Original work

TITLE PAGE

Deep learning application to automatic classification of pharmacist interventions.

Ahmad Alkanj,[1,*] Julien Godet, [2,3,4,*] Erin Johns, [1,2] Bénédicte Gourieux,[1,5] and Bruno Michel [1,4,5]

[1]Laboratoire de Pharmacologie et de Toxicologie Neurocardiovasculaire, Faculté de médecine, Strasbourg, [2]ICube - IMAGeS, UMR 7357, Université de Strasbourg, France, [3]Groupe Méthode Recherche Clinique, Pôle de Santé Publique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France, [4]Faculté de Pharmacie, Université de Strasbourg, France, [5]Service de Pharmacie, Hôpitaux Universitaires de Strasbourg.

*Contributed equally

Corresponding author: Ahmad Alkanj, MSc, PharmD, Laboratoire de Pharmacologie et de Toxicologie Neurocardiovasculaire, Faculté de médecine, Strasbourg, 1 Rue Eugène Boeckel, 67000 Strasbourg, France; ahmad.alkanj@etu.unistra.fr; telephone number : +33698300258

Keywords: Clinical pharmacy, Drug related problem, Pharmacist intervention, Classification, Deep neural network.

ABSTRACT

Objective: Medication review (MR) is the systematic assessment of a patient's medicines - a critical step in preventing of drug adverse events. MR aims to identify drug-related problems (DRP) that trigger documented pharmacist interventions (PI). The information-rich data documenting PI, produced daily, provides a unique opportunity to develop a deep learning algorithm to automatically categorize PI.

Materials and Methods: The study was conducted at the University Hospital of Strasbourg. Text data documenting PI were collected over the year 2017. Data from the first six-months of 2017 were reviewed by pharmacists who manually assigned to each PI the main class of the 29 possible classes of the French Society of Clinical Pharmacy classification. A deep neural network classifier was then trained to learn to automatically predict the main PI class from processed text data. Accuracy, specificity and sensitivity metrics were used to evaluate performance.

Results: 27,699 PI (first six-months of 2017) were extracted, processed and used to train and evaluate a classifier. Class prediction accuracy calculated on the validation dataset was 78.0%. Class specific sensitivities and specificities ranged from 0.31 to 0.96 and from 0.94 to 1.00, respectively. To demonstrate the classification ability of the algorithm, we predicted the PI class for documents collected during the second semester of 2017. Of the 4,460 predictions checked, only 67 required corrections. The latter data was concatenated with the original dataset to create an extended dataset to re-train the neural network. The updated global accuracy reached 81.0% showing that the prediction process can still improve with the increase in the amount of data.

Conclusion: PI classification is beneficial for assessing and improving pharmaceutical care practice. Here we report a high performance automatic PI classification based on deep learning. This application could find an essential place to highlight the clinical relevance of the review of drug prescriptions performed daily by hospital pharmacists.

Keywords: Deep neural network, Pharmacist Interventions, Drug-related problems, Medication review, Clinical pharmacist

1. INTRODUCTION

Among the hospital clinical pharmacy activities, medication review (MR) of medication prescription is a critical step in the prevention of adverse drug events [1-4]. MR enables to identify drug-related problems (DRP). DRP are a composite set of events (prescriptions errors, inadequate monitoring …) and can be broadly defined as *"... events or circumstances involving drug therapy that actually or potentially interfere with desired health outcomes"* [7]. DRP should triggers pharmacist interventions (PI) defined as *"any proposals for treatment modification initiated by the pharmacist"* [5]. These PI constitute recommendations for best practices in prescribing and administration [6]. While classifying PI has potential important implications for pharmaceutical care practice, a universally accepted classification system has not yet been adopted. Historically, a first coding system for DRP was proposed in the USA in 1993[8]. Subsequently, many countries have developed their own DRP classification system [9-12]. In France, the French Society of Clinical Pharmacy (SFPC) has adopted a 10 main categories classification of DRP [13]. This is currently the preferred method for recording DRP used both in daily pharmaceutical care practice and research in France [13].

Over the past decade, the implementation of electronic prescription software in French hospitals has facilitated the deployment of MR as prescriptions have been integrated into massive electronic medical record data. This information-rich data, including PI made by clinical pharmacists, provides a unique opportunity to develop expert artificial intelligence-based tools to help pharmacists classify DRP and alert them to certain risky prescribing and administration attitudes.

