Gaining Insight from Patient Journey Data using Process-Oriented Analysis Approach

Lua Perimal-Lewis
School of Computer Science, Engineering and Mathematics
Flinders University of South Australia
GPO Box 2100, Adelaide 5001, South Australia

lua.perimal-lewis@flinders.edu.au

Shaowen Qin
School of Computer Science, Engineering and Mathematics
Flinders University of South Australia
GPO Box 2100, Adelaide 5001, South Australia

shaowen.qin@flinders.edu.au

Campbell Thompson
Medicine
The University of Adelaide
5005 South Australia

campbell.thompson@adelaide.edu.au

Paul Hakendorf
Clinical Epidemiology Unit,
Flinders Medical Centre
Flinders Drive, Bedford Park 5042 South Australia

Paul.Hakendorf@health.sa.gov.au

Abstract

Hospitals are continually struggling to cater for the increasing demand for inpatient services. This is due to increased population, aging, and the rising incidence of chronic diseases associated with modern life. The high demand for hospital services leads to unpredictable bed availability, longer waiting period for acute admission, difficulties in keeping planned admission, stressed hospital staff, undesirable patient and family experience, as well as unclear long term impact on health care capacity. This study aims to derive some correlation between various factors contributing to ward occupancy rate and operation efficiency. The aim is also to discover the inpatient flow process model proposing to use process mining techniques combined with data analysis to depict the relationships among inpatients, wards and Length of Stay (LOS) in an effort to gain insight into factors that could be focused to relieve access block. Open source process mining software - ProM is used for this study. The study is done in collaboration with Flinders Medical Centre (FMC) using data from their Patient Journey Database as case study.

Keywords: *inpatient journey analysis, process mining, length of stay*

1 Introduction

Patients are the most important entity of any hospital. Hospital's businesses are centred on patients and increasing number of hospitals are joining in the ideology of patient-centred care. Therefore understanding how patients flow through the hospital system is crucial in any effort to improve the way hospitals function which would directly contribute to formulating solutions addressing the challenges facing hospitals. One such challenge which is often in media's limelight is Emergency Department (ED) overcrowding also known as access block.

(Fatovich 2002) describes ED overcrowding as an international problem where an ED is unable to provide timely emergency care therefore consequently instructing ambulances to divert to another facility simply because of the lack of capacity to safely attend to newly arrived patients. Access block in ED is not an ED problem, therefore looking at patient flow through the hospital especially of the inpatient flow to understand bottlenecks would provide relief to ED congestion.

(O'Connell, Bassham et al. 2008) asserts that ED congestions are intensified by regular failure to manage processes involved in progressing patients through the hospital. According to (O'Connell, Bassham et al. 2008) some of the contributing factors for these failures are:

- disorganised nature of discharge planning
- the reliance on visits by medical staff for decisions
- the lack of a shared understanding among staff, patients and carers of the probable patient path
- accumulated impact on the occurrence of undesirable events
- increased length of stay if patients are not admitted and treated in their "home ward"

The authors mentioned above emphasised the need for better inpatient management for a better flow which in turn will reduce issues faced within ED.

Physician autonomy is one of the issues contributing to uncoordinated and often disengaged functions at hospitals. (Abraham and Reddy 2010) specifies the challenges faced by hospitals for patient transfer are contributed by uncoordinated inter-departmental interactions based on superficial status differences between clinical and non-clinical staff and professional hierarchies.

Hospitals are continually looking to improve efficiency. The term Clinical Process Re-engineering has become a common term when referring to clinical process improvements. Clinical Process Re-engineering could be considered similar to Business Process Re-engineering (BPR). Both initiatives focus on continuous improvements of business processes or clinical processes to gain "competitive advantage". However, this "competitive advantage" is service oriented and in cases of hospitals the aim is to service the health care needs of its population. BPR aims at improving core business process and in the hospital setting the core business process can be regarded as the patient journey.

Clinical process redesign, according to (Ben-Tovim, Dougherty et al. 2008) is the use of process redesign and change management to health care. The activities in the redesign process focuses on the patient's perspective. The aim of clinical process redesign is to harmonise the poorly coordinated patient journey as they move across multiple departments, making them simpler whilst looking at the overall design of the clinical processes (Ben-Tovim, Dougherty et al. 2008). Clinical process redesign demands a drastic and sometimes unusual change to the overall functioning and processes. Clinical Process Redesign takes a holistic approach by looking at a wider area in the redesign process which requires continuous fine-tuning and adjustments to continuously adapt to the ever transient nature of the hospital system.

