Medication Pattern Mining Considering Unbiased Frequent Use by Doctors

Yuji Morita¹, Masatoshi Yoshikawa¹*, Noboru Kada¹, Akihiro Hamasaki², Osamu Sugiyama³, Kazuya Okamoto⁴ and Tomohiro Kuroda⁴

¹ Graduate School of Informatics, Kyoto University, Kyoto, Japan
² Center for Diabetes & Endocrinology, Kitano Hospital, Osaka, Japan
³ Preemptive Medical and Lifestyle Disease Research Center, Kyoto University Hospital, Kyoto, Japan
⁴ Division of Medical IT & Administration Planning, Kyoto University Hospital, Kyoto, Japan

1 Introduction

Electronic medical records (hereafter called “EMRs”) are designed for the primary use of data, such as the efficiency of medical work by order entry systems and medical accounting systems. Secondary data, such as medical research, decision support, and educational support for young doctors, are expected to accumulate via promoting the computerization of medical data. Especially, computer-based clinical decision support systems (hereafter called “CDSS”) to support doctor’s clinical decision making is an important example of secondary use of EMRs [1]. The CDSS is a system that provides better medical care to patients by reducing decision mistakes and sharing medical evidence when medical staff makes decisions, such as diagnosis and prescription [2].
Clinical decision support in primary care has attracted attention recently, and many previous studies have aimed to improve the quality of medical care by presenting appropriate treatment procedures to doctors using sequential data mining [3]. For example, when conducting medication therapy (one of the therapies used in primary care), the doctor appropriately selects the type and dose of the next prescription medicine based on the relationship among patient’s specific information, the current medication and the next medication. Then, many clinical experiences can be shared, and the next prescription medicine can be efficiently decided within a range of adaptations by suggesting frequent sequences extracted from prescription data under the same conditions to the doctor as well as potential medications that can be prescribed. Many previous studies have extracted prescription sequences based only on frequency information, such as the number of prescriptions and the total number of patients issued the prescription [4, 5, 6].

However, simply applying sequential data mining techniques based on frequency information to EMRs are not always appropriate. Because each doctor has his/her own strategy for medication, consideration of difference among medication patterns of doctors is necessary. In cases where a very small number of doctors issue a prescription representative of a medication pattern to many patients many times, the prescribing intention of this very small number of doctors has a large influence on pattern extraction. This problem introduces bias to the final extracted prescription sequences. Therefore, we attempt to extract prescription sequences from more diverse perspectives by considering unbiased frequent use by doctors as an additional factor. Our hypothesis is that prescriptions performed by many doctors unbiasedly are highly reliable.

2 Methods

2.1 Formulation of Frequent Pattern Extraction Considering Unbiased Frequent Use by Doctors

We start with definition of basic terms. A prescription is a pair of medicine and dose such as (thiamazole, 5 mg). A sequence (pattern) is a consecutive sequence of two prescriptions such as (thiamazole, 5 mg) → (thiamazole, 10 mg). We say a doctor performed a sequence when the doctor performed the prescription which appears in the consequent of the sequence when the previous prescription was the antecedent of the sequence. For example, let \( \alpha \) be the sequence (thiamazole, 5 mg) → (thiamazole, 10 mg). If a doctor performed the prescription (thiamazole, 10 mg) when the previous prescription was (thiamazole, 5 mg), we say the doctor performed \( \alpha \).

We propose the concept that the unbiased frequent use by doctors is a factor in addition to frequency information. Thus, we formulated a prescription sequence mining method that considered unbiased frequent use by doctors.

More specifically, we propose the DF-score as a new score for a sequence calculated by the following formula based on the Support and D-rank value.

\[
DF - score(\alpha) = Support(\alpha) \times D - rank(\alpha)
\]

Here, \( \alpha \) is a sequence, \( Support(\alpha) \) is \( \alpha \)'s support value, and \( D - rank(\alpha) \) is the value of unbiased frequency of \( \alpha \) performed by doctors. In other words, the \( DF - score(\alpha) \) is an index that combines information based on the patient-based frequency and the doctors’ unbiased usage.

We define the unbiased frequent use by doctors as an indicator based on the following factors.

