Extracting Nursing Practice Patterns from Structured Labor and Delivery Data Sets

Eric S. Hall, MS and Sidney N. Thornton, PhD
Department of Biomedical Informatics, University of Utah and Intermountain Healthcare, Salt Lake City, UT

Abstract

This study was designed to demonstrate the feasibility of a computerized care process model that provides real-time case profiling and outcome forecasting. A methodology was defined for extracting nursing practice patterns from structured point-of-care data collected using the labor and delivery information system at Intermountain Healthcare. Data collected during January 2006 were retrieved from Intermountain Healthcare's enterprise data warehouse for use in the study. The knowledge discovery in databases process provided a framework for data analysis including data selection, preprocessing, data-mining, and evaluation. Development of an interactive data-mining tool and construction of a data model for stratification of patient records into profiles supported the goals of the study. Five benefits of the practice pattern extraction capability, which extend to other clinical domains, are listed with supporting examples.

Introduction

When providing care, clinicians with immediate access to outcomes from similar patient cases could make better informed point-of-care decisions. Stratification of data from similar cases into profiles based on patient characteristics would allow clinicians to more easily anticipate complications and predict likely outcomes. This paper presents a methodology for extracting patterns of clinical practice from structured point-of-care data sets. Through the identification of practice pattern attributes and the analysis of variation in patterns depending on patient characteristics, this research demonstrates the feasibility of a computerized care process model that provides real-time case profiling and outcome forecasting. This model has the potential to support patient safety by improving the quality of clinical decisions and reducing unnecessary variation in clinical practice.[1]

Though vast stores of documentation and charted measurements reside in clinical databases, excesses of these data combined with innate cognitive constraints contribute to unnecessary variation and errors in clinical practice.[2] Panels of experts assemble to distill clinical knowledge into protocols, alerts, and guidelines that provide mechanisms whereby various decision makers follow the same standards of practice to achieve desired outcomes when confronted with complex decisions and distracting information.[3] Even when decision support tools are computerized and integrated into clinical information systems, they are not a panacea for decision-making. Effective evidence-based decision support requires tools be able to incorporate acquired information into clinical processes, thereby impacting clinical care.[4]

Previous efforts have focused on integrating ICD-9 coded data into practice. While these codes are available as they facilitate hospital billing processes, studies have found abstracted ICD-9 codes provide insufficient granularity for adequate encoding of clinical details.[5] ICD-9 codes offer a limited binary view of clinical information where conditions are documented as either present or absent. Furthermore, ICD-9 coding follows clinical diagnosis and can be biased by clinician interpretation or requisites for billing reimbursement.[6] Contrastingly, point-of-care data provide details of condition severity and supply information prior to clinical interpretation. When structured, point-of-care elements allow standard interpretation by computerized tools enabling automated decision support. Efforts to develop an antibiotic consultant used computerized point-of-care data to construct patient profiles allowing clinicians to prescribe more timely and cost-appropriate antibiotic regimens.[7] Structured point-of-care data provide the fine granularity of unbiased clinical information essential for generating patient profiles and representing patterns of nursing practice.

Each year, approximately 29,000 labor and delivery patients are admitted to one of Intermountain Healthcare's 15 facilities at which the Storkbytes information system is in use. Storkbytes acquires and records fetal monitor measurements and supports computerized nurse charting.[8] The system enables structured documentation of patient interventions and observations using a menu-based interface. The time of entry and an identification of the documenting nurse are recorded with each data point. Storkbytes data is extracted to Intermountain Healthcare's Enterprise Data Warehouse (EDW) along with demographic, financial, case mix, and other clinical

data relating to each patient. The EDW contains data for over 450,000 labor and delivery patients admitted to Intermountain facilities since 1995.[9] In this study, data retrieved from the EDW served as the source for practice pattern extraction and analysis.

Methods

This descriptive study incorporates structured retrospective data into the methodological framework provided by the Knowledge Discovery in Databases (KDD) process.[10] The KDD process, which includes data selection, preprocessing, data-mining, and evaluation, efficiently managed the large volumes of highly dimensional data in this study.