In this context, the objective of this study was to train a pipeline processing PI text data and feeding a deep neural network classifier able to predict the most appropriate PI code according to the SFPC classification.

2. MATERIALS AND METHODS

2.1. Dataset

The study was conducted in a large, public, University Hospital (over 2,400 beds) in Strasbourg, France that provides surgical, medical, research and teaching activities. Data were collected for adult in-patients over a full calendar year period (January - December 2017) directly from the hospital information system Dxcare® (exclusion of intensive care units, obstetrics and gynaecology, and psychiatry). The extracted

data were limited to prescription lines that resulted in PI and included text comments (in French) describing the reason for the intervention.

2.2. Classification of drug-related problems

DRP classification was based on the externally validated SFPC classification [13]. The hierarchical structure of the SFPC classification allowed us to define 29 classes and subclasses for the DRP (see table 1). To ensure that all DRP could be categorized, an additional class ("other") was introduced for situations that do not fall into one of the 28 categories of the SFPC classification.

Table 1: French Society of Clinical Pharmacy classification of drug-related problems [13]

Code	Drug-related problems
1	Non conformity to guidelines or contra-indication
1.1	Non conformity of the drug choice compared to the Formulary
1.2	Non conformity of the drug choice compared to the guidelines
1.3	Physio-pathologic contra-indication
2	Untreated indication
2.1	Valid indication without a drug
2.2	A drug is missing after transfer
2.3	The patient is missing a pre-medication or a prophylactic treatment
2.4	A synergic or corrective drug should be associated
3	Subtherapeutic dosage
3.1	Dose too low for this specific patient
3.2	Length of the treatment too short
4	Overdosage
4.1	Supra-therapeutic posology
4.2	Duplicate prescription
5	Drug without indication
5.1	No justified indication for the drug
5.2	The drug is being prescribed for a too long period without any risks
5.3	Therapeutic redundancy
6	Drug interaction
6.1	Take into account
6.2	Precaution for use
6.3	A drug interferes with another drug and can lead to a non-adapted pharmacological impact
6.4	Contra-indication between 2 drugs
6.5	Interaction published but not integrated into Vidal © dictionary
7	Adverse drug reaction
8	Improper administration
8.1	Other route more effective or less costly for the same efficacy
8.2	The method for administration is not adequate
8.3	Inappropriate drug form
8.4	Incomplete formulation
8.5	Inappropriate timing of administration and/or repartition of doses
9	Failure to receive drug
9.1	Physico-chemical incompatibility between several injectable drugs
9.2	Compliance problem
10	Drug monitoring
11	Other

2.3. Expert labelling of pharmacist' interventions data

Data related to prescriptions with PI over the first six-months of 2017 were exhaustively labeled by two experts. Data were reviewed independently by two pharmacists who assigned a PI code number to each text documenting a PI. In case of a discrepancy between the 2 pharmacists, the coding was reconsidered in order to reach a final consensus. A new dataset including DRP coding was thus obtained.

2.4. Pharmacist interventions processing

Documents describing PI were written as French plain text - often in the form of technical notes for prescribers and needed to be processed. Natural language processing was done using R (version 4.1.0) [14]. Briefly, abbreviations and keywords were replaced by their corresponding expressions. Accents, punctuations, numbers were removed. The Porter's stemming algorithm (package tm) was used to stem French words in the corpus of comments. Terms were tokenized as n-grams (n from 1 to 3) and a document term matrix was built using term frequency weighting t generate a numerical representation of the data. To take into account the fact that the data is not balanced for a correct performance evaluation, a stratified split of the document-term matrix was done (caret package) to build the training (85%) and the validation dataset (15%) [15].

2.5. Deep neural network classifier

A deep neural network classifier was built from R using Keras library package (2.4.0) and running on Tensorflow backend (2.5.0) [16,17]. The model was composed of three dense layers with Rectified Linear Unit (ReLU) activation functions and regularized using dropout to avoid overfitting. The 29 units' output layer was activated with a softmax function [18]. The model had about 3×10^6 trainable parameters. The weighted (inverse of the class frequencies) categorical cross entropy loss function was minimized using Adaptative Moment Estimation (Adam) optimizer. The model was trained using batch sizes of 64 over about 10 epochs. The training time was fast and the algorithm converged smoothly. We used Accuracy (Acc = (TP+TN)/N), Specificity (Sp =TN/(TN+FP)) and Sensitivity (or recall) (Se = TP/(TP+FN) metrics for performance evaluation. For classes and subclasses where the number of cases was less than 5, performance indicators were not reported, as the interpretation of these results are challenging.

