Data Mining Medication Administration Incident Data to Identify Opportunities for Improving Patient Safety

by

Michael David Gray

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 18, 2009

Keywords: medication administration errors, data mining, medication error, healthcare informatics, human factors, nursing

Approved by

Robert E. Thomas, Jr., Chair, Professor of Industrial and Systems Engineering
Juan E. Gilbert, Professor of Computer Science and Software Engineering
Kathleen C. Brown, Professor of Nursing
Kathy Jo Ellison, Associate Professor of Nursing
Bill G. Felkey, Professor of Pharmacy Care Systems

Abstract

This research analyzed historical data related to medication administration errors at a 340 bed regional medical center. The objective was to determine if data mining techniques could identify relationships within the error data that point to processes and circumstances that enable medication administration errors. The Cross Industry Standard Process for Data Mining (CRISP-DM) was used to determine if data mining techniques applied to medication administration error data could yield information that could improve the systems and processes supporting medication administration at a regional medical center. Data sources from the point of medication dispensing to the patient's response were investigated. Base data over a one year period were queried to obtain all available information relating to acknowledged medication administration errors. These data were analyzed using Microsoft SQL Server 2005 - Clustering Algorithm. The clustering algorithm results confirm the limitations of self reporting as a means of medication administration error measurement. Further, the research identifies cultural, process, and policy inconsistencies that drive self reporting behavior and subsequently lead to marginalized error event knowledge capture. These findings contribute to the development of recommendations for design improvements for medication error reporting systems. Additionally, the difficulty of deriving information from multiple Healthcare IT systems that are not integrated is demonstrated. The results provide practical guidance for organizations evaluating Clinical Decision Support Systems designed to support the medication use process.

Acknowledgments

The author would like to thank Dr. Rob Thomas for his insight and direction as dissertation Chair. Thanks to Dr. Juan Gilbert, Dr. Kathy Jo Ellison, Dr. Kathleen Brown and Professor Bill Felkey for their responsiveness and expert advice. In addition, thanks are due to the Deep South Center for Occupational Health and Safety and the National Institute for Occupational Safety and Health (NIOSH) for their funding support. Finally, I wish to thank my family and friends for their constant support and encouragement, especially my parents, sister, and brother.

The author wishes to express deepest gratitude and offer special thanks for my former wife Dabsey Maxwell who has dedicated her life to the care and development of our children Monroe and Rainey. Dabsey, without you this life accomplishment would not have been realized.

Table of Contents

List of Figures

List of Abbreviations

AHRQ	Agency for Healthcare Research and Quality
IOM	Institute of Medicine
HRA	Human Reliability Analysis
CRISP- DM	Cross Industry University of Georgia
KDD	Knowledge Discovery in Databases
JCAHO	Joint Commission on Accreditation of Healthcare Organizations
ADE	Adverse Drug Event
NCC MERP	National Coordinating Council for Medication Error Reporting and Prevention
ISMP	Institute for Safe Medical Practices
RN	Registered Nurse
MC	Medical Center
IRB	Institutional Review Board
IHI	Institute for Healthcare Improvement
PSO	Patient Safety Officers
ISMP	Institute for Safe Medical Practices
MAR	Medication Administration Record
eMAR	Electronic Medication Administration Record
SASS	Microsoft® SQL Server™ 2005 Analysis Services

Chapter 1

Introduction and Background

Much has been done to address the patient safety crisis that was heralded by the Institute of Medicine's landmark report "To Err is Human: Building a Safer Health System", which estimated that between 44,000 to 96,000 people die each year as a result of medical errors (Institute of Medicine, 2000). Many studies conducted in the last decade attempting to determine medical errors frequency, type, and contributing factors, have been added to the last half century of medical error research. Error evaluation methods, analysis techniques, process improvement methodologies, and technology solutions from healthcare informatics, computer science, human factors engineering, risk management and many other disciplines are being used to help solve a complex quality problem that is pervasive. The Institute of Medicine called for a 50% reduction in medical error within ten years (Institute of Medicine, 2000). Unfortunately, healthcare in the United States has had difficulty achieving double digit improvement in medical error reduction (Leape & Berwick, 2005). Despite the elevated awareness of patient safety as an issue and the programs that are being implemented to improve quality, the perception of progress remains in doubt. For example in one study over half of U.S. physicians felt their ability to deliver quality care had decreased over a five year period. Further 30% of physicians rated their hospitals as fair to poor at finding and addressing medical errors (Blendon et al., 2001).

Perhaps the technological advances in medicine are counteracting measures designed to make healthcare delivery safer. Over the last 20 years, the United States healthcare industry has seen

dramatic innovation due to science and technology. Knowledge, skills, care interventions, devices, and drugs may have advanced more rapidly than our ability to deliver them safely, effectively, and efficiently (Institute of Medicine, 2001)

Ironically, this point-of-view does not appear to be shared by all healthcare professionals. Survey results indicated that when queried about their awareness of medication administration errors occurring at their institutions, 91 percent of all respondents, which included CEOs, chief nursing officers and pharmacy directors, believed that they were well informed. However, when the same leaders were asked to estimate the number of medication errors that occurred at their institutions in the last month, 43 percent of CEOs and 25 percent of heads of nursing did not know and 34 percent of pharmacy managers estimated that as many as 21 medication errors occur at their institutions over a 4-week period (Bruskin-Goldring Research, 1999).

In contrast, research involving 36 hospitals and skilled nursing facilities indicated that medication errors occur at a rate that approaches 1 out of every 5 doses with 7% of all doses considered to be harmful to the patient (Barker, Flynn, Pepper, Bates, & Mikeal, 2002). Medication administration is considered one of the areas of the medication use process associated with the greatest risk for errors. One study of Adverse Drug Events (ADE) in two hospitals over a 6-month period found that 38% of the errors were attributed to drug administration by the nursing staff (Pepper, 1995). A study analyzing 3 years of medication error records reported to the U.S. Pharmacopeia (USP) MEDMARX medication error reporting program, which included over 154,000 records, found that 37% of errors were attributed to medication administration (Santell, Hicks, McMeekin, & Cousins, 2003). The risk lies in the complexity of selecting the correct drug, dose, route, patient, and time while being cognizant of prescribing or dispensing errors (Institute of Medicine, 2004a). Clearly there is a disconnect

between what the healthcare community is able to acknowledge regarding its error rate and what the error rate actually appears to be. Perhaps the error reporting methods themselves contribute to the misperception of error prevalence and create a false sense of security in the patient safety programs that are based on error analysis.

Holden and Karsh (2007), in their review of medical error reporting system literature, acknowledge that although medical error and incident reporting systems are being advocated by national and international public and private sector healthcare policy and patient safety advocates as an important component for improving patient safety: it is only one of many tools in an effective patient safety program (Holden & Karsh, 2007). The Institute of Medicine is among the proponents recommending the use of both mandatory and voluntary reporting systems for identifying and learning from medical errors and near misses (Institute of Medicine, 2000). Incident reporting systems may have direct relevance within the context of the healthcare organization that uses them, but their findings do not have broad application to the healthcare sector in general. This is because the various public and private incident reporting systems do not conform to a standard taxonomy for classifying error attributes (Institute of Medicine, 2004b). Therefore valid comparisons of different studies on medication errors are extremely difficult because of differences in variables, measurements, populations and methods (Manasse Hr, 1989).

Although incident reporting systems seem to be pervasive in their use, they have limited ability to quantify the frequency and magnitude of medical errors being committed (Flynn, Barker, Pepper, Bates, & Mikeal, 2002a). Cullen et al. caution the use of voluntary incident reporting for use as a quality assurance/quality improvement tool as it leads to bias for assessing

quality of care. Evidence from their study revealed that only a small fraction of Adverse Drug Events (ADE) was detected using voluntary incident reports (Cullen et al., 1995).

There are a number of cognitive, social, and administrative reasons for medication errors being under reported thereby reinforcing the perception that the error rate and error magnitude are significantly less than they actually are. In support of this assertion, Barker and McConnell point out that the primary disadvantage of using incident reports for medication administration errors is that a nurse must be aware that an error has occurred. However, research has shown that nurses are rarely aware of errors (Barker & McConnell, 1962).

Nurses are not the only clinicians that under report medication errors. Physicians working at a pediatric hospital acknowledged that they reported less than 20% of their perceived medical errors in the incident reporting system; whereas nurses in the same study reported they reported more than 80% of their perceived errors. Beyond the findings that suggest multiple deterrents to incident report system use, both physicians and nurses indicated a likelihood for increased system use should specific design factors improve (Taylor et al., 2004).

Despite its inadequacies incident reporting of one version or another has almost become ubiquitous in U.S. hospitals. Therefore in the context of patient safety improvement programs, incident reporting systems remain as an important and relatively inexpensive means of capturing data on errors in any or all of three basic categories: adverse events, "no harm events," and "near misses" (Agency for Healthcare Research and Quality, 2001).

