Annotation Guidelines for the Adverse Drug Event (ADE) and Medication Extraction Challenge

Written by Kevin Buchan
Edited by Marilena Di Bari and Amber Stubbs

1. Overview

The task hereby proposed is aimed at creating a set of medical documents in which different information related to drugs are annotated. The task actually consists of two subtasks:

1. Identifying drug names, dosages, durations and other entities.
2. Create two different relations:
   - the relation of drugs with specific symptoms and diseases;
   - the relation of drugs with adverse drugs events (ADEs).

Both entity tags and attributes used to indicate the presence of drug information are in accordance to those found in discharge summaries for patients who experienced an ADE while in the hospital. The data was collected from the MIMIC-III (Medical Information Mart for Intensive Care III) clinical care database[1].

2. Background and definitions

The World Health Organization (WHO) gives the following definitions:

- An adverse drug event (ADE) as "an injury resulting from medical intervention related to a drug"
- An adverse drug reaction (ADR) as, "a response to a drug that is noxious and unintended and occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or for modification of physiological function[2]."

This means that ADRs are a subset of ADEs, which only include appropriate use of drugs. Therefore, drug administration errors are excluded in the accepted ADR definition.

These well-established definitions for ADE and ADR, which have been used for almost 50 years[3], continue to be used in a variety of recent research studies[4]-[6].

One example demonstrating the difference between ADE and ADR definitions was published by Bates et al. in a well-cited research article:

"Oversedation and aspiration pneumonia resulting from a 10-fold overdose of a drug would not be considered an ADR according to the WHO definition, but would be an ADE[7]."

In their publication, Bates et al. state that the ADE definition is preferred because it is both "more comprehensive and clinically significant than the ADR"[7]. For this reason, our aim is to automatically detect ADEs, defined as injuries resulting from medical interventions related to a drug. In addition to ADEs, we are interested in accurately detecting drug names, dosages, durations, routes and other drug relations.

3. Tags

Each drug has a set of potential Entity Tags that will be used to describe information that we are interested in capturing. This section discusses each Entity Tag type that will be used in this annotation project. Note that Entity Tag types should only be used for individual events and not patterns of events.

In the following examples the underlined terms or phrases highlighted in yellow should be annotated, whereas the underlined terms or phrases in red text should not be annotated. As for as the medical context goes, please note that the premise for all the examples below is that the patient is experiencing muscle pain, secondary to statin therapy for coronary artery disease.

To speed up annotation, an index with the most common medical abbreviations follows.

IV = Route
IVF = Drug
IV Fluids = Drug
PRBCs = Drug

PRN= Frequency
QD = every day = Frequency
bid = twice a day = Frequency

PO = by mouth = Route
Gtt = drops = Route

3.1. Drug

This Entity Tag type should be selected to specify a drug. The Drug Entity Tag type should include classes of drugs, as well. However, illicit drugs and alcohol should not be annotated. For example:

1. *Patient is experiencing muscle pain, secondary to statin therapy for coronary artery disease.*
2. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg Prednisone tablet daily for 5 days.*
3. *Patient has been switched to lisinopril tablet 10mg 1 tablet PO QD.*

4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*
5. *Sulfa (sulfonamide antibiotics)*
6. *Tylenol (Acetaminophen)*
7. *B.i.d. (twice a day)*

3.2. Strength

This Entity Tag type should be selected to specify the strength of a drug. For example[1]:

1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg Prednisone tablet daily for 5 days.*
2. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
3. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*
4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

3.3. Dosage

This Entity Tag type should be selected to specify a drug dosage. For example:

1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20mg Prednisone tablet daily for 5 days.*
2. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
3. *Patient has been switched to lisinopril tablet 10mg 1 tablet PO QD.*
4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*
5. *2ml PO bid*[2]
6. *Sliding scale*
7. *Bolus*
8. *Taper (as in steroid taper)*

[1]

[2] Note: The tags ml and units should be included in the annotation for liquid meds only. If it is a tablet or a pill form of drug, just 2 needs to be annotated as dose.