Rate of doctors who performed a certain medication pattern: \( D_{\text{rate}} \)

The unbiased variance representing the usage frequency dispersion for a certain prescription sequence: \( V_{\text{UB}}(\alpha) \)

Then, the unbiased frequent rank of \( \alpha \) is calculated by the following formula:

\[
D_{\text{rate}}(\alpha) = \frac{\text{count}(D(\alpha))}{|D|}
\]

\( count(D(\alpha)) \) indicates the number of doctors who performed sequence \( \alpha \) and \( |D| \) indicates the total number of all doctors. Thus, \( n_{\text{rate}} \) represents the rate of doctors who have performed sequence \( \alpha \) out of all doctors.

\( V_{\text{UB}}(\alpha) \) is defined as the unbiased variance representing the usage frequency dispersion for sequence \( \alpha \) for each doctor as follows:

\[
V_{\text{UB}}(\alpha) = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2
\]

Here, let \( \sum = \{D_1, D_2, \ldots, D_n\} \) be the set of all doctors. The proportion of patients who are prescribed sequence \( \alpha \) out of all prescribed patients by doctor \( d_i \) is \( d_i \), and \( \bar{d} \) shows the average value of the proportion of patients prescribed by all doctors using sequence \( \alpha \). We obtain \( D - rank(\alpha) \) by dividing \( D_{\text{rate}}(\alpha) \) by \( V_{\text{UB}}(\alpha) \). Therefore, a sequence performed by many doctors unbiasedly yields higher \( D - rank(\alpha) \) value.

Based on these calculations, Figure 1 shows the entire process from data acquisition to calculation of the DF-score of prescriptions. The procedure is as follows:

1. A correspondence table showing the relationship between the product name and the generic name of medication is prepared.
The prescription data and clinical laboratory test value data from the patients are extracted from the EMRs for analysis based on the prepared correspondence table.

To preprocess the data for sequential pattern mining, the data extracted from the EMRs is converted into a transaction form after grouping each patient based on the clinical laboratory test result.

Sequential pattern mining is applied to the transformed data set to extract frequent patterns.

The unbiased frequent use by doctors (D-rank) and the DF-score of each extracted frequent pattern are calculated. Finally, frequent prescription patterns are extracted based on the DF-score.

3 Experiment

3.1 Data Set

The information in the EMR includes the patient’s name, age, progress record, and medical imaging. Under the approval of ethical committee of Kyoto University School of medicine, we applied prescription data and clinical laboratory test value data from patients with thyroid disease. These data sets contain anonymized data acquired from September 20, 2000, to December 22, 2015, by the EMR system KING (Kyoto University Hospital Information Galaxy) used at Kyoto University Hospital.

3.2 Evaluation Method

To evaluate the objective effectiveness of the extracted medication pattern, we assessed the degree to which the clinical laboratory test values improved after prescribing the medication pattern. For example, after prescribing a medication pattern, the clinical laboratory test values were individually considered on a case-by-case basis, some cases improved and others deteriorated. On this occasion, countless evaluation approaches can be used depending on the aspect to be interpreted from the clinical laboratory test value (i.e., defining improvement and deterioration based on different criteria). Therefore, we classify individual cases into several groups and propose an evaluation method from the perspective of the improvement rate of clinical laboratory test values for each group. For this purpose, we propose the following two evaluations.

- Evaluation method 1: An evaluation method based on the normal value range.
- Evaluation method 2: An evaluation method based on the median value of the normal range.

We describe specific definitions and procedures for the above two evaluation methods. First, we explain evaluation method 1. In medication therapy for many diseases, including thyroid disease, prescribed medication is generally selected in consideration of maintaining the clinical laboratory test values within the normal value range as much as possible. Therefore, the normal value range (which is the medical standard) is considered very useful information for the measurement of the improvement of the clinical laboratory test value. Evaluation method 1 is a method based on this normal value range. A specific score is assigned to each case based on the clinical laboratory test value state, such as a transition of one, and the final evaluation value is calculated by averaging the scores of all cases. The method used to assign the score is defined based on the clinical opinion of the specialist. Figure 2 shows an example of a method used to assign a score according to each clinical laboratory test value state. The method gives a score when the examination value state at the time of prescription is within the normal value range (Figure 2a), has an abnormal value (Low) (Figure 2b) or has an abnormal value (High) (Figure 2c). The transition of the clinical laboratory test value after the prescription of a patient in a certain clinical laboratory test value state in which a medication pattern is performed is classified as described. Then, a score is assigned for each case according to the classification result. Let the clinical laboratory test value at prescription be $V_i$. The next clinical laboratory test value of $V_i$ will be $V_i'$. The upper limit value of the normal value range will be $N_u$, and the lower limit value of the normal value range will be $N_l$.
• Improvement
  • \( N_L \leq V' \leq N_U \)
  • \( V < V' < N_L \)
  • \( N_U < V' < V \)
  ⇒ Assign score "1"
  • Remain unchanged
  • \((N_L > V) \land (V = V')\)
  • \((N_U < V) \land (V = V')\)
  ⇒ Assign score "0"
  • Exceedance
  • \((N_L > V) \land (V = V')\)
  • \((N_U < V) \land (N_L > V')\)
  ⇒ Assign score "0"
  • Deleterious change
  • \((N_L \leq V \leq N_U) \land (N_U < V')\)
  • \((N_L \leq V \leq N_U) \land (N_L > V')\)
  • \((N_L > V) \land (V > V')\)
  • \((N_U < V) \land (V < V')\)
  ⇒ Assign score "-1"

