# ADEQUACY OF USING WARFARIN AMONG ELDERLY INPATIENTS

*Ya-Han Hu, Department of Information Management, National Chung Cheng University, 168 University Road, Min-Hsiung, Chia-Yi 62102, Taiwan, 886-5-2720411 ext. 34622, yahan.hu@mis.ccu.edu.tw*
*Chia-Lun Lo, Department of Information Management, National Chung Cheng University, 168 University Road, Min-Hsiung, Chia-Yi 62102, Taiwan, 886-5-2720411 ext. 24614, ntallen1@gmail.com*
*Kang Ernest Liu, Department of Economics, National Chung Cheng University, 168 University Road, Min-Hsiung, Chia-Yi 62102, Taiwan, 886-5-2720411 ext. 34120, ECDKL@ccu.edu.tw*

## ABSTRACT

Considering a rapidly aging society, improvement in high-alert medication safety becomes crucial in clinical decision making. Effectiveness of using these drugs is complicated due to drug-to-drug interactions for the elderly. In our study, adequacy of using warfarin among the elderly is investigated using 360 inpatient cases in Taiwan. Our evaluation results show that incorporation of a learning-based classification approach can improve predicting performance and thus decision support systems using decision-tree induction may be constructive in the clinical practice of these high-alert medications.

## INTRODUCTION

Aging of the population has increased significantly worldwide. The elderly population has reached 15% in the world and is predicted to surpass 20% in 2020 and climb to 31.86% in 2050 [7]. This aging phenomenon is even worse in the developed countries and has raised attention to issues like health care, economy, and education. One of the most important issues for the elderly is medical problems. Due to physiological degradation, the increasing probability of having chronic diseases and disability drives up the demand for medical treatment and drug consumption [6].

Confronted with multiple chronic diseases, the elderly often need to take a variety of drugs and may face risks from polypharmacy. Many studies have indicated that inappropriate medication for the elderly has attracted increasing attention. In the US, related studies show that the proportion of inappropriate prescription drugs can reach 10% to 27% and that the waste of medical resources should be reduced [11][13]. Among the drug safety problems for the elderly, one of the most serious issues is the use of narrow therapeutic range medications. Because of aging and multi-drug use, effects of these high-alert drugs become difficult to predict for elderly inpatients and thus can create serious adverse drug problems; adequate usage of these high-alert medications is extremely important to elderly patients, clinicians, pharmacists, hospital managers and society as a whole.

Warfarin is one of the most commonly used anticoagulants to prevent diseases like thromboembolism [1]. Among the cardiovascular prescription drugs used in the US, warfarin is ranked the fourth most frequently-used anticoagulant [4]. Oral absorption makes warfarin one of the most effective anti-coagulants [10]; however, warfarin is also at the top list of adverse drug events in the US [2], mainly because of its narrow therapeutic range and high drug-to-drug interaction (DDI) with nearly 250 medicines [12]. Compared with young patients, this issue becomes more serious to the elderly because elderly patients are more likely to take multiple-drugs to control chronic diseases and their physical conditions may be worse than younger ones. Therefore, improving the management of clinical use of warfarin for the elderly becomes crucially important.

Even though data mining techniques have been successfully applied to the study of vancomycin, one of the high-alert drugs in antibiotic medication [5], investigation of the adequacy of warfarin usage is rare in the medication-related decision support research [8]. Colombet et al. [2] incorporated dosing nomograms for the physician making the dosage; however, this decision is based on an inpatient's age and the INR (internalization normalized ratio) value. Without considering other factors such as DDI, the decision to use warfarin can be hazardous. Therefore, the purpose of this study is to examine whether the adequacy of using warfarin among elderly inpatients can be improved when data mining techniques are incorporated. Our experimental results demonstrate that prediction accuracy can be significantly improved when an efficient artificial intelligence classification model is employed, showing a promising tool for clinical practice.