Many hospitals around the world have adopted "lean thinking" methodology to improve patient care and have seen significant improvements in their clinical processes. Whilst "lean thinking" has contributed to better understanding of the patient flow process and therefore better co-ordination, implementing this concept alone in isolation might not be enough to reduce the hospital wide crisis. As shown in hospitals which have adopted this method, the initial improvements need to be sustainable and should continue to give results. FMC's Redesigning Care program based on the concept of "lean thinking" has produced positive results which have enabled the hospital to provide safer and more accessible care (Ben-Tovim, Bassham et al. 2008).

The use of Decision Support System (DSS) in health care is wide spread. DSS in Health Care industry could be divided into 2 broad categories. One category of DSS is used to help physicians with their day-to-day decision making for example DSS based on clinical practice guideline in the management of diabetic patients (Lobach and Hammond 1997). The other category of DSS is used by hospital management to make decisions for better hospital resource management. The fundamental information needed for such a system is based on outcomes of some method of data analysis and modelling. The closer the outcome is in depicting the real scenario the better the DSS output.

Public Hospitals are required to conform to certain Key Process Indicators (KPIs). Conformance to KPIs is essential because of the competitive nature of government funding and the need to provide justified information to tax-payers. Hospitals do already have mature processes and the ability to report on these measurable statistics. An in-depth evaluation of the processes behind these measurable statistics would aid in knowledge discovery hidden within the process.

The study aims to investigate the process of Patient Flow or Careflow within the hospital with the understanding that smooth flow would reduce bottlenecks in hospitals and this directly influences the overall capacity that would consequently reduce over-crowding and other systematic issues such as bed availability. Flinders Medical Centre's (FMC) Patient Journey Database is used as a case study.

2 Background

Health Care modelling is traditionally done using various mathematical modelling techniques based on statistics. The results of these modelling are used for forecasting and predicting in order to improve health care performances.

New approaches in health care modelling are emerging where more than one technique and approach are used to discover hidden information that might not be easily discovered from one approach. Combinations of techniques are used to complement each other.

(Ceglowski, Churilov et al. 2007) proposed combining Data Mining techniques and discrete event simulation for identifying bottlenecks in the patient flow between ED and a hospital ward by providing insight into the complex relationship between patient urgency, treatment and disposal and the occurrence of queues for treatment.

Process mining in healthcare is still in its infancy. (Mans, Schonenberg et al. 2008) used process mining techniques to better understand different clinical pathways taken by different groups of patients and used the technique to identify bottlenecks. (Rebuge and Ferreira 2011) concluded that although process mining techniques have been proven in some instances as being successful in mining health data, there are still room for improvements to identify the right algorithm to handle noise in the data, complexity of data and the ad hoc nature of health data.

3 Proposed Methodology - Process Oriented Analysis Approach

3.1 Process Mining

Process as defined by (Cambridge-University-Dictionary-Online 2011) is a "*a series of actions that you take in order to achieve a result*". Process is embedded in every aspect and at every level of one's daily routines. The output of a process is a result which could be either favourable or not so favourable. Exploring the various activities within a process contributes to deeper

knowledge, understanding and discoveries of the intricacies of what actually happens within a process that finally produces the result.

Health care industries have embraced the notion of paperless system in a big scale making them data rich which have introduced many positive challenges to researchers to discover knowledge from the use of this data. The notion of efficient patient care providing patient-centred approach has seen the emergence of various Health Information Systems (Vezyridis, Timmons et al. 2011). Electronic Patient Management or Tracking Systems have all become not only common but essential systems for any hospital. These information systems store invaluable information which can be used for knowledge discoveries. Process mining enables the discovery of knowledge regarding a process. Process mining uses event or process logs to extract information regarding a process as it takes place (van der Aalst and Weijters 2004). As noted by (van der Aalst and Weijters 2004), these process / event logs do not have to necessarily originate from a Workflow Management System. A process log could be derived from a data set that contains an order of events which could be used to assemble a process model that portrays the activity of the subject matter. This concept forms the basis of this methodology.