The final evaluation value is calculated by the following formula based on the score assigned to each case.

\[
\frac{1}{n} \sum_{i=1}^{n} \text{Score}_i
\]

\(\text{Score}_i\) indicates the score value in case \(i\). The final evaluation value calculated by the above formula represents the average value of the scores of \(n\) cases. Evaluation method 1 interprets the pattern with the final evaluation value closer to 1 as the medication pattern that results in better improvement of the clinical laboratory test value.

Second, we explain evaluation method 2. Since the above-described evaluation method 1 is a method based only on the normal value range, concrete values from the clinical laboratory test values are not taken into consideration. Therefore, even though the score values assigned in evaluation method 1 are the same, the actual clinical laboratory test values may differ. Conversely, evaluation method 2 focuses on the median of the normal range using a different approach than evaluation method 1. We compare quantitative values between case groups before and after prescription using the distance from the median of the normal range. We evaluate the transition of the clinical laboratory test value of each case by RMS (Root Mean Square) using the median of the normal range. A large weight can be given to a case located at a distance from the median of the normal range using RMS, and the influence of the deteriorated case is considered to more appropriately reflect the final evaluation value. The evaluation formula is defined as follows:

\[
\sqrt{n \sum_{i=1}^{n} (d_i')^2} - \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_i)^2}
\]

\(d_i\) is the difference between the clinical laboratory test value and the median value of the normal range at time \(t\) of a case \(i\). \(d'_i\) is the difference between the clinical laboratory test value and the median of the normal range of the clinical laboratory test at time \(t+1\) performed next to the time \(t\). Here, if the examination value at each time \(t\) is within the normal value range, set the value of \(d_i, d'_i\) to 0 to give discontinuity between the distance within the normal value range and the distance outside of the range. Thus, evaluation method 2 interprets a medication pattern with a smaller final evaluation value as better improvement.
value as a pattern that can better improve the inspection value. For example, Figure 3 shows that the final evaluation value is calculated by the following formula:

\[ \sqrt{\frac{d'_1 + d'_2 + 0 + d'_4}{4}} - \sqrt{\frac{d_1 + d_2 + d_3 + d_4}{4}} \quad (\because d'_1 = 0) \]

3.3 Procedure

Figure 4 shows the evaluation procedure. The procedure is outlined below.

1. Patient data for a specific condition for analysis are extracted from the EMRs based on the information from the current clinical laboratory test value state and the previous prescription medicine.

2. Prescription sequences are extracted by applying the proposed method and the existing method to the patient data; then, the next medicine to be prescribed is analyzed.

3. Regarding the next prescribed medicine extracted by each method, the evaluation value for the improvement of the clinical laboratory test value is calculated by the two evaluation methods described in Section 3.2. Then, the average value is compared with the top \( k \) evaluation values.

We describe each of the above steps in detail. In step (1), patient data for a specific condition were first extracted based on clinical laboratory test value state patterns. The clinical laboratory tests we use are TSH (thyroid stimulating hormone), FT3 (free triiodothyronine) and FT4 (free thyroxine). In our experiment, we focus on the following three patterns of the clinical laboratory test value state:

- TSH, FT3, and FT4 are normal values.
- TSH is an abnormal value (High); FT3 and FT4 are normal values.
- TSH is an abnormal value (Low); FT3 and FT4 are normal values.