Data Selection: As this research was not intended to be a comprehensive analysis of all collected data points, representative elements were selected for investigation. Administrative data included patient admission time and date as well as admission facility. Patient age, length of gestation, and time of delivery were also selected. Patient conditions and outcomes were chosen from a set of perinatal quality indicators as well as from a list of common labor and delivery events.[11] Clinical condition data identified whether or not the patient experienced induced or augmented labor, a previous C-Section, the administration of epidural anesthesia, or the delivery of twins. Outcomes selected for analysis included the occurrence of fetal distress, delivery by C-Section, the number of complications occurring during labor, patient length of stay, and raw cost of care. Lastly, several point-of-care elements were selected to represent various classes of nursing activities.

To select a range of representative nursing activities, occurrences of each documented activity were tabulated and ranked by frequency for the entire year of 2006. From the 50 most common documented activities, 15 were selected for inclusion in this study. The most frequent activity was strip review – review of fetal monitor tracings including fetal heart rate and maternal contraction data. Recording blood pressure and heart rate were the next most common activities. The inclusion of these elements represented the monitoring and recording of vital signs. The remaining activities included: noting accelerations and variability of the fetal heart rate, recording frequency of uterine contractions, recording oxygen saturation level, pain score appraisal, assessment of learning needs, monitoring Pitocin administration, observation of the patient out of bed to use the restroom, nurse in the room to give general care, evaluation of epidural effectiveness, evaluation of the patient's comfort level, and recording that the patient verbalized knowledge.

Patients admitted to one of Intermountain Healthcare's ten largest facilities, each averaging more than 100 monthly admissions, were included in the study. Using a one-month sample of approximately 2,000 patients provided statistical power $(1-\beta)$ of 90% for detection of an effect size of 0.1 with a two-tailed α of 0.05. Records from the month of January 2006 were used to identify patterns that were validated with data from August 2006.

Preprocessing: Preprocessing of the selected elements provided additional information for use in the subsequent analysis. For every patient, documentation attributes were tabulated representing the number of occurrences of each nursing activity and a rate of documentation – the average time between each activity type. For example, the number of times a strip review occurred for each patient was calculated, as was the average number minutes between each strip review for each patient. To support analysis of nursing patient load, the number of patients for which a nurse charted each hour was calculated, as was the hourly documentation rate of each activity.

Patients admitted more than 24 hours before delivery or who delivered prior to admission were excluded from the analysis set. From the January data set of 2,099 patients, 2,038 remained after preprocessing. Approximately 275,000 instances of the 15 selected point-of-care elements were included in the study.

Data-Mining: To facilitate the data-mining element of KDD, we developed the interactive Pattern Extraction and Analysis Tool (PEAT). PEAT allows semiautomatic exploration of relevant data points enabling identification of nursing activity patterns, attributes of those activities, and variability in activity attributes according to patient characteristics. The application interprets structured data retrieved from Intermountain Healthcare's EDW.

Evaluation: To demonstrate the feasibility of practice pattern extraction from structured data, a profile stratification model was constructed into which patient records could be assigned to particular profiles based on constellations of patient conditions and admission facility. PEAT's interface supports navigation among patient profiles allowing the user to view activity attributes and outcomes associated with clusters of user-specified patient characteristics. In this manner, the application enables outcome forecasting for patients represented by the currently selected profile. PEAT also provides a comparison of selected profiles to the overall patient set. These calculations highlight significant variations in outcomes and activity rates among the assorted profiles. Review of the compiled profiles by domain

experts enables identification of significant patterns and actionable findings.[12]

