PREDICTION ON UNUSED MEDICINE IN DIABETIC BY DECISION TREE

Sirapob Tayommai
Mathematics Department,
Faculty of Science,
Mahasarakham University, Maha Sarakham, Thailand

Monchaya Chiangpradit
Mathematics Department,
Faculty of Science,
Mahasarakham University, Maha Sarakham, Thailand

Parimoke Kerdhantuk
Faculty of Pharmacy,
Mahasarakham University, Maha Sarakham, Thailand

ABSTRACT

The main proposes of this paper were 1) to predict the unused medicine using data mining technique, decision tree 2) to find the factors effecting to have unused medicine in order to plan for solving problem about the drug reconcile in diabetes patients appropriately. The sample groups used in this research by simple random sampling were 486 diabetes patients taking Metformin and Glibenclamide in Na-doon district. By data mining technique, decision tree, there are 10 variables effecting to have unused medicine and 18 rules to predict unused medicine. The rules were correct 81.58 %.

KEYWORDS— Drug Reconcile, Unused medicine, Decision Tree

1. INTRODUCTION

Recently, the cost of medicine has increased continually every year. Each year, Thai government has waste huge amount of budget in medicine only. The study from the Bureau of Drug, Food and Drug Administration, found that Thailand has increased trend of import medicine since 2007 to 2012. In 2007, Thailand has imported medicine total value of 52.6 billion Baht and continually increased annually. In 2012 has a total value of 116.4 billion Baht, which is 121.29% increased from 2007 [1]. One of the reasons that make the value of medicine used in this country increase is the unreasonable issued of medicine to patients, which result the high amount of unused medicine left with patient. This unused medicine is hidden issue of Thailand healthcare system, which has bad affect to both economic and public health of this country, as well as other countries. Hataikan [2] revealed that in each country has lots of unused medicine, such as 570 million Baht per year in UK, 175 million Baht per year in USA, 750 million Baht per year in Iran and 150 million Baht per year in Thailand.

Unused medicine has become big issue that found in many of the patient with chronic illness, especially Diabetic. Diabetic is a disease that occurs from body’s disorder in generating not enough insulin hormones, which make the body unable to use of sugar efficiency. In long term, it can damage atherosclerosis if they have not been cure correctly, which could lead to critical complication such as Diabetic Retinopathy, Diabetic Nephropathy, Diabetic Neuropathy, Coronary Vascular Disease, Cerebrovascular Disease, Peripheral Vascular Disease and Diabetic Ulcer that can cause the death. In 2007, found the amount of patients with Diabetic as much as 246 million people, 4 in 5 of patients with Diabetic around the world are Asian.

Ampur Na-doon is a tiny Ampur, has more than 1,300 patients with Diabetic. If compare the patients with Diabetic’s number with population of Ampur Na-doon found that it has prevalence of 3,233 patients with Diabetic per 100K population, which is the most prevalence in Maha Sarakham province and in North East. Moreover, information of year 2013 from special clinic, Na-doon hospital found that patients with Diabetic in Ampur Na-doon have complication condition and increase the use of medicine at rate of 2.88% in 2012 and 43.33% in 2013, this is very high rate. However, it has been accepted that if able to control the sugar level in patients with Diabetic to normal level, then the chance of complication condition reduce and less death of patients.

The key issue that make patients with Diabetic cannot control level of sugar is because the inappropriate use of medicine, which can be notice from the unused medicine at patient household. There is research stated that from the survey of unused medicine of patients with Diabetic, more than 90% has unused medicine, which reason may be incorrect use of medicine by patient, reduce use of medicine by patients or other health condition of patients such as cannot read, sight issue. Moreover, another issue that causes incorrect use of medicine by Patients with Diabetic is that they do not have care giver to look after.
Therefore, researcher is interested in study factors of unused medicine by patients and create prediction model of unused medicine to use in practice, by data mining technique, to forecast unused medicine by decision tree. This would lead to plan an appropriate guideline for problem solving of unused medicine by patients. To reduce the amount and issue from unused medicine, in order to make patients use medicine correctly and appropriate as a result of less complication conditions and disease severe.

II. DATA UNDERSTANDING

This research were to study factors of unused medicine by patients with Diabetic in Ampur Na-doon, Maha Sarakham province, and create prediction model of unused medicine by using data mining, decision tree, with scope of research as follow:

(1) Population is 598 patients with Diabetic in Ampur Na-doon, Maha Sarakham province that only use Glibenclamide and Metformin, not included patients who use insulin injection [3].

(2) Sample used in this study, by simple random sampling, was 486 patients with Diabetic in Ampur Na-doon, Maha Sarakham province that only use Glibenclamide and Metformin, not included patients who use insulin injection.

