Contextualized Medication Event Dataset (CMED)

Diwakar Mahajan, Jennifer Liang and Ching-Huei Tsou

IBM Research

Oct 28, 2020

The Problem

- Healthcare providers need a full understanding of a patient's medication history to provide appropriate treatment recommendations.
- However, many medication events are documented only in unstructured clinical notes and therefore difficult to find and easy to miss, especially at the point-of-care where time is limited.
- Several attempts have been made to identify and classify medication change events in clinical text.
- However, most of these works only focus on dosage adjustment action without placing it in the context of documented clinical discussion.

Related Works

- Previous work in medication change classification has been largely driven by use-cases and have led to **varied label definitions** -
  - Change status for warfarin (labels - *on* and *stop*)[1]
  - Status for heart failure medications (labels - *active, discontinued*, *negative*)[2]
  - Status for dietary supplements (labels - *continuing*, *discontinued*, *started* and *unclassified*)[3]
  - Sohn et. al [4] (labels - *start*, *stop*, *increase*, *decrease*, *no-change*)
  - Pakhomov et. al [5] (labels - *past*, *continuing*, *stop*, *start*, not classified)
- Other works like assertions on *other* clinical concepts like problems and test (Uzuner et al [6]).

Bridging the Gap

- Previous label definitions are driven by use cases and lack structure. Further many of the labels are not mutually exclusive
  - e.g. *past* event can be a *stop* as well.
- No previous attempts have been made to capture:
  - certainty of a medication event (e.g. certain, hypothetical, conditional, etc.)
  - actor initiating the clinical event (e.g. physician, patient, etc.)
- There is a need for more organized schema of label definitions
- Solution – Multi-dimensional context extraction for medication change events

Annotation Guidelines - Multidimensional Context Classification for Medication Change Events

We organize this task as a two-step process:

STEP 1: Medication Event Classification

Given a medication mention in a note, determine if a medication change event is being discussed:

- **NoDisposition:** no action is being discussed.
  - e.g. status statements (“doing well on 10mg lisinopril”)
- **Disposition**: presence of a medication change action being discussed
- **Undetermined**: unclear if a change is being discussed
  - e.g. “Plan: Lasix” – unclear if just stating a medication patient is on (NoDisposition) or starting a new med (Disposition)

Annotation Guidelines - Multidimensional Context Classification for Medication Change Events

STEP 2: Multi-dimensional Context Classification

Classify Disposition events along the five context dimensions i.e. Action, Negation, Temporality, Certainty and Actor

Dimension	Definition	Labels*
Action	What is the change being discussed?	Start, Stop, Increase, Decrease, OtherChange, UniqueDose, Unknown
Negation	Is the change being discussed negated?	Negated, NotNegated
Temporality	When is this change intended to occur?	Past, Present, Future, Unknown
Certainty	How likely is this change to have occurred / will occur?	Certain, Hypothetical, Conditional, Unknown
Actor	Who initiated the change?	Physician, Patient, Unknown

*Detailed label definitions presented at the end

Examples

Text	Event	Action	Negation	Temporality	Certainty	Actor
Pt currently on *lisinopril*	NoDisposition	-	-	-	-	-
Plan: incr *losartan* from 1 tab qd to bid.	Disposition	Increase	NotNegated	Present	Certain	Physician
If BP<100 hold off on taking *hctz*	Disposition	Stop	NotNegated	Future	Conditional	Physician
In ED, given *ativan* 1 mg IV x 1	Disposition	UniqueDose	NotNegated	Past	Certain	Physician
She was experiencing a bad episode of dry cough so stopped taking lisinopril	Disposition	Stop	NotNegated	Past	Certain	Patient
On Zocor, pt wants to discuss switching to generic to save money & Disposition	Disposition	OtherChange	NotNegated	Present	Hypothetical	Patient
Will hold off on empirically starting abx based on urinalysis	Disposition	Start	Negated	Present	Certain	Physician

Dataset

- Clinical notes in i2b2 2014 Heart Disease Risk Factor Challenge DataSet were utilized
- Annotation team of 3 annotators lead by a Physician
- 500 notes with 9,013 medication mentions annotated

Inter-Annotator Agreement

- IAA was calculated using Cohen's Kappa
- IAA calculated on 120 notes (double annotated)
  - Medication event classification (2,495 medication mentions)
    - Disposition vs NoDisposition vs Undetermined 0.88
  - Multi-dimensional context classification on agreed Disposition events (367 instances)
    - Action 0.87
    - Negation 0.83
    - Temporality 0.94
    - Certainty 0.75
    - Actor 0.72

Data Statistics – Train/Test Split

- We split the dataset into 400 notes for training and 100 for test.
- Subtask A: Medication Extraction –

Label	Training Dataset	Test Dataset
Medication	7230	1783

- Subtask B: Medication Event Classification -

Label	Training Dataset	Test Dataset
Disposition	1413	335
NoDisposition	5260	1326
Undetermined	557	122