Results

The profile stratification data model in Figure 1 accommodates real-time stratification of all 2,038 records selected for analysis in this study.
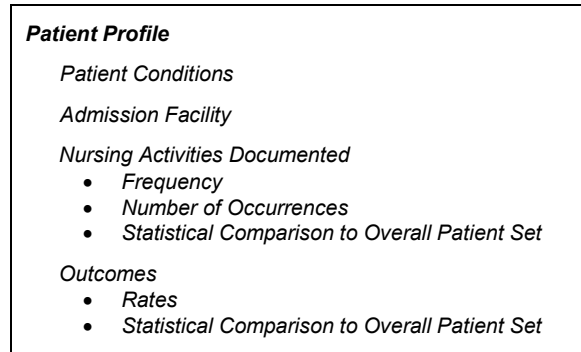
```
Patient Profile

    Patient Conditions

    Admission Facility

    Nursing Activities Documented
        • Frequency
        • Number of Occurrences
        • Statistical Comparison to Overall Patient Set

    Outcomes
        • Rates
        • Statistical Comparison to Overall Patient Set
```

Figure 1. Profile stratification data model

PEAT presents summary calculations for the entire set of patients in the *Overall Outcomes* grouping. Calculations presented in the *Current Patient Set* category are based on a user-specified subset of patient records determined by selections in the *Facilities*, *Conditions*, and *Point-of-Care Elements* groups. Listed in the *Current Patient Set* are outcomes associated with the patient subset and calculations representing the variability between outcomes of the current subset and the overall patient set. Statistical scores show the significance of those differences by means of the Satterthwaite t-test and chi-squared test as appropriate. During analysis, the overall set of patients (the 2,038 record set) may be reduced to the *Current Patient Set*, thus enabling further analysis of specific patient subgroups.

The *Facilities* grouping permits restriction of the analysis subset to patients admitted to user-designated facilities. Likewise, the *Conditions* category allows the user to limit the subset based on age range or clusters of patient conditions.

The *Point-of-Care Elements* category allows the user to view patterns associated with the 15 supported nursing activities. PEAT calculates the rate of documentation for each activity along with the average number of times each activity was documented for each patient. The user is able to enter a documentation interval for a selected activity and then view, or forecast the outcomes associated with those rates of documentation in the *Current Patient Set* group. Statistical calculations compare those results to the outcomes calculated for the entire patient set. For example, the user might select the activity *fetal heart rate variability noted*. PEAT

would then display that on average this activity occurs every 36 minutes for a given patient. A user specified interval for the activity of *20 to 30 minutes* would cause PEAT to identify a subset of 602 patients who had fetal heart rate variability noted on average every 20 to 30 minutes. PEAT would then present a comparison of the outcomes calculated for the subset of 602 patients to those of the complete 2,038 patient set. Aside each outcome calculation, PEAT would display scores that establish the statistical significance of the differences between the current patient subset and the overall set. In this example, PEAT displays a 5.9% more frequent occurrence of fetal distress for patients whom this rate of fetal heart rate variability was documented – a significant difference (p < 0.0001). Other outcome comparisons are presented in the same manner.

Lastly, the *Hourly Nurse Statistics* group presents average activity-specific documentation rates for a user specified range of nurse-patient load.

Discussion

Once established, a mechanism for practice pattern extraction and analysis provides a number of opportunities for harnessing knowledge from collected data. Five of these benefits are illustrated below with examples from the study data set.

I. Starter set generation of best practice guidelines

Summarizing data generated from normal practice and stratified into patient profiles enables clinical scrutiny and analysis of current patterns of practice. This initial step in the guideline development process is facilitated by clustering cases with similar profiles and examining variation in nursing activities and activity attributes. The associated tasks and task attributes serve as a baseline set for the profiled condition set. Table 1 presents outcomes corresponding to various intervals of recording blood pressure for patients sharing the profile of not having a previous C- Section.

Table 1. Comparison of outcomes by documented interval of blood pressure assessment

Outcomes	0-10 Minute Interval	10-25 Minute Interval	25-40 Minute Interval
Number of Patients	146	964	448
Length of Stay (Hours)	70.8	44.6	48.6
Duration of Labor (Hours)	2.0	6.8	9.5
C-Section Rate	21.9%	8.6%	12.1%
Fetal Distress Rate	6.8%	21.1%	25.7%
Number of Complications	0.37	0.31	0.36
Cost	$3,229	$2,522	$2,829

Exploration of a narrower patient profile consisting of a more complex condition constellation would support the development of a guideline for higher acuity cases that require the most frequent interval (0-10 minute) of nurse interaction. Additionally, too infrequent nurse interaction is consistent with less favorable outcomes (25-40 minute). For this profile, patients with the best outcomes experienced a 10-25 minute interaction interval.

