Discovery of Outpatient Care Process of a Tertiary University Hospital Using Process Mining

Eunhye Kim, RN, BS[1], Seok Kim, MS[1], Minseok Song, PhD[2], Seongjoo Kim, MS[3], Donghyun Yoo, MS[4], Hee Hwang, MD, PhD[1], Sooyoung Yoo, PhD[1]

[1]Center for Medical Informatics, Seoul National University Bundang Hospital, Seongnam; [2]School of Technology Management, Ulsan National Institute of Science and Technology, Ulsan; [3]Cyberdigm Co., Seoul; [4]Patient's Affairs, Seoul National University Bundang Hospital, Seongnam, Korea

Objectives: There is a need for effective processes in healthcare clinics, especially in tertiary hospitals, that consist of a set of complex steps for outpatient care, in order to provide high quality care and reduce the time cost. This study aimed to discover the potential of a process mining technique to determine an outpatient care process that can be utilized for further improvements. **Methods:** The outpatient event log was defined, and the log data for a month was extracted from the hospital information system of a tertiary university hospital. That data was used in process mining to discover an outpatient care process model, and then the machine-driven model was compared with a domain expert-driven process model in terms of the accuracy of the matching rate. **Results:** From a total of 698,158 event logs, the most frequent pattern was found to be "Consultation registration > Consultation > Consultation scheduling > Payment > Outside-hospital prescription printing" (11.05% from a total cases). The matching rate between the expert-driven process model and the machine-driven model was found to be approximately 89.01%, and most of the processes occurred with relative accuracy in accordance with the expert-driven process model. **Conclusions:** Knowledge regarding the process that occurs most frequently in the pattern is expected to be useful for hospital resource assignments. Through this research, we confirmed that process mining techniques can be applied in the healthcare area, and through detailed and customized analysis in the future, it can be expected to be used to improve actual outpatient care processes.

Keywords: Process Mining, Hospital Information Systems, Outpatients, Workflow Analysis

Corresponding Author
Sooyoung Yoo, PhD
Center for Medical Informatics, Seoul National University Bundang Hospital, 82 Gumi-ro 173beon-gil, Bundang-gu, Seongnam 463-707, Korea. Tel: +82-31-787-1151, Fax: +82-31-787-4004, E-mail: yoosoo0@snu.ac.kr

I. Introduction

To receive outpatient care service in a general hospital, patient should follow multi-step process such as reception, consultation, treatment, test, and payment. An error in the process not only causes discomfort for patients, but also becomes a factor that may interfere with their treatment [1]. However, difficulties arise when a standardized process flow is applied to all patients who visit a hospital because there are significant differences in the hospital process of individuals and hospitals. It can also be challenging to estimate which hospital process is applicable even among the patients with the same diagnosis due to many complex factors. Under such circumstances, there is a definite need to discover and

provide effective hospital processes to patients to reduce the time cost and to provide patients with high-quality service [2]. Furthermore, it is rapidly becoming a goal for hospitals to provide the best service at the lowest cost by evaluating existing infrastructure and services offered by the hospital [3]. Accordingly, this study aimed to discover a complex hospital process using a process mining technique, and to find ways to take advantage of it.

Process mining is defined as simply analyzing the event log, which contains the recorded workflow progress, and extracting the useful knowledge [4]. Since the analysis only pertains to database logs already in existence, the time and cost of collecting the data is low, and since the analysis is only performed on recorded data, there is no room for distortion of information, resulting in an objective process analysis [5-7]. Furthermore, since process mining can analyze the recorded log data of various systems, many corporations, hospitals, banks, and public institutions, are beginning to utilize it for process analysis [8-10]. In the healthcare area, a process mining technique was used to analyze the treatment process for stroke patients [3] and gynecologic oncology [10]. However, no study has tried to evaluate the feasibility of a process mining technique to analyze a complex hospital process. Therefore, in this study, we tried to discover the outpatient care process of a tertiary university hospital, which consists of a set of complex steps, using a process mining technique. By comparing the outpatient care process model derived from this process mining technique with an expert-driven process model recognized by hospital specialists, we investigated the applicability of a process mining technique in a hospital setting.

II. Methods

1. Modeling of Outpatient Care Mega–Process
To model an outpatient care mega-process, nurses and professors from the Center for Medical Informatics, nurses from the Outpatient team, and staff members from the Patient's Affairs team and Quality Assurance team participated and joined in discussion. Figure 1 shows the outpatient care mega-process model derived from the domain experts which was then used as our baseline model and as a comparison with a process model discovered by a process mining technique.