III. METHODOLOGY

This research is to study factors of unused medicine by patients with Diabetic in Ampur Na-doon, Maha Sarakham province, and create prediction model of unused medicine by using data mining, decision tree.

Data mining prediction models were constructed using WAS 3.7.4, a tool for data mining. To predict a binary categorical outcome variable in this study, we used classification as a prediction model: classification, one of the major predictive data mining methods for categorical target variables is a two-step process. In the first step, a classifier is built based on predetermined sets of data (learning or training) and in the second step, the model is used for classification of test data. A test dataset is randomly selected from the general data and is independent of the training data [4].

In medicine and many other areas, decision trees are the most popular and powerful classification algorithm with the growth of data mining methods [5]. Decision tree is popular for flow chart-like tree structure and transparency. And allows the decision-maker to examine and understand the decision model. The goal of decision tree construction to create a model that predicts the value of a target variable base on several independent of predictor variables.

Independent variables consist of age, number of care giver, list of medicine, number of medicine time, years in Diabetic, gender, education, occupation, married status, average income, number of common diseases, personal hygiene, other healthcare problem such as eyesight, hearing, memory and understanding of medicine/ disease of patients. Dependent variable is unused medicine of patient, means amount of unused medicine value per total medicine value in a month that is more or equal to average ratio of unused medicine by patients with Diabetic in Ampur Na-doon, Maha Sarakham province, which is 11.56%. In other word, if patient has unused medicine’s rate more or equal to 11.56% is mean that patient has unused medicine, but if patients has unused medicine’s rate less than 11.56%, that is mean that patient do not have unused medicine or completely taken medicine.

Research time: since 1st of September 2013 to 28th of February 2014 (7 months).

Research tool is a 6 parts of questionnaire about dispense of unused medicine by patients with Diabetic in Ampur Na-doon. Test validity and reliability by using Kuder Richardson (KR-20) = 0.673.

To find cut point for reference that patient has unused medicine, because there is no document or research support about Cut point of unused medicine. Therefore, in order to find Cut point we had performed as below:

(1) Collect the unused medicine information of all patients with Diabetic, 1,348 people of Ampur Na-doon. Officer of 12 health services record unused medicine of every patient.

(2) Data collected since August 2013 until November 2013 (total of 4 months).

(3) Return of unused medicine was calculated as a ratio between dispense and unused in percentage.

(4) Bring all 4 months of data to find mean ratio between dispense and unused medicine in order to use it as a Cut point of unused medicine.

IV. RESULTS

From the 486 data of sampling patients that answered questionnaire found that, most of samples are around 60-69 years old, 38.07%, 50-59 years old, 26.75%, and more than 70 years old is 14.43%, respectively. Samples average age is 58.45 year old; oldest is 77
years old and youngest is 32 years old. Most of them, 36.01%, stay alone with no care giver, at least a care giver at 35.80% respectively. Most of patients took 3 lists of other common diseases medicine with Diabetic medicine is 73.4%, 4 lists is 23.66% and 5 lists is 18.93%, respectively. 38.68% of patients took medicine twice a day. 32.92% of patients took medicine 3 times a day and 27.37%, respectively. Most of them, 42.80%, had had Diabetic for 3-5 years, 31.69% had had Diabetic for 6-10 years and 13.99% only had Diabetic for 1-2 years. Average of patients with Diabetic is 5.80 years, the longest is 20 years. Most of patients, 70.78%, are female. Education at lower elementary school (grade 3) is 42.59%, no education is 34.36% and complete elementary school (grade 6) is 13.17%, respectively. Most of samples, 42.80%, are in agriculture. 27.37% is merchants and 22.63% is unemployed. 45.68% is single, 37.65% is married (from analysis of questionnaires found that the person who answered the question understand that staying alone mean single which made the questionnaire full of single status. 79.22% have income less than 1,000 Baht per months, 18.72% have income between 1,000-4,999 Baht per month. 54.32% only have Diabetic disease, 24.28% have 1 other common disease. 52.67% have a bad habits, 59.05% do not have hearing problem, 54.12% do not have eyesight problem, 51.03% has memory problem. The patients who answered questionnaire had knowledge point about disease and medicine, average of 6.05 point. The highest point is 9 and lowest is 4.

