

Development of a Medical Incident Report Corpus with Intention and Factuality Annotation

Hongkuan Zhang[†], Ryohei Sasano[†], Koichi Takeda[†], Zoie Shui-Yee Wong[‡]

[†] Graduate School of Informatics, Nagoya University, Nagoya, Japan

[‡] Graduate School of Public Health, St.Luke’s International University, Tokyo, Japan
zhang.hongkuan@a.mbox.nagoya-u.ac.jp

Abstract

Medical incident reports (MIRs) are documents that record what happened in a medical incident. A typical MIR consists of two sections: a structured categorical part and an unstructured text part. Most texts in MIRs describe what medication was intended to be given and what was actually given, because what happened in an incident is largely due to discrepancies between intended and actual medications. Recognizing the intention of clinicians and the factuality of medication is essential to understand the causes of medical incidents and avoid similar incidents in the future. Therefore, we are developing an MIR corpus with annotation of intention and factuality as well as of medication entities and their relations. In this paper, we present our annotation scheme with respect to the definition of medication entities that we take into account, the method to annotate the relations between entities, and the details of the intention and factuality annotation. We then report the annotated corpus consisting of 349 Japanese medical incident reports.

Keywords: medical incident reports, intention and factuality annotation, incident comparative table

1. Introduction

Medical incidents have been shown to cause significant harm to patients (Kohn et al., 2000). To prevent similar incidents from happening again, several countries have established national reporting systems to collect medical incident reports (MIRs). Among the gathered incident types, those caused by medication errors are one of the most frequently occurring types (Grissinger, 2010). Figure 1 shows an example of a medication-related MIR that contains two sections: a structured categorical information part including patient age, sex, incident date, type, severity level, etc., and an unstructured text part that records what happened in the incident. While both the structured and unstructured parts contain important information on the incidents, it is relatively difficult to extract information from the unstructured part because it is written in natural language. Therefore, accurate techniques for information extraction from the unstructured part of medication-related incident reports are needed.

Recognizing the intention of clinicians and the factuality of medication¹ is essential to extract important information from an MIR. For example, the text shown in the bottom half of Figure 1 describes that clinicians intended to use *Rebamipide 100mg* but *Sennoside 12mg* were actually delivered, which indicates that the drug and its amount were wrong. Table 1 summarizes the intention and factuality of this incident. Our goal is to automatically construct this kind of comparative table from incident reports. To achieve this, we think that three different technologies are required: 1) medication entity recognition, 2) entity relation detection, and 3) intention/factuality analysis. Let us consider the example in Figure 1 again. First, we have to recognize the drug names (i.e., *Loxoprofen*, *Rebamipide*, and *Sennoside*) and their attributes including dosages (i.e., *1 tablet*

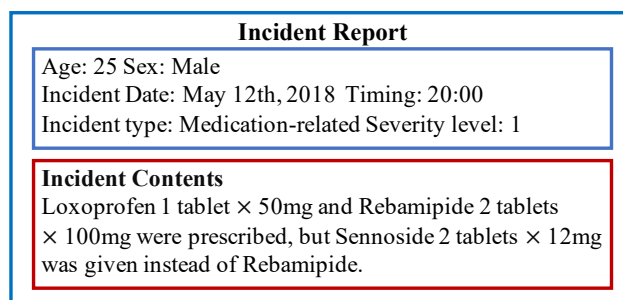


Figure 1: Example of medication-related incident report.

Entities	Intended	Actual	Implication
Drug	Loxoprofen	Loxoprofen	
Amount	50mg	50mg	
Dosage	1 tablet	1 tablet	
Drug	Rebamipide	Sennoside	Wrong drug
Amount	100mg	12mg	Wrong amount
Dosage	2 tablets	2 tablets	

Table 1: Incident comparative table generated from the report in Figure 1.

and *2 tablets*) and amounts (i.e., *50mg* and *100mg*). Next, we should detect the relations between entities, e.g., *Rebamipide* and *Sennoside* are corresponding and the amount *12mg* is an attribute of *Sennoside*. Then, we have to analyze the intention and factuality to clarify that *Rebamipide* was the intended drug and *Sennoside* was the actually delivered drug.

In this paper, we report our development of a Japanese intention and factuality annotated medical incident report (IFMIR) corpus as the first step of information extraction from medical incident². The corpus contains three types of annotations: medication entity, entity relation, and inten-

¹In this paper, the factuality of medication indicates whether a medication was actually given or not.