3.4. Duration

This Entity Tag type should be selected to specify a drug duration. For example:

1. *The patient suffers from steroid-induced hyperglycemia. The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20mg Prednisone tablet daily for 5 days*[3].
2. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*
3. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

3.5. Frequency

This Entity Tag type should be selected to specify a drug frequency. For example:

1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20mg Prednisone tablet daily for 5 days.*
2. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
3. *Patient has been switched to lisinopril tablet 10mg 1 tablet PO QD.*
4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*
5. *tylenol PRN*
6. *tylenol Q8H PRN*[4]
7. B.I.D
8. Around the clock

3.6. Form

This Entity Tag type should be selected to specify a drug form. For example:

1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20mg Prednisone tablet daily for 5 days.*
2. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
3. *Patient has been switched to lisinopril 10mg 1 tablet PO QD. Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

3.7. Route

[3] Note: The preposition "for" needs to be included in the annotation only when related to the Duration tag. See difference with the Reason tag.

[4] PRN = frequency. If it is by itself like in sentence 7, just PNR has to be tagged. If, instead, it is attached to another frequency such as Q8H in sentence 8, the entire 'Q8H PRN' has to be tagged as frequency

This Entity Tag type should be selected to specify the route of administration for a drug. For example:

1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20mg Prednisone tablet daily for 5 days.*
2. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
3. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*
4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*
5. *Gtt*
6. *Drip*
7. *Inhalation*
8. *Topical*

3.8. Reason

This Entity Tag type should be selected to specify reasons for drug administration. Reason Entity Tags include signs, symptom and diseases. For example:

1. *Patient is experiencing muscle pain, secondary to statin therapy for coronary artery disease.*
2. *The patient suffers from steroid-induced hyperglycemia.*
3. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
4. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*

Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.

3.9. ADE

This entity tag type should be selected to specify an ADE. For example:

1. *Patient is experiencing muscle pain, secondary to statin therapy for coronary artery disease.*
2. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20mg Prednisone tablet daily for 5 days.*
3. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*
4. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*
5. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

4. How to annotate real instances: some examples

Magnesium hydroxide 400mg/5ml suspension Sig: 30 ml

- Magnesium hydroxide = Drug
- 400mg/5ml = Strength
- Suspension = Form
- 30 ml = Dosage

O2 Sat 95% on 4L

- O2 = drug
- 4L = dosage

2 puffs

- 2 = dosage
- puffs = form

2.5mg for inhalation (of a 2.5mg/3mL solution)

- 2.5mg = strength
- inhalation = route
- 2.5mg/3mL = strength
- solution = form

Patient is on docusate sodium 5mg/ml take 5mg bid.

- docusate sodium = drug
- 5mg/ML = strength
- 5mg = strength
- bid = frequency

Oxycontin 40 mg/day

- Oxycontin = drug
- 40 mg = strength
- day = frequency

6 cycles of RCHOP

- RCHOP = drug
- 6 cycles = dosage

5. Relations

After Entity Tags have been annotated, it is possible to create relations between Entity Tags as described in **Section 3.3 – Creating Relation annotations**. This section discusses each Relation type that will be used in this annotation project. Underlined terms or phrases highlighted in yellow should be annotated as relation annotations.

1. Strength-Drug – this is a relationship between the drug strength and its name. For example:

   1.1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x **20 mg Prednisone** tablet daily for 5 days.*

   Strength-Drug relation*: 20 mg → Prednisone*

   1.2. *Patient prescribed 1-2 **325 mg / 10 mg Norco** pills every 4-6 hours as needed for pain.*

   Strength-Drug relation*: 325 mg / 10 mg → Norco*

   1.3. *Patient has been switched to **lisinopril 10mg** 1 tablet PO QD.*

   Strength-Drug relation*: 10mg → Lisinopril*

   1.4. *Patient received **100 Units/kg** IV **heparin sodium** injection for treatment of deep vein thrombosis.*

   Strength-Drug relation*: 100 Units/kg → heparin sodium*

2. Dosage-Drug – this is a relationship between the drug dosage and its name. For example:

   2.1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed **1** x 20 mg **Prednisone** tablet daily for 5 days.*

   Dosage-Drug relation*: 1 → Prednisone*

   2.2. *Patient prescribed **1-2** 325 mg / 10 mg **Norco** pills every 4-6 hours as needed for pain.*