A number of other methods exist for detecting medication errors. One method of detecting and enumerating medication errors using observation began in the early 1960s (Barker & McConnell, 1962). A study conducted in 1999 compared incident reporting, medical chart reviews, and direct observation for efficiency and accuracy in detecting medication errors

(Barker, Flynn, & Pepper, 2002). The results of the study determined that direct observation was more efficient and accurate than reviewing charts and incident reports in detecting medication errors (Flynn, Barker, Pepper, Bates, & Mikeal, 2002a). The effectiveness of direct observation has been demonstrated in other medical settings. For example, in a prospective study of surgical units, almost 80% of the errors identified by trained observers were not officially recognized or recorded by the individual or the institution that made the error (Krizek, 2000). Despite its error capturing capabilities, the observation method is an expensive detection technique when compared to incident reporting and medical chart reviews. Flynn et al. calculated the mean cost per examined dose for incident reports, assuming that the error report is processed by an third party, and found that these costs were $.067 for medical chart review, and $4.82 for observation (Flynn, Barker, Pepper, Bates, & Mikeal, 2002b).

Barker et al. acknowledge that "since the beginning of medication error research no single method of error detection will work in all situations" (Barker, Flynn, & Pepper, 2002). For that reason, the observation method in combination with incident reports and medical chart reviews should be used to help determine the "clues to cause." These results should then be used to direct the detailed hazard analysis such as Root Cause Analysis (RCA) and Failure Modes and Effects Analysis (FMEA) to the areas of the medication use process most likely in need of redesign to prevent "latent error" conditions (Reason, 1990).

Data Mining in Healthcare

By some estimates, the amount of data stored in the world's databases doubles every 18 months. According to IDC, the amount new digital information created in 2008 totaled 487 billion gigabytes (John Gantz, 2009). It is evident that the rate in which data are created has out paced our ability to understand what it tells us. As the volume of data generated continues to

grow exponentially, our ability as humans to become aware of the new data's existence, determine its relevance, comprehend its meaning, and use it to make decisions is becoming exponentially more difficult. Unfortunately, important decisions are being made in the absence of information that remains hidden in data stores. If the available data could be analyzed to expose patterns that, when evaluated, aid the decision making process, the resulting knowledge could have a powerful impact in our daily lives. Empowered with this capability we would reduce uncertainty in the decision outcomes. This is the opportunity that data mining presents. Data mining is about solving problems by analyzing data already present in databases (Ian H. Witten, 2000).

Growth in healthcare data contributes to the overall digital expansion as a result of the use of higher resolution diagnostic and imaging systems, electronic medical records, computer prescriber order entry, enterprise hospital management systems, electronic laboratory test results and development of new drugs and their associated pharmacological data. Healthcare, much like any other industry, is creating massive amounts of data but cannot keep pace with the need to understand what the data means beyond its immediate intended use. However, unlike most industries that measure the lost opportunity resulting from uninformed decisions, in terms of reduced market share or shrinking profit margins, healthcare in addition to these measures also accounts for unintended results from its decisions in terms of human suffering and lives lost. Decision support systems and other health information technology innovations have significantly improved the standard of care by enabling greater compliance to clinical best practices, and providing alerts to key care parameters that are outside of nominal or expected values. These technologies have had an impact on improving therapeutic processes for patients as well as achieving improvements in protecting patients' safety.

Generally, the studies that connect Information Technology (IT) with improved outcomes in the clinical setting, have been conducted in select academic medical centers within settings that include substantial resources and a long standing trend of IT implementation and adoption not typical for most U.S. hospitals. Therefore the generalizability of such findings are considered limited (Chaudhry et al.).

IT does not guarantee improvements in quality of care. Obviously, when IT is not implemented properly it can have the opposite effect by compounding problems it was originally intended to solve. Further, IT can introduce other problems that were not previously experienced using the incumbent systems or processes. However, IT systems can produce better results and improve the care delivery process. Bates and his colleagues determined that implementing a computer physician order entry system with decision support capabilities resulted in an 83% reduction in serious medication errors (Bates et al., 1999).

Within the context of data analysis methods, data mining can be considered to be an exploratory approach in which a hypothesis is specified and the validity of the hypothesis is tested against the data (Osei-Bryson & Rayward-Smith). Hospitals collect data on patients, doctors, medications, and procedures. These data are often untidy, incomplete, and sometimes erroneous, and yet if used properly, data analysis can be a valuable asset for management (Osei-Bryson & Rayward-Smith).

Data mining uses a variety of methods drawn from statistics and machine learning. These include approaches such as tree and rule induction, k-nearest neighbor algorithms, Bayesian classifiers, neural networks, support vector machines, logistic regression, discriminant analysis, clustering, multidimensional scaling, and association rule mining. While the objective for all of these methods is to determine relationships among variables, classify instances, and to

predict new instances, tree- and rule-based approaches are particularly useful for classification problems (Apté & Weiss, 1997).

Chapter 2

Research Objectives and Methodology

Data mining is aimed at the development and application of models that can improve the quality of organizational decision-making. The first objective of this research is to determine if data mining can provide medication error situational awareness beyond what is currently provided from incident report analysis. The second objective is to determine if the results from data mining analysis can identify areas in the medication administration process that require redesign to improve patient safety. This exploratory research will attempt to address these objectives by answering the following three research questions.

By applying appropriate data mining techniques to a medication administration error dataset;

1) is it possible to find a subset of input attributes that differentiate medication administration errors that occur within a distinct step of the overall medication administration process?
2) is it possible to find a subset of input attributes that are associated with known medication administration error event types which are not documented in the incident reporting dataset?
3) is it possible to predict the occurrence of medication administration error types that are not currently being document in the incident reporting system?

This research considered medication error incident report data spanning a 30 month period from a 340-bed regional medical center. The Medical Center is fully accredited and

offers programs and services in nearly every medical specialty, including a full range of inpatient and outpatient procedures. In total, the Medical Center employs 2500 people with an active medical staff of more than 145 physicians.

The scope of this project was limited to the medication administration process. An assumption is made that the medication was correctly ordered by the physician, received and accurately filled by the hospital pharmacy, was verified as an appropriate therapeutic dose by a clinical pharmacist and loaded properly into an automated dispensing machine or dispensed from controlled pharmacy location. Although it is conceivable that errors in the medication ordering and dispensing process could be identified, only the ordering and dispensing errors that contribute to a medication administration error were accounted for in this research.

Given the limitations of incident reports for capturing the medication error frequency, severity, and type, the purpose of this exploratory research is to determine if data mining can offer additional information about the error environment. The relationships that may exist, could point to components of the process where redesign would most effectively improve medication administration safety. An ultimate outcome of this research was to provide the sponsor hospital with the insight to most effectively determine the key areas within the medication administration process that contribute to the majority of medication administration errors.

Additionally, this data mining research attempted to determine the feasibility of predicting the dominant medication error conditions that are associated with acknowledged errors but are not apparent to the clinician when the errors are originally discovered. The ability to predict medication administration error conditions associated with one or more error types can be useful to guide further systems safety analyses. Also the results may suggest where in the process other types of errors could be occurring due to patterns in the error environment but are

not currently being documented in the incident reporting system. With this knowledge the sponsor hospital can focus clinical process improvements on the sub-processes which contribute to the majority of documented and predicted errors. The results could give rise to a heightened awareness that medication errors are occurring more frequently than is currently being recognized or acknowledged and that other methods of assessing the error rate and type should be considered.

Research Method

Although this research protocol only considered historical data, which did not require the construction of data gathering instruments for administration to human subjects, Institutional Review Board approval was sought and received from both the sponsor hospital and from Auburn University. The data used to conduct this study included health information which is protected under The Health Insurance Portability and Accountability Act of 1996 (HIPAA). Therefore strict data privacy and security measures were adhered to in accordance with the Medical Center's data privacy and security policies and procedures. The data gathering process was conducted onsite at the Hospital's facilities on its secured network. The datasets were sanitized, removing all ability to associate a patient's identity within the data. Once de-identified datasets were constructed, the data mining analysis was conducted offsite.

The Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology was used as a means for managing the knowledge discovery in data (KDD) process for this research effort. The CRISP-DM framework provided a systematic approach for understanding the medication use process within the collaborating hospital. In addition, it provided structure to the data gathering phase as information surrounding the medication administration process was voluminous. After the necessary data aggregation and data cleansing phases were completed, the

methodology helped organize the modeling process where the appropriate data mining techniques were tested and the final data mining algorithm and its supporting software tool were selected. During the evaluation phase, the data mining results were analyzed to determine if any previously unknown relationships emerged and what those relationships imply about the error conditions latent in the hospital's medication administration process.

