dHealth 2019 – From eHealth to dHealth
D. Hayn et al. (Eds.)
© 2019 The authors, AIT Austrian Institute of Technology and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms
of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/978-1-61499-971-3-128

Design and Evaluation of a Smart Medication Recommendation System for the Electronic Prescription

Seyed Hadi GHASEMI[a], Kobra ETMINANI[a,1], Hamidreza DEHGHAN[a], Saeid ESLAMI[a], Mohammad Reza HASIBIAN[a], Hasan VAKILI ARKI[a], Mohammad Reza SABERI[a], Mahdi AGHABAGHERI[b], Seyedeh Mahdieh NAMAYANDEH[b]
[a]*Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran*
[b]*Shahid Sadoughi University of Medical Sciences, Yazd, Iran*

Abstract. Background: electronic prescription is shown to have many benefits in terms of reducing medication errors, improving patient safety, productivity, and resource management, but it may cause new errors and physician frustration if not designed and implemented properly. Improving usability and user-centered design is essential for physicians' adoption. Objectives: To enhance the efficiency of the e-prescribing system by reducing the risk of inappropriate selection of the medication and also to reduce the prescribing time and effort to reach the desired drug. Methods: Important data fields for predicting medications were determined through interviews with pharmacists. Among those, fields which were available in a claims dataset of 16 million prescriptions were extracted and were used to develop a neural network model to be used by a recommender system that displays the most probable medications on top of the drop-down list in the e-prescription application. Results: Offline and field evaluations both showed that this model could improve performance. Conclusion: smart recommenders systems can improve e-prescription usability, safety, and enhanced physicians' adoption.

Keywords. electronic prescribing, recommender system, usability

1. Introduction

Medical errors are the third leading cause of death in the United States[1], and about 20% of these errors are related to medication errors [2, 3]. Electronic prescribing (e-prescribing) has been offered as a solution to this problem and has been shown to have many benefits [4-7]. But on the other hand, if electronic prescribing is not implemented properly, it can add new errors (e-iatrogenesis) [8-11]

Usability and User-Centered Design (UCD) [12] are critical elements in the design and development of electronic prescription and electronic health record (EHR) systems in general, which can enhance patient's safety and encourage physicians' adoption and reduce their dissatisfaction with these systems [13].

[1] Corresponding Author: Kobra Etminani, Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran, E-Mail: etminanik@mums.ac.ir

Poor usability not only results in an increased level of clinician frustration but also can lead to errors, posing serious threats to patient safety. [14-16]

This study is aimed to enhance the efficiency of the e-prescribing system by reducing the risk of inappropriate selection of the medication and also to reduce the prescribing time of the physicians. We propose a model that recommends the most commonly prescribed medication on top of the drop-down list in the e-prescription application and showed that it can improve performance.

2. Methods

The main idea was based on the observation that pharmacists can read physicians' handwritten prescriptions just with a few legible letters. It was obvious that they use some complementary information such as the physicians' specialty to narrow down the search space which contains all possible alternatives to the correct medication, eventually reaching a very probable result and confirming that with other clues on the prescription. We decided to use this approach to design a smart system to "guess" what might be in the physician's mind when he/she has entered only a few characters from the beginning of drug's name. So that the application could display the search results based on their probability, instead of sorting them alphabetically.

At the first phase of the study, in order to extract the pharmacists' tacit knowledge and find out which information fields they use to reach the conclusion about a specific drug, we conducted semi-structured deep interviews with pharmacists. They were provided an initial list of eight fields and asked to try to describe how they think and talk about the information fields they may use if the case of facing an illegible prescription. These fields were identified through two brain-storming sessions and include "physician specialty", "frequency of drug use in general", "frequency of drug use among physicians with same specialty", "other drugs in the prescription", etc. The pharmacists were allowed to modify the list, add new fields and/or remove useless fields from the list. The viewpoints of the interviewees where noted and also all the sessions' voice were recorded with the interviewees' permission. Recorded voices were rechecked by the interviewers to ensure nothing is missed from the notes. Interviews were continued until information saturation and no new fields were added to the list in two consequent interviews. A total of sixteen pharmacists were interviewed.