²<https://github.com/HongkuanZhang/IFMIR-Corpus>

当事者以外の関連職種 (複数回答可) ←	Related occupation except the person who made a mistake
薬剤師	
関連医薬品 1 ←	Related medication
【販売名】 りくしリクシアナ60mg 【製造販売業者】 未記入	
発生要因 (複数回答可) ←	Incident cause
確認を怠った	
事例概要 ←	Summary of incident
【事例の内容】 ←	Incident contents
リクシアナ60mgを調剤しなければならないところ、直前に調剤したこともあってかりクシアナ30mgを誤って調剤。鑑査役の薬剤師が間違いを発見し、正しいリクシアナ60mgを患者にお渡しした。	
【事例の背景要因の概要】 ←	Summary of incident background
直前の調剤で30mgの規格の方を調剤しており、「リクシアナ」と見た瞬間に30mgと思い込んで調剤してしまったこと。また薬袋に入れる前に確認を怠ったこと。	
【改善策】 ←	Improvement plan
ピッキングをしてそのまま薬袋に入れる前に、全て一度集めてきてから、薬袋の薬剤名と見比べながら入れるようにする。	

Figure 2: Example of JQ incident report.

tion/factuality. Among these, intention/factuality is a distinctive characteristic of our corpus, which has not been considered in the previous work.

2. Related Work

There are several annotated biomedical resources that have addressed named entities and their attributes for information retrieval. For instance, the i2b2 corpus from the 2010 clinical NLP challenge (Uzuner et al., 2011) focuses on recognizing three conceptions, i.e., clinical problem, test, and treatment, as well as extracting their assertions and relations. The ShARe corpus in ShARe/CLEF (Suominen et al., 2013) and SemEval (Pradhan et al., 2014; Elhadad et al., 2015) focuses on identifying disorder mentions and UMLS concept identifiers (CUIs).

Several researchers have concentrated on drug-related named entities. These include Segura-Bedmar et al. (2013), who organized a Drug named entity recognition (NER) shared task for recognition of drugs and extraction of drug-drug interactions (DDIs), Herrero-Zazo et al. (2013), who developed the DDI Corpus focusing on pharmacological substances and DDI, and Henry et al. (2019), who developed the n2c2 2018 track 2 shared task focused on adverse drug events (ADEs) and medication extraction. The n2c2 data consists of 500 discharge summaries from the MIMIC (Medical Information Mart for Intensive Care) III database (Johnson et al., 2016), and the data was labeled with 8 types of conceptions and 7 types of relations. Morita et al. (2013) and Aramaki et al. (2014) held the shared task MedNLP and MedNLP-2. They annotated complaint and diagnosis in 50 collected Japanese medical history summaries as well as their modalities and ICD-10 codes.

As for MIR, there have been several studies on incident type classification (Ong et al., 2010; McKnight, 2012; Gupta et al., 2015; Wong, 2016; Wang et al., 2017; Liang and Gong, 2017). Most of the corpora used in these studies are annotated with document-level category or use the incident type from the unstructured part of incident reports as a label. In

contrast, our corpus focuses on annotating entities in the incident reports with a word-level label. We also annotate intention/factuality on entities, which has not been considered in the previous work.

3. Report Selection and Pre-processing

The Japan Council for Quality Health (JQ) has collected reports on medical near-misses and adverse events that occurred in Japan since 2004. Figure 2 shows an example of a JQ incident report. All the gathered reports, which are written in Japanese, have been anonymized and opened to the public³. In this study, we use JQ incident reports as the target of the annotation.

JQ incident reports contain several types of incidents. To enhance the annotation efficiency, in this study we only annotate the incident contents in medication-related incident reports comprising 30 to 120 characters. We focus on incident contents for annotation because they often summarize the important details of the incident without redundancy. We introduced the length constraint because very short incident contents do not include sufficient details and very long incident contents tend to contain overly complicated descriptions.

Furthermore, in this study, we focus on MIRs that report *wrong medication* incidents. There are various types of MIRs and we classify them into four categories:

- **Wrong medication:** MIRs that report wrong drugs, dosages, forms, amounts, etc. The following example reports the wrong dosage incident.
 - (1) Regular medication Narusus 2mg should be given 2 tablets at a time, but only 1 tablet was given.
- **Drug omission:** MIRs that report the omission of drugs, as in the following example.