(van der Aalst, Reijers et al. 2007) stated that process mining aims to construct a process model from observed behaviour from a process perspective, organisational perspective and case perspective.

(van der Aalst, Reijers et al. 2007) concludes that the most significant output of process mining is the discovery of main process flow. (Rozinat, Wynn et al. 2009) stated that process mining normally creates a static model that could be used by the users of the systems to reflect on the process.

In the context of this study, the aim of process mining is to discover the various paths taken by inpatients moving from ED to ward/wards. It also hopes to discover the most frequent path taken and the associated parameters concerned with this path. Another aim is to gather socio network information of Doctors and Ward where the information presented will indicate a doctor's ward pattern. (Braitberg 2007) argues that improving operational efficiency based on average bed occupancy is too weak to predict a complex hospital system and the dynamic nature of patient flow. Hence by using process mining the aim is to measure operation efficiency by calculating the throughput time for each ward, doctors and a cluster or group of patient. The aim is to analyse inpatient journey or flow and to indentify the streams that often cause delays for patients to flow outside of ED.

3.2 Inliers vs. Outliers LOS Analysis

In-patient Length of Stay (LOS) has become one of the many ways used to measure performance of a hospital. (Thomas, Guire et al. 1997) states that patient mean LOS has been used to measure quality of care and hospital efficiency in terms of resource usage. (Thomas, Guire et al. 1997) further asserts that lower than normal LOS could indicate that hospitals are discharging patients early possibly sacrificing quality of care. Hospitals react differently to the continuous rising cost of health care. One way is to reduce the average in-patient LOS and unfortunately some hospitals reduce the number of beds in the hospital as a direct response to increasing cost of healthcare.

In the context of this study, the aim is to firstly use statistical data analysis method to clearly understand and define the boundaries of what is termed as Inliers and Outliers.

3.3 Data

The Patient Journey database from (FMC) records information on the journey or movement of a patient from the time of admission to the time of discharge. Therefore, it only contains information on inpatients or officially admitted patients.

Each admission is given a unique journey number that would remain the same until discharge. Each movement of the patient from one ward to another is recorded with a timestamp, so at any point the 'start time" in a ward and the 'end time' in a ward is known together with the ward name. Each journey is also linked to the doctor who is treating the patient. The wards are a subset of Units and the Units are a subset of Division. Each journey is also given a status of Inliers, Outliers or Inliers/Outliers. One of the notable criteria of this set of data is the ability to expand or link the patient journey with any other data from another database which collects some other information of interest. For example, the journeys could be linked to a separate database within the hospital which might only collect disease, drug or cost related information. The individual patients are not identifiable at any point.

The repository of data starts from 1st of April 2003. New or latest data can be easily added to the analysis as or when needed if deemed necessary for a particular analysis. The patient journey data will be clustered or grouped using appropriate parameters best for the particular analysis at hand. The clinicians play a major role in identifying these clusters or pool for a particular analysis that would best represent a particular cluster of interest.

3.4 Collaboration with Clinicians

Regular contact with the clinicians enabled the portrayal of the actual undertakings at the hospital which enhanced the quality of understanding and study as this is based on a real world problem. Clinicians' insight was also invaluable in explaining the processes within the hospital's context and the ways to interpret result of the analysis.

4 Preliminary Work

4.1 ProM (Process Mining) Toolkit

The Patient Journey database contains all the essential information required to construct a process model. Information contained in this database is not only useful to construct a process model but it also contains information which will be used to discover some hidden knowledge regarding a Patient's Journey. The following section demonstrates some of the knowledge which could be revealed from this database using an open-source

process mining tool called ProM. One of the benefits of this study is the notion of gaining insights into Patient Journey using well maintained database containing historical records of the various movements of a patient during an admission episode.

Records from Patient Journey Database were pre-processed and formatted into MXML format so the data could be read by the ProM toolkit. The paragraphs below show some of the output of ProM upon processing the input data from the Patient Journey Database. For the purpose of ensuring that the pre-processing exercise is done accurately and the output of ProM is in accordance to the information that is aimed to be achieved, a small subset of data has been used for the ease of manual calculation to confirm the results. Five patient journeys have been included in this analysis. A journey starts at the time of admission and concludes at the time of discharge.