2.6. Ethics committee approval

The local Ethics committee (Comité d'éthique des Facultés de Médecine, d'Odontologie, de Pharmacie, des Écoles d'Infirmières, de Kinésithérapie, de Maïeutique et des Hôpitaux) approved this non-interventional and retrospective study (reference CE-2022-21).

3. RESULTS

3.1. Manual encoding of pharmacists' interventions

A total of 27,699 unique text documents associated to PI were extracted for the first six-months of 2017. For each document, a specific PI code was attributed independently by the two pharmacists. In case of divergent coding between the two raters (n= 1,006 discrepancies; 3.6% of the total comments) a consensus label was determined. It is important to notice that a single code was attributed to every document. Three subclasses (codes 2.3, 6.5, 9.2) were not observed in this dataset. In order to make the algorithm able to learn to classify these situations, we manually added credible synthetic comments proposed by clinical pharmacists, resulting finally in 27,721 unique documents (table 2).

3.2. Pharmacists' interventions processing

The 27,721 PI documents were split (stratification on classes and subclasses) and to form a 23,576 (85%) training dataset and a 4,145 (15%) validation dataset. In the training set, processed comments (stemmed and cleaned) were typically short with a median number of 82 characters (IQR = [59-115], min=3, max= 1,153), corresponding to a median number of 13 words or expressions. Comments were tokenized using n-grams (n from 1 to 3). We used the full training set of documents to generate a (178,680 x 23,576) term-document matrix based on term frequency weighting. The size of this initial matrix was reduced to (6,356 x 23,576) after removing sparse terms.

3.3. Deep neural network classifier

The training dataset and its corresponding 23,576 ground truth PI classes and subclasses were fed to a deep neural network composed of 3 dense layers including a 29- classes and subclasses output layer activated by a softmax function. The loss function was weighted by the inverse of the class frequencies of the training labels to account for unbalanced class representations in the training dataset. The model was then evaluated on the validation set processed similarly to the training set. The class prediction

accuracy calculated on the unseen validation dataset was 78.0%. A tabulated description of data by PI classes and subclasses is presented in Table 2 (initial algorithm). Classes and subclasses sensibilities and specificities ranged from 0.31 to 0.96 and from 0.94 to 1.0, respectively. The output of the softmax function, normalizing the output of the network last layer to a probability distribution, was used to calculate a probability score of attribution for each class label. The probability scores of the retained classes were largely asymmetric, (mean = 0.879, median = 0.999) with 2,626 out of the 4,145 predicted scores (63.3%) being larger than 0.95. As expected, the global accuracy associated to predictions with scores larger than 0.95 was better than the overall prediction, reaching 94.7% for the selected corresponding documents.

Table 2: Performance of the initial algorithm in terms of average predictive score, sensitivity, and specificity for the different classes and subclasses of drug-related problems

Code	Drug-related problems	Initial algorithm (Accuracy: **78.0%**)		
		Initial dataset n	**Validation dataset n (% ot the DRP)**	**Prediction Performance**
1	Non conformity to guidelines or contra-indication			
1.1	Non conformity of the drug choice compared to the Formulary	11,935	1,790 (43.2%)	Avg pred score : 0.96 Se : 0.93 Sp : 0.94
1.2	Non conformity of the drug choice compared to the guidelines	1,434	215 (5.0%)	Avg pred score : 0.71 Se : 0.48 Sp : 0.97
1.3	Physio-pathologic contra-indication	166	24 (0.6%)	Avg pred score : 0.69 Se : 0.37 Sp : 0.99
2	Untreated indication			
2.1	Valid indication without a drug	90	13 (0.3%)	Avg pred score : 0.66 Se : 0.31 Sp : 0.99
2.2	A drug is missing after transfer	328	49 (1.2%)	Avg pred score : 0.92 Se : 0.96 Sp : 0.99
2.3	The patient is missing a pre-medication or a prophylactic treatment	0*(12)	1 (0.0%)	Avg pred score : 0.63 Se : NA Sp : NA
2.4	A synergic or corrective drug should be associated	467	70 (1.7%)	Avg pred score :0.88 Se : 0.78 Sp :0.99
3	Subtherapeutic dosage			
3.1	Dose too low for this specific patient	1,108	166 (4.0%)	Avg pred score : 0.72 Se : 0.31 Sp : 0.98
3.2	Length of the treatment too short	29	4 (0.1%)	Avg pred score : 0.41 Se : NA Sp : NA
4	Overdosage			
4.1	Supra-therapeutic posology	2,824	423	Avg pred score : 0.85