   Dosage-Drug relation*: 1-2 → Norco*

2.3. *Patient has been switched to **lisinopril** 10mg **1** tablet PO QD.*

Dosage-Drug relation*: 1 → lisinopril*

2.4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

3. Duration-Drug – this is a relationship between a drug duration and its name. For example:

3.1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg **Prednisone** tablet daily for **5 days**.*

Duration-Drug relation*: 5 days → Prednisone*

3.2. *Patient prescribed 1-2 325 mg / 10 mg **Norco** pills every **4-6 hours** as needed for pain.*

Duration-Drug relation*: 4-6 hours → Norco*

3.3. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*

3.4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

4. Frequency-Drug – this is a relationship between a drug frequency and its name. For example:

4.1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg **Prednisone** tablet **daily** for 5 days.*

Frequency-Drug relation*: daily → Prednisone*

4.2. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*

4.3. *Patient has been switched to **lisinopril** 10mg 1 tablet PO **QD**.*

Frequency-Drug relation*: QD → lisinopril*

4.4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

5. Form-Drug – this is a relationship between a drug form and its name. For example:

    5.1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg **Prednisone tablet** daily for 5 days.*

    Form-Drug relation: *tablet* → *Prednisone*

    5.2. *Patient prescribed 1-2 325 mg / 10 mg **Norco pills** every 4-6 hours as needed for pain.*

    Form-Drug relation: *pills* → *Norco*

    5.3. *Patient has been switched to **lisinopril** 10mg 1 **tablet** PO QD.*

    Form-Drug relation. *tablet* → *lisinopril*

    5.4. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

6. Route-Drug – this is a relationship between the route of administration for a drug and its name. For example:

    6.1. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg Prednisone tablet daily for 5 days.*

    6.2. *Patient has been switched to **lisinopril** 10mg 1 tablet **PO** QD.*

    Route-Drug relation: *PO* → *lisinopril*

    6.3. *Patient received 100 Units/kg IV **heparin sodium** **injection** for treatment of deep vein thrombosis.*

    Route-Drug relation: *injection* → *heparin*

7. Reason-Drug – this is a relationship between the reason for which a drug was administered (e.g., symptoms, diseases, etc.) and a drug name. For example:

    7.1. *Patient is experiencing muscle pain, secondary to **statin** therapy for **coronary artery disease.***

    Reason-Drug relation: *coronary artery disease* → *statin*

7.2. *The patient suffers from steroid-induced hyperglycemia. Patient prescribed 1 x 20 mg Prednisone tablet daily for 5 days.*

7.3. *Patient prescribed 1-2 325 mg / 10 mg **Norco** pills every 4-6 hours as needed for **pain**.*

Reason-Drug relation*: pain → Norco*

7.4. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*

7.5. *Patient received 100 Units/kg IV **heparin sodium** injection for treatment of **deep vein thrombosis**.*

Reason-Drug relation*: deep vein thrombosis → heparin sodium*

8. ADE-Drug – this is a relationship between an adverse drug event (ADE) and a drug name. For example:

8.1. *Patient is experiencing **muscle pain**, secondary to **statin** therapy for coronary artery disease.*

ADE-Drug relation*: muscle pain → statin*

8.2. *The patient suffers from **steroid**-induced **hyperglycemia**. Patient prescribed 1 x 20 mg Prednisone tablet daily for 5 days.*

ADE-Drug relation*: hyperglycemia → steroid*

8.3. *Patient prescribed 1-2 325 mg / 10 mg Norco pills every 4-6 hours as needed for pain.*

8.4. *Patient has been switched to lisinopril 10mg 1 tablet PO QD.*

8.5. *Patient received 100 Units/kg IV heparin sodium injection for treatment of deep vein thrombosis.*

Distance between annotations

When annotating a non-drug entity's relationship to a drug entity, simply relate it to the closest occurrence of the drug mention. To determine the "closest" occurence, first check the same sentence; then look at nearby sentences and paragraphs. See the example below for clarification

For example:

Patient is on aspirin(1) for headache(2). Patient is feeling better, as aspirin(3) treated the headache(4) effectively.