The derived mega-process model was designed to cover the entire spectrum of what a patient may experience upon visiting a hospital regardless of the patient types, such as first visitor or return visitor to hospital; thus, it was designed to encompass many different cases that may occur from a patient's perspective.

2. Collection of Event Log Data
Based on the outpatient care process shown in Figure 1, the event log was defined as seen in Table 1, and data was collected from the hospital information system of Seoul National University Bundang Hospital from May 1, 2012 to May 31, 2013.

During the data collection period, all event logs from the Health Promotion Center were excluded because that department is responsible for only independent health screenings, which is not involved with the outpatient care process. In addition, because tests can have a long completion time, the activity completion time was set to the time when a patient began the examination. Regarding treatments, there was no information available on the activity completion time; therefore, the treatment issued date was collected instead, and the treatment process was assumed to be the final step of all processes. We also note that the log data of education and home care scheduling was excluded because the activity completion time was not managed in some cases in the hospital information system.

3. Process Inference Algorithm
Song et al. [11] analyzed and benchmarked two process analysis products, DISCO (Fluxicon Process Laboratories, Eindhoven, The Netherlands) and the open-source ProM (Process Mining Group, Eindhoven Technical University, Eindhoven, The Netherlands). He adjusted the threshold value and extracted a simplified process model to produce an optimized algorithm, which only shows the main flow of a model. The optimized algorithm was then used on the data collected in this study to continue the process mining. To increase the validity of the analysis, the process mining derivation included heuristic mining and fuzzy mining derived using the algorithm that was developed from the benchmarking (Figure 2).

The process models derived from process mining and the expert-driven process model were compared, and the common workflow of both models as well as the unique workflow of each model were compared and contrasted. To estimate the importance of certain activities, the total number of activities and the frequency of each activity were shown with the corresponding activity and adjusted using a pre/post activity relationship table to analyze the process flow more realistically and in greater depth.

III. Results

1. Statistical Outline of Log Data
During the month of May 2012, 123,299 outpatient cases occurred out of a total of 698,158 event logs, and the total event
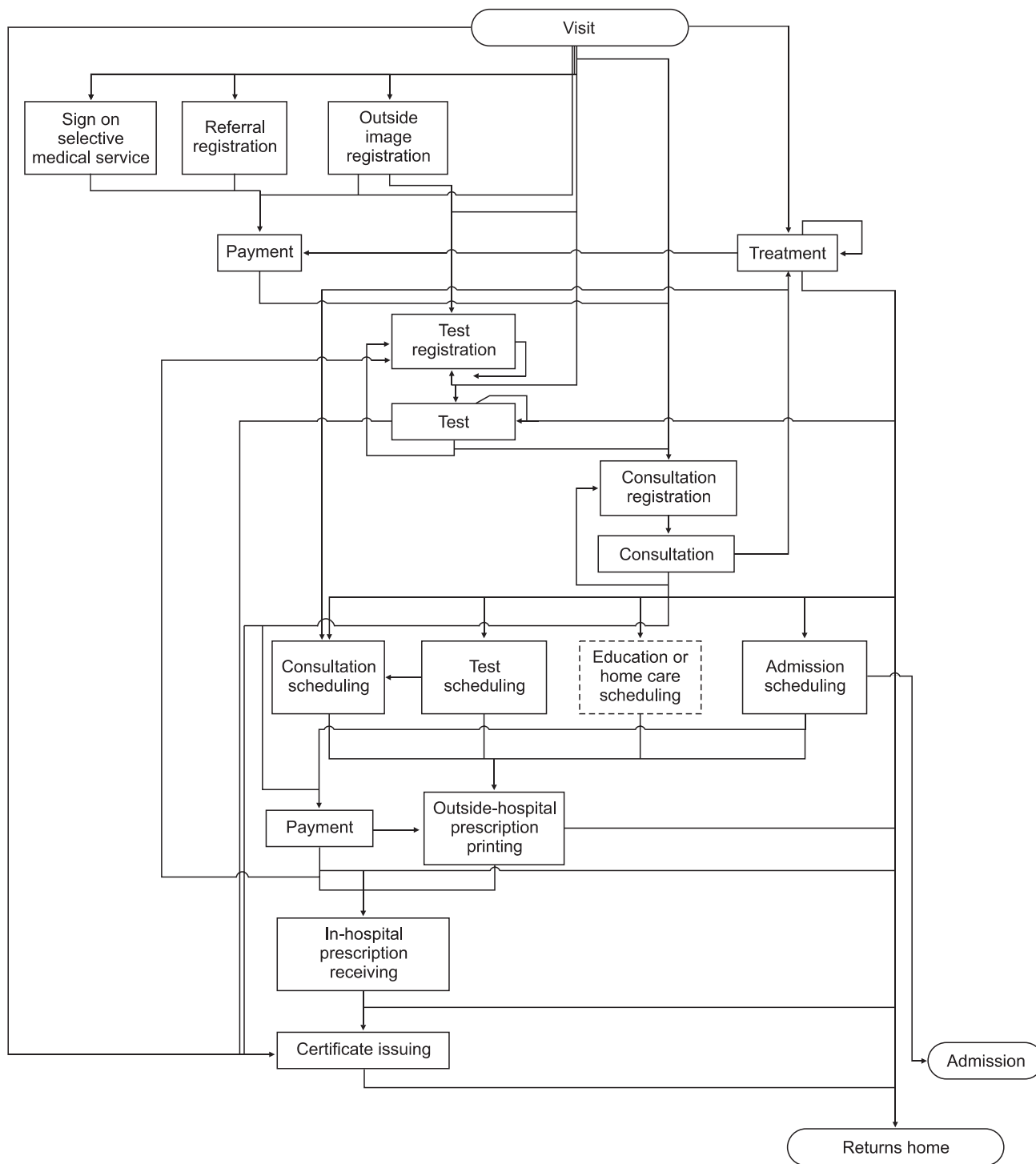
Figure 1. Outpatient care mega-process model derived from domain experts.