Data Statistics – Train/Test Split

- Subtask C - Multi-Dimensional Classification for Medication Disposition Events

Dimension	Label	Training Dataset	Test Dataset
Action	Start	568	131
	Stop	341	67
	Increase	129	22
	Decrease	54	13
	UniqueDose	285	88
	OtherChange	1	0
	Unknown	35	14
Negated	Negated	32	6
	NotNegated	1381	329

Dimension	Label	Training Dataset	Test Dataset
Temporality	Past	745	173
	Present	494	132
	Future	145	29
	Unknown	29	1
Certainty	Certain	1177	281
	Hypothetical	134	33
	Conditional	100	15
	Unknown	2	6
Actor	Physician	1278	311
	Patient	107	17
	Unknown	28	7

Proposal - n2c2 Shared Task

TASK: Identify and classify medication change events in clinical notes.

- Subtask A: **Medication Extraction**

  Given a clinical note, identify medications in text. This is a Named Entity Recognition task.

- Subtask B: **Medication Event Classification**

  Given a medication mention in a note, determine the type of medication event Disposition, NoDisposition or Undetermined.

- Subtask C: **Multi-Dimensional Classification for Medication Disposition Events**

  Given an identified Disposition event, capture its context across orthogonal dimensions Action, Negation, Temporality, Certainty and Actor.

- Subtask D: **End-to-End Evaluation**

  End-to-end task of classifying medication into specific event type and classifying the multi-dimensional context for Disposition events.

Detailed label definition

Action indicates the type of change being discussed, and can take one of seven labels:

- Start: indicates start of a medication patient is not already on
- Stop: indicates stop of a medication patient is already on
- Increase: indicates an increase in daily dose
- Decrease: indicates a decrease in daily dose
- OtherChange: indicates a non-dosage related change, such as changes in timing (e.g. take in am instead of pm), change from brand name to generic, or change in formulation (e.g. oral tab to oral solution).
- UniqueDose: indicates single administration, where patient has taken a medication but it is unclear from the text whether it's part of a longer planned regimen; often applicable in inpatient or emergency room settings (e.g. “In the ED, patient received vancomycin 1 gram”
- Unknown: used when it is unclear which of the other labels are appropriate to use, for example, “Will change to Lasix to bid'' -- unclear if the change is from another medication to Lasix (Start), or if the patient is already on Lasix and is changing to a different frequency (OtherChange) or different dose (Increase or Decrease).

Detailed label definition

Negation indicates if the change action is being negated:

- Negated: indicates the change action is being negated (e.g. “will not initiate beta-blocker given sinus bradycardia'')
- NotNegated: indicates the change action is NOT negated, this is the default option if there is no negation present (e.g. “Start NPH 10 Units qAM’’)

Detailed label definition

Temporality indicates when the change action is intended to occur:

- Past: indicates the action has already taken place
- Present: indicates an action intended for the present time
- Future: indicates an action that will take place in the future
- Unknown: used when it is unclear which of the other labels are appropriate to use

Detailed label definition

Certainty indicates whether the change action was implemented or just discussed:

- Certain: indicates definitive action that will take place or has already occurred
- Hypothetical: indicates an action being considered but not yet decided upon
- Conditional: indicates an action that is dependent upon a specified condition being met
- Unknown: used when it is unclear which of the other labels are appropriate to use

Detailed label definition

Actor indicates the individual who initiated the change action:

- Physician: indicates a recommendation by the healthcare provider, including physicians, nurse practitioners, or other providers participating in the patient's care
- Patient: indicates an action initiated by the patient or their caretaker without consulting their healthcare provider
- Unknown: used when it is unclear which of the other labels are appropriate to use

References

[1] Liu, M., Jiang, M., Kawai, V.K., Stein, C.M., Roden, D.M., Denny, J.C. and Xu, H., 2011. Modeling drug exposure data in electronic medical records: an application to warfarin. In AMIA annual symposium proceedings (Vol. 2011, p. 815). American Medical Informatics Association.

[2] Meystre, S.M., Kim, Y., Heavirland, J., Williams, J., Bray, B.E. and Garvin, J., 2015. Heart failure medications detection and prescription status classification in clinical narrative documents. *Studies in health technology and informatics*, *216*, p.609.

[3] Fan, Y., He, L. and Zhang, R., 2016, December. Classification of use status for dietary supplements in clinical notes. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1054-1061). IEEE.

[4] Sohn, S., Murphy, S.P., Masanz, J.J., Kocher, J.P.A. and Savova, G.K., 2010. Classification of medication status change in clinical narratives. In AMIA Annual Symposium Proceedings (Vol. 2010, p. 762). American Medical Informatics Association.

[5] Pakhomov, S.V., Ruggieri, A. and Chute, C.G., 2002. Maximum entropy modeling for mining patient medication status from free text. In *Proceedings of the AMIA Symposium* (p. 587). American Medical Informatics Association.

[6] Uzuner, Ö., South, B.R., Shen, S. and DuVall, S.L., 2011. 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. Journal of the American Medical Informatics Association, 18(5), pp.552-556.