A checklist containing all information fields we found in the interviews was prepared and sent to twenty pharmacists via email, asking them to rank the fields by their relative importance in reading illegible prescriptions. Then the fields were sorted by their average rank, making the final list ready for the next phase.

Some of the identified fields were not available (for example, the "diagnosis" field is not recorded in claims data), and also some are not applicable in the context of electronic prescribing, such as the number of ordered drug which won't be available to the system before the drug itself is known.

We used a drug claims database of over 16 million prescriptions containing 46 million drug items, prescribed over two consequent years in a large province of Iran, to build and train a model for predicting drug names in the context of an electronic prescription system.

To build a model, we used the "Lift" concept which is commonly used in the data-mining approach: "Association Rule Mining". It is defined as the ratio of two probabilities. The ratio of the probability of an event might occur in a specific condition

to the probability of that event might occur in general. We used this value to rank drugs matching the user's input and sort the search result based on those ranks.

For example, when a drug name starts with letters "Ac", it can be "Acetaminophen", "Acetazolamide" or some other drugs. In our database, there were 471.000 prescriptions (out of 16 million) which contain "ACETAMINOPHEN 325MG TAB". So the probability of this drug to be prescribed in general is 471,000/16,000,000 which equals to 0.029. But when we have another piece of information, the probability may change. If we know that the physician is an ophthalmologist, the database says that there are about 186000 prescriptions from ophthalmologists among which "ACETAMINOPHEN 325MG TAB" is prescribed 700 times. So the probability changes to 700/186000 which equals to 0.00382. In this example, knowing the specialty of the physician changed the probability to about 1/7 of its previous value. On the other hand, the same calculations for "ACETAZOLAMIDE 250MG TAB" changes its general probability of 0.00097 to 0.03614, leading to a lift of 37.25. This means that ophthalmologists prescribe this drug 37 times more often than general.

In the same way, other pieces of available information change the probability upward or downward. With the available data we could access 7 information fields:

- Doctor profile
- Patient
- Specialty
- Previous Drug (i.e. other drugs in the same prescription)
- Previous Drug in the Same Specialty as the doctor
- Drug Simple Name (I.e. all dosage-forms and strengths of the drug)
- Drug Simple Name in the Same Specialty as the doctor

To combine these effects in a weighed manner we constructed a "simulations" table by selecting random prescriptions from the database, simulating their prescribing and finding all matching drugs (beginning with first N letters of the drug name) and calculating lifts for those 7 fields for all matches. The "class label" field for the actual prescribed drug was set to 1 and for other matching drugs set to 0. We tested 45 different combinations of those fields and the number of matching characters, N. In this way, the simulations table was filled with 24 million simulated rows for different configurations.

Then we used Brain.js open-source neural network library (https://github.com/BrainJS) to set up and train a 5-layer neural network with the back-propagation learning algorithm.

The neural network was fed with 45 sets of 10000 randomly selected cases in each set (with almost equal distribution of rows with class labels 0 and 1, 30% for train and 70% for test).

The trained network was exported as a JavaScript function which gets an array of 8 lifts (1st field was drugs general probability and the rest were lifts for 7 different information fields) and returns a number in (0,1) range denoting the probability of the drug being prescribed given those lifts.

2.1. Offline Evaluation (Lab study)

2.1.1. Matching-drug level

After training the neural network, the model output was calculated for all rows in the simulations table. To calculate the model's performance we used ROC Curve analysis on the simulations table and set appropriate cut-off points for each configuration group. This evaluation measures the performance of the model in its lowest level: assigning the correct label to each of drugs in search result set.

2.1.2. Selected-Drug level

We evaluated the model performance in "selected-drug" level by calculating the rank of the actually prescribed drug in each set of prescription items in the simulations table between three sorting methods: sorting by drug name alphabetically, simple sorting by drugs frequency in general, and sorting by the model output. Then compared the percentage of the actual drug being in the first rank, top 3, top 5 or top 10 in these three sorting methods.