In this study, decision tree is based on program Wega 3.7.4 (Waikato Environment for Knowledge Analysis), developed by University of Waikato, New Zealand. Analyzed the data by decision tree, C4.5 (J48) algorithm, to find forecasting model has shown in Fig 1-5 and Table I. Table I showed that the best model was T80R5B which highest correctly classified instances is 81.6%, highest F-Measure is 80.9%, precision is 86.5%. 35 nodes and 18 leaves. This model was conducted by EBP technique. The rule of the model showed in table II.
**Fig. 3** Recall of decision tree

**Fig. 4** F-Measure of decision tree
Table I: Efficiency of Decision Tree for Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Train:Test</th>
<th>Leaves</th>
<th>Size</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>T80R5B</td>
<td>D1(90:10)</td>
<td>18</td>
<td>35</td>
<td>.865</td>
<td>.816</td>
<td>.809</td>
<td>81.58%</td>
</tr>
</tbody>
</table>

Table II: Rules of Unused Medicine in Diabetic by Decision Tree

<table>
<thead>
<tr>
<th>Rule</th>
<th>Predict</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>memory problem and 1 list of medicine</td>
</tr>
<tr>
<td>2</td>
<td>memory problem, more than 1 list of medicine and common disease more than 2 disease</td>
</tr>
<tr>
<td>3</td>
<td>memory problem, more than 1 list of medicine, common disease less than 2 disease, income 1000-4999 baht/month</td>
</tr>
<tr>
<td>4</td>
<td>memory problem, more than 1 list of medicine, common disease less than 2 disease, no income 1000-4999 baht/month, diploma</td>
</tr>
<tr>
<td>5</td>
<td>memory problem, more than 1 list of medicine, common disease less than 2 disease, no income 1000-4999 baht/month, not diploma, divorce</td>
</tr>
<tr>
<td>6</td>
<td>memory problem, more than 1 list of medicine, common disease less than 2 disease, no income 1000-4999 baht/month, not divorce, not diploma or middle school</td>
</tr>
<tr>
<td>7</td>
<td>memory problem, more than 1 list of medicine, common disease less than 2 disease, no income 1000-4999 baht/month, not divorce, middle school, 1 care giver</td>
</tr>
<tr>
<td>8</td>
<td>memory problem, more than 1 list of medicine, common disease less than 2 disease, no income 1000-4999 baht/month, not divorce, middle school, no care giver or more than 2</td>
</tr>
<tr>
<td>9</td>
<td>No memory problem and more than 1 care giver</td>
</tr>
<tr>
<td>10</td>
<td>No memory problem, no care giver or 1 care giver and age less than 40</td>
</tr>
<tr>
<td>11</td>
<td>No memory problem, no care giver or 1 care giver and age less than 40</td>
</tr>
<tr>
<td>12</td>
<td>No memory problem, no care giver or 1 care giver, age equal or more than 40, high school</td>
</tr>
<tr>
<td>13</td>
<td>No memory problem, no care giver or 1 care giver, age 60-69, under high school, no occupation</td>
</tr>
<tr>
<td>14</td>
<td>No memory problem, no care giver or 1 care giver, age 40-49, under high school, no occupation</td>
</tr>
<tr>
<td>15</td>
<td>No memory problem, no care giver or 1 care giver, age equal of more than 40, under high school, employed, low knowledge</td>
</tr>
<tr>
<td>16</td>
<td>No memory problem, no care giver or 1 care giver, age 50-59, under high school, employed, high knowledge</td>
</tr>
<tr>
<td>17</td>
<td>No memory problem, no care giver or 1 care giver, age 50-59, under high school, agriculturist, high knowledge</td>
</tr>
<tr>
<td>18</td>
<td>No memory problem, no care giver or 1 care giver, age 50-59, under high school, employed but not agriculturist, high knowledge</td>
</tr>
</tbody>
</table>
From Table II, there are 10 factors that influence unused medicine. 10 factors consist of memory problem, list of medicine, other common disease, income, education, status, number of caregiver, age, occupation and knowledge/understanding about Diabetic. There are 18 rules from this model.

V. CONCLUSIONS

In conclusion, the decision tree analysis in the current study has created a simple and easy to apply tool in clinical settings and public health. By decision tree analysis, 16 independent variables such as age, number of caregiver, list of medicine, number of medicine time, years in Diabetic, gender, education, occupation, married status, average income, number of common diseases, personal hygiene, other healthcare problem such as eyesight, hearing, memory and understanding of medicine/disease of patients. Dependent variable is unused medicine of patient. The suitable algorithm was by technique C4.5 (J48). Model was binary, CF = 0.25, leaf =2, pruning fold = 3. An interesting finding is shortest rule of unused medicine as shown in rule 2, Table II. That is memory problem, 1 list of medicine and 2 other common diseases.

ACKNOWLEDGMENT

The authors would like to thank reviewers for their comments that help improve the manuscript. This Research was financially Supported By Mahasarakham University 2015 Copyright of Mahasarakham University.

REFERENCES

