



## King's Research Portal

*Document Version*  
Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Wu, H., Ibrahim, Z., Iqbal, E., & Dobson, R. J. B. (2016). Predicting Adverse Events from Multiple and Dynamic Medication Episodes – a preliminary result in a large mental health registry. In *IJCAI 2016 - Workshop on Knowledge Discovery in Healthcare Data*

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# Predicting Adverse Events from Multiple and Dynamic Medication Episodes - a preliminary result in a large mental health registry

Honghan Wu      Zina M. Ibrahim      Ehtesham Iqbal      Richard JB Dobson

King's College London  
London, United Kingdom

{honghan.wu, zina.ibrahim, ehtesham.iqbal, richard.j.dobson}@kcl.ac.uk

## Abstract

Adverse drug reactions (ADRs) are undesirable and potentially fatal outcomes resulting from the use of medications. The possibility of experiencing an ADR varies between individuals owing to disease heterogeneity, genetic and demographic factors, patient treatment history and disease trajectories. Therefore, understanding the interplay among the multiple factors leading to ADRs is crucial to increasing drug effectiveness, individualising drug therapy and reducing incurred cost.

In this paper, we present the first step towards predicting ADRs based on patient profiles and treatment trajectories hidden within the Electronic Health Records (EHRs). We propose a flexible encoding mechanism that can effectively capture the dynamics of multiple medication episodes of a patient at any given time. We enrich the encoding with a drug ontology and patient demographics data and use it as a base for an ADR prediction model. We evaluate the resulting predictive approach under different settings using real anonymised patient data obtained from the EHR of the South London and Maudsley (SLaM), the largest mental health provider in Europe. Using the profiles of 38,000 mental health patients, we identified 240,000 affirmative mentions of dry mouth, constipation and enuresis and 44,000 negative ones. Our approach achieved 93% prediction accuracy and 93% F-Measure. Overall, we found that using our encoding can improve prediction accuracy by 10% compared to static medication modelling settings.

## 1 Introduction

Adversities associated with prescribed medication can seriously affect the patient's wellbeing [Edwards and Aronson, 2000] and present a real financial burden on healthcare providers (estimated to lead to an annual cost of 466 million pounds in the United Kingdom alone [Pirmohamed *et al.*, 2004]). The problem arises from the fact that trial-tested prescribed drugs are not evaluated for long-term effectiveness

in diverse (and possibly comorbid) patient populations, resulting in many unknowns with respect to the possibilities of the onset of adverse drug reactions (ADRs). Currently, discovering the factors affecting patients' response to treatment relies on spontaneous reporting systems that rely on patient and clinician data entry and resulting highly under-reported ADR instances [Tatonetti *et al.*, 2012]. However, there is a large body of unutilised knowledge embedded in the Electronic Health Records (EHRs) of hospitals, containing valuable information regarding treatment responses as well as patient profiles and disease trajectories. Therefore, there is an immediate use case for automating the ability to mine, and eventually predict, the onset of ADRs from EHRs in order to enable healthcare professionals to choose the most appropriate medication for a particular patient.

Many factors can lead to the onset of ADRs, ranging from patient habits (such as smoking or alcohol intake) to demographics, comorbidities, genetic dispositions and other prescribed (and non-prescribed) medications. For example, Clozapine, a typical antipsychotic drug, might cause dizziness or nausea among other 240 side effects, whereby the pathology of these side effects is multifactorial and mostly unknown [Kuhn *et al.*, 2015]). Analysing the factors associated with ADRs using EHRs resolves to the task of uncovering the temporal associations connecting the different factors. Things are more complicated in long-term debilitating and comorbidity illnesses whereby ADRs are associated with multiple, overlapping and long medication episodes. From a technical perspective, a challenging task lies in modelling the time series of medication episodes in an effective way so that they can be fed with other influencing factors into predictive algorithms.