## METHOD AND MATERIAL

### Classification Techniques

To build a dosage adequacy evaluation system for the elderly, this study investigates two well-known single classification techniques, C4.5 [9] and logistic regression (LGR). C4.5 is a decision-tree-based classification technique. The tree generation process of C4.5 consists of two phases: the growing and the pruning phases. In the growing phase, a divide-and-conquer approach is utilized to select a suitable variable as an internal node of the decision tree and thus to partition the training dataset into subsets. This process is recursively applied to each internal node (i.e., a subset of the training dataset) until any of the stopping criteria is satisfied. Meanwhile, a class label is given for a leaf node based on majority voting. As in the pruning phase, C4.5 adopts the pre-pruning approach to reduce the size of a tree so that problems resulting from both noise data and the over-fitting can be alleviated.

On the other hand, the LGR is a widely used statistical technique for modeling a dependent variable by a linear combination of one and several independent variables. The main difference between LGR and linear regression is that LGR deals with binomial or multinomial classification problems while linear regression requires the dependent variable to be of interval or ratio scales. Meanwhile, the LGR aims to predict the occurrence probability of an event by fitting data into a logistic function, thereby allowing inputs with any values to be transformed and confined to values between 0 and 1. The form of the logistic regression formula is defined as: $p = \frac{1}{1+e^{-(\beta_0+\Sigma_{i=1}^{k}\beta_i x_i)}}$, where $\beta_0$ is the intercept and $\beta_i$ is the regression coefficient of independent variable $x_i$. The probability of the chosen class is expressed by $p$.

Each regression coefficient represents the degree of contribution of the corresponding variable. A positive regression coefficient increases the probability of the chosen class; while a negative one decreases it.

In addition to the investigations of C4.5 and LGR, this study further employs classifier ensembles to enhance the predictive power of the classical classification techniques. Adaptive Boosting [3], or AdaBoost in short, is one of the most popular classifier ensembles and can be integrated with several supervised learning algorithms. AdaBoost iteratively applies a selected classification algorithm and evaluates each instance in the training dataset. For the instances incorrectly classified by the current classifier, their weight will increase for the next round of learning. In other words, AdaBoost encourages a new classifier to learn from instances misclassified by earlier ones by assigning larger weight to those instances. After a sequence of classifiers is built, AdaBoost utilizes a weighted majority vote to make predictions. Although the concept of AdaBoost is simple, previous studies have shown that several classification algorithms in conjunction with AdaBoost achieve higher classification accuracy than individual base classifiers.

**Data**

In this study, we collected complete records of the inpatients who were 65 years old and above and had received warfarin therapy in a medical center in Taiwan from January 2005 to December 2009. A washout period of three month is considered to eliminate the influence of warfarin from the previous period of treatment; therefore, inpatients' records were excluded if they had any warfarin treatment before 1-Apr-2005. Each clinical record contains demographics, such as gender, age, and weight. In addition, the adequate dosage of warfarin for inpatients varies if they have symptoms like congestive heart failure and thyrotoxicosis [10]. Both of them can be identified by the inpatient's historical diagnosis codes in physician orders. As to inpatients' functions of liver and kidney, such information can be effectively collected from laboratory information system (LIS); the two most well-known indicators, alanine aminotransferase (ALT) and serum creatinine (SCr), are utilized in this study.

To build a classification model, each inpatient in the dataset is associated with a particular class label, i.e. adequate or inadequate warfarin therapy. An adequate warfarin therapy is given if the INR value falls within the target range (that is, 1–3) after the first warfarin dose; otherwise, the inpatient is classified as inadequate. Furthermore, the DDI could create potential risks to the elderly due to polypharmacy; warfarin is recognized to have DDIs with more than 250 medicines. It is indispensible to consider DDI on warfarin dosage determination [12]. In our study, we found a list of forty medicines causing severe DDI with warfarin in the case medical center. Based on the aforementioned definitions of variables, this study collects, filters, and preprocesses clinical data of all inpatients; a total of 360 validated clinical cases having encountered the DDI are considered in our research. Their descriptive statistics are shown in Table 1.