2.2. User Evaluation (Prescription level)

For real-user evaluation (prescription level), we implemented a recommender system into an existing laboratory electronic prescription software. The recommender system was based on the exported model from the neural network.

Then we asked 26 physicians to write prescriptions for 10 scenarios of patients with common complications (high blood pressure, migraine, sinusitis, etc.) in a cross-over design. Participants were randomly assigned to two groups. In the first round, one of which worked with unmodified software (alphabetical sort) and the other worked with software plus enhanced sorting of search results. In the next round, the software for groups was swapped.

In this step, we have to select one of the 45 configurations. Since the patients were not real patients, there was no patient profile available. Also, physicians who participate in our study were not among those we had their claims data, so physician profiles were not available either. For the number of typed characters, we choose 3. So the neural network model used to enhance the search results in this evaluation step was based on physician's specialty and co-occurrence of drugs in same prescription and assuming that user has entered first three characters of the drugs' name (model code: Spec-Drug-3).

All physicians' activities in the software were logged and analyzed. Time to find the desired drug, number of typed characters to reach the desired drug, and the position of the selected drug in the list were extracted from the log for analysis. The multi-regression analysis was used to test the efficacy of the model with control of intervention order.

3. Results

The first phase resulted in an ordered list of 22 information fields. The most important fields where drug's form, asking the patient about his/her medication history, considering other legible items that may exist in the prescription, physician's specialty, drug's dosage,

Table 1. Evaluation steps for the model. In the matching-drug level, the result of the model for each search result, and its concordance with the drug that actually prescribed is the basis of the performance measurement. In the selected-drug level, the rank of the desired drug within the result set determines the performance. In the prescription level, some measures of overall performance are recorded and compared between groups.

1. Select random prescriptions from the claims database		
2. For each drug item ("desired drug") in each selected prescription		
2.1. Extract first N letters of drugs name => user query		
2.2. Find All Drug Names Starting with "user query" => matched drugs		
2.3. For each "Matched Drug":	Matching Level	Offline Evaluation (Lab Study)
2.3.1. Calculate Lifts in the actual prescription context		
2.3.2. Feed lifts into the model and calculate result => "model result"		
2.3.3. Assign a class label: If "model result">= threshold, class="Yes", Otherwise class = "No"		
2.3.4. Assign "classification result":		
	"matching drug" = "desired drug" / "matching drug" ≠ "desired drug"	
Class label = "Yes"	True Positive (TP) / False Positive (FP)	
Class label = "No"	False Negative (FN) / True Negative (TN)	
3. Count the number of TPs, TNs, FPs, and FNs		
4. **Calculate the model performance measures: *Sensitivity = TP / (TP + FN), Specify = TN/TN+FP***		
5. For each drug item ("desired drug") in selected prescription	Selected Drug Level	
5.1. Sort "matching drugs" by name of drugs alphabetically.		
5.2. Assign "Alphabet-Top-N" label to "Yes" or "No", by checking whether the "desired drug" is among top N drugs of the list or not. (N=1,3,5,10)		
5.3. Sort "matching drugs" by "model result" in descending order.		
5.4. Assign "Frequency-Top-N" label to "Yes" or "No", by checking whether the "desired drug" is among top N drugs of the list or not. (N=1,3,5,10)		
5.5. Sort "matching drugs" by "model result" in descending order.		
5.6. Assign "Model-Top-N" label to "Yes" or "No", by checking whether the "desired drug" is among top N drugs of the list or not. (N=1,3,5,10)		
6. Count the number of "Yes"s and "No"s in each sorting method, for each level of N.		
7. **Compare differences between each pair of sorting method, for each level of N. (Using statistical methods, such as *Chi-square* test)**		
8. Implement the model into a laboratory e-prescribing application, define some test scenarios.	Prescription Level	User Experience
9. Ask physicians to prescribe drugs for test cases, record all users' activities in the application. Set the sorting method to "alphabet" or "model", in a cross-over design.		
10. Measure "time to find the desired drug", "number of entered characters" to reach the desired drug and "the position of the selected drug".		
11. **Compare differences of measures in the previous step, between the two sorting methods, using statistical methods such as "multi-regression analysis".**		

asking the patient about his/her symptoms, usage instructions, and patient's age, respectively.