In this paper, we propose a bitwise encoding mechanism that can capture the temporal precedence and duration information of medication episodes as well as their distance from the time of inspection. The encoding method presented is flexible for enrichment with additional mined knowledge and external ontologies and can be readily used by off-the-shelf machine learning algorithms. We show how a predictive model can be formed using the presented encoding, combined with mined patient demographics (age, gender and ethnicity) as well as an external (drug) ontology to anticipate ADR onset (Figure 1). We present an initial study of using the resulting model to predict the onset of three ADRs, mainly dry mouth,

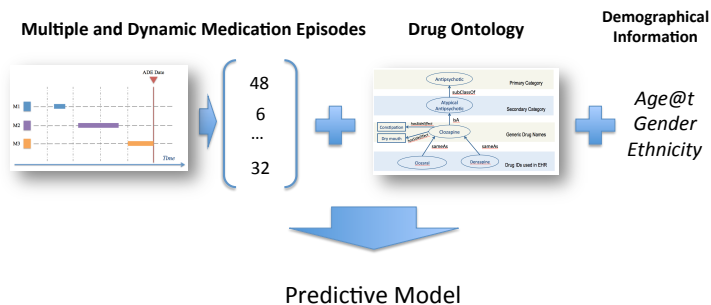


Figure 1: Building Blocks of the Prediction Model

constipation and enuresis, from a large mental health register.

Various approaches have been proposed for modelling similar healthcare events. [Sun *et al.*, 2012] proposed counting the frequencies in certain time periods to capture discrete temporal features. While such models are simple and efficient enough for later analysis, they are missing some critical information for ADR models such as the continuous nature of medication episodes and the distances of events to the time of inspection. In contrast to these simple approaches, graph-based models [Liu *et al.*, 2015] capture the temporal orders corresponding intervals. While these models can convey most needed information, they are not easy to digest for many off-the-shelf learning algorithms.

In addition, EHR-embedded temporal events have been studied for various reasons: [Monroe *et al.*, 2013] studied the temporal event sequence simplification for visualisation purposes; [Zhou *et al.*, 2014; Liu *et al.*, 2015; Sun *et al.*, 2012] examined various approaches to model temporal events for patients stratification or personalised medicine. Our encoding mechanism complements existing solutions with features making it more effective in modelling medication episodes for adverse event prediction.

The paper is structured as follows. After detailing the encoding system and proposed multi-factorial predictive model in Section 2, we show the results of applying the model to predict the three ADRs. We conclude by summarising our work and outlining future directions in Section 4.

## 2 Temporal Encoding for Patient’s Medication Episodes

Here, we present the novel vector-based encoding of dynamic and multiple medication episodes. Figure 2 illustrates an example encoding scenario. The upper part shows the sample medication episodes of a patient that are related to a certain inspecting time spot (marked as AE date in the figure). The encoding is realised through the following steps.

1. From the the inspecting time spot, look back to a certain period of time, e.g., 20 days. In our experiments, we selected 30 days.
2. Split the time period into intervals using a unit of time, e.g., a day or a week. Number each interval in an ascen-

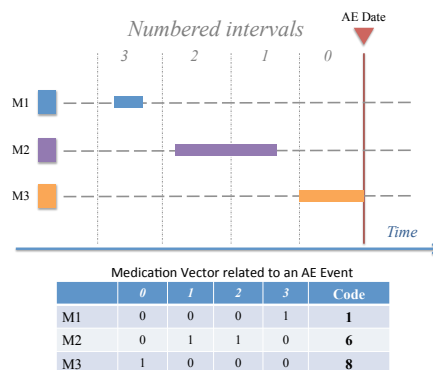


Figure 2: The bitwise encoding of dynamic medication episodes for a given adverse event

dent order (started from zero) starting from the inspecting time spot. For example, 4 intervals are identified and numbered in Figure 2

3. Add the patient’s medication episodes to the timeline.
4. For each medication episode, allocate a sequence of bits aligned with the interval order we obtained in step 2. The interval numbered zero is aligned with the most significant bit in the sequence. For each interval, set the corresponding bit to 1 if the interval intersects with the episode, and to 0 otherwise. For example, Medication Episode 2 (M2) in Figure 2 spans across interval number 1 and 2. Its encoding is 0110 in binary code.
5. Repeat step 4 for all medication episodes. We will get a vector representing all the relevant medication episodes of the inspecting time spot.