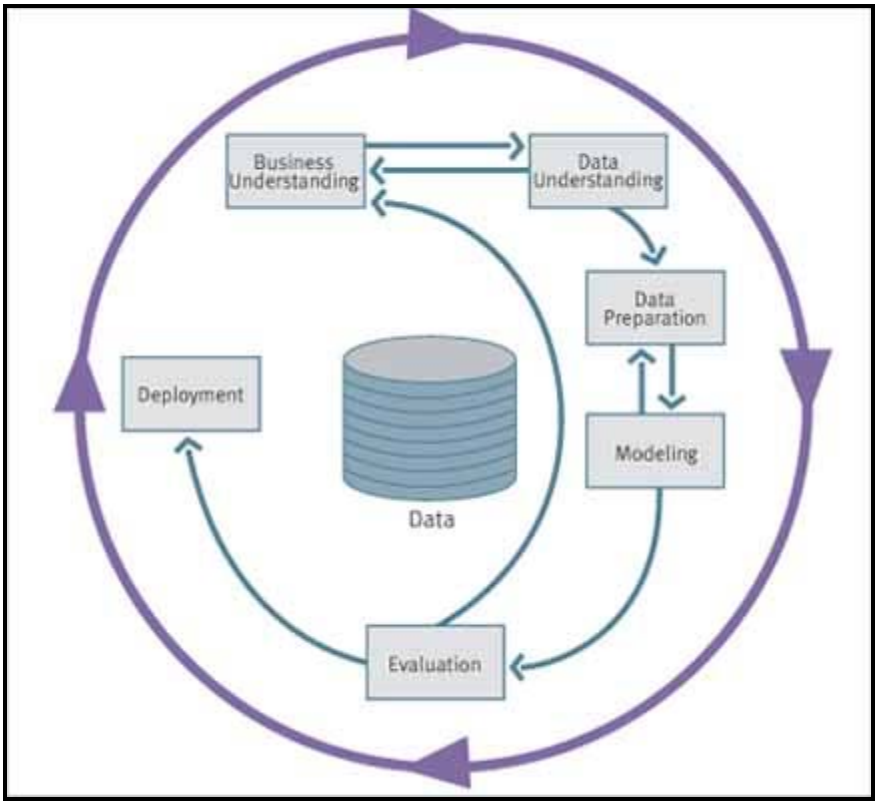


Figure 1: Phases of the CRISP-DM Reference Model (CRISP-DM, 2000)

When referring to the medication use process, the researcher used the United States Pharmacopeia's (USP) definition which begins with the prescriber evaluating the patient and ends with monitoring the patient's response (Pharmacopeia, 2004). The initial desire was to capture data from all sources supporting the medication use process as part of this study. However, due to data access constraints required by the hospital, the scope of this research effort

concentrated on the medication administration process exclusively. Medication administration processes may vary slightly between healthcare organizations. Therefore, the medication administration process for the purposes of this study begins at the point of dispensing and ends with the nurse monitoring the patient's response. The point of dispensing could be from the hospital pharmacy or from an automated medication dispensing system located in the care unit.

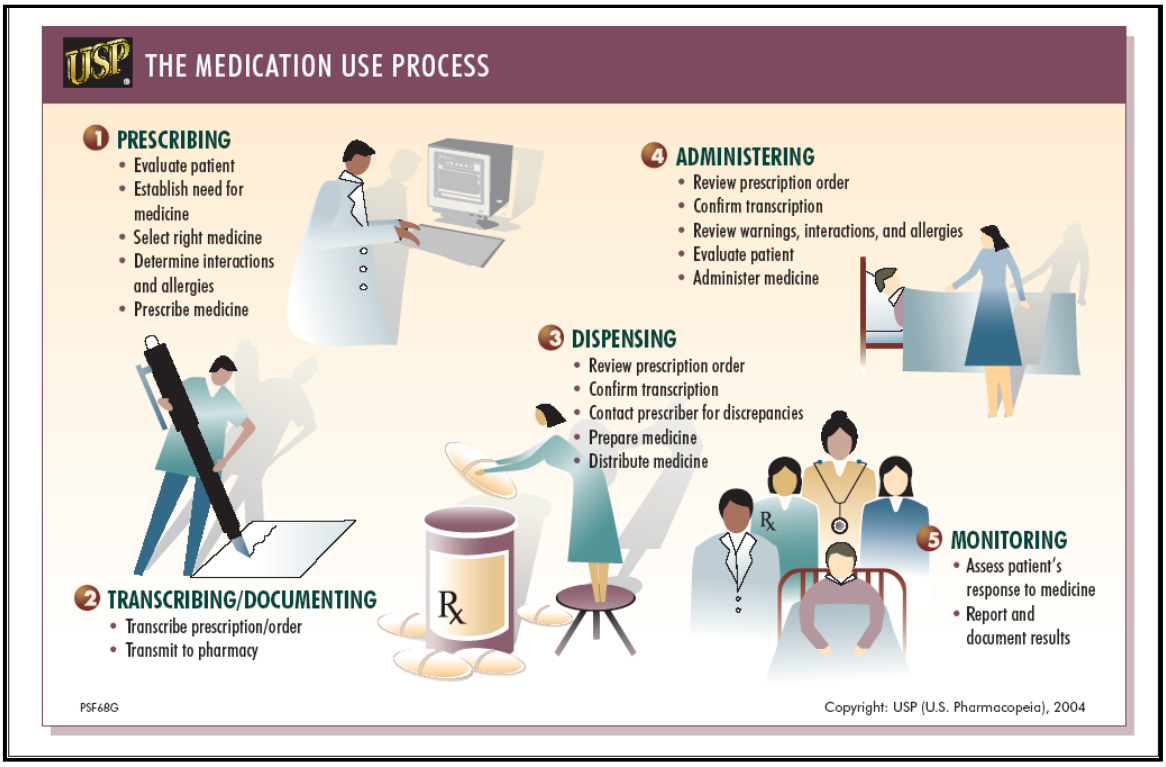


Figure 2: U.S. Pharmacopeia – The Medication Use Process (Pharmacopeia, 2004)

The method used to understand the Medical Center's medication administration process consisted of four primary activities. The first activity was to review the Medical Center's internal policy and procedure manuals for the medication administration process. This review provided an understanding of the intended safe medication administration steps and how they should be performed. Part of the documentation review included a familiarization with the

Medical Center's patient safety program guidelines and its patient safety improvement methodology. An awareness and understanding of these written protocols provided a basis for communicating with clinicians and administrators regarding the patient safety environment and the perceived effectiveness of its mechanisms.

Conducting individual interviews with members of the Medical Center who work within or frequently interact with the medication administration process provided essential perspectives about the environment where systems, clinicians, pharmaceuticals, and the patient converge. The first round of interviews involved hospital administrators. Meeting with the Directors of Nursing, Pharmacy, and Information Technology provided the necessary administrative and logistical support to complete this research. With their authorization and approval the researcher was granted access to the facility, permitted to interview staff, and gather data from the Medical Center computer systems. The Medical Center administrators also shared their approach to and philosophy of patient safety improvement and how they managed the patient safety culture within their organizations.

Meeting with the clinical leadership was important to understanding the patient safety infrastructure supporting the medication administration process. However, the most insightful interviews detailing day-to-day operational activities of the medication administration process came from discussions with nurses, pharmacists, pharmacy technicians, and nurse trainers. Conversations with the Medical Center staff representing these roles proved invaluable for obtaining undocumented information regarding the medication administration process. The staff level interviews illustrated the need to conduct detailed job shadowing of nurses and pharmacy technicians. Job shadowing for nurses began with the researcher participating in the Medical Center's nurse employee orientation and concluded with observing a new nurse and her

trainer/mentor performing various patient care and administrative tasks in the Medical/Surgical nursing unit. These interviews and daily observations demonstrated that the medication administration process is complex and requires reliance on a mutual understanding of how the process steps are executed. However, the documented medication procedure was not sufficiently reflected by the observed operational reality within the nursing units. Therefore, an effort to revise and extend the medication administration process documentation was initiated.

Each nursing unit at the Medical Center has a nurse that serves as the Patient Safety Officer in addition to their normal clinical duties. Patient Safety Officers play a key role in creating, implementing, and monitoring patient safety initiatives for their nursing unit. Consequently, the researcher conducted focus groups with Patient Safety Officers to receive their observations and experiences with the medication administration process and to discover opportunities for improving patient safety. Two group sessions each one hour in length were facilitated over two consecutive days with 15 Patient Safety Officers participating in each session. The most valuable output from the focus group sessions was a complete process map for how medication administration is actually being performed. The process map exposed the inconsistencies in the medication administration procedure manual. The new process map was the foundation for updating the medication administration reference manual and was made available to the nurse orientation and preceptor training programs. The "map" was subsequently used during the data mining analysis phase to associate results to the part of the process where the error conditions began. The process map created by the Patient Safety Officers is illustrated in figures 3 and 4.