logs were cataloged into 15 different types of events and 162 different departments. Excluding treatment, the frequency and ratio of the start/end of the 14 types of events are shown in Table 2.

The type of event that occurred most frequently was the payment step, which accounted for approximately 18% of the total number of events. The frequency of the events that occurred in the beginning stages of the process were con-

sultation registration, test registration, and payment, which comprised approximately 80% of the total number of events during the beginning stages. The frequency of the events that occurred in the ending stages of the process were outside-hospital prescription printing, testing, and payment, which comprised approximately 73% of the total number of events. We note that the cases for test registration and consultation registration appeared in the ending stage in the Table 2 can

Table 1. The types and attributes of event logs to analyze outpatient care process

Event type (activity)	Attribute
Sign on selective medical service	Case ID, Activity completion time, Resource ID, Resource department code
Referral registration	Case ID, Activity completion time, Resource ID, Resource department code
Outside image registration	Case ID, Activity completion time, Resource ID, Resource department code
Payment	Case ID, Activity completion time, Resource ID, Resource department code
Test registration	Case ID, Activity completion time, Resource ID, Resource department code, Test code, Type of test, Scheduled test date
Test	Case ID, Activity completion time, Resource ID, Resource department code, Test code, Type of test, Scheduled test date
Consultation registration	Case ID, Activity completion time, Resource ID, Resource department code, Patient type, Department code, Appointment method, Appointment Date
Consultation	Case ID, Activity completion time, Resource ID, Resource Department code, Patient type, Department code, Appointment method, Appointment Date
Consultation scheduling	Case ID, Activity completion time, Resource ID, Resource department code, Patient type, Practitioner ID, Scheduled department code, Scheduled consultation date
Test scheduling	Case ID, Activity completion time, Resource ID, Resource department code, Test code, Type of test, Scheduled test date
Admission scheduling	Case ID, Activity completion time, Resource ID, Resource department code
Outside-hospital prescription printing	Case ID, Activity completion time, Resource ID
In-hospital prescription receiving	Case ID, Activity completion time, Resource ID, Resource department code
Certificate issuing	Case ID, Activity completion time, Resource ID, Resource department code
Treatment	Case ID, Treatment start date, Resource ID, Resource department code, Treatment code

Case ID: unique ID for identification of outpatient patients of the day, Resource ID: unique ID for identification of someone or something that performed the specific activity, Resource department code: unique code for identification of the resource ID's departments.

be regarded as the activities performed by the staff rather than those requested from the patients.