³<https://jcqhc.or.jp/>

(2) While checking the remaining amount of medicines today, it was discovered that the number was the same as yesterday. We confirmed the drug was not taken yesterday.

- **Extra drug:** MIRs that report duplicate or unnecessary drugs, as in the following example.

(3) The doctor noticed the number of PTP sheets is not right this morning. The patient then admitted to taking a duplicate drug after questioning.

- **Others:** MIRs not included in the above three classes (e.g., the wrong patient, confirmation of the amount, etc.). The following example reports an incorrect drug storage method.

(4) Desmopressin was found not stored in a cool place when the clinician went to take it.

The reason we focus on *wrong medication* is as follows. Most MIRs of *wrong medication* explicitly describe the pair of intended but not executed medication and the corresponding mistakenly executed medication. In the case of Example (1), 2 tablets are the intended dosage and 1 tablet is the corresponding actual administered dosage. In contrast, with the other three types, recognizing the intention and factuality for MIRs often requires inference or common sense. Hence, we consider that the MIRs of *wrong medication* are suitable for the first stage of intention/factuality annotation.

The overall procedure of report selection and processing is as follows.

1. Collect a set of medication-related MIRs from the case search page of the Japan Council for Quality Health Care⁴.
2. Extract incident contents comprising 30 to 120 characters from the collected MIRs.
3. Manually classify the extracted incident contents into one of the four classes and select *wrong medication* as the annotation target.

To estimate the consistency of the classification between individuals, we asked two annotators with a background in natural language processing to classify the same 125 extracted incident contents from MIRs on June 2018 into four classes. Table 2 shows the inter-annotator agreement confusion matrix. The observed agreement on classifying four types was 82%. The agreement on classifying *wrong medication* or not was 96%.

4. Our Annotation Scheme

The IFMIR corpus contains three types of annotations: medication entity, entity relations, and intention/factuality. We present the annotation scheme for each type in this section.

	WM	DO	ED	Others	Total
WM	34	0	0	2	36
DO	1	43	2	11	57
ED	1	1	3	1	5
Others	1	1	2	23	27
Total	37	44	7	37	125

Table 2: Inter-annotator agreement confusion matrix. WM, DO, ED and Others stand for *wrong medication*, *drug omission*, *extra drug*, and *others*, respectively.

4.1. Medication Entity Annotation

Inspired by the annotation scheme in the n2c2 2018 shared task (Buchan, 2018), we defined 12 types of medication named entities as the targets of the annotation, as shown in Table 3. Entities other than drugs are attributes of a drug. Note that form and strength included in the brand name of the drug are annotated separately. We also annotate the amount with specific attributes of *amount per unit*, *amount per consumption*, *total amount per day*, and *total amount for the whole medication* to differentiate various indicators. For example, 250mg in “Metyrapone 250mg was given to the patient” is annotated as *amount per unit* and 500mg in “2 tablets Metyrapone (total: 500mg) was taken today” is annotated as *total amount per day*.

4.2. Event Relation Annotation

We give event indexes to each medication entity to annotate relations among entities. Typically, the same drug and its attributes share an index. However, in the case of wrong drug incidents, the corresponding intended drug and actually delivered drug share the same index. In addition, we give the special index 0 to all medication entities that are not related to the incident. For incident-related entities, we apply the first-come-first-served basis for indexing.

- (5) *Loxoprofen*₁ 1 tablet × 50mg and *Rebamipide*₁ 2 tablets × 100mg were prescribed. *Loxoprofen*₂ was delivered correctly but *Sennoside* 2 tablets × 12mg was given as instead of *Rebamipide*₂.

For explanation, we show the text in Figure 1 again as Example (5). We assign indexes 1 to *Loxoprofen*₁ and *Loxoprofen*₂, and index 2 to *Rebamipide*₁ and *Rebamipide*₂, since *Loxoprofen* and *Rebamipide* are used in the separate events. *Sennoside* is assigned with index 2 because it is mistakenly used instead of *Rebamipide* and thus *Rebamipide* and *Sennoside* are related to the same wrong drug event. As for non-drug entities, index 1 is assigned to the 1 tablet and 50mg, and index 2 is assigned to the 2 tablets, 100mg, and 12mg.