			(10.2%)	Se : 0.74 Sp : 0.96
4.2	Duplicate prescription	787	118 (2.8%)	Avg pred score : 0.86 Se : 0.81 Sp : 0.99
5	Drug without indication			
5.1	No justified indication for the drug	585	87 (2.1%)	Avg pred score : 0.79 Se : 0.67 Sp : 0.99
5.2	The drug is being prescribed for a too long period without any risks	39	5 (0.1%)	Avg pred score : 0.44 Se : 0.80 Sp :0.99
5.3	Therapeutic redundancy	340	51 (1.2%)	Avg pred score : 0.75 Se : 0.43 Sp : 0.99
6	Drug interaction			
6.1	Take into account	4	0 (0.0%)	Avg pred score : NA Se : NA Sp NA
6.2	Precaution for use	409	61 (1.5%)	Avg pred score : 0.84 Se : 0.79 Sp : 0.99
6.3	A drug interferes with another drug and can lead to a non-adapted pharmacological impact	108	16 (0.4%)	Avg pred score :0.76 Se : 0.75 Sp : 0.99
6.4	Contra-indication between 2 drugs	647	97 (2.3%)	Avg pred score : 0.87 Se : 0.81 Sp : 0.99
6.5	Interaction published but not integrated into Vidal © dictionary	0*(5)	0 (0.0%)	Avg pred score : NA Se : NA Sp : NA
7	Adverse drug reaction	12	1 (0.0%)	Avg pred score : 0.34 Se : NA Sp : NA
8	Improper administration			
8.1	Other route more effective or less costly for the same efficacy	68	10 (0.2%)	Avg pred score : 0.92 Se : 0.90 Sp : 1.0
8.2	The method for administration is not adequate	449	67 (1.6%)	Avg pred score : 0.86 Se : 0.78 Sp : 0.99
8.3	Inappropriate drug form	1,246	186 (4.5%)	Avg pred score : 0.86 Se : 0.77 Sp : 0.99
8.4	Incomplete formulation	460	69 (1.7%)	Avg pred score :0.86 Se :0.77 Sp : 0.99
8.5	Inappropriate timing of administration and/or repartition of doses	1,298	194 (4.7%)	Avg pred score : 0.85 Se : 0.74 Sp : 0.98
9	Failure to receive drug			
9.1	Physico-chemical incompatibility between several injectable drugs	26	3 (0.1%)	Avg pred score : 0.51 Se : NA Sp : NA
9.2	Compliance problem	0*(5)	0 (0.0%)	Avg pred score : NA Se : NA Sp : NA
10	Drug monitoring	875	131 (3.2%)	Avg pred score : 0.85 Se : 0.66 Sp : 0.99
11	Other	1,965	294 (7.1%)	Avg pred score : 0.77 Se : 0.59 Sp : 0.97

Avg pred score: Average predictive score; Se: Sensitivity; Sp: Specificity.
0(n)*: Drug related problems (DRP) not observed in the dataset (n= number of credible synthetic comments added)

In order to further demonstrate the classification capacity of this algorithm, we predicted the classes and subclasses for all the PI collected between July and December 2017. A term-document matrix was generated and fed to the trained neural network to retrieve classification predictions. All predictions with a probability larger or equal to 0.95 and with a code differing from the subcategory "non-conformity of the drug choice compared to guidelines – code 1.1" (this subclass can be easily identified from Rx using drugs dictionaries) were extracted and manually checked by a pharmacist. Amongst the 4,460 predictions checked, 67 corrections were done, corresponding to an error rate of about 1.5%.