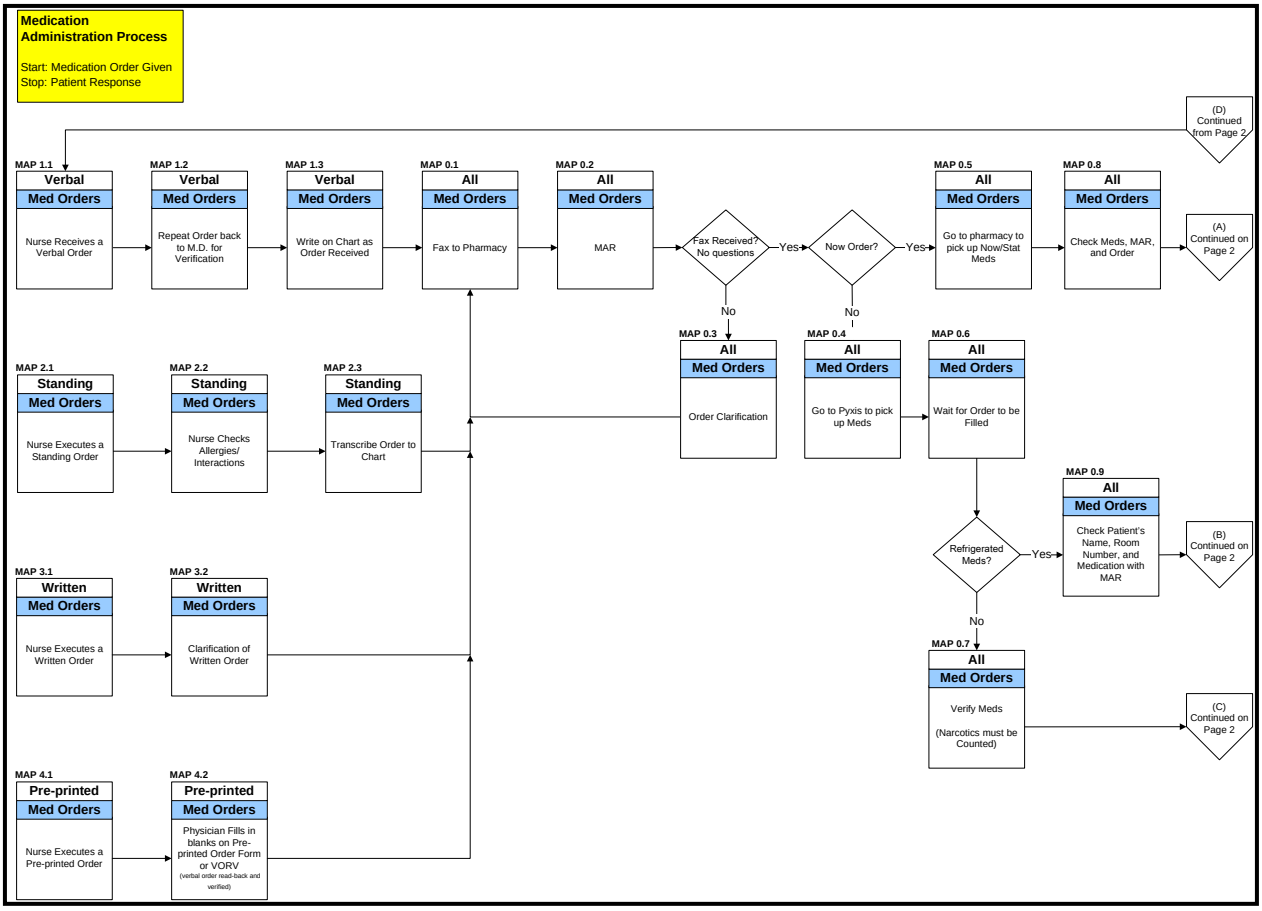


Figure 3: Medical Center's Medication Administration Process Map – Page 1

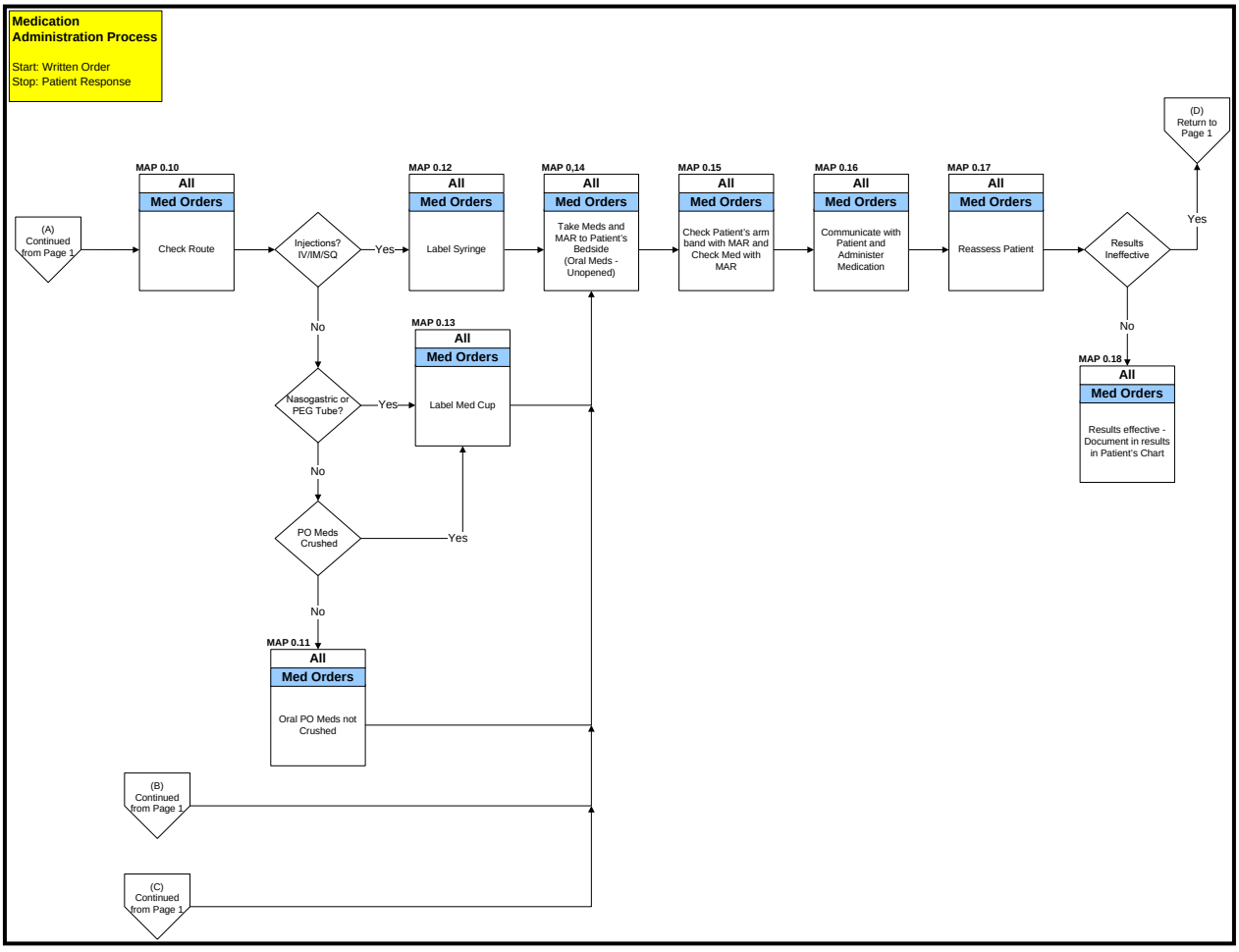


Figure 4: Medical Center's Medication Administration Process Map – Page 2

With an accurate process map the researcher was able to understand how a nurse functions within the medication administration process. The activity of producing the process map exposed other potential data sources that could further characterize the environment and circumstances where medication errors occur.

Finally, the last activity used to gain understanding of the medication administration environment was involvement in the patient safety improvement efforts initiated when an error occurred. Most often this activity consisted of participating in the Medical Center's medication

error review board meetings. In these monthly meetings the board members review a summary of the medication errors reported and discuss findings from error investigations initiated in prior meetings. All reported medication errors are evaluated for "system errors" to determine whether future errors can be prevented through system or process changes. The board determines whether an error requires Root Cause Analysis (RCA) or Failure Modes and Effects Analysis (FMEA) to identify the process breakdowns which allowed the error to occur. Additionally, the review board discusses proposed changes to the medication administration surveillance programs or considers adopting additional nurse education programs that emphasize practical solutions for protecting patient safety.