The resulting bitwise encoding has the following features:

- *Simple and Succinct*: Each medication episode is represented as a numeric value which can be used as the feature value in a predictive model which can be consumed by most machine learning algorithms.
- *Informative*: The encoded number conveys both the duration information of an episode (i.e. occurrences across

multiple time intervals are recorded in their corresponding bits) and also the distances of each occurrence (i.e., the orders of bits in the number). For example, for duration, M2 crosses two intervals and its encoded number has two bits set as 1 accordingly; for distances, M1 and M3 are both present in only one interval. However, M3 is encoded as 8 because it is closer to the event, while M1 is represented as 1.

- *Flexible* Different units of time can be used in step 2, enabling variable levels of granularity. For example, if the unit used is 4 times bigger in Figure 2, all three medication episodes will be encoded as 1. The same number of bits can cover medication periods that are four folds longer. The obvious loss of information however must be taken into account when choosing a time unit.

## 2.1 Predictive Model

Having illustrated the encoding of the temporal information embedded within medication episodes, we now show how this information can be combined with additional background knowledge as well as external ontologies into a feature vector for prediction.

Our predictive model considers two sets of features: 1) demographics, 2) the generated dynamic medication episodes which are further enriched by an ontology. Equation 1 shows the features of demographic information of a patient  $p$  at a given time  $t$  where  $age@t$  means the patient’s age at the inspecting time  $t$ .

$$F_d(p, t) = (age@t, gender, ethnicity) \quad (1)$$

Equation 2 gives the feature calculation from dynamic medication episodes given a patient  $p$ , an adverse event type  $a$  and a time  $t$ .

$$F_m(p, a, t) = (\alpha \cdot \text{diag}(\vec{S}_a) + \text{diag}(\vec{1})) \cdot \vec{M}_{p,t} \quad (2)$$

$\vec{S}_a$  is a binary vector describing drugs’ relation with side effect  $a$ . For example, suppose there are two drugs ( $d_1, d_2$ ), and  $d_1$  has the side effect  $a$  but  $d_2$  does not. Then, the  $\vec{S}_a$  is  $[1, 0]$ . Such vector is derived from the drug ontology.  $\vec{M}_{p,t}$  is the vector generated from the medication episodes of the patient  $p$  at a given time  $t$  using our encoding method.  $\alpha$  is the weighting factor for adjusting side effect knowledge importance in the final prediction model.

The final feature vector for prediction is a concatenation of the above two as shown in equation 3. In this paper, we view the ADR prediction as a classification problem. Therefore, the feature vector is used as the inputs for training and testing classification algorithms on our datasets. Various classifiers have been tested. The experiment section will give the detail about model selection.

$$F(p, a, t) = (F_m(p, a, t), F_d(p, t)) \quad (3)$$

## 3 Implementation and Evaluation

### 3.1 Data Source

We used data extracted from the Clinical Record Interactive Search System (CRIS) [Stewart *et al.*, 2009] to evaluate our encoding. CRIS is a database containing a de-identified replica of the EHRs used by the South London

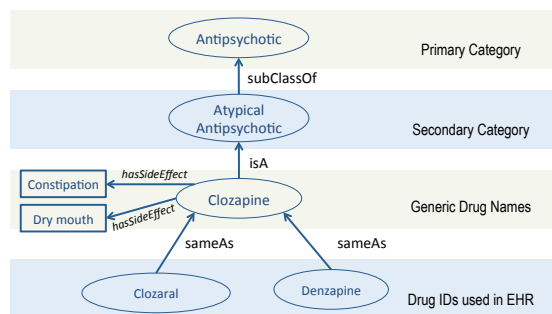


Figure 3: Drug Ontology for AE Prediction

and Maudsley Foundation Trust (SLaM), the largest mental health provider in Europe. SLaM serves over 1.2 million patients and stores much of its clinical records and prescribing information in unstructured free text format. As of October 2015, CRIS contains over 240,000 patient records comprised of over 22,270,000 free text documents including correspondence, discharge summaries, events, mental health care plan and mental state formulations [Stewart *et al.*, 2009].