4.3. Intention/Factuality Annotation

Most texts in MIR describe what medication was intended to be given and what was actually given, as what happened in an incident is largely due to the discrepancies between intended and actual medications. Thus, we also annotate the attribute regarding the intention and factuality of each medication entity. Here we defined three types of labels

⁴<http://www.med-safe.jp/mpsearch/SearchReport.action>

Entity type	Definition	Examples
Drug	The intended to deliver or actual delivered drug name, or entities described as drugs.	<i>Arimidex</i> tablet 1 mg were mistakenly delivered as <i>Letrozole</i> in the morning. <i>Antibiotics</i> was missed during night shift.
Form-form	The form of a drug (e.g., tablet, subcutaneous injection).	Should have given Antilate <i>lotion</i> , but Antebate <i>ointment</i> was given.
Form-mode	The mode is a drug mode of action that is associated with pharmacodynamic action.	Sodium Valproate was prescribed, but Sodium Valproate <i>ER</i> was dispensed.
Strength-amount	The amount is defined as medication dose or IV fluid volume.	A doctor ordered Vancomycin <i>500 mg</i> diluted in <i>100 ml</i> normal saline, but the nurse used Vancomycin <i>500 mg</i> diluted in <i>10 ml</i> normal saline.
Strength-rate	Rate typically represents one measure against another quantity or measure.	Soldem3A 200 ml was set on pump and started at <i>100 ml/hr</i> .
Strength-concentration	Concentration is defined as diluted medication concentration with nominator and denominator or presented as percentage or IV fluid concentration.	<i>20%</i> glucose was injected. Precedex <i>200µg/2ml</i> was dissolved in 48 ml of physiological saline.
Route	Route is defined as the route of drug administration to the patient, which may include the infusion sites, routes and pumps.	On April 6, Alpiny <i>suppository</i> 100 mg should be prescribed, but 200 mg was delivered.
Duration	Duration is defined as period during which a drug is administered to the patient.	Inhaled medical products for <i>14 days</i> were finished within <i>5 days</i> .
Timing	Timing is defined as a scheduled administration time that is predefined as time interval.	Nurse forgot to give oxycodone to patient at <i>8 a.m.</i> . After discovery, the nurse administrated oxycodone at <i>11 a.m.</i> .
Date	Date is defined as a time unit including a date and time unit longer than one day.	On <i>April 6</i> , Alpiny suppository 100 mg was given to the patient. Predonin should have been given <i>tomorrow</i> morning, but was administered at noon <i>today</i> .
Frequency	Frequency is defined as how many times a drug is given per unit of time.	A doctor ordered heparin calcium <i>3 times/day</i> for the prevention of deep thrombosis after myoma operation. Nurse administered it <i>1 time/day</i> .
Dosage	Dosage is defined as the number of units (e.g., tables, bottles, ampules) given to the patient as a single dose.	Doctor ordered <i>3×100</i> mg aspirin tablets 2 times/day, but the nurse gave <i>1×100</i> mg aspirin tablet.

Table 3: Definitions and examples of 12 medication named entity categories.

for intention/factuality annotation. The definitions are as follows:

- **Intended & Actual:** The entity was intended to be given and was actually given. This indicates no error has occurred as to this entity.
- **Intended & Not-actual:** The entity was intended to be given but actually was not given. This indicates the intended medication was not delivered.
- **Not-intended & Actual:** The entity was not intended to be given but actually was. This indicates the not intended medication was mistakenly delivered.

In Example (5) in Section 4.2, *Loxoprofen₁* and *Loxoprofen₂* are labeled as *Intended & Actual*, and *Rebamipide₁* and *Rebamipide₂* are labeled as *Intended & Not-actual*, and *Sennoside* is labeled as *Not-intended & Actual*. Note that we do not annotate any intention/factuality label to those medication entities that are labeled as index 0.

Another alternative annotation method that annotates intention and factuality of medication separately may seem more intuitive. For instance, the *Loxoprofen₁* is annotated as *Intended* and *Loxoprofen₂* is annotated as *Actual*. But this method is not suitable for some situations as shown in the Example (6).

- (6) LIXIANA 30mg was prescribed, but 50mg was given to the patient.

If we annotate the intention and factuality of medications separately in the Example (6), we will annotate *LIXIANA* and *30mg* as *Intended*, and annotate *50mg* as *Actual*. By comparing the intended and actual amount, we can understand the *30mg* is not the actual medication and the *50mg* is not the intended medication. However, it is not obvious whether LIXIANA was actually given or not from this annotation, while we can conjecture that LIXIANA was actually given. Therefore, we thought that labeling the intention and factuality together can provide more explicit and complete information and thus decided to annotate them together.