Basic Log Statistics reveal statistical information from the data set. Each unique ward's execution or processing information is represented. The statistical information that can be obtained is on minimum time a patient spent in the ward, maximum time a patient spent in a ward, arithmetic mean, standard deviation, geometric mean, sum and number of times the ward has been used.

Pattern Analysis output for the 5 patient journeys is as per Figure 1. Two distinct patterns were established from the data set supplied. The pattern reveals the flow or movement of a patient from ward to ward. It could be concluded from the output below, that ward "FMC" or ED, is frequently used. The aim is to derive common patterns used by a cluster of patients so this information could be used for better capacity planning.

Pattern Analysis Result

Table view

Name	Activities
Pattern_0	"6D":"ANG":"FMC":
Pattern_1	"3D":"CCU":"TL":"ANG":"FMC":

Figure 1: Pattern Analysis – Patient Journey Flow

The frequency of usage for each Ward could be shown as per Figure 2. It shows the performance of each of the Wards involved in the 5 patient journeys used as input. A quick glance at the diagram shows that one particular ward is being used exponentially more than the other wards involved in these journeys. Again to benefit from this result, it is important to have a pre-defined cluster or group of patients to be studied.

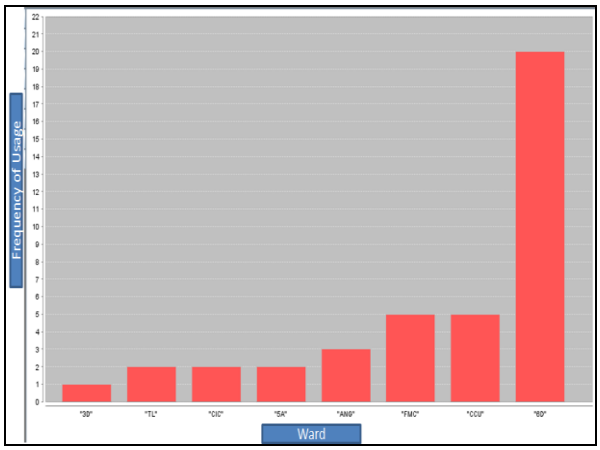

Figure 2: Frequency of Ward Usage

Individual Journeys could also be further analysed to depict the Length of Stay (LOS). This information combined with diagnostic information could be used to reveal commonalities in overall LOS and disease which would contribute to better capacity planning. (see Figure 3).

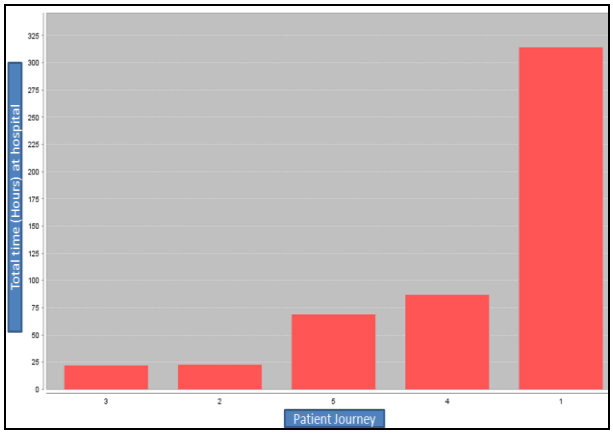

Figure 3: Journey Length of Stay

Performance Sequence Analysis facilitates the assessment of performance of the flow of journey categorised as patterns and performance of each ward involved in the pattern. Mean throughput time for each block under a pattern could be discovered. This analysis will aid is discovery of process patterns that could potentially cause issues to the system. For example identifying wards with high throughput, would enable investigation into the cause of such behaviour. Similar analysis could be carried out on doctors, which will show the transfer of work between 2 doctors, throughput time as well as the frequency of certain behaviour or pattern relating to a doctor.

Patient Journey flow discovery is as depicted in Figure 4. Each of the square boxes represents a Ward. The number below each Ward indicates the frequency of the Ward usage. The integers next to arrows represent the number of journeys that have used that path. The decimal numbers next to arrows represents Dependency Relationship, between the 2 Wards involved. A decimal number close to 1 indicates a strong dependency relationship between the 2 wards. Strong dependency relationship shows the flow in that path is likely to happen.