Exposure to the post-error review and corrective action planning process advanced this research by providing information about how the data collected in the incident reporting system was used by pharmacists and nurses to make improvements to their part of the medication use process. Most importantly, it highlighted the fact that the Medical Center was limited to using administrative interventions for corrective measures to improve the medication use process. The Medical Center rarely used quantitative methods to perform system and process evaluation, redesign, and outcomes measurement as part of their patient safety improvement efforts. Although the capability for more sophisticated analysis and improvement measuring techniques were desired, the Medical Center acknowledged that they did not have access to reliable data in a meaningful format to make informed decisions at the point of care. The result is a patient safety improvement process that relies disproportionately on the professionalism and vigilance of clinicians to compensate for system and process inadequacies with latent error potential.

Chapter 3

Data Gathering and Data Analysis

The data-gathering methodology for this research project is one that focused on accumulating all available data corresponding to a medication administration error event. As originally conceived, this research attempted to collect and aggregate all data elements stored in the Medical Center's pharmacy management system, automated dispensing machines' database, medical records, patient billing records, patient census data, nurse staffing data, and nurse credential data.

Typically only the data elements recorded in the Medical Center's incident reporting system were used to identify medication administration error causality (See Appendix 1). Besides wanting to consider all available data surrounding the error event, this research expands the analysis window of time to test if other potential contributing factors were present in the days prior to an error event occurring. To accomplish this goal, the researcher attempted to accumulate data associated with the entire length of stay for a patient involved in a medication administration error. At the time of this research the Medical Center did not use an enterprise-wide database integration structure, commonly referred to as a Data Warehouse. Therefore, data had to be acquired from individual databases for each transactional system supporting the medication administration process.

A high-level summary of the automated systems and paper based records requested for the purpose of assembling a comprehensive view of the medication error environment at the Medical Center included the following:

- Medical Incident Report Data – self reported
- Nurse Staffing Data – hours, shift, patient nurse ratio, staff mix
- Nurse Credentials – education, certification, length of service, time in profession
- Patient Medical Record Data – medical chart
- Automated Drug Dispensing Machine Data – Pyxis MedStation 2000
- Other – Patient Census, Admissions, Discharges, Transfers

A conceptual illustration of the technical environment from which the data was gathered is depicted in figure 5.

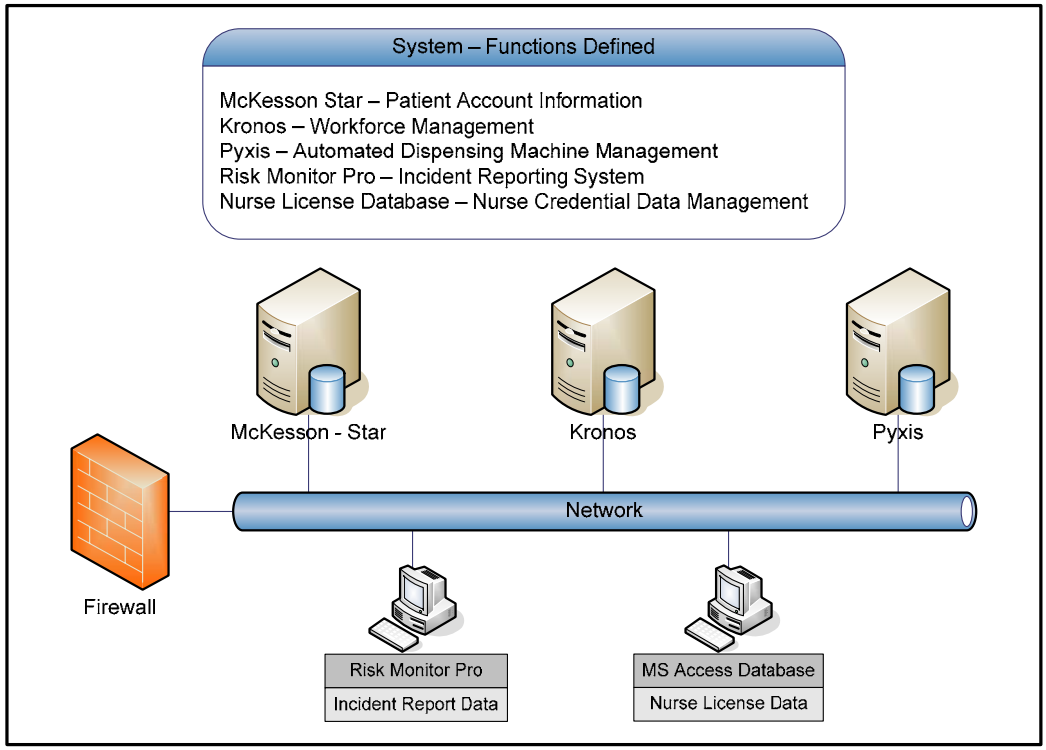


Figure 5: Conceptual representation of data sources

The first step in the data gathering phase was to determine the key data source and its attributes that would uniquely identify and isolate an individual medication error event. This data source was provided by the data stored in the Medical Center's incident reporting system, Risk Monitor Pro, which is a commercially available software system licensed by rL Solutions. This hosted internet browser based incident reporting system is used by the Medical Center to record medication errors, IV infiltrations, and blood transfusion errors. For the purposes of this research, only medication errors were considered.

During the initial data gathering process a software version upgrade for Risk Monitor Pro was completed at the beginning of July, 2005. Therefore, to eliminate the risk of data integrity issues associated with pre-upgrade data and post-upgrade data, the researcher deferred the data extraction from the Risk Monitor Pro until the software upgrade was completed and a successful data integrity check was confirmed. The data extract from the Risk Monitor Pro covered all reported medication errors from January, 2003 through June, 2005 – a 30 month incident report history. The data from Risk Monitor Pro were exported to a comma delimited file, and then were uploaded into a Microsoft Excel Spreadsheet for data organization and preliminary analysis. The initial data were sorted by individual Medical Center departments. Every department at the Medical Center used this system to record medical error incidents. Therefore to determine if error reporting trends were associated with a particular department, it was important to preserve the department identifier and find other department level data sources to associate with the incident report data. Risk Monitor Pro has many fields to capture information associated with the error event. Some of the fields are mandatory, such as patient identifying information (i.e. patient medical record number) while others are optional. For example "Actions" is an optional field which is used by the person creating the incident report to specify

the action taken in response to an error. In the case of "Actions" the reporter may choose an action from a list of predefined actions. If the reporter does not choose to change an optional field in Risk Monitor Pro the default response "Not Specified" will be recorded. In addition to mandatory and optional fields, the last field in the data output was a free form text field where the person creating the incident report described the error event in narrative form.

Once the incident reporting data were organized and reviewed for data continuity/consistency, the construction of the initial base dataset began. The "base dataset" is the organized collection of data directly linking a patient to a uniquely identified medication error report. The first version of the base dataset was an extract from the Risk Monitor Pro incident reporting system for the previously mentioned 30 month period.

The unique identifier for a row of data in the base dataset constructed from the Risk Monitor Pro incident reporting system was the "Incident ID" field. The "Incident ID" is a number automatically generated by Risk Monitor Pro when a new incident report is completed. The Medical Center uses a unique six digit medical record number to identify a patient within many of its clinical support and business support systems. The medical record number is required to complete an incident report in the Risk Monitor Pro system. Therefore, the base dataset uses the medical record number to identify patients. However, in other systems the patient is uniquely identified by a nine digit account number. In order to retrieve the admission date and the discharge date for a patient involved in a medication error, from the Medical Center's McKesson Star system, a translation from the medical record number to its corresponding patient account number must be performed. With the patient account numbers, a query was made using McKesson TrendStar application to return the admission date and discharge date for a patient involved in a medication error. The patient's length of stay in

number of days was derived from the admission date and the discharge date. The length of stay variable was added to the base dataset along with the patient's admission date and discharge date to determine if reported medication errors were correlated to the length of time that the patient was in the hospital. Additionally, TrendStar was queried for the total number of admissions, discharges, and transfers that each nursing unit experienced during the date span of this research so that a measure of department activity or volatility could be assessed as a contributing factor for error. These data were added to the base dataset and the first round of data cleansing began. The base dataset was cleansed to remove incident report identification anomalies. For instance, some rows in the base dataset contained erroneous patient medical record numbers. This prevented correlation between the patient's medical record number and their account number, resulting in an inability to correctly add the admission data and discharge date for that patient. Although this situation was rare, it required eliminating the incident report from the base dataset. After this round of data cleansing the base dataset which totaled 512 incident reports was imported into Microsoft® SQL Server™ 2005 becoming "Base Dataset A". (See Appendix 1)

During the data gathering phase the Medical Center was implementing the Cerner Millennium, which is a healthcare information technology computing platform with multiple integrated systems that support almost every aspect of a healthcare organization's functions. However the implementation did not directly impact the medication administration process. The Emergency Department was the first clinical area to be supported by the Cerner Millennium information system. The Emergency Department was converted to the new system and was in active daily use at the end July 2005. Theoretically the Emergency Department could have reported more errors in the months after the implementation was completed. Therefore, the data

collection phase for this research was scheduled to avoid any conflicts with the enterprise implementation. The last month of data considered for data mining analysis was July 2005.