Annotation file

No.	Entity	Entity type	Intended/factuality	Index
T1	Loxoprofen	Drug	Intended & Actual	Index 1
T2	1 tablet	Dosage	Intended & Actual	Index 1
T3	50mg	Amount	Intended & Actual	Index 1
T4	Rebamipide	Drug	Intended & Not-actual	Index 2
T5	2 tablets	Dosage	Intended & Actual	Index 2
T6	100mg	Amount	Intended & Not-Actual	Index 2
T7	Loxoprofen	Drug	Intended & Actual	Index 1
T8	Sennoside	Drug	Not-intended & Actual	Index 2
T9	2 tablets	Dosage	Intended & Actual	Index 2
T10	12mg	Amount	Not-intended & Actual	Index 2
T11	Rebamipide	Drug	Intended & Not-actual	Index 2

Figure 3: Example of BRAT annotation with resulting annotation file.

5. Statistics of Annotated MIRs

The annotation process consists of two phases: a pilot annotation phase for fixing the annotation scheme and a main annotation phase for annotating on more data according to the fixed scheme. For the pilot annotation, we first collected 489 medication-related MIRs on April 2018 from the search page of the Japan Council for Quality Health Care. We then extracted the incident contents consisting of between 30 and 120 characters that had been manually classified as *wrong medication*. As a result, we obtained 49 texts for annotation. We used the BRAT Annotation Tool⁵, a system designed for annotating corpora through a browser, for annotation. Figure 3 shows an example of BRAT annotation with the resulting annotation file. Note that, our target texts are actually written in Japanese but we show an English example in Figure 3 for better understanding. We created the gold standard annotation on 49 MIRs, and fixed our annotation scheme during the pilot annotation. We trained the annotators according to our scheme and used the 49 gold standard data as the reference. For the main annotation, we first collect 300 MIRs of *wrong medication* that are extracted from 1315 JQ MIRs from May 2018 to January 2019. Next, we ask an annotator for the first annotation and then ask another annotator to check the first annotation. Eventually, we made a corpus consisting of 349 MIRs in total, i.e., 49 MIRs with the pilot annotation and 300 MIRs with the main annotation. The corpus contains 25148 characters. The statistics of the annotated medication entities

⁵<https://brat.nlplab.org/>

Entity type	NPA	NMA	Total
Drug	102	649	751
Strength-amount	65	376	441
Timing	14	226	240
Dosage	32	138	170
Form-form	27	107	134
Date	14	115	129
Route	0	97	97
Duration	4	60	64
Frequency	14	42	56
Strength-rate	0	55	55
Strength-concentration	0	16	16
Form-mode	0	11	11
Total	272	1892	2164

Table 4: Number of entities for each entity type. NPA and NMA stand for the number for the pilot annotation results and the number for the main annotation results respectively.

Intention/Factuality	NPA	NMA	Total
Intended & Actual	142	1038	1180
Intended & Not-actual	58	345	403
Not-intended & Actual	43	384	427
Total	243	1767	2010

Table 5: Number of intention and factuality attributes.

are listed in Table 4. The most common entity type is *Drug* followed by *Strength-amount*, *Timing*, and *Dosage*. Table 5 shows the statistics of the intention and factuality labels. The total number of intention and factuality labels was

Precision	Recall	F-measure
0.957	0.963	0.960

Table 6: Precision, recall, and F-measure of inter-annotator agreement on medication entity.

154 fewer than that of entity type labels shown in Table 4 because we did not annotate intention/factuality labels on entities that are labeled as index 0. The reason the numbers of entities labeled as *Intended & Not-actual* (403) and *Not-intended & Actual* (427) are close is that these entities often co-occur. We confirmed that each text has at least one entity that is labeled as *Intended & Not-actual* and one entity that is labeled as *Not-intended & Actual*.

To measure the inter-annotator agreement, we randomly select 30 MIRs and ask two groups of annotators to annotate these 30 MIRs independently. Each group consists of the first annotator and the checker, thus each MIR are annotated by four annotators. We treat results from one group as correct annotation and results from another group as system output, and we estimate the inter-annotator agreement on medication entity annotation by calculating recall, precision, and F-measure. Table 6 shows the results. The F-measure was 0.960. We then estimated the inter-annotator agreement on intention/factuality annotation with the same setting using the medication entities that are equally annotated by two groups and are not labeled as index 0. The agreement rate was 0.928.