Table 1. Summary Statistics of Variables

| Variables | Range | Descriptive Statistics |
|---|---|---|
| Gender | Male/Female | Male: 149 / Female: 211 |
| Age | 65 to 98 | $\mu$=77.98, $\sigma$=7.17 |
| Congestive heart failure | Yes/No | Yes:13 / No:347 |
| Weight | 34 to 97 | $\mu$=62.09, $\sigma$=11.58 |
| ALT | 6 to 689 | $\mu$=40.3, $\sigma$=60.7 |
| SCr | 0.3 to 10.6 | $\mu$=1.35, $\sigma$=1.12 |
| Warfarin dose | 0.25 to 6 | $\mu$=1.183, $\sigma$=0.68 |
| Adequacy | Yes/No | Yes:285 / No:75 |

**Experimental design**

This study adopts Weka 3.7.3 open-source data mining software (www.cs.waikato.ac.nz/ml/weka) to construct the classification systems based on J48 (C4.5 in Weka), simple logistic (LGR in Weka), and AdaBoostM1 (AdaBoost in Weka). Several parameters are chosen: as for J48, the confidence factor and the minimum number of instances per leaf are set as 0.25 and 2, respectively; whereas in AdaBoostM1, number of iterations and the weight threshold for pruning are selected to be 10 and 100.

Inadequacy rate of warfarin therapy is about 20.8% (Table 1), resulting in a serious class imbalance problem in the dataset. A resample module in Weka is thus adopted to modify the distribution of instances of two classes by oversampling the inadequate class and undersampling the adequate class; as a result, the distribution within each class is modified to be almost identical. In addition, some useful instances in the adequate class may not be chosen by the resample method, resulting in the loss of valuable information for classifications. Therefore, the random resample technique is applied thirty times to construct datasets; for each generated dataset, ten-fold cross-validation is then applied in all the experimental evaluations.

To evaluate our model performance, we consider four performance metrics in this study, i.e., *precision*, *recall*, *F*1, and *accuracy*. Using the confusion matrix in Figure 1, these metrics can be calculated as: $precision = \frac{a}{a+c}$, $recall = \frac{a}{a+b}$, $F1 = \frac{2 \times precision \times recall}{precision + recall}$, and $accuracy = \frac{a+d}{a+b+c+d}$, respectively.

| | | Predicted class | |
|---|---|---|---|
| | | adequate | inadequate |
| Actual class | adequate | *a* | *b* |
| | inadequate | *c* | *d* |

Figure 1. Confusion matrix

## EVALUATION RESULTS

Table 2 summarizes the evaluation results on precision, recall, $F_1$ and overall accuracy for the clinical cases with interaction effects. As shown, C4.5 exhibited its $F_1$ of 0.781, significantly higher 0.21 than that of the LGR (0.571). On the overall accuracy experimental result, LGR exhibited an overall accuracy of merely 0.6, significantly lower than that of the experimental result of C4.5, which was statistically significant at the 0.05 level ($t=22.516$, $p<0.001$). The C4.5 system appeared to have a higher predictive value than the LGR system.

Next, both the C4.5 and LGR classifiers became increasingly effective when supplemented by AdaBoost. Also shown in Table 2, the AdaBoost technique caused improvement in overall accuracy and $F_1$. The C4.5 system performed a higher overall accuracy to 0.875 (improved about 0.076). On the other hand, the overall prediction accuracy of the LGR is not changed even though the AdaBoost technique is added. Based on results of two ensemble classifiers, both C4.5 and C4.5+Adaboost classifiers outperform the LGR and LGR+AdaBoost.