These fields were categorized into five groups: fields related to (1) patient's profile, (2) physician's profile, (3) physician's specialty, (4) the medication's properties and (5) other medications in the prescription.

In the "matching drug" level, 45 ROC curve analyses were made. Figure 1 shows a sample of 3 ROC curves for configurations that previous drug and physician's profile is used, and the user has entered 2, 3 or 4 letters of the drug name. In this sample, choosing cut-off points of 0.77 for the first curve (2 letters) results in a sensitivity of 0.931 and specificity of 0.921. Table 2 shows these performance measures.

In the "selected-drug" level, as shown in figure 2, in alphabetical sort, the desired drug is in top of the list in only 12% of cases, but when sorting by results of the model based on the physician's profile, the first suggested drug is the desired one in about two-thirds of cases. Sorting on the general frequency better results than alphabetical sort, but performance is lower than the model. The same difference is seen when comparing the desired drug being in top 3, top 5 or top 10 suggestions.

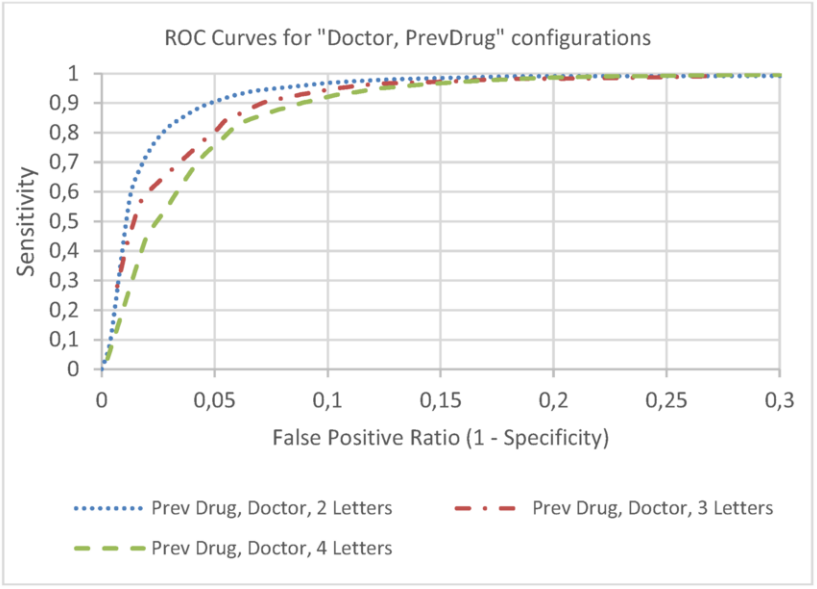

Figure 1. ROC Curves for model accuracy, when the physician's profile and previous drug in the prescription are known, and the user has entered 2, 3 or 4 letters of the drug name. Note that for better illustration, the horizontal axis range is changed to (0,0.3).

Table 2. Model performance measures in "matching drugs" level, for different combinations of known fields and the number of entered characters. Minimum and maximum values for each measure are in bold. SEN: Sensitivity, SPC: Specificity, spec: doctor's specialty, prevdrug: previous drug in the same prescription.