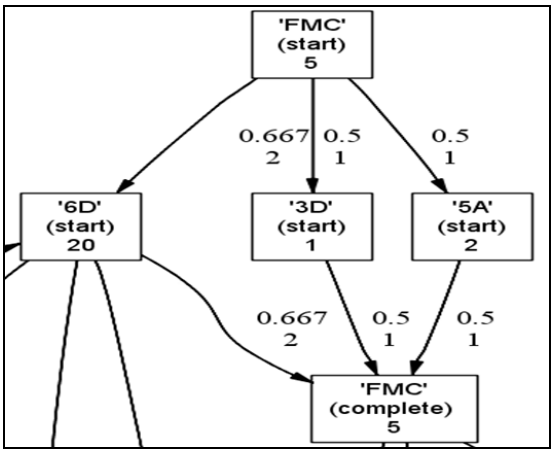

Figure 4: Patient Journey Control Flow Discovery

The outputs as described and shown are the sort of results and analysis we intend to discover from this study which will enable an enactment of a process model.

The extremely complex flow of patient journeys requires sound knowledge on the best way to derive a pool of data to be analysed which would contribute to meaningful knowledge discovery.

Patient Journey Control Flow (see Figure 4) shows the control flow of 5 journeys only. This indicates the complexities involved in patient journey analysis as an output such as the above would prove meaningless when there are more journeys involved. Therefore it is vital to invest time in deriving an appropriate pool size which will ultimately reveal the knowledge and information that could be used for decision making.

4.2 Data Analysis

The aim of this data analysis is to analyse inpatient journeys to reveal correlation in the LOS of Outliers and Inliers. An Outlier is an inpatient who is admitted to Wards other than the Home Ward. An Inlier is an inpatient who is admitted to a Home Ward. Home Ward is a ward which is equipped with appropriate medical team and specialised equipment to treat the patient's primary disease at the time of admission.

The data consist of patient journeys with 3 groups of Statuses which are Inliers, Outliers and Inliers/Outliers. Inliers are patient journeys with 100% of the time at hospital are spent in an Inlier ward. Outliers are patient journeys with 100% of the time at hospital are spent in an Outlier ward. Inliers/Outliers group consist of patient journeys with part of the hospital time is spent in an Inlier ward and the other part in an Outlier ward.

The hypothesis for this study is that Inliers have shorter LOS compared to Outliers. The aim is to investigate if outlier patients end up staying longer in the hospital as they might be treated in a ward that is not the speciality unit for the patient's condition. In other words, the effect on LOS for patients that do not stay in their Homewards. It was decided to include General Medicine patient journeys only.

The challenge experienced so far has been to find the correct cluster of data or patient journey which would appropriately represent the characteristics needed to give a holistic representation taking into consideration of the various outside factors that influence this status.

The result of the very first analysis from the cluster selected revealed figures which were contrary to the hypothesis. Upon consultation with the clinicians and looking at the data closely, it was established that LOS study from grouping the journeys using status field alone was not an accurate representation of each group. Therefore, after the first filter where non General Medicine patient journeys were excluded, the second challenge was to ensure the journeys flowed or transitioned exclusively within the General Medicine units. This cluster of data has been confirmed by the Clinicians to be the best representation of data for the analysis.

The next stage is to separate the time spent at ward "FMC" or ED from the overall LOS. It has to be noted here, that these are inpatient records and in theory ward "FMC" or ED time should be zero however, this is not the case for most of the records, which indicates that many inpatients are spending time at ED when they should already be placed in a ward.

This brings another question to be clarified before LOS results could be finalised. *"The question is whether time spent in ward "FMC" should be classified as Inliers time or Outliers time?"* Following on from this clarification, LOS might be able to be calculated. Before this decision could be made, the cluster or the pool criteria has to be checked with the Clinicians.

5 Discussion

The experience in carrying out this study thus far has been iterative. The process undertaken has proven to be a viable approach to analysing the inpatient journey. The main challenge has been in defining the boundaries of the parameters needed and defining the parameters in accordance to the actual practice rather than analysing or undertaking process mining based on the field value only.

According to Clinician's view the possibility of an Inlier staying longer is very viable as often when there are lack of beds the patients who are less sick are transferred to Outlier wards and patients who need more specialised care wait in ED before ending in an Inlier ward and consequently contributing to a longer Inlier LOS. The other scenario could be that a sicker patient staying longer at the hospital might eventually end up in an Inlier ward after being at various outlier wards.