Table 2 compares outcomes of two patient sets that met the criterion of never having a previous C-Section. The 964 patients had blood pressure recorded at intervals of 10-25 minutes while 792 patients had blood pressure recorded at rates other than every 10-25 minutes. Five of the six outcomes associated with the 10-25 minute interval subset demonstrate statistically significant improvement when compared to outcomes generated from the other set of patients. A guideline authoring committee could use this interval as a starting point to establish a standard of practice for patients with this particular profile. This process can be repeated to identify favorable intervals for any activity of interest relating to any specific patient profile generated by PEAT.

Table 2. Comparison of outcomes by documented interval of blood pressure assessment

Outcomes	10-25 Minute Interval	All Other Intervals	P-value
Length of Stay (Hours)	44.6	52.8	0.006
Duration of Labor (Hours)	6.8	8.8	< 0.0001
C-Section Rate	8.6%	15.3%	< 0.0001
Fetal Distress Rate	21.1%	22.1%	0.59
Number of Complications	0.31	0.38	0.009
Cost	$2,522	$2,932	0.0002

The practice intervals identified in this analysis, though correlated with favorable outcomes, should not be interpreted as causal to specific results. While appropriate rates of nurse-patient interaction are expected to support favorable patient outcomes, activity intervals are not the sole contributor to those outcomes. Documented nursing activity rates provide a single dimension for data exploration. Patient conditions such as drug use, advanced maternal age, or other characteristics may be equal or greater contributors to patient outcomes as are nursing interaction intervals. However, when consistent patterns are identified for patients with a common profile, significant interval findings warrant additional research and provide a foundational starting point for the development of best practice guidelines to be subsequently reviewed by domain experts.

II. Analysis of nurses' patient loads
Calculating hourly patient load of each nurse allows an analysis of the effect of patient load on nursing activity documentation rates. Table 3 lists the average number of times a nurse documented each of the listed activities for each patient according to load size. Though many of the rates remain steady as patient load increases from one patient to two, all but one of the per patient activity rates decline as patient load increases to three. Follow-up research could focus on nursing activity rates among patients of varying acuity within a single nurse's patient load. Also, analysis of patient outcomes resulting from decreased activity rates corresponding to increased patient load would support staffing decision makers in identifying manageable patient load size and maintaining patient loads within predetermined bounds to support the delivery of quality patient care.

Table 3. Hourly per patient nursing activity rates by patient load

Hourly Activities	Load of 1 Patient	Load of 2 Patients	Load of 3 Patients
Blood Pressure	1.81	1.38	1.17
Strip Review	1.75	1.76	1.38
Accelerations Noted	0.99	1.07	0.92
RN in Room to Give Care	0.36	0.36	0.43
Patient Comfortable	0.25	0.22	0.18

III. Analysis of variation in practice
Identification of facility specific nursing patterns associated with a particular patient profile enables the comparison of practices and outcomes for similar patient cases at various hospitals. Table 4 compares the average time between documentation of selected activities calculated from 293 records at one hospital to 1,745 patient records from the other nine hospitals in this study.

Table 4. Comparison of nursing activity documentation rates by facility

Activities	Specified Hospital (Minutes)	All Other Facilities (Minutes)	P-value
Blood Pressure Entered	27.4	23.6	< 0.0001
Heart Rate Entered	29.1	24.5	< 0.0001
Learning Needs Assessed	120.6	165.3	< 0.0001
RN in Room to Give Care	121.2	146.7	0.003
Pain Score Assessed	140.6	127.4	0.05

Table 5 compares the outcomes for these two patient sets and demonstrates significant differences for four of the six outcomes. This examination identifies

variations in practice and outcomes by facility and allows more effective targeting of efforts to increase compliance to established guidelines.