You do not need to relate all instances of headache to all instance of aspirin. Simply relate the entity in question to the **closest mention** of the drug. For example:

Annotation 1: Aspirin (1) = Drug entity
Annotation 2: Headache (2) = Reason entity
Relation 1: Headache(2) --> Aspirin(1) = Reason-Drug relation

Annotation 3: Aspirin (3) = Drug entity
Annotation 4: Headache (4) = Reason entity
Relation 2: Headache(4) --> Aspirin(3) = Reason-Drug relation

Note that you do not need to annotate the relations between 1 and 4 or 2 and 3. Linking to the correct entity within the sentence is sufficient.

Appendix 1: Additional Annotation Examples and FAQs

1. Should the same drug be annotated multiple times?

Yes, every mention of a drug should be annotated. This is true even if the same drug is mentioned multiple times in the same paragraph or sentence.

2. Is "DVT prophylaxis" a reason? Same with other types of prophylaxis

Yes, DVT prophylaxis and other types of prophylaxis would be a “reason” someone is prescribed heparin, for example. However, only annotate the “reason entity” if it is attached to a drug (e.g., heparin). Remember also to annotate the relation “Reason-Drug”.

3. Is radiation considered a drug or a therapy? What about chemotherapy?

Radiation is not considered a drug (it is a procedure, but we are not annotating those).
Chemotherapy should be annotated as a “drug entity”.

4. Mark "Tablet sustained release" all as "form" or just "Tablet"?

“Tablet sustained release” should all be tagged as “form”.

5. How to deal with specific start or specific end dates when only one of the other is given for duration?

We are only interested in explicitly mentioned durations. Do not annotate specific dates.

6. Should TPN (tube feeding) be marked as a drug?

No.

7. In situations where there are multiple ADEs associated with a drug, how should they be marked, and how should the relationships be marked?

Each ADE should be marked separately, and each should be related to each chemotherapy drug.

8. Do we mark a dosage, route, etc for O2 if the drug is never mentioned but it implied? i.e.: pt is on 3L NC, bipap, etc?

No, you should never annotate anything if a drug is not explicitly mentioned. *Bipap* and *NC* are **routes** through which *O2* (i.e., a **drug**) is delivered; if *O2* is not mentioned, the routes should not be annotated.

9. Are stem cells or bone marrow transplants "drugs?"

Stem cells, bone marrow, packed red blood cells (PRBCs) and fresh frozen plasma (FFP), are of similar broad type (they're all human blood products--they all come from the blood bank), and they should all be marked as "drugs". We include this category on the advice of our expert annotators.

References

[1] Johnson A E *et al.*, "MIMIC-III, a freely accessible critical care database," *Scientific data*, vol. 3, 2016.

[2] WHO. International drug monitoring: the role of national centres. *Tech Rep Ser WHO* 1972, no 498.

[3] Edwards I R and Aronson J K, "Adverse drug reactions: definitions, diagnosis, and management," *The Lancet*, vol. 356, no. 9237, pp. 1255–1259, Oct. 2000.

[4] Reddy V L, Pasha S J, Rathinavelu M, and Reddy Y P, "Assessment of knowledge, attitude and perception of pharmacovigilance and adverse drug reaction (ADR) reporting among the pharmacy students in south India," *IOSR J Pharm Biol Sci*, vol. 9, no. 2, pp. 34–43, 2014.

[5] Petrovic M, Tangiisuran B, Rajkumar C, van der Cammen T, and Onder G, "Predicting the Risk of Adverse Drug Reactions in Older Inpatients: External Validation of the GerontoNet ADR Risk Score Using the CRIME Cohort," *Drugs & Aging*, pp. 1–8, 2016.

[6] Toklu H Z *et al.*, "The knowledge and attitude of the healthcare professionals towards pharmacovigilance and adverse drug reaction reporting in Northern Cyprus," *Journal of Pharmacovigilance*, 2016.

[7] Bates D W *et al.*, "Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group," *JAMA*, vol. 274, no. 1, pp. 29–34, Jul. 1995.

[8] Stenetorp P, Pyysalo S, Topić G, Ohta T, Ananiadou S, and 'ichi Tsujii J, "BRAT: a web-based tool for NLP-assisted text annotation," in *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 2012, pp. 102–107.