2. Matching Rate Between the Expert–Driven Process Model and Machine–Driven Process Model

The outpatient care process model derived by using this process mining technique was compared to the mega-process model suggested by domain experts (Figure 1) in order to understand major flows within the model. Figure 3 shows the major workflow which was discovered by the process mining technique.

Regarding the total flow frequency, the matching rate, which is the matching rate with the expert-driven process model, was found to be 89.01%. Moreover, most of the processes occurred with relative accuracy in accordance with the expert-driven process model.

3. Frequency Analysis by the Number of Outpatient Events

In the application of the clustering technique, the log data was cataloged by the number of events, and an analysis was conducted with respect to the frequency of outpatient events. Figure 4 shows the number of cases and ratio per the number of outpatient events.

The cumulative percentage of the seven types of events comprised 80% of the total frequency, and in the case of outpatient care services, the majority of the processes were completed within the seven events. Cases with several events, such as one case with 15 events (Test registration > Test > Test registration > Test > Test registration > Test > Test registration > Test > Test registration > Test > Consultation registration > Consultation > Consultation scheduling > Payment > Outside-hospital prescription printing) comprised 7.06% of the events. This event process was found to be applicable for patients who require many screenings.

4. Frequent Outpatient Pattern Analysis

Using process mining techniques, the patterns primarily perceived by outpatients were identified. Each pattern is shown together with the activity time and frequency, and patterns with a higher frequency had a greater importance. Table 3
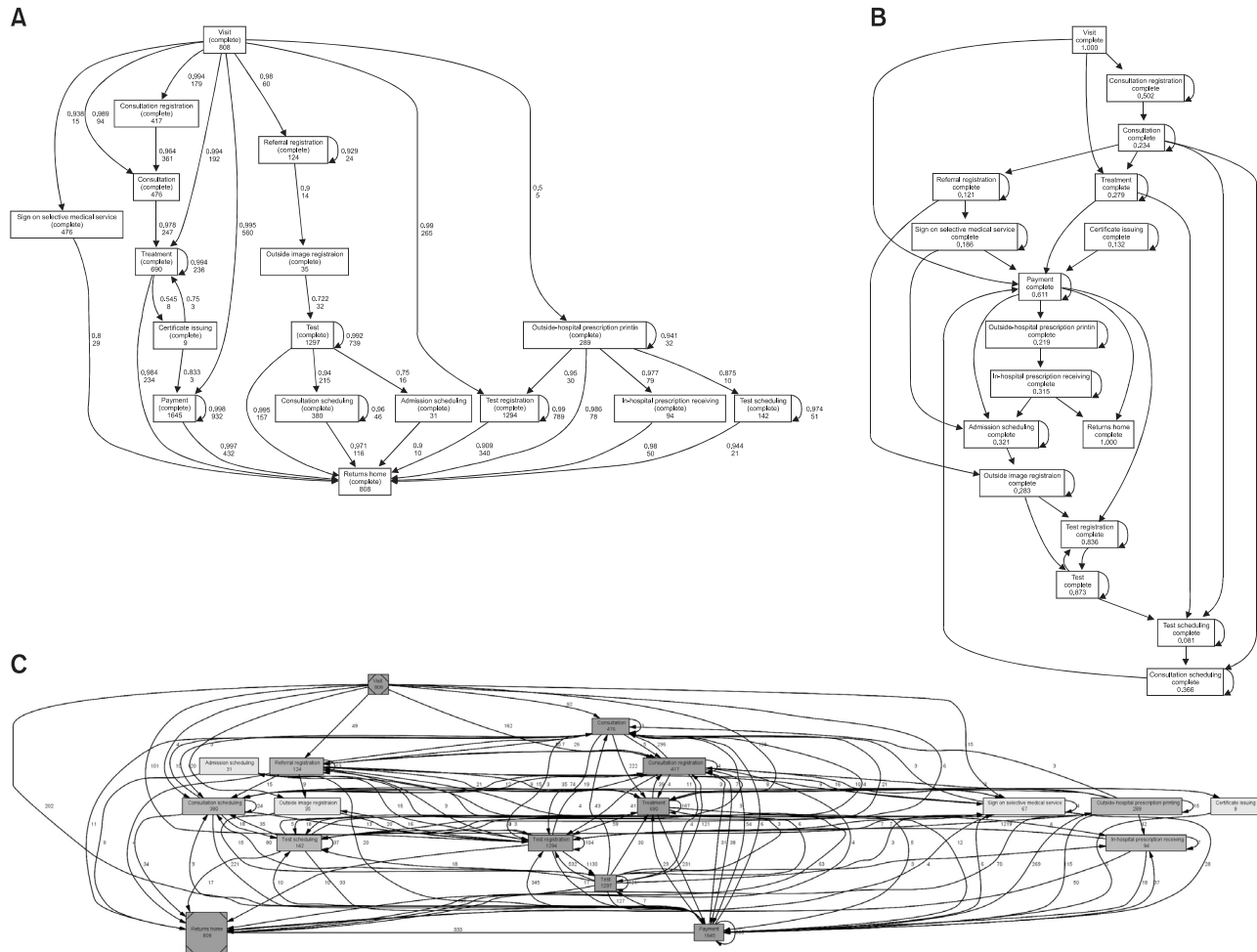
Figure 2. Various process mining algorithms: (A) heuristic miner, (B) fuzzy miner, and (C) comp miner.