Finally, this dataset was concatenated with the original one to create an extended dataset (about 15% larger than the original) that we split as before into training and validation datasets. The network was trained again from scratch on this expanded dataset. The updated global accuracy reached 81.0% (+3.0 percent points) (see table 3 - improved algorithm).

Table 3: Performance of improved algorithm in terms of average predictive score, sensitivity, and specificity for the different classes and subclasses of drug-related problems

Code	Drug-related problems	Improved algorithm (Accuracy: 81.0%)		
		Extended dataset (n)	Validation dataset (n)	Prediction Performance
1	Non conformity to guidelines or contra-indication			
1.1	Non conformity of the drug choice compared to the Formulary	11935	2.388 (37.2%)	Avg pred score : 0.95 Se : 0.93 Sp : 0.96
1.2	Non conformity of the drug choice compared to the guidelines	1515	302 (4.7%)	Avg pred score : 0.70 Se : 0.52 Sp : 0.97
1.3	Physio-pathologic contra-indication	166	33 (0.5%)	Avg pred score : 0.62 Se : 0.42 Sp : 0.99
2	Untreated indication			
2.1	Valid indication without a drug	90	18 (0.3%)	Avg pred score : 0.66 Se : NA Sp : 0.99
2.2	A drug is missing after transfer	563	112 (1.7%)	Avg pred score : 0.95 Se : 0.95 Sp : 0.99
2.3	The patient is missing a pre-medication or a prophylactic treatment	12	2 (0.0%)	Avg pred score : 0.68 Se : NA Sp : NA
2.4	A synergic or corrective drug should be associated	599	119 (1.9%)	Avg pred score : 0.89 Se : 0.82 Sp : 0.99
3	Subtherapeutic dosage			
3.1	Dose too low for this specific patient	1208	242 (3.8%)	Avg pred score : 0.72 Se : 0.52 Sp : 0.98

3.2	Length of the treatment too short	29	5 (0.1%)	Avg pred score : 0.56 Se : NA Sp : NA
4	Overdosage			
4.1	Supra-therapeutic posology	3849	767 (11.9%)	Avg pred score : 0.84 Se : 0.72 Sp : 0.98
4.2	Duplicate prescription	1399	277 (4.3%)	Avg pred score : 0.90 Se : 0.89 Sp : 0.99
5	Drug without indication			
5.1	No justified indication for the drug	679	136 (2.1%)	Avg pred score : 0.73 Se : 0.65 Sp : 0.99
5.2	The drug is being prescribed for a too long period without any risks	39	7 (0.1%)	Avg pred score : 0.46 Se : 0.86 Sp : 1.0
5.3	Therapeutic redundancy	343	70 (1.1%)	Avg pred score : 0.73 Se : 0.44 Sp : 0.99
6	Drug interaction			
6.1	Take into account	4	0 (0.0%)	Avg pred score : NA Se : NA Sp : NA
6.2	Precaution for use	738	147 (2.3%)	Avg pred score : 0.91 Se : 0.87 Sp : 0.99
6.3	A drug interferes with another drug and can lead to a non-adapted pharmacological impact	108	21 (0.3%)	Avg pred score : 0.70 Se : 0.48 Sp : 0.99
6.4	Contra-indication between 2 drugs	865	173 (2.7)	Avg pred score : 0.95 Se : 0.92 Sp : 0.99
6.5	Interaction published but not integrated into Vidal © dictionary	5	1 (0.0%)	Avg pred score : 0.98 Se : NA Sp : NA
7	Adverse drug reaction	12	2 (0.0%)	Avg pred score : 0.40 Se : NA Sp : NA
8	Improper administration			
8.1	Other route more effective or less costly for the same efficacy	132	26 (0.4%)	Avg pred score : 0.95 Se : 0.88 Sp : 0.99
8.2	The method for administration is not adequate	589	117 (1.8%)	Avg pred score : 0.86 Se : 0.87 Sp : 0.99
8.3	Inappropriate drug form	1527	304 (4.7%)	Avg pred score : 0.87 Se : 0.81 Sp : 0.99
8.4	Incomplete formulation	596	119 (1.9%)	Avg pred score : 0.87 Se : 0.84 Sp : 0.99
8.5	Inappropriate timing of administration and/or repartition of doses	1821	364 (5.7%)	Avg pred score : 0.89 Se : 0.83 Sp : 0.98
9	Failure to receive drug			
9.1	Physico-chemical incompatibility between several injectable drugs	26	5 (0.1%)	Avg pred score : 0.53 Se : NA Sp : NA
9.2	Compliance problem	5	1 (0.0%)	Avg pred score : 0.89 Se : NA Sp : NA
10	Drug monitoring	1158	231 (3.6%)	Avg pred score : 0.85 Se : 0.77 Sp : 0.99
11	Other	2169	436	Avg pred score : 0.76

			(6.8%)	Se : 0.66 Sp : 0.98

Avg pred score: Average predictive score; Se: Sensitivity; Sp: Specificity.