As most of the ADR and drug episode information is recorded unstructured free-text, we used in-house developed natural language processing tools in conjunction with manual annotation to extract and structure relevant information [Iqbal *et al.*, 2015]. As we will show later, over 62 per cent of identified ADRs in our data are associated with multiple medication episodes.

### 3.2 Drug Ontology

In the EHRs, various drug naming or ID systems are simultaneously used. For example, in our data, a single drug maybe referred to by its generic drug name (e.g., Clozapine) or trade names (e.g., Clozaral or Denzapine). Using a consistent representation of drug names will give us the ability to reduce unnecessary (semantically) duplicated features in our prediction model, which can lead to significant performance differences in (at least certain) machine learning algorithms. Furthermore, grouping drugs based on their functions or effects will enhance our ability to further abstract the medications.

There exist many publicly available resources containing drug knowledge, such as DrugBank [Wishart *et al.*, 2006] and Drug Ontology<sup>1</sup>, as well as mental health-specific drug information, e.g., Psychology Ontology<sup>2</sup>. By inspecting these resources with clinicians and pharmacists, we found it unavoidable to populate our own drug ontology based on (some of) the available resources. This is because off-the-shelf resources provide insufficient information to map psychiatric drug IDs used in the EHRs to their generic names as these resources: a) do not contain some of the trade names used in the EHRs; b) are not populated with information about new antipsychotic drugs.

<sup>1</sup><https://ontology.atlassian.net/wiki/display/DRON/Drug+Ontology+Home>

<sup>2</sup><http://bioportal.bioontology.org/ontologies/APAONTO>

We identified 369 drug IDs in our EHR system, which we mapped to 240 generic names. We further categorised these drugs via a two-layer system comprising of a primary and a secondary category as shown in Figure 3. When a generic drug does not have a secondary category, it is typed directly by its primary category, e.g., Lithium is typed as Mood Stabiliser that is a primary category.

In addition to drug hierarchy, the known side effects of a drug are important knowledge that is crucial to ADR prediction. We further enriched our ontology with side effect knowledge extracted from SIDER [Kuhn *et al.*, 2015].

### 3.3 Dataset Preparation and Classifier Selection

In this preliminary study, we focus on three types of adverse events identified in the CRIS registry: dry mouth (#event 58,347), constipation (#event 86,602) and enuresis (#93,366), which involves 20,795 distinct patients with mental health disorders. For each identified AE, we pick up the patient’s past 30 days medication episodes. Since we are particularly interested in AEs that are associated with dynamic and multiple medication episodes, it would be helpful to understand how often such situations happen in real data. As shown in Figure 4, a fairly large proportion ( $\geq 62\%$  on average) of AEs are associated with multiple medications.

To generate negative data items (none adverse event), we picked 18,038 patients who have medication episode data in CRIS registry but never had the three types of adverse events reported. We generate a negative data item for each medication episode. For the same patient, if there are multiple episodes with the same end date, these data items will be merged into one. But how to set the event dates for these negative data items is very interesting considering that an adverse event might happen days after a medication episode. We will show some experiment results on different ways of setting such dates. About 44,900 negative data items were generated in this study.

The dataset was split into a proportion of 80/20 for training and testing. 10-folds cross validation has also been used. The former setting achieved better performance across three AE types.

We have tried various classification models including Naïve Bayesian, Bayesian Network and different decision trees. Among these models, Bayesian Network achieved the best performance in our dataset considering the accuracy and efficiency. For Bayesian Network training, different settings have been tried on search algorithms used in the model including: a) the inclusion or exclusion of using Markov Blanket correction on the learned network and b) various setting of maximum number of parent nodes. No performance difference was observed between options in setting a. For setting b, the classification accuracy increases along with the maximum number of parents from 1 to 4. In all three AE case, the performance climbs to its highest when the number is set as 4 and it decreases thereafter. However, the improvements from 1 to 4 are not significant. The biggest difference in all three cases is 0.9% in Enuresis case.