	2 Letters		3 Letters		4 Letters	
Known Fields	**SEN**	**SPC**	**SEN**	**SPC**	**SEN**	**SPC**
patient	**0.999**	0.990	0.999	0.985	0.998	0.986
doctor	0.863	0.945	0.919	0.900	0.930	0.874
prevdrug	0.933	0.865	0.902	0.870	**0.811**	0.895
Spec	0.902	0.876	0.860	0.853	0.903	**0.821**
doctor patient	0.995	0.992	0.997	0.988	0.996	0.982
prevdrug patient	0.997	0.992	0.998	0.980	0.998	0.987
prevdrug doctor	0.925	0.942	0.919	0.919	0.917	0.902
spec doctor	0.931	0.921	0.942	0.881	0.950	0.863
spec patient	0.996	0.992	0.997	0.986	0.997	0.986
spec prevdrug	0.926	0.901	0.930	0.860	0.905	0.860
prevdrug doctor patient	0.993	**0.993**	0.997	0.985	0.996	0.982
spec doctor patient	0.996	0.991	0.996	0.987	0.996	0.987
spec prevdrug doctor	0.950	0.927	0.933	0.910	0.941	0.882
spec prevdrug patient	0.996	0.992	0.997	0.988	0.998	0.984
spec prevdrug doctor patient	0.997	0.991	0.996	0.987	0.993	0.987

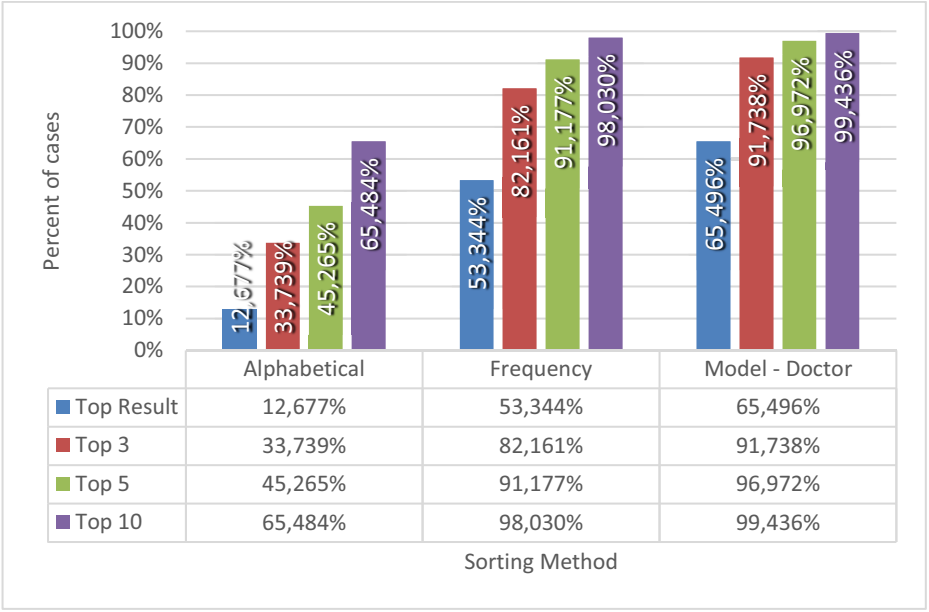

Figure 2. Percentage of cases that the desired drug is in 1st rank, top 3, top 5 or top 10 suggestions. Comparing alphabetical sorting with sorting by frequency and sorting by model results.

Chi-square tests showed that differences observed between each pair of sorting methods, in all these four levels, and in all configurations of models were statistically significant ($p < 0.001$, df=1).

In the last evaluation step (user experience, prescription-level), the multi-regression analysis showed that sorting by the model is significantly better than alphabetical sort in terms of less time to find the desired drug ($p<0.001$) and fewer number of entered characters ($p<0.01$). The position of the selected drug was not significantly different between sorting methods ($p>0.05$).

4. Discussion

Recommender systems are widely used in commercial and e-commerce sites, and many methods for implementing and evaluating these systems are developed [17].
In this project, we used collaborative filtering[18] method to enhance user usability of an e-prescribing system.

In 2014, Syed-Abdul et al. proposed a smart model that recommends most commonly prescribed medications in the drop-down menu for a given disease. They used association between diagnosis and prescribed drugs to calculate Mean Prescription Rank (MPR) of prescriptions and Coverage Rate (CR) of prescriptions and developed a model to compute a proactive medication list using these concepts. They showed that this system can shorten the length of the medication drop-down menu in the electronic

prescription application and concluded that this could improve safety and save time. They showed that "diagnosis" field can be used in developing recommender systems.