The Cerner implementation consumed most of the Medical Center's IT staff resources. Therefore all data gathering was conducted by the researcher which required querying hospital IT systems to retrieve and extract data associated with the patients, nurses, and medications related to an acknowledged error. Although the data gathering process was protracted with limited direct support from the hospital IT staff and system administrators, this circumstance enabled the data to be in the direct control of the researcher. This degree of control provided a level of confidence in the data's authenticity and integrity.

In addition to considering the Cerner implementation, the data gathering process considered other IT system replacement lifecycles. For example, the Medical Center converted to a new time reporting and resource management system in June of 2004. As of this date all employee time and scheduling data were being generated and tracked in the Kronos system. The old time reporting and resource management system was decommissioned and no longer supported by the IT organization. Therefore, retrieving archived data from the old time management database and correlating it to the Base Dataset A was not feasible. The research team made the decision to only use the data available in the Kronos system, thus reducing the date span of the research data to 13 months – June 2004 through July 2005.

From the Kronos system, the nursing hours by shift for each nursing unit/department on any specified day could be obtained. The Kronos system was linked to the hospital's time clocks, which is the official record for accumulating and calculating the hours that a nurse has worked for payroll purposes. In order to extract all nursing hours charged during a day to include nurses who were "floating" between departments, the "job codes" that included all

nurses were included into the query request. For example, if a nurse clocked into Kronos system to work first shift on the Medical/Surgical unit and at the end of that shift, the nurse clocked in as a "floater" to into another department to work part of the 2nd shift, Kronos recorded that nurse's hours to the respective department by shift. Data extracted from Kronos allowed the researchers to expand the dataset to include the hours an individual nurse worked by shift for the entire length of stay of a patient that experienced a medication administration error. This information was important because it provided the ability to calculate nursing hours of care per shift for any day during the 13 month span of the research study. It also indicated which nurses worked during the days prior to a medication error being reported.

The Medical Center participates in the American Nursing Association's National Database of Nursing Quality Indicators (NDNQI) program. This program provides the Medical Center with benchmarking reports related to key hospital metrics aggregated from all participating member hospitals. According to independent NDNQI reports the Medical Center scores better than its peers (based on number of beds) for NDNQI measurement parameters. Of these measurement parameters reported to NDQI the metric of relevance for this research is the number of nursing hours per patient day. In other words, the Medical Center manages their nursing hours per patient day metric through their resource staffing plans and patient census forecasts to maintain an above peer group rating. As an independent measure for this research, nursing hours to patient ratio by day for each department was derived from Kronos data and patient census data from TrendStar and added to the base dataset. Also added to the dataset using the same sources were nursing hours by shift, total nursing hours of care, the number of Registered Nurses (RN), the average daily census, and nurse-to-patient ratio. After the Kronos data and TrendStar data were extracted and cleansed, these data were appended to the current

version of the base dataset within the Microsoft® SQL Server™ 2005 database creating "Base Dataset B" which was reduced to 182 incident reports. (See Appendix 1)

At the time of this study, the Medical Center was not using an Electronic Medical Record (EMR). The medical charts were paper based and were scanned as a static electronic image after the patient was discharged. Historical medical records were imaged using the "Document Acquisition and Storage Suite" by SoftMed Systems Inc. Medical diagnosis was an important data element to add to base dataset. The reason for including medical diagnosis as a data element was to verify whether a patient's therapeutic complexity as indicated by their diagnosis was a determining factor in the error environment. Since the researcher is not a clinician, it was not feasible to accurately and efficiently record relevant patient diagnosis information manually from medical charts. Therefore, chart reviews were not conducted as part of this study. An alternate approach to obtain patient medical diagnosis information from the Medical Center's billing system was proposed. Unfortunately, researcher access to the billing system was not granted.

Nevertheless, system access to the database storing all of the medication dispensing data from the Medical Center's automated dispensing machines was authorized. At the time of this research, the Medical Center used Pyxis MedStation 2000s for their automated medication dispensing machines. These machines were located in the departments as close to the nurses' station as the floor plan permitted. All medications for a patient that experienced a medication error as reported in the incident reporting system for the entire length of stay was obtained by querying the Pyxis database. This data enabled the dataset to consider medications that were given before, during, and after the day that the error was reported to have occurred. Also the Pyxis data indicates which nurse retrieved the medication for the patient involved in an error.

Therefore, the opportunity for finding patterns relating to the medications administered and the nurses who administered them prior to or on the day of the error exists within the Pyxis data retrieved. Once the Pyxis data were secured and evaluated for continuity and integrity, it was combined with other data elements to form the final base dataset prior to data mining and titled "View B" in the Microsoft® SQL Server™ 2005 database. With the Pyxis data added View B consisted of 21,696 rows. (Appendix 1)

Data Cleansing

The data cleansing phase of this research was by far the most time-consuming and labor-intensive part of the process. Initially data cleansing was an iterative process. When a series of data was extracted from a database, these data were evaluated for missing or erroneous data elements prior to being joined with the base dataset. However, when data gathering was complete and all of the incremental components of data were added to the base dataset, one final round of data evaluation was performed prior to data mining. The time spent on this final round of data cleansing was essential for truly understanding the data and its limitations. This final scan of the dataset was the basis for selecting the appropriate data mining algorithm used for the data analysis phase. The arduous task of data cleansing also allowed the researcher to become intimate with the apparent trends within the data and notice indicators which suggest how the data was entered into the source system. For example, human inspection and observation of the incident reporting system data indicated system features that allowed data entry to be performed in a certain way. Observed within the incident reporting system data were clues to human behavior or human-system interaction factors for how the error reporter used the system. Also observation of the data in the incident reporting system led to assertions relating to a difference between the medication error reporting policies and actual reporting practices.

Data Analysis

Before discussing how the base dataset was evaluated to determine the appropriate data mining technique used for analysis, it is useful to summarize View B's structural configuration. Up to this point the result of data gathering, aggregation, and cleansing yielded a large table of data consisting of rows and columns. Each row in the View B database table was an instance of all data gathered that directly relates to a specific medication error incident report. The columns of the View B table constitute a defined set of features or attributes that characterize the medication error environment.

Clustering techniques are used when class prediction is not possible or not desired and the intent is to place the instances into natural groups. "These clusters presumably reflect some mechanism at work in the domain from which instances are drawn, a mechanism that causes some instances to bear a stronger resemblance to one another than they do to the remaining instances."(Ian H. Witten, 2000) Probabilistic clustering was chosen as the data mining method because the mechanisms within the data that influence the clustering for the base dataset were unknown. The decision to use probabilistic clustering was partially based upon the clustering tools available to the researcher.

The tool used to analyze the View B table was Microsoft® SQL Server™ 2005 Analysis Services. Other proprietary clustering algorithms were pilot tested with the base dataset prior to final selection of the data mining tool (Gilbert, 2006). However, given the number of instances and attributes to mine and the desire to quickly add additional database tables and join them for subsequent rounds of data mining analysis, Microsoft® SQL Server™ 2005 Analysis Services (SASS) was capable of accommodating current and future inquires. Also, knowing the results from this exploratory research would ultimately be presented to Medical Center clinicians who

may not be familiar with data mining or database functions. The choice of using a tool that graphically displayed output results was also an important factor in the decision to use the SASS data mining tool.

The clustering algorithm used in Microsoft® SQL Server™ 2005 Analysis Services is based upon the K-means algorithm. This algorithm clusters probabilistically rather than categorically. With probabilistic clusters every instance is evaluated and assigned to each cluster with a degree of certainty or probability.

Using Microsoft® Business Intelligence Development Studio, a data mining project was created using the View B table stored in the MS SQL 2005 database. The mining structure illustrated in Appendix 2 shows that all attributes or columns within View B were used as the source inputs for building the data mining models to analyze the medication error related data. Three successive data mining models were created, each one using the clustering algorithm native to SASS. When evaluating the clusters generated from the data mining models, a 50% probability threshold for considering relevance for cluster membership was used. Therefore if an attribute value was calculated to have a 49% probability of inclusion in Cluster 1 and a 51% inclusion probability for Cluster 2, the attribute value would be considered to be a member of Cluster 2 and not a member of Cluster 1.