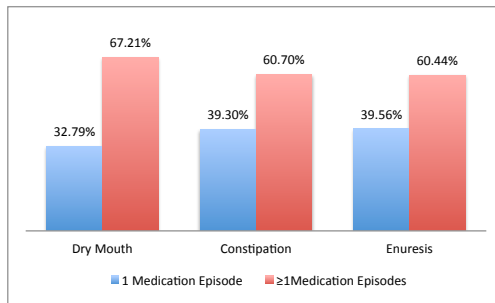


Figure 4: #AEs: Single vs Multiple Medication-episodes

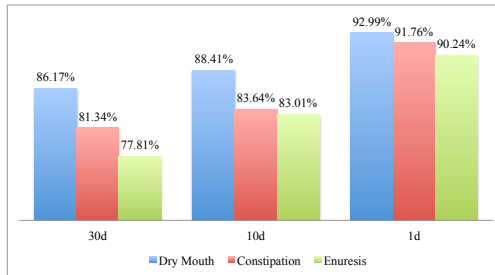


Figure 5: Encoding performance (average classification accuracy): the effect of time length represented by a bit (compression ratio)

### 3.4 Demographic vs Medication Episode Features

We have two types of feature sets - demographic features and medication episode ones. Comparing their performances will reveal some insights about what factors are more likely to be associated with adverse events. Table 1 shows the detail of demographic feature only results. In general, the F-Measure is around 70%. But for AE class of dry mouth type, the performance is extremely low - 12% True Positive and 20% F-Measure. Table 2 gives the performance of dynamic medication episode features. In all cases, about 90% F-Measure has been achieved with the lowest at 85% at AE class of dry mouth and the highest at 96% at none AE class of the same case. In summary, it is quite obvious that dynamic medication episode features are much better indicators for predicting Adverse Events. The combination of the two sets of features can achieve the best results as shown in Table 3. Both average Precision and F-Measure are around 93%.

### 3.5 Different Settings in Dynamic Medication Episode Encoding

In the bitwise encoding of medication episodes, the length of time interval represented by a bit might lead to different levels of abstractions that encoding represents. In other words, this setting is a kind of compression ratio of the encoding - the longer the more information loss. In the extreme setting, the whole period is represented by a bit. In this situation, the dynamic information (length of episode and their distances to the inspecting time) of medications is lost. In Figure 5, we compared three settings - a bit representing 1 day, 10 days and 30 days. 1-day setting got the best performance. In general,

ADE Type	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Dry Mouth	0.12	0.02	0.64	0.12	0.2	0.731	AE
	0.98	0.88	0.79	0.98	0.87	0.73	N-AE
Weighted Avg.	0.78	0.68	0.75	0.78	0.72	0.73	
Constipation	0.74	0.40	0.63	0.74	0.68	0.74	AE
	0.60	0.26	0.72	0.60	0.65	0.74	N-AE
Weighted Avg.	0.67	0.33	0.68	0.67	0.67	0.74	
Enuresis	0.82	0.40	0.70	0.82	0.75	0.78	AE
	0.60	0.18	0.76	0.60	0.67	0.78	N-AE
Weighted Avg.	0.72	0.29	0.72	0.72	0.71	0.78	

Table 1: Demographic-Feature Only Results

ADE Type	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Dry Mouth	0.83	0.04	0.86	0.83	0.85	0.97	AE
	0.96	0.17	0.95	0.96	0.96	0.97	N-AE
Weighted Avg.	0.93	0.14	0.93	0.93	0.93	0.97	
Constipation	0.90	0.07	0.92	0.90	0.91	0.97	AE
	0.93	0.10	0.91	0.93	0.92	0.97	N-AE
Weighted Avg.	0.92	0.08	0.92	0.92	0.92	0.97	
Enuresis	0.88	0.07	0.93	0.88	0.90	0.96	AE
	0.93	0.12	0.88	0.93	0.90	0.96	N-AE
Weighted Avg.	0.90	0.10	0.90	0.90	0.90	0.96	