Table 2. Event distribution

| Activity	Total		Start		End	
	Frequency	Ratio (%)	Frequency	Ratio (%)	Frequency	Ratio (%)
Payment	122,311	18.3	26,815	21.8	20,481	21.3
Test registration	107,235	16.0	22,377	18.2	273	3.6
Test	108,976	16.3	30	0.0	24,637	25.7
Consultation registration	76,064	11.4	49,380	40.1	270	0.3
Consultation	79,772	11.9	3,534	2.9	3,425	3.6
Consultation scheduling	61,965	9.3	496	0.4	4,303	4.5
Outside-hospital prescription printing	48,000	7.2	149	0.1	24,694	25.7
Test scheduling	21,771	3.3	4,013	3.3	4,003	4.2
Referral registration	14,928	2.2	6,552	5.3	2,560	2.7
Sign on selective medical service	14,724	2.2	8,602	7.0	6,845	7.1
In-hospital prescription receiving	6,846	1.0	419	0.3	3,486	3.6
Outside image registration	3,011	0.5	514	0.4	0	0.0
Admission scheduling	2,503	0.4	84	0.1	1,087	1.1
Certificate issuing	1,030	0.2	334	0.3	5	0.0
Total	669,136	100	123,299	100	123,299	100
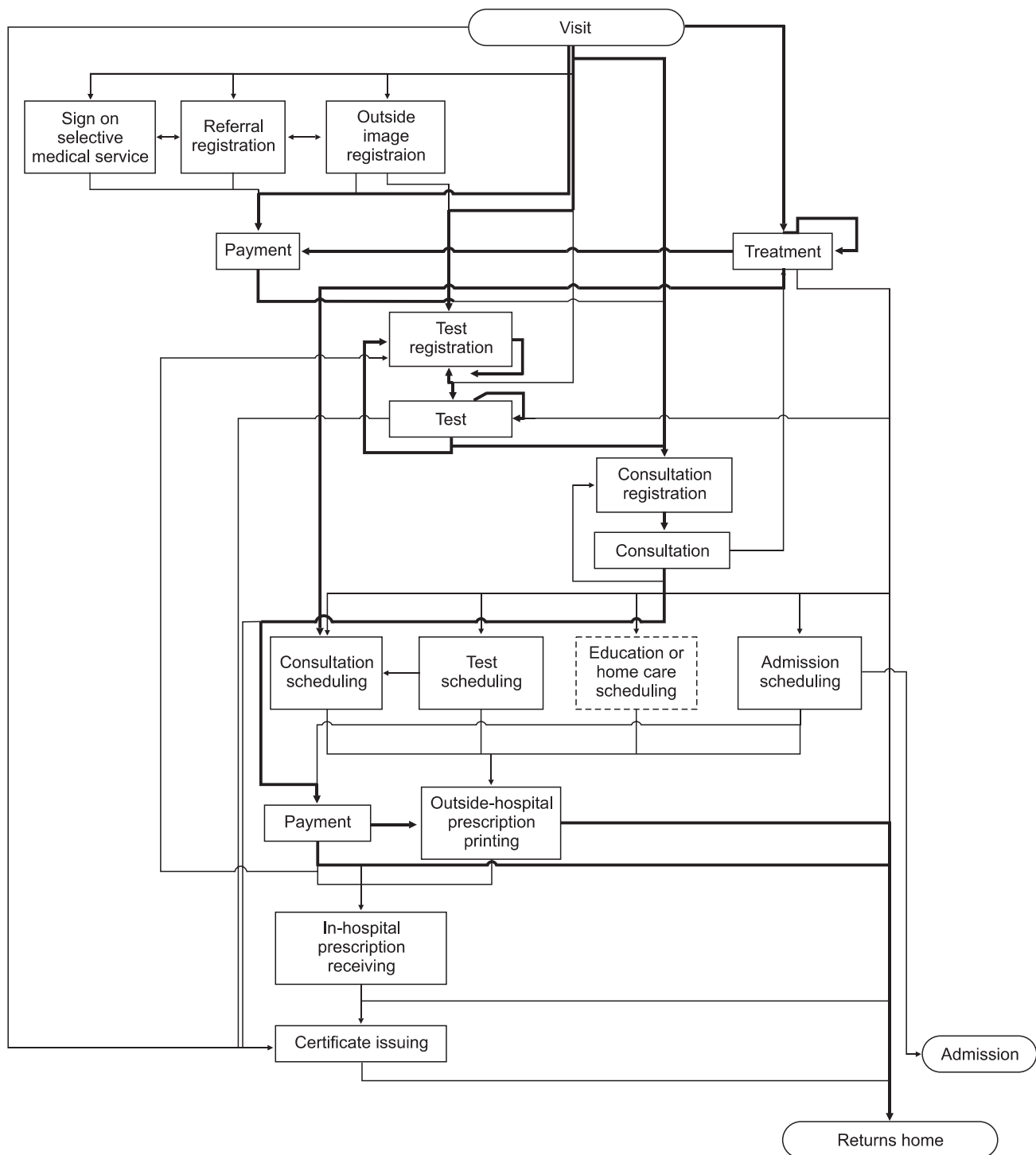
Figure 3. A major workflow (bold line) discovered by the process mining algorithm.