Mining model 0 was specified to cluster with the "Incident ID" field as the key, so that clustering with respect to the reported medication errors would be the desired output. All other attributes in the View B table were used as inputs for cluster development and for prediction of cluster membership. The clustering results from Model 0 are detailed in Appendix 3. Data mining Model 0 helped identify attributes that were not contributing significant information about the reported medication error instances and were dominating cluster creation. For

example, the attribute “InjuryBodyPart” contained the value - “Not Specified” for all instances in the mined dataset. “Not Specified” is the default value that Medical Center’s incident reporting system records for data fields that are not completed by the person creating an incident report. Therefore many of the optional fields from the Risk Monitor Pro incident reporting system have recorded default values because they were not changed during the incident report creation process by the person reporting the medication error. Additionally, analysis of Model 0 clusters helped identified redundant or duplicative attributes. One example is the variations of the date and time information recorded when an incident report is created in the Risk Monitor Pro incident reporting system – “Entered_Date”, “Entered_Time”, and “Reported_By_Date”, “Reported_By_Time.” Many of these date and time stamp fields are populated using the system date and time from Risk Monitor Pro when an incident report is generated. From a clustering perspective, duplicative date and time information is not necessary.

The results of the Model 0 clustering suggest that attributes which do not provide significant information about medication error incidents and attributes representing redundant information disproportionally influencing the output. Therefore, these attributes were excluded from subsequent mining models. Mining Model 1was created from Model 0 with some of the attributes excluded. The input attributes and the predictive attributes were changed for Model 1 as well. A complete inventory of attributes included for constructing Model 1 is listed in Appendix 4.

Identical to Model 0, mining Model 1 used the “Incident ID” attribute as the key field and generated 10 clusters. Using 50% probability as the lower bound for considering an attribute to be a member of a cluster was not an adequate measure for Model 1 cluster analysis. Many attributes had high probability of membership in all clusters. For example, the attribute

"Factor28MAR Unclear" with a value of "No", had a probability of cluster membership in all 10 clusters ranging from 98%-100%. Attribute "Factor14MAR Misinterpretation" with a value of "No" was 75%-100% likely to be associated with all clusters. Clearly the output clusters from running Model 1 on the dataset were clustering on what was not reported as part of medication error incident report. Clustering continued to be dominated by default values from the attributes originating from the incident reporting system in addition to "null" or "missing" values from other attributes.

Data mining Model 2 benefited from iterations of the two previous models. The "Incident ID" attribute remained the key field for Model 2 clustering. The attribute list used for input and prediction for clustering was reduced in response to the observations of the first two clustering attempts. The definitive list of attributes incorporated in the third mining model can be found in Appendix 5. The cluster diagram resulting from running Model 2 on the View B table is represented in figure 6.

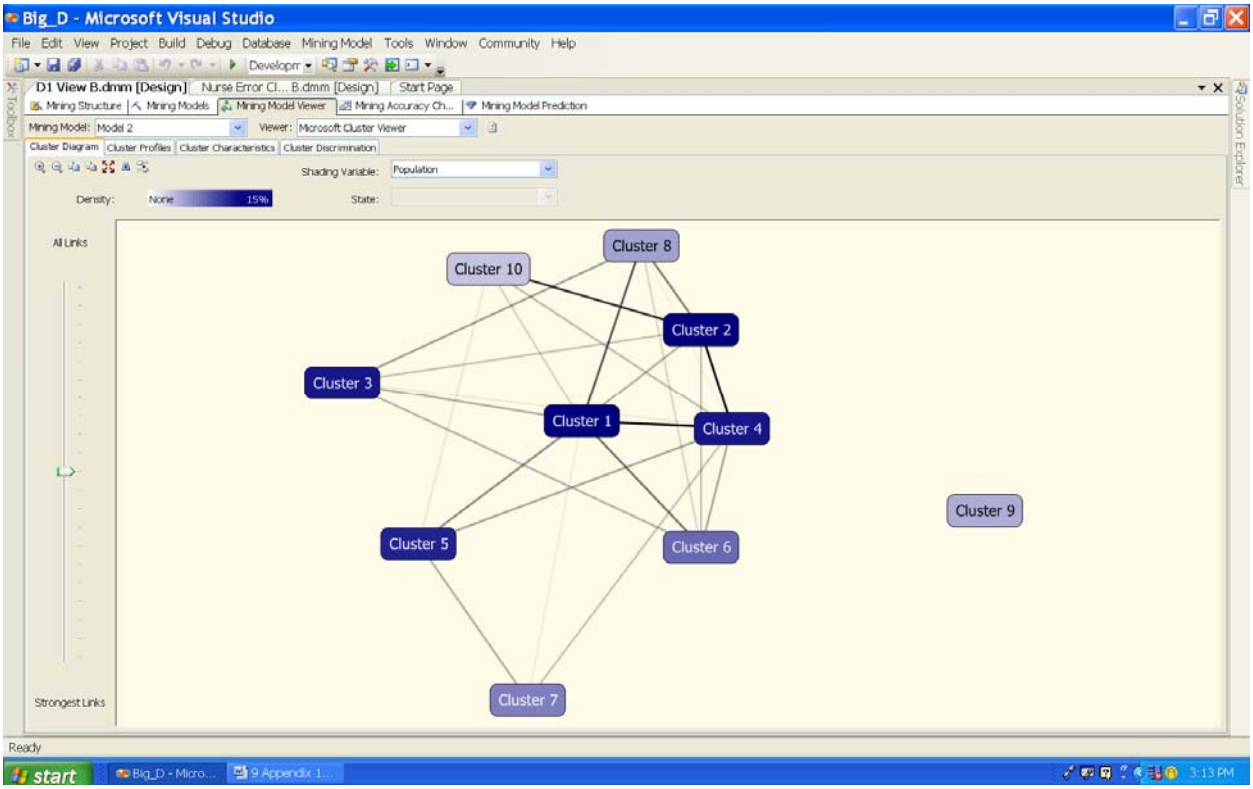


Figure 6: Clustering Diagram – Model 2

Model 2 generated 10 clusters. The individual cluster characteristics and the results for the population of data generated from Mining Model 2 are illustrated in Appendix 5. A review of the clustering results clearly demonstrated that the data mining Model 3 confirms the trend of the previous two clustering trials. Clustering continued to highlight the obvious pattern of attributes with values that reflect incomplete incident reports. Each of the 10 clusters has high probability attribute values equaling "Null" or "Missing." Other high probability attribute values across all clusters contain the default responses from the incident report system. Obviously the default, null or missing values provide no insight into the circumstances surrounding the reported medication error. From the 147 instances Model 2 used to generate clusters some summary information can be learned.

Although Model 2 returned clustering results that did not provide conclusive information about the medication error environment beyond the contents of the incident report, the relationship between the nurse and the types of incidents reported remained to be evaluated. Therefore the process of linking the nurse credential data with the data from View B was initiated. The first step in the process of creating a new clustering model testing the nurse-error relationship was accomplished by querying the View B table for all nurses that administered medications to a patient involved in the error for the day that the error occurred. Since the time of the error was not accurately recorded in the incident reporting system and could not be obtained from alternate sources, the error could not be isolated to a particular shift. Therefore the query results from View B returned all 1st, 2nd and 3rd shift nurses who administered medications to the patient on the day of the error. The query results were imported into a new database table and joined with the nurse credential data to create the "Nurse Error Close Association B" table containing 365 rows. (See Appendix 6) The nurse credential data included such attributes as the nurse's name, sex, hire date, birth date, date of licensure, the assigned department, and highest degree earned. A complete listing of the nurse credential attributes are outlined in Nurse Error Association Model in Appendix 6. The mining model denotes which attributes were used as input and predictive clustering variables with "Incident ID" as the key field. The cluster diagram resulting from running Nurse Error Close Association Model 2 is represented in figure 7.