6. Conclusion and Future Direction

In this paper, we presented an annotation scheme for medication-related medical incident reports. Our annotation consists of three types: medication entity, entity relations, and intention/factuality annotations. We also reported the the statistics of annotation on 349 Japanese MIRs.

There are three major directions for future work. First, we will develop an automatic information extraction system that recognizes not only the medication entities in medical incident reports but also the intention and factuality of these entities. Second, we will extend the target incident report type for annotation. In this study, we focused only on wrong medication type. However, drug omission and extra drug incidents are also major types of medical incidents. Therefore, our next step will be to annotate the medical incident of these types. Lastly, we plan to extend the medication entities for annotation. We defined 12 types of medication entities but found that several other entity types, such as process and patient, can be essential information to understand the incident reports.

7. Bibliographical References

Aramaki, E., Morita, M., Kano, Y., and Ohkuma, T. (2014). Overview of the NTCIR-11 MedNLP-2 task. In *NTCIR*.
 Buchan, K. (2018). Annotation guidelines for the adverse drug event (ADE) and medication extraction challenge. *n2c2, US*.
 Elhadad, N., Pradhan, S., Gorman, S., Manandhar, S., Chapman, W., and Savova, G. (2015). Semeval-2015

task 14: Analysis of clinical text. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*.
 Grissinger, M. (2010). The five rights: a destination without a map. *Pharmacy and Therapeutics*, 35(10):542.
 Gupta, J., Koprinska, I., and Patrick, J. (2015). Automated classification of clinical incident types. *Studies in health technology and informatics*, 214:87–93.
 Henry, S., Buchan, K., Filannino, M., Stubbs, A., and Uzuner, O. (2019). 2018 n2c2 shared task on adverse drug events and medication extraction in electronic health records. *Journal of the American Medical Informatics Association*.
 Herrero-Zazo, M., Segura-Bedmar, I., Martínez, P., and Declerck, T. (2013). The DDI corpus: An annotated corpus with pharmacological substances and drug–drug interactions. *Journal of biomedical informatics*, 46(5):914–920.
 Johnson, A. E., Pollard, T. J., Shen, L., Li-wei, H. L., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., and Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. *Scientific data*, 3:160035.
 Kohn, L. T., Corrigan, J. M., and Donaldson, M. S. (2000). *To Err Is Human: Building a Safer Health System*. The National Academies Press, Washington, DC.
 Liang, C. and Gong, Y. (2017). Automated classification of multi-labeled patient safety reports: A shift from quantity to quality measure. In *MedInfo*.
 McKnight, S. D. (2012). Semi-supervised classification of patient safety event reports. *Journal of patient safety*, 8(2):60–64.
 Morita, M., Kano, Y., Ohkuma, T., Miyabe, M., and Aramaki, E. (2013). Overview of the NTCIR-10 MedNLP task. In *NTCIR*.
 Ong, M.-S., Magrabi, F., and Coiera, E. (2010). Automated categorisation of clinical incident reports using statistical text classification. *Qual Saf Health Care*, 19(6):e55–e55.
 Pradhan, S., Elhadad, N., Chapman, W., Manandhar, S., and Savova, G. (2014). Semeval-2014 task 7: Analysis of clinical text. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*.
 Segura-Bedmar, I., Martínez, P., and Herrero-Zazo, M. (2013). SemEval-2013 task 9 : Extraction of drug–drug interactions from biomedical texts (DDIExtraction 2013). In *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*.
 Suominen, H., Salanterä, S., Velupillai, S., Chapman, W. W., Savova, G., Elhadad, N., Pradhan, S., South, B. R., Mowery, D. L., Jones, G. J., et al. (2013). Overview of the ShARe/CLEF eHealth evaluation lab 2013. In *International Conference of the Cross-Language Evaluation Forum for European Languages*.
 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):552–556.

- Wang, Y., Coiera, E., Runciman, W., and Magrabi, F. (2017). Using multiclass classification to automate the identification of patient safety incident reports by type and severity. *BMC medical informatics and decision making*, 17(1):84.
- Wong, Z. S. Y. (2016). Statistical classification of drug incidents due to look-alike sound-alike mix-ups. *Health informatics journal*, 22(2):276–292.