shows the 10 most frequent processes ranked in order of decreasing frequency.

These processes in Table 3 comprised 48.57% of the total frequencies. From the analysis results, the outpatient care process was in good agreement with the expert-driven process model, and although there may be situations with numerous processes within a hospital, there was no pattern in which a process cannot happen in a hospital.

IV. Discussion

There have been a few studies investigating process mining techniques in the healthcare domain. Mans et al. [3] applied process mining to two dataset of the clinical course data and
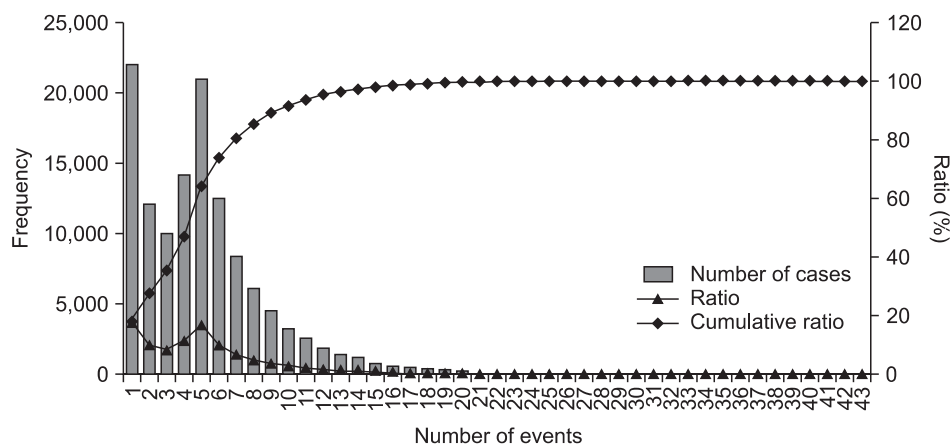
Figure 4. Number of cases and ratio per number of outpatient events.