For each of these three patterns of the clinical laboratory test value state, we consider two previous prescription medicines patterns: thiamazole 5 mg and thiamazole 10 mg. The evaluation was conducted based on the combination of the patient's condition using six (3 \( \times \) 2) patterns in total. In step (2), prescription sequence mining was performed for the six total patterns described above using the existing method based only on frequency information and the proposed method that considered unbiased frequent use by doctors. Then, the medicine to be prescribed next in the same situation was analyzed based on the prescription sequences extracted by each method. In step (3), the evaluation values for the improvement of the clinical laboratory test values were calculated for each prescription sequence extracted in step (2) using the two evaluation methods described in Section 3.2. Then, by comparing the average value of the top \( k \) evaluation values of the prescription sequence extracted by each method, we examined how the proposed and existing methods properly extracted patterns that resulted in improvement of the clinical laboratory test values and evaluated the objective effectiveness of the extracted patterns by considering unbiased frequent use by doctors. Moreover, to measure the performance of top \( k \) ranking of the existing and proposed method, we used a weighted average value based on the rank. This formula is defined as follows:

\[ \frac{1}{k} \sum_{i=1}^{k} \text{SCORE}_i \]

\( \text{SCORE}_i \) shows the evaluation value for the improvement of the clinical laboratory test value in the medication pattern of ranking \( i \) calculated by each evaluation method. For a value of \( k \), we considered the eventual recommendation of the pattern to the doctor; because the maximum number of

![Figure 3: An example of a case group with an abnormal value (high).](image-url)
Table 1: Average values of the top k (Evaluation method 1).

<table>
<thead>
<tr>
<th></th>
<th>Evaluation method 1</th>
<th>Evaluation method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TSH, FT3, FT4=Normal/value</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Existing method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamazole_5 mg</td>
<td>k = 1: 0.840</td>
<td>k = 2: 0.604</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.365</td>
<td>k = 4: 0.259</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.180</td>
<td></td>
</tr>
<tr>
<td>Thiamazole_10 mg</td>
<td>k = 1: 0.772</td>
<td>k = 2: 0.563</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.470</td>
<td>k = 4: 0.428</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.360</td>
<td></td>
</tr>
<tr>
<td><strong>Proposed method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamazole_5 mg</td>
<td>k = 1: 0.840</td>
<td>k = 2: 0.310</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.304</td>
<td>k = 4: 0.183</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.129</td>
<td></td>
</tr>
<tr>
<td>Thiamazole_10 mg</td>
<td>k = 1: 0.772</td>
<td>k = 2: 0.608</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.442</td>
<td>k = 4: 0.334</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.311</td>
<td></td>
</tr>
</tbody>
</table>

(TSH=Abnormal value(Low)) \land (FT3, FT4=Normal value)

<table>
<thead>
<tr>
<th></th>
<th>Evaluation method 1</th>
<th>Evaluation method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Existing method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamazole_5 mg</td>
<td>k = 1: 0.232</td>
<td>k = 2: 0.173</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.159</td>
<td>k = 4: 0.162</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.143</td>
<td></td>
</tr>
<tr>
<td>Thiamazole_10 mg</td>
<td>k = 1: 0.231</td>
<td>k = 2: 0.080</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.080</td>
<td>k = 4: 0.057</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.060</td>
<td></td>
</tr>
<tr>
<td><strong>Proposed method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamazole_5 mg</td>
<td>k = 1: 0.397</td>
<td>k = 2: 0.280</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.234</td>
<td>k = 4: 0.195</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.174</td>
<td></td>
</tr>
<tr>
<td>Thiamazole_10 mg</td>
<td>k = 1: 0.189</td>
<td>k = 2: 0.086</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.095</td>
<td>k = 4: 0.059</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.058</td>
<td></td>
</tr>
</tbody>
</table>

(TSH=Abnormal value(High)) \land (FT3, FT4=Normal value)