4. DISCUSSION

The study presents a simple pipeline including a robust deep neural network classifier to categorize DRP from PI. To our knowledge, no other deep learning DRP classification system has been described in the literature to date. The classification relied on the external validated SFPC classification[13]. This is currently the most widely used DRP classification method in France [19]. It contains explicit definitions for each category of DRP to reduce ambiguity in coding a problem.

The 29 classes and subclasses of the initial dataset were unbalanced, with an over-representation of the situations corresponding to “non-conformity of the drug choice compared to the formulary” and “non-conformity of the drug choice compared to the guidelines” and an under-representation of classes such as “adverse drug reaction” or “failure to receive drug”. This is in full line with the fact that these events are rare - as described by Bedouch et al. [19]. In their study, "adverse drug reaction" or "failure to receive a drug" accounted for 4.3% and 0.8% of their 34,522 interventions, respectively [19]. To note also that the class named “other” represented 7.1% of all PI of the initial dataset, showing possible coding difficulties of the experts due to a lack of detail in comments or complexity of interpretations. This has to be compared with some studies in which the "other" class represented up to 39% of all DRP.[19]

Regarding the initial training, 78% accuracy was achieved on the validation dataset. The model performed worst for subclasses such as, “physio-pathologic contra-indication”, "valid indication without drug", “dose too low for this specific patient”. Several explanations can be given - including insufficient number of cases in the initial dataset or more linguistically complex documents that may require some inference to understand which class the entities involved are related to.

Algorithmic prediction errors in the additional dataset (second six-months of 2017) spotted by an expert pharmacist were low (1.5% of the predictions with a probability larger or equal to 0.95 and a code differing from the subcategory “non-conformity of the drug choice compared to guidelines – code 1.1”). Detailed manual analysis of these errors showed that they corresponded to unique clinical situations, misspellings, or grouping of multiple PI.

Incorporating additional validated data from the second half of 2017 (see figure 1) into an expanded dataset improved performance (accuracy reached 81%). This clearly shows that the prediction process has the potential to improve further with an increase in the amount of data and strengthens the validity of the results. Algorithmic prediction errors in the additional data (second six-months of 2017) spotted by the expert pharmacist were few (1.6%). They corresponded to unique clinical situations, misspellings, or grouping of multiple DRPs in a single comment.

This study has several limitations. Our findings are currently limited in scope as the study was conducted in a single French hospital setting. Its generalization properties to other institutions has not been explored. Adult patients in intensive care units, obstetrics and gynaecology, and psychiatry and children were not included. Consequently, evidence of the accuracy of the algorithm to classify PI in these targeted populations has yet to be demonstrated. In addition, MR can be influenced by the pharmacists' experience. To limit this bias, the hospital's pharmacy quality policy requires that all residents and newly hired pharmacists undergo specific training to be authorized to perform PI and write comments.

The future objectives will be to exploit this deep neural network algorithm to perform large descriptive analysis of PI and DRP intercepted through MR with the aim to value the work of clinical pharmacists. The description of DRP classes, automatically generated by the algorithm, will also be crucial to increase knowledge on the types and frequencies of DRP observed in different clinical settings. Particular attention will be paid to the prevalence rates of DRP by drug classes, by patient profiles (age, number of drugs prescribed) or medical specialities. Although this type of analysis will focus on the drug related problems spoted by pharmacists but not on their cause or consequence, it will be useful to alert medical and pharmaceutical teams about certain prescribing patterns and draw the attention on practices that could be improved.