The collaboration with clinicians has been an invaluable experience in this process. Hospitals are already undertaking various statistical data analysis for various reporting purposes to conform to KPIs and the approach taken in this study is to further break down the information and data to discover hidden knowledge.

6 Conclusion & Future Work

A smooth patient journey is an important aspect of a patient's experience in the hospital. It is as important as the actual treatment given to the patient as the patient experiencing a smooth journey will have the added bonus of the placebo effect of physiological tranquillity working hand in hand aiding the healing process.

After the finalisation of the pool or sample as an output of the data analysis, process mining techniques will be used for knowledge discovery of the inpatient journey. Length of stay of Inliers versus Outliers will be

analysed in terms of process-oriented approach with the notion to discover if Outliers have or have not followed and optimum flow. Patient journey control flow will be analysed using various process mining algorithms that are within ProM which at the same time will reveal an effective or appropriate algorithm to use for this sort of analysis.

The complex movement of patients show that patient journey analysis with data analysis approach combined with using process mining techniques where the intricacies of the process with each journeys are discovered would give further insight into a patient's journey.

Process mining results in conjunction with the usage of the proposed methods are perceived to offer added benefit to the already successful implementation of "lean thinking" and possibly enhance the areas where "lean thinking" approach alone is inadequate to investigate to reveal insight to access block.

7 References

Abraham, J. and M. C. Reddy (2010). "Challenges to inter-departmental coordination of patient transfers: A workflow perspective." International Journal of Medical Informatics **79**(2): 112-122.

Ben-Tovim, D. I., J. E. Bassham, et al. (2008). "Redesigning care at the Flinders Medical Centre: clinical process redesign using "lean thinking"." Medical Journal of Australia **188**(6): S27-S27-31.

Ben-Tovim, D. I., M. L. Dougherty, et al. (2008). "Patient journeys: the process of clinical redesign." Medical Journal of Australia **188**(6): S14-S14-17.

Braitberg, G. (2007). "Emergency department overcrowding: dying to get in?" Medical Journal of Australia **187**(11/12): 624-624-625.

Cambridge-University-Dictionary-Online (2011). Cambridge Dictionary Online. Advanced Learner's. http://dictionary.cambridge.org/dictionary/british/process_1?q=process. Accessed 14 August 2011.

Ceglowski, R., L. Churilov, et al. (2007). "Combining Data Mining and Discrete Event Simulation for a value-added view of a hospital emergency department." The Journal of the Operational Research Society **58**(2): 246-246-254.

Fatovich, D. M. (2002). "Emergency medicine." British Medical Journal **324**(7343): 958-958-962.

Lobach, M. D. P. M. S. D. F. and P. W. E. Hammond (1997). "Computerized Decision Support Based on a Clinical Practice Guideline Improves Compliance with Care Standards." The American Journal of Medicine **102**(1): 89-98.

Mans, R., H. Schonenberg, et al. (2008). Process mining techniques: An application to stroke care.

O'Connell, T. J., J. E. Bassham, et al. (2008). "Clinical process redesign for unplanned arrivals in hospitals." Medical Journal of Australia **188**(6): S18-S18-22.

Rebuge, Á. and D. R. Ferreira (2011). "Business process analysis in healthcare environments: A methodology based on process mining." Information Systems **In Press, Corrected Proof**.

Rozinat, A., M. T. Wynn, et al. (2009). "Workflow simulation for operational decision support." Data & Knowledge Engineering **68**(9): 834-850.

Thomas, J. W., K. E. Guire, et al. (1997). "Is patient length of stay related to quality of care?" Journal of Healthcare Management **42**(4): 489-489-507.

van der Aalst, W. M. P., H. A. Reijers, et al. (2007). "Business process mining: An industrial application." Information Systems **32**(5): 713-732.

van der Aalst, W. M. P. and A. J. M. M. Weijters (2004). "Process mining: a research agenda." Computers in Industry **53**(3): 231-244.

Vezyridis, P., S. Timmons, et al. (2011). "Going paperless at the emergency department: A socio-technical study of an information system for patient tracking." International Journal of Medical Informatics **80**(7): 455-465.