<table>
<thead>
<tr>
<th></th>
<th>Evaluation method 1</th>
<th>Evaluation method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Existing method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamazole_5 mg</td>
<td>k = 1: 0.612</td>
<td>k = 2: 0.428</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.367</td>
<td>k = 4: 0.343</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.312</td>
<td></td>
</tr>
<tr>
<td>Thiamazole_10 mg</td>
<td>k = 1: 0.536</td>
<td>k = 2: 0.426</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.386</td>
<td>k = 4: 0.342</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.305</td>
<td></td>
</tr>
<tr>
<td><strong>Proposed method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiamazole_5 mg</td>
<td>k = 1: 0.612</td>
<td>k = 2: 0.494</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.446</td>
<td>k = 4: 0.381</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.335</td>
<td></td>
</tr>
<tr>
<td>Thiamazole_10 mg</td>
<td>k = 1: 0.501</td>
<td>k = 2: 0.419</td>
</tr>
<tr>
<td></td>
<td>k = 3: 0.349</td>
<td>k = 4: 0.327</td>
</tr>
<tr>
<td></td>
<td>k = 5: 0.299</td>
<td></td>
</tr>
</tbody>
</table>
elements that humans can recognize in short-term memory is five [7], we use $k$ values ranging from 1 to 5. In this experiment, the clinical laboratory test value item of the analysis target is TSH, because TSH reacts more sensitively even with a small amount of thyroid hormone excess or deficiency than FT3 and FT4 and can accurately capture small changes in the thyroid hormone.

### 4 Results

Table 1 and 2 show the performance result calculated by the formula (1) for the top $k$ ($k=1,2,\ldots,5$) using the evaluation method 1 and 2, respectively. Each table shows the performance results of the above-mentioned six total pattern combinations. The lower table is the results of the proposed method, and the upper table is the results of the existing method. Recall that larger the value the better the method in the evaluation method 1, and vice versa in the evaluation method 2.

![Table 2: Average values of the top $k$ (Evaluation method 2).](image)

![Figure 5: An example of user interface of predictions for the next medication for a disease [8].](image)
5 Discussion and Conclusion

First, when TSH, FT3, and FT4 are within the normal value range, the existing method shows better performance in almost all cases, in the evaluation method 1. In the evaluation method 2, the existing method shows better performance when the previous prescription medicine was thiamazole (5 mg), while the proposed method gives better performance when the previous medicine was thiamazole (10 mg). Under these circumstances, we suggest that better improvement of the clinical laboratory test value can be expected in the existing method than the proposed method when TSH, FT3, and FT4 are within the normal value range.

Second, when TSH has an abnormal value (Low) and the FT3 and FT4 values are in the normal range, the proposed method has a better evaluation value in almost all cases in both evaluation methods 1 and 2. Thus, improvement of the clinical laboratory test value can be expected in the proposed method in this case.

Third, when TSH has an abnormal value (High) and the FT3 and FT4 values are in the normal range in evaluation method 1, the evaluation value of the existing method is better when the previous prescription medicine is thiamazole (10 mg), and the evaluation value of the proposed method is better when the previous prescription medicine is thiamazole (5 mg). In evaluation method 2, the proposed method has a better evaluation value in almost all cases. Thus, improvement of the clinical laboratory test value can be expected in the proposed method in this case.

In conclusion, the present study suggested that our prescription mining method is useful in certain situations such as when patients are not in normal condition (i.e. the clinical laboratory test value is outside of the normal value range). This result suggests our method supports medical doctors more efficiently to decide a reasonable choice for clinical improvement especially when such decision is difficult.

Finally, we discuss possible future extension of our work. In this study, the data mining methods used only prescription data and clinical laboratory test value data. However, various other information recorded in the EMRs including personal information, such as age, gender, weight, and history of disorder, can be effectively used for data mining. Progress notes written in natural language, various medical imaging data, and information from outside of the hospital (i.e., a life log) can also be effectively used in combination with our study. Other possible extension includes evaluation experiments using data other than thyroid diseases and data from multiple medical institutions. A prototype of user interface for suggestions of medications to prescribe next is shown in Figure 5 [8].

6 Acknowledgement

This work was partially supported by The Japan Society for the Promotion of Science (JSPS) KAKENHI grant number 15H02705 and 15K00466. The funding body had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

7 Authors’ Contributions

M. Yoshikawa, A. Hamasaki, O. Sugiyama, K. Okamoto, and T. Kuroda have contributed to the conception and design of research; T. Kuroda, K. Okamoto and Y. Morita collected and managed the research data. Y. Morita has drafted the article and performed experiments. N. Kada developed a user interface of the system. M. Yoshikawa, A. Hamasaki, O. Sugiyama, K. Okamoto, and T. Kuroda performed the interpretation the data and results. M. Yoshikawa revised the article critically. All authors have given approval of the final version to be published.

8 Conflict of Interest

The authors declare no conflicts of interest associated with this manuscript.

References