5. CONCLUSION

PI classification is beneficial for assessing and improving pharmaceutical care practice. It should be systematized as part of the care process. Here we report a high performance automatic PI classification

based on deep learning that could find a crucial place for highlighting the clinical relevance of drug prescription reviews performed on a daily basis by hospital pharmacists.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS

AA, JG, EJ, BG and BM all made substantial contributions to the conception of the work and the acquisition, analysis, and interpretation of data for the work; drafted the work and revised it critically for important intellectual content; approved the final version of the work to be published; and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICT OF INTEREST STATEMENT

AA, JG, EJ, BG and BM have no conflicts of interest that are directly relevant to the content of this study.

SUMMARY TABLE

what was already known on the topic

- Medication Review enables to identify drug-related problems (DRP) and triggers pharmacists' interventions (PI). PI aims to adhere to best practices in prescribing.
- DRP and PI classification is an important part of pharmaceutical care practice.

what this study added to our knowledge

- The deep neural network classifier we developed highlights the potential contribution of artificial intelligence to improving clinical pharmacy practices.
- The classifier developed can help pharmacists to classify PI and alert them on certain risky prescribing.

REFERENCES

1. Renaudin R, Baumstarck K, Daumas A, *et al*. Impact of a pharmacist-led medication review on hospital readmission in a pediatric and elderly population: study protocol for a randomized open-label controlled trial. ***Trials*** 2017;18(1):65.

2. Holland R, Smith R, Harvey I. Where now for pharmacist led medication review?. ***J Epidemiol Community Health*** 2006;60(2):92-93.

3. Vinks TH, Egberts TC, de Lange TM, *et al*. Pharmacist-based medication review reduces potential drug-related problems in the elderly. ***Drugs Aging*** 2009;26(2):123-33.

4. Cheung KC, van der Veen W, Bouvy ML, *et al*. Classification of medication incidents associated with information technology. ***J Am Med Inform Assoc*** 2014;21(e1):e63–e70.

5. Bright JM, Tenni PC. The Clinical Services Documentation (CSD) System for Documenting Clinical Pharmacists' Services. ***Aust J Hosp Pharm*** 2000;30(1):1015.

6. Pfister B, Jonsson J, Gustafsson M. Drug-related problems and medication reviews among old people with dementia. ***BMC Pharmacol Toxicol*** 2017;18(1):52.

7. Pharmaceutical Care Network Europe, PCNE. Working group on drug-related problems. http://www.pcne.org/sig/drp/drug-related-problems.php (accessed 26 July 2022).

8. ASHP statement on pharmaceutical care. ***Am J Hosp Pharm*** 1993;50:1720-1723.

9. Barber ND, Batty R, Ridout DA. Predicting the rate of physician-accepted interventions by hospital pharmacists in the United Kingdom. ***Am J Health Syst Pharm*** 1997;54(4):397-405.

10. Mutnick AH, Sterba KJ, Peroutka JA, *et al*. Cost savings and avoidance from clinical interventions. ***Am J Health Syst Pharm*** 1997;54(4):392-6.

11. Galindo C, Olivé M, Lacasa C, *et al*. Pharmaceutical care: pharmacy involvement in prescribing in an acute-care hospital. ***Pharm World Sci.*** 2003;25(2):56-64.

12. van Mil, J. F, Westerlund, L. T, Hersberger, K. E, *et al*. Drug-related problem classification systems. ***Ann Pharmacother*** 2004;38(5):859-867.

13. Allenet B, Bedouch P, Rose F X, *et al*. Validation of an instrument for the documentation of clinical pharmacists' interventions. ***Pharm World Sci*** 2006;28(4):181-188.

14. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL. http://www.R-project.org (accessed 26 July 2022).

15. Kuhn M, Wing J, Weston S, *et al*. Package 'caret' - Classification and Regression Training. April 19, 2022. https://cran.r-project.org/web/packages/caret/caret.pdf (accessed 26 July 2022).

16. Keras. https://keras.io/ (accessed 26 July 2022).

17. TensorFlow. https://www.tensorflow.org/ (accessed 26 July 2022).

18. Javier Rubio F. The softmax Function. 2020. https://rpubs.com/FJRubio/softmax#:~:text=The%20softmax%20function%20is%20the,1exp(xj). (accessed 26 July 2022)

19. Bedouch P, Sylvoz N, Charpiat B, *et al*. Trends in pharmacists' medication order review in French hospitals from 2006 to 2009: analysis of pharmacists' interventions from the Act-IP© website observatory. ***J Clin Pharm Ther*** 2015;40(1):32-40.