Table 2. Classification performance of the four models

| Evaluation | C4.5 | LGR | C4.5+AdaBoost | LGR+AdaBoost |
|---|---|---|---|---|
| Precision | 0.839 | 0.591 | 0.922 | 0.589 |
| Recall | 0.732 | 0.557 | 0.814 | 0.567 |
| F1 | 0.781 | 0.571 | 0.864 | 0.576 |
| Accuracy | 0.799 | 0.595 | 0.875 | 0.595 |

## CONCLUSION

This study responds to the challenge of predicting appropriate warfarin prescriptions for the elderly by developing decision support systems. Specifically, we applied the supervised learning techniques, including C4.5 and LGR, as well as their extensions using AdaBoost techniques for improving predictive performance. According to our analysis of 360 inpatient cases in Taiwan, both systems predict the adequacy of warfarin more accuracy than does of the clinical physicians' subjective decision. The overall evaluation results verify that DDI is a critical factor in warfarin dosage decision-making. Besides, the C4.5 classifier with AdaBoost is suggested as the most effective prediction model in this study.

Considering the complicated characteristics of warfarin, this study shows that the decision support systems incorporating learning-based classification approaches can serve as a supplementary tool due to the superior performance in predicting adequacy. Even though drug-to-drug interactions are so common for the elderly that they complicate the effectiveness of using these drugs, our study provides sufficient evidence to support the assumption that risks from inadequate use of high-alert medications can be dramatically reduced, and thus, the improvement in the safe use of high-alter drugs will be of benefit to clinicians and can avoid wastes of health care resources.

## REFERENCES

[1] ASHP, ASHP statement on the pharmacist's role in clinical pharmacokinetic monitoring. *American*

*Journal of Health-System Pharmacy*, 1998, 55 (16), 1726-1727.

[2] Colombet, I., Bura-Rivière, A., Chatila, R., Chatellier, G. & Durieux, P. Personalized versus non-personalized computerized decision support system to increase therapeutic quality control of oral anticoagulant therapy: an alternating time series analysis. *BMC Health Service Research*, 4 (27), 1-8.

[3] Freund, Y. & Schapire, R.E. A Short Introduction to Boosting. *Journal of Japanese Society for Artificial Intelligence*, 1999, 14 (5), 771-780.

[4] Horton, J.D. & Bushwick, B.M. Warfarin therapy: evolving strategies in anticoagulation. *American family physician*, 1999, 59 (3), 635-646.

[5] Hu, P., Wei, C., Cheng, T.H. & Chen, J.X. Predicting Adequacy of Vancomycin Regimens: A Learning-based Classification Approach for Improving Clinical Decision Making. *Decision Support Systems*, 2007, 43 (4), 1226-1241.

[6] Laroche, M.L., Charmes, J.P., Nouaille, Y., Picard, N. & Merle, L. Is inappropriate medication use a major cause of adverse drug reactions in the elderly?. *British Journal of Clinical Pharmacology*, 2007, 63 (2), 177-186.

[7] Lutz, W., Sanderson, W. & Scherbov, S. The coming acceleration of global population ageing. *Nature*, 2008, 451, 716-719.

[8] Page, R.L. II, Linnebur, S.A., Bryant, L.L. & Ruscin, J.M. Inappropriate prescribing in the hospitalized elderly patient: Defining the problem, evaluation tools, and possible solutions. *Clinical Interventions in Aging*, 2010, 5, 75-87.

[9] Quinlan, J. R. *C4.5*: *Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.

[10] Setiabudi, E., Alwi, I. & Setiati, S., Oral anticoagulant treatment in management of elderly patients with atrial fibrillation: Is it beneficial or detrimental?. *Acta Medica Indonesiana*, 2008, 40 (1), 40-47.

[11] Spinewine, A., Schmader, K.E., Barber, N., Hughes, C., Lapane, K.L., Swine, C. & Hanlon, J.T. Appropriate prescribing in elderly people: how well can it be measured and optimised?. *The Lancet*, 2007, 370 (9582), 173-184.

[12] Tadros, R. & Shakib, S. Warfarin-indications, risks and drug interactions. *Australian Family Physician*, 2010, 39 (6), 476-479.

[13] Zhan, C., Sangl, J., Bierman, A.S., Miller, M.R., Friedman, R., Wickizer, S.W. & Meyer, G.S. Potentially inappropriate medication use in the community-dwelling elderly. Findings from the 1996 Medical Expenditure Panel Survey. *Journal of American Medical Association*, 2001, 286 (22), 2823-2829.