Our study showed that patient's profile, physician's profile, physician's specialty and other prescribed drugs can also be used alone or in combination with each other to develop recommender systems for electronic prescribing.

Future researches may combine these fields with diagnosis and reach better results.

Although we could show that recommender system can improve usability by reducing time and effort to find the desired drug, it's efficacy to enhance patient safety should be studied in future researches in physicians' routine practice.

References

[1] Makary, M.A. and M. Daniel, *Medical error-the third leading cause of death in the US.* BMJ, 2016. **353**: p. i2139.

[2] Bates, D.W., et al., Incidence of Adverse Drug Events and Potential Adverse Drug Events: *Implications for Prevention.* JAMA, 1995. **274**(1): p. 29-34.

[3] Tamblyn, R., et al., *The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care.* CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne, 2003. **169**(6): p. 549-556.

[4] Eslami, S., N.F. de Keizer, and A. Abu-Hanna, *The impact of computerized physician medication order entry in hospitalized patients--a systematic review.* Int J Med Inform, 2008. **77**(6): p. 365-76.

[5] Meisenberg, B.R., R.R. Wright, and C.J. Brady-Copertino, *Reduction in chemotherapy order errors with computerized physician order entry.* J Oncol Pract, 2014. **10**(1): p. e5-9.

[6] Porterfield, A., K. Engelbert, and A. Coustasse, *Electronic prescribing: improving the efficiency and accuracy of prescribing in the ambulatory care setting.* Perspectives in Health Information Management, 2014. **11**(Spring).

[7] Reckmann, M.H., et al., *Does computerized provider order entry reduce prescribing errors for hospital inpatients? A systematic review.* J Am Med Inform Assoc, 2009. **16**(5): p. 613-23.

[8] Weiner, J.P., et al., *"e-Iatrogenesis": the most critical unintended consequence of CPOE and other HIT.* 2007. **14**(3): p. 387-388.

[9] Nanji, K.C., et al., *Errors associated with outpatient computerized prescribing systems.* J Am Med Inform Assoc, 2011. **18**(6): p. 767-73.

[10] Campbell, E.M., et al., *Types of unintended consequences related to computerized provider order entry.* J Am Med Inform Assoc, 2006. **13**(5): p. 547-56.

[11] Koppel, R., et al., *Role of computerized physician order entry systems in facilitating medication errors.* Jama, 2005. **293**(10): p. 1197-203.

[12] Ratwani, R.M., et al., *Electronic health record usability: analysis of the user-centered design processes of eleven electronic health record vendors.* Journal of the American Medical Informatics Association, 2015. **22**(6): p. 1179-1182.

[13] Cohen, J.F., J.-M. Bancilhon, and M.J.S.A.C.J. Jones, *South African physicians' acceptance of e-prescribing technology: An empirical test of a modified UTAUT model.* 2013. **50**(1): p. 43-54.

[14] Tamblyn, R., et al., *The development and evaluation of an integrated electronic prescribing and drug management system for primary care.* Journal of the American Medical Informatics Association: JAMIA, 2006. **13**(2): p. 148-159.

[15] Johnson, K.B., et al., *Showing Your Work: Impact of annotating electronic prescriptions with decision support results.* J Biomed Inform, 2010. **43**(2): p. 321-5.

[16] Halamka, J., et al., *E-Prescribing collaboration in Massachusetts: early experiences from regional prescribing projects.* Journal of the American Medical Informatics Association: JAMIA, 2006. **13**(3): p. 239-244.

[17] Portugal, I., P. Alencar, and D.J.E.S.w.A. Cowan, *The use of machine learning algorithms in recommender systems: a systematic review.* 2018. **97**: p. 205-227.

[18] Ekstrand, M.D., et al., Collaborative filtering recommender systems. 2011. 4(2): p. 81-173.