We do not annotate mentions for a disorder that are disjoint.

Example 1.8.a

A tumor was found in the left ovary.

We do NOT annotate the disjoint span "tumor … ovary" even though it corresponds to C0919267 (preferred term: ovarian neoplasm).

1.9 A disorder mention is annotated even if it does not pertain to a specific patient or to the author of the post

Any mention of a disorder should be annotated in a post, even if a particular disorder mention does not pertain to the author of the post.

Example 1.9.a

My son has <mark>schizophrenia</mark> .

The span "schizophrenia" is a disorder mention (C0036341, preferred term: Schizophrenia). The fact that the author's son does not change the annotation.

Example 1.9.b

He told me to return to the ED immediately if any <mark>rash</mark> occurs.

The span "rash" is a disorder mention (C0015230, preferred term: Exanthema), even though the author never have any, does not have it currently, and may never have any.

1.10 All the disorder mentions in a post must be annotated

Every disorder mention must be annotated, no matter (i) what section of the post it appears in and (ii) whether a previous mention of the same disorder was already mentioned.

1.11 A span of text should be annotated as disorder mention only if it is being used as a disorder, and not if it's being used as other purpose

We do not talk about a "disease" - we talk about "a mention of a disease in context" - that is, a word that refers to an actual disease that someone is suffering from.

Example 1.11.a

Heat regulation
ויסות חום

There isn't a disorder mention in this text. There is a mention "heat feeling" in the UMLS (C0234657, preferred term: heat feeling), however this text doesn't mean to mention heat as a disorder.

Example 1.11.b

Most patients state that they do not feel anything
רוב החולים מדווחים שהם לא הרגישו שום דבר

There isn't a disorder mention in this text. There is a mention "Feeling Sick" in the UMLS (C0857027, preferred term: Feeling Sick). The Hebrew translation of "Feeling Sick" is "חולים", and it is the same as the Hebrew translating of "Patients". However this text doesn't mean to mention "חולים" as a disorder, but as the noun: patients.

Example 1.11.c

Diabetes clinic
מרפאת סכרת

There isn't a disorder mention in this text. Diabetes is a disorder in the UMLS, but in this context is isn't meant as disease. Nobody suffers from the disease in this context.

Example 1.11.d

He felt mental stress
הוא חש לחץ נפשי

There is a disorder mention. Preferred term: mental stress, CUI: C0038443. In this context the mention mental stress describes disorder that someone is suffering from.

Example 1.11.e

John suffers from diabetes
ג'ון סובל מסכרת

There is a disorder mention. Preferred term: Diabetes, CUI: C0011847. In this context the mention diabetes describes disorder that someone is suffering from.

Example 1.11.f

John works in a diabetes hospital
ג'ון עובד בבית חולים לסכרת

There is no disorder mention here. The term diabetes in this context is NOT describing a disorder that someone is suffering from.

2. Annotating Chemical and Drugs Mentions

2.1. A Drug mention is a concept in UMLS part of the Chemical and Drugs Semantic Group

The following semantic types are covered by this UMLS semantic group:

- T116   Amino Acid, Peptide, or Protein
- T195   Antibiotic
- T123   Biologically Active Substance
- T122   Biomedical or Dental Material
- T103   Chemical
- T120   Chemical Viewed Functionally
- T104   Chemical Viewed Structurally
- T200   Clinical Drug
- T196   Element, Ion, or Isotope
- T126   Enzyme
- T131   Hazardous or Poisonous Substance
- T125   Hormone
- T129   Immunologic Factor
- T130   Indicator, Reagent, or Diagnostic Aid
- T197   Inorganic Chemical
- T114   Nucleic Acid, Nucleoside, or Nucleotide
- T109   Organic Chemical
- T121   Pharmacologic Substance
- T192   Receptor
- T127   Vitamin

2.2 Mention Detection Rules
The principles of detecting mentions of Drugs are similar to those that apply for Disease mentions annotation:

- A drug mention is a reasonable synonym of the lexical form of a UMLS Chemical and Drugs Group CUI
- A drug mention is explicit
- A drug mention does not include the dose
- A drug mention does not contain any negation in its span
- A drug mention does not contain any mention of temporality in its span
- The span for a drug mention cannot be disjoint
- A drug mention is annotated even if it does not pertain to the author of the post
- All the drug mentions in a post must be annotated
- A span of text should be annotated as drug mention only if it is being used as a drug, and not if it's being used as other purpose

2.11 A span of text should be annotated as chemical or drug mention only if it is being used as a chemical or drug, and not if it's being used as other purpose

Example 2.11.a

I want to inject the new medicine
אני רוצה להזריק את התרופה החדשה

There isn't a chemical or drug mention for "inject", there is for "medicine". There is a mention "injection" in the UMLS (C1272883, preferred term: Injection), however this text doesn't mean to mention injection as a chemical or drug, but as the verb: inject.

Example 2.11.b

Wrote on my tablet
כתבתי על הטאבלט שלי

There isn't a chemical or drug mention in this text. There is a mention "tablet" in the UMLS (C0039225, preferred term: tablet), however in this text the word tablet refers to a tablet computer, and not a chemical or drug mention.