Table 3. Analysis of the 10 most frequent process patterns

Process	Frequency	Ratio (%)	Average duration (min)
Consultation registration > Consultation > Consultation scheduling > Payment > Outside-hospital prescription printing	9,825	11.05	39.5
Payment	6,590	7.41	0
Sing on selective medical service	6,184	6.95	0
Test registration > Test	5,580	6.28	14.2
Consultation registration > Consultation > Consultation scheduling > Payment > Outside-hospital prescription printing > Treatment	3,011	3.39	46.2
Test scheduling	2,662	2.99	0
Consultation registration > Consultation > Payment > Outside-hospital prescription printing	2,439	2.74	39.3
Test registration > Test > Test registration > Test	2,368	2.66	52.7
Consultation registration > Consultation > Consultation scheduling > Payment > Treatment	2,277	2.56	51.8
Payment > Test registration > Test	2,259	2.54	44.5

the pre-hospital behavior data for stroke patients to discover clinical pathways and identify bottlenecks of pre-hospitalisation pathways. They also demonstrated the applicability of process mining for a gynecological oncology process, and showed the possibility of mining a complex hospital process. Rebuge and Ferreira [12] applied their process mining methodology in the analysis of a hospital emergency services to gain insight into the healthcare process flow, their performance, and their compliance to institutional guidelines. Yang and Hwang [8] studied a process mining technique for identifying fraudulent and abusive healthcare cases.

In this study, we applied process mining techniques to discover outpatient care process and confirmed that process mining techniques can be useful even in a healthcare environment where a variety of devices and systems are used. In particular, unlike previous studies that did not compare the machine-driven process model with an expert-driven process

model, we tested whether there is a recognizable difference in practices within hospitals by analyzing the machine-driven process and comparing it to the expert-driven process model. Proper application of the process mining technique only pertains to the log data stored in a system, which limits the application of the technique for processes that do not record the time duration, such as treatment activity. Thus, future studies should determine which data to utilize in such cases.

Based on the findings of this study, frequent process patterns can be utilized for re-assignment of resources, such as hospital staff and departments. In addition, it may also be useful for relocating resources based on distance and department relationships. We propose that the average time spent per process pattern can be utilized to analyze bottleneck points and extract delayed time for testing or consultation.

In the future, we plan to conduct an analysis of outpatient processes per practicing professors in order to select and

optimally utilize staff for outpatient care services. We expect that further analysis using process mining techniques can be useful for real process improvements regarding hospital outpatient care processes.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

References

1. Lenz R, Elstner T, Siegele H, Kuhn KA. A practical approach to process support in health information systems. J Am Med Inform Assoc 2002;9(6):571-85.

2. Anyanwu K, Sheth A, Cardoso J, Miller J, Kochut K. Healthcare enterprise process development and integration. J Res Pract Inf Technol 2003;35(2):83-98.

3. Mans R, Schonenberg H, Leonardi G, Panzarasa S, Cavallini A, Quaglini S, et al. Process mining techniques: an application to stroke care. Stud Health Technol Inform 2008;136:573-8.

4. van der Aalst WM, van Dongen BF, Herbst J, Maruster L, Schimm G, Weijters AJ. Workflow mining: a survey of issues and approaches. Data Knowl Eng 2003;47(2):237-67.

5. van der Aalst WM, Reijers HA, Weijters AJ, van Dongen BF, Alves de Medeiros AK, Song M, et al. Business process mining: an industrial application. Inf Syst 2007;32(5):713-32.

6. Van der Aalst WM, Weijters T, Maruster L. Workflow mining: discovering process models from event logs. IEEE Trans Knowl Data Eng 2004;16(9):1128-42.

7. Gunther CW, van der Aalst WM. Fuzzy mining: adaptive process simplification based on multi-perspective metrics. In: Proceedings of the 5th International Conference on Business Process Management; 2007 Sep 24-28; Brisbane, Australia. p. 328-43.

8. Yang WS, Hwang SY. A process-mining framework for the detection of healthcare fraud and abuse. Expert Syst Appl 2006;31(1):56-68.

9. Lin F, Chou S, Pan S, Chen Y. Mining time dependency patterns in clinical pathways. Int J Med Inform 2001;62(1):11-25.

10. Mans RS, Schonenberg MH, Song M, van der Aalst WM, Bakker PJ. Application of process mining in healthcare: a case study in a Dutch hospital. In: Fred A, Filipe J, Gamboa H, editors. Biomedical engineering systems and technologies. Heidelberg: Springer; 2008. p. 425-38.

11. Song M, Yang H, Siadat S, Pechenizkiy M. A comparative study of dimensionality reduction techniques to enhance trace clustering performances. Expert Syst Appl 2013;40(9):3722-37.

12. Rebuge A, Ferreira DR. Business process analysis in healthcare environments: a methodology based on process mining. Inf Syst 2012;37(2):99-116.