Table 2: Medication Episode-Feature Only Results

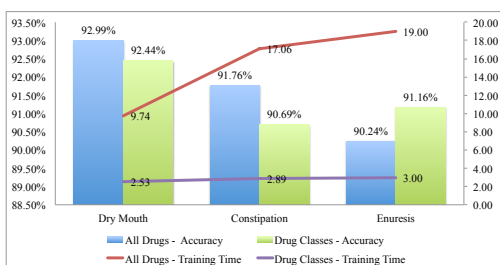


Figure 6: Feature Dimension Reduction by Drug Ontology

the smaller the granularity the better performance. Compared to the 30-day setting, 1-day setting can achieve up to 12% better performance. This concludes that the dynamic nature of the medication episodes is a very important factor associated with adverse events and the bitwise encoding can capture such information effectively.

### 3.6 Ontology-based Feature Dimension Reduction

The semantics in the drug ontologies can be utilised to combine semantically similar drugs so that the dimensions of the feature vector  $F_m$  (Equation 2) can be reduced significantly (from 226 for all drugs to 47 for using direct parent category). The performance differences between these two settings are illustrated in Figure 6. For dry mouth and constipation cases, the performances (average classification accuracy) decreased in about 1 per cent in average when using drug classes, while the accuracy did improve in the enuresis case when the feature dimension was reduced. This is a very interesting observation worth further investigation to see whether it is a single

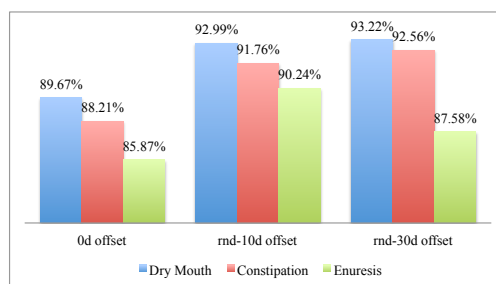


Figure 7: The Effect of Date Settings of None AE data items

special case or the drugs related to enuresis make the case special. Figure 6 also reveals that using less feature dimension can significantly increase the training speed, which is not surprising.

### 3.7 The event time settings for negative data items

As discussed in section 3.3, when populating the negative data items (none AE data items) we need to decide which date to use for setting the date of the item. Figure 7 shows the results of different settings we tried. The date settings are based on the offset of the medication episode that is used to generate the data item. Three settings have been tested: 0-offset means using the end date of the medication episode, rnd-10 offset means randomly picking a number from 1 to 10 then adding it (as days) to the medication end date to get a new date and using it, and finally rnd-30 offset is the same but using a random number from 1 to 30. Rnd-30 offset gets best performance (classification accuracy) in dry mouth and con-

ADE Type	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Dry Mouth	0.89	0.05	0.85	0.89	0.87	0.98	AE
	0.95	0.11	0.97	0.95	0.96	0.98	N-AE
Weighted Avg.	0.94	0.10	0.94	0.94	0.94	0.98	
Constipation	0.93	0.08	0.91	0.93	0.92	0.97	AE
	0.92	0.08	0.93	0.92	0.93	0.97	N-AE
Weighted Avg.	0.92	0.08	0.92	0.92	0.92	0.97	
Enuresis	0.92	0.09	0.92	0.92	0.92	0.96	AE
	0.91	0.08	0.91	0.91	0.91	0.96	N-AE
Weighted Avg.	0.92	0.08	0.92	0.918	0.92	0.96	

Table 3: Medication Episode and Demographic Combined Results

stipation, while the enuresis case gets the best performance at rnd-10 setting. Due to the significant performance decrease in enuresis in rnd-30 setting, we generally use rnd-10 offset in our model.