20. Lesar TS. Prescribing errors involving medication dosage forms. J Gen Intern Med. 2002 Aug;17(8):579-87.

21. Basger BJ, Moles RJ, Chen TF. Application of drug-related problem (DRP) classification systems: a review of the literature. ***Eur J Clin Pharmacol*** 2014;70(7):799-815.

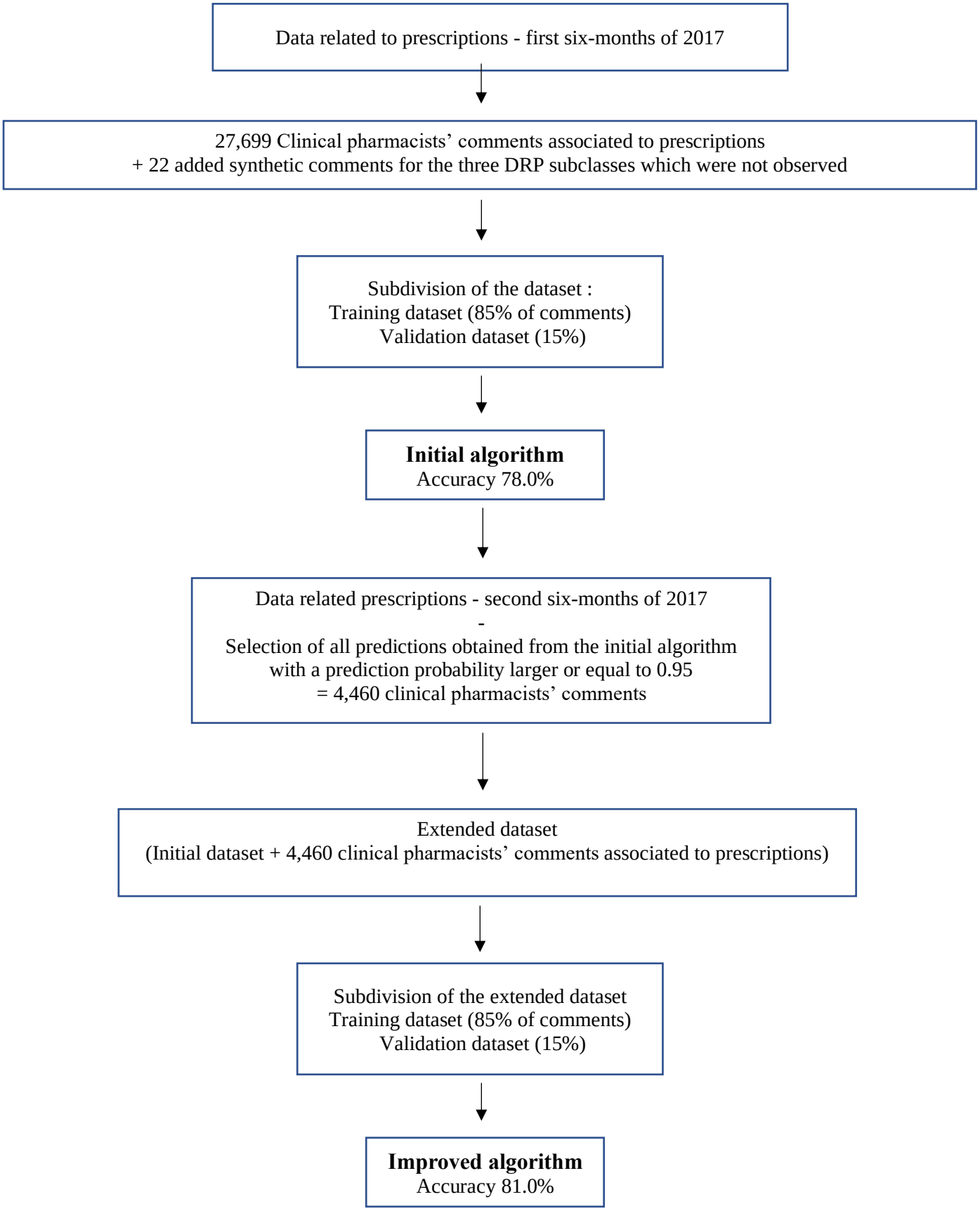

Figure 1: Flowchart summarizing the different steps that led to the developpement of the improved algorithm