Table 5. Comparison of outcomes by facility

Outcomes	Specified Hospital	All Other Facilities	P-value
Length of Stay (Hours)	68.0	48.2	0.001
Duration of Labor (Hours)	7.8	7.1	0.025
C-Section Rate	24.2%	20.2%	0.11
Fetal Distress Rate	20.1%	19.4%	0.76
Number of Complications	0.47	0.29	< 0.0001
Cost	$3,304	$2,731	0.035

IV. Hypothesis generation

Exploration of practice patterns yields testable hypotheses for selected patient subpopulations. The data in this study suggest that patients whose labor was induced experienced average labor duration of 3.2 hours longer ($p < 0.0001$) than patients whose labor was not induced. This finding has been supported by a separate recent study at Intermountain Healthcare.[13] By producing experimental hypotheses for follow-up research, findings can serve as catalysts for studies to improve outcomes within the domain.

V. Candidate selection for future data collection

As results are limited by the type and accuracy of data collected, acquisition tools and techniques must be refined. The exercise of pattern analysis reveals that some desired measurements are absent from the collected data set. In this study, estimated nurse patient load was calculated from the number of patients for which charting occurred in a given hour. Though adequate for this study, this approximation requires that a nurse chart for each of her patients every hour. Consideration of acuity in patient load calculations would also improve analysis and enable examination of findings within like-acuity subgroups. Thus, future studies would benefit from collection of hourly nurse load and patient acuity data.

Conclusion

This research focused on extracting and analyzing nursing practice patterns from structured labor and delivery data; however, the methods presented are extensible to other clinical domains for which structured point-of-care data are available. Use of the KDD process and a customized data-mining tool demonstrate the feasibility of constructing a model that supports case profiling and outcome forecasting. Once in place, a mechanism for automating and expediting the process of extracting practice patterns extends the benefits listed above supporting improved decision-making and higher quality care.

References

1. Bates DW, Gawande AA. Improving safety with information technology. N Engl J Med 2003;348(25):2526-34.
2. Morris AH. Developing and implementing computerized protocols for standardization of clinical decisions. Ann Intern Med 2000;132(5):373-83.
3. Tourangeau AE, Doran DM, Hall LM, O'Brien LP, Pringle D, Tu JV, et al. Impact of hospital nursing care on 30-day mortality for acute medical patients. Journal of Advanced Nursing 2007;57(1):32-44.
4. Wan TT. Healthcare informatics research: from data to evidence-based management. J Med Syst 2006;30(1):3-7.
5. Bui AA, Taira RK, El-Saden S, Dordoni A, Aberle DR. Automated medical problem list generation: towards a patient timeline. Medinfo 2004:587-91.
6. Cheung NT, Fung V, Chow YY, Tung Y. Structured data entry of clinical information for documentation and data collection. Medinfo 2001;10(Pt 1):609-13.
7. Evans RS, Pestotnik SL, Classen DC, Burke JP. Development of an automated antibiotic consultant. MD Comput 1993;10(1):17-22.
8. Twede M, Gardner RM, Hebertson RM. A PC-based system for intrapartum monitoring. Contemporary OB/GYN, "Special Issue -- Technology 1985" Oct 1984; 24:13-17.
9. Lau LM, Lam SH, Barlow S, Lyon C, Sanders D. Enhancing an enterprise data warehouse with a data dictionary. Proc AMIA Fall Symposium 2001.
10. Fayyad U, Piatetsky-Shapiro G, Smith P. Advances in knowledge discovery and data mining. 1996:1-36.
11. Suresh GK, Ferguson LA, Tomlinson M, Campbell BB, Ohlinger J, Prochnicki B, et al. Identification and collection of quality indicators for perinatal care. J Nurs Care Qual 2007;22(1):73-9.
12. Berger AM, Berger CR. Data mining as a tool for research and knowledge development in nursing. Comput Inform Nurs 2004;22(3):123-31.
13. IHC Health Services. Management of elective labor induction. Intermountain Healthcare White Paper 2006.