## 4 Conclusion

We reported a retrospective study of predicting adverse events in EHRs from South London and Maudsley Foundation Trust (SLaM), the largest mental health provider in Europe. In addition to demographic patient information (e.g., age, gender and ethnicity) and drug related knowledge (e.g., categorisation and side effects), the prediction model puts a special focus on effective approaches of modelling the dynamic and multiple medication episodes, which are very common among patients with mental health disorders (observed among  $\geq 62\%$  events in our data). Specifically, a novel bit-wise encoding approach is introduced for capturing the medication trajectories, which has proved to be more effective (10% accuracy improvement) than static medication modelling settings. Experimental results show that the medication episodes are much better indicators than demographic information for predicting adverse events, while the best performance (93% in both precision and F-Measure) was obtained by combining these two. From the evaluation, we also observed mixed performance changes when reducing the feature dimension by using the drug ontology hierarchy - prediction accuracies decreased slightly in two cases but increased in another. This is an interesting phenomenon that we plan to investigate in detail in our ongoing work.

**Acknowledgement** The authors would like to acknowledge the National Institute for Health Research (NIHR) Biomedical Research Centre and Dementia Unit at South London and Maudsley NHS Foundation Trust and Kings College London. This project has also received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 644753 (KConnect).

## References

[Edwards and Aronson, 2000] IR Edwards and JK Aronson. Adverse drug reactions: definitions, diagnosis, and management. *The Lancet*, 356(9237):1255-1259, 2000.

[Iqbal *et al.*, 2015] Ehtesham Iqbal, Robbie Mallah, Richard George Jackson, Michael Ball, Zina M. Ibrahim, Matthew Broadbent, Olubanke Dzahini, Robert Stewart, Caroline Johnston, and Richard J. B. Dobson. Identification of adverse drug events from free text electronic patient records and information in a large mental health case register. *PLoS one*, 10(8):e0134208, 2015.

[Kuhn *et al.*, 2015] Michael Kuhn, Ivica Letunic, Lars Juhl Jensen, and Peer Bork. The sider database of drugs and side effects. *Nucleic acids research*, page gkv1075, 2015.

[Liu *et al.*, 2015] Chuanren Liu, Fei Wang, Jianying Hu, and Hui Xiong. Temporal phenotyping from longitudinal electronic health records: A graph based framework. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 705–714. ACM, 2015.

[Monroe *et al.*, 2013] Megan Monroe, Rongjian Lan, Hanseung Lee, Catherine Plaisant, and Ben Shneiderman. Temporal event sequence simplification. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2227–2236, 2013.

[Pirmohamed *et al.*, 2004] Munir Pirmohamed, Sally James, Shaun Meakin, Chris Green, Andrew K Scott, Thomas J Walley, Keith Farrar, B Kevin Park, and Alasdair M Breckenridge. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *Bmj*, 329(7456):15–19, 2004.

[Stewart *et al.*, 2009] Robert Stewart, Mishael Soremekun, Gayan Perera, Matthew Broadbent, Felicity Callard, Mike Denis, Matthew Hotopf, Graham Thornicroft, and Simon Lovestone. The South London and Maudsley NHS foundation trust biomedical research centre (SLaM BRC) case register: development and descriptive data. *BMC psychiatry*, 9(1):1, 2009.

[Sun *et al.*, 2012] Jimeng Sun, Fei Wang, Jianying Hu, and Shahram Edabollahi. Supervised patient similarity measure of heterogeneous patient records. *ACM SIGKDD Explorations Newsletter*, 14(1):16–24, 2012.

[Tatonetti *et al.*, 2012] Nicholas P Tatonetti, P Ye Patrick, Roxana Daneshjou, and Russ B Altman. Data-driven prediction of drug effects and interactions. *Science translational medicine*, 4(125):125ra31–125ra31, 2012.

[Wishart *et al.*, 2006] David S Wishart, Craig Knox, An Chi Guo, Savita Shrivastava, Murtaza Hassanali, Paul Stothard, Zhan Chang, and Jennifer Woolsey. Drugbank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*, 34(suppl 1):D668–D672, 2006.

[Zhou *et al.*, 2014] Jiayu Zhou, Fei Wang, Jianying Hu, and Jieping Ye. From micro to macro: data driven phenotyping by densification of longitudinal electronic medical records. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 135–144. ACM, 2014.