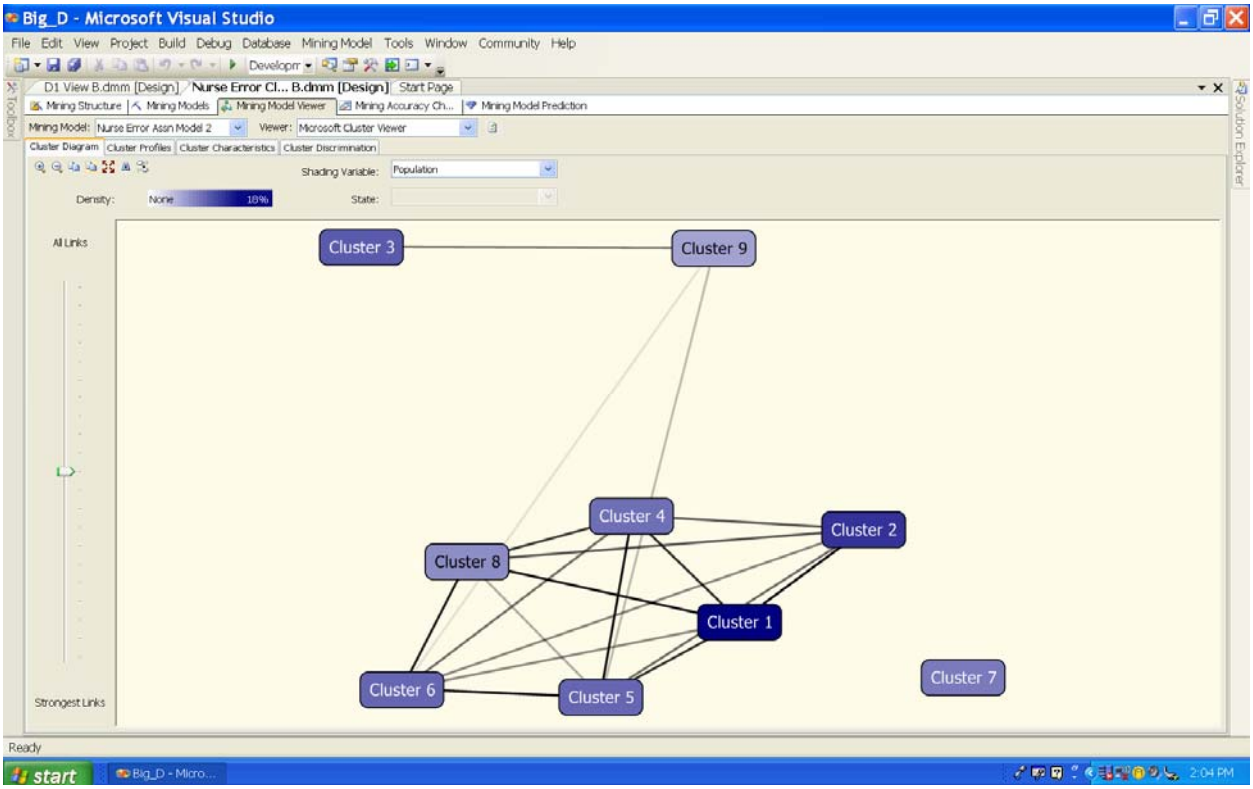


Figure 7: Clustering Diagram – Nurse Error Close Association Model 2

Of the 9 clusters that were generated, no cluster indicated a relationship with a set of Incident IDs. Therefore the results of this mining model does not demonstrate a connection with a nurse's demographic/credential attributes and the reported medication administration error. In essence this analysis could not determine if nurse with less tenure either in the nursing profession or at the Medical Center is more likely to commit an error than a nurse with more experience. Interpreting the mining model further reveals the absence of any relationship between the nursing department and the reported administration error. Values for the "Nurse Dept" attribute scored 50.4% probability for inclusion in cluster 3 and 65.19% probability for cluster 11 which suggest the department attribute has some relationship to the other dominant attribute values of those clusters. Cluster 3 contains Registered Nurses (RN) assigned to Unit Surgical/Medical

ICU with a Bachelors of Nursing degree as their highest degree attained. Cluster 11 contains RNs with Associates of Nursing degrees as their highest degree earned working in Medical/Surgical unit. The cluster 3 and cluster 11 examples typify the clustering behavior of the entire Nurse Error Close Association Model. The model clustered well on the nurse demographic and credential data, but with the absence of any association to Incident IDs the model has limited use for answering questions about relationships between error and nurse characteristics.

Chapter 4

Results and Conclusions

From data aggregation and analysis it was not possible to associate an individual nurse with an error event. It was possible however, to isolate the nurses assigned to a patient on the day the error was reported to have occurred. Therefore, it is highly unlikely that anyone could identify the participant in a reported medication error using the data sources obtained in this research study. In other words, if a nurse who committed a medication administration error used a proxy such as the nursing supervisor, charge nurse, or patient safety officer to report the error event in the incident reporting system, this level of anonymity is preserved. However, it is possible that other notes and informal records could exist elsewhere that identified the individual nurse who committed the error. For example, a nurse who committed an error may choose to document the situation via handwritten or electronic format such as in an email and submit the error information to the nurse supervisor or patient safety officer in the unit for entry into the incident reporting system at a later time. Although it may not be a consistent practice for all departments, interviews with Medical/Surgical nurses revealed that they regularly record more incident related information in the patient's chart than is actually logged into the Risk Monitor Pro incident reporting system.

In general, the Medical Center's nurses related that wrong time errors are easy to make during the 9 am medication distribution time. The Medical Center technically considers wrong time medication administration errors to be any medication that was received by the patient 30

minutes before the scheduled time as per the Medication Administration Record (MAR) or 30 minutes after the scheduled time. In practice, however wrong time errors are infrequently captured as an error in the incident reporting system. The incident reporting system's interface design and general usability features may be a factor contributing to incomplete incident reports. Reporting behavior appears to be influenced by organizational and social factors with underreporting of wrong time errors being an example. However, few wrong time errors appear to pose a significant risk to the patient. In fact, some wrong time errors could be an unavoidable by-product of good clinical judgment made by the nurse. For example, consider nauseated patients who are unable to keep anything in their stomach. In this case, administering an oral solid medication before the nausea passes will provide no therapeutic benefit. Is this a wrong time error or an example of good clinical judgment? Thus organizationally supported and socially accepted justifications for wrong time errors at the Medical Center appear related to reporting behavior with respect to wrong time errors. Therefore, it is logical to assume that other medication error reporting norms beyond wrong time errors exist which further limits the use of incident reporting as the basis for broad safety analysis investigations. Medical Center nurses appeared convinced that IV medications, which include IV fluids, is the dose form that experiences more errors than any other dose form. However the data clustering results do not support this assertion, nor does a manual inspection of the incident report data. This suggests that though the nurses may know they are making more errors with IV medications, they are not reporting these errors in the incident reporting system more than any other dose form. Perhaps the contrast between what is "thought" to be a prevalent medication error condition and the actual reporting behavior associated with it is another indicator that more needs to be understood about the barriers to reporting and underlying patient safety culture at the Medical Center.

The benefits of integrated systems or the consolidation of key data from independent systems into a data warehouse cannot be overstated. Although the Medical Center generates and stores large volumes of data related to the medication administration process, the data resides in separate databases. This architecture does not allow data to be extracted, aggregated, and analyzed in a way that is useful to the day-to-day process of improving patient care. For example, the pharmacy will not know if the Medical/Surgical unit's current nurse-to-patient ratio is below the hospital's acceptable range, meaning that nurses could be experiencing a high task load situation where more human errors are likely to occur. However the pharmacy will notice that it is dispensing a greater number of high alert medications to patients in the Medical/Surgical unit. Theoretically, if the pharmacy had both pieces of information, they could make the nursing supervisor aware of the situation so that action could be taken to emphasize patient safety vigilance during that shift and make staffing changes for subsequent shifts.

This research demonstrated the difficulty of data gathering and data cleansing in preparation for data mining analysis when multiple, disparate systems are involved. Hospitals do not have the time or resources to dedicate to data mining analyses using non-integrated data sources. Underfunded hospitals struggle to advance their IT infrastructure and must prioritize their IT investments. In this situation, the hospital will need to utilize its incumbent systems and data structures along with any new technology such as Computerize Prescriber Order Entry (CPOE), Electronic Medical Records (EMR), or Bar coding systems to make full use of its data sources in evaluating/monitoring the medication administration process. To leverage and extend the decision support capabilities imbedded or overlaying these new systems the hospital should assess the value of its data assets across the enterprise and prioritize which data elements need to be available in a data warehouse for near real-time situational awareness. With a data warehouse

receiving data from multiple systems, online application processing (OLAP) routines can push known patient safety management parameters to decision makers such that if the conditions for an error are identified, timely action can be taken.

Conclusions

The Medical Center's incident reporting system is used by a limited number of nurses. Of all the entries recorded in the incident reporting system not a single incident report was submitted by a physician or a pharmacist. Most of the incident reports were submitted by nursing supervisors, the on-shift charge nurse, or nursing unit patient safety officers. The fact that more nurses are not submitting incident reports suggests that barriers exist to reporting medication administration errors at the Medical Center. Inspection of the incident report output data indicates most incident reports are submitted with only the required data fields completed. In particular, the "Actions" field, and "Factors" fields are often recording the default value "Not Specified". The "Actions" field is used to annotate the action that was taken in response to the medication error. One example on the list of selectable items for this field is "Treatment Provided" and another example is "Physician Notified". In total, there are 18 different actions to choose from. The "Factors" field has 29 different selections to use to explain the contributing factors to the medication error, but they are rarely used. Yet, the utility of the incident report information for improving medication administration process and reducing the potential for error depends on this information being recorded in as much detail and as close in time to the error event as possible. From the available data, it is not conclusive that the incident reporting systems user interface design is a barrier to completing the report effectively. The time required to complete an incident report using the web based system is unknown, but if the time required to complete an incident report is perceived to be too long, the user will avoid using the system.

This will equate to fewer incident reports being submitted or the user will complete only the required fields necessary for creating a record. This will result in reports that are of limited value for follow up investigation or error trend analysis. Although a human computer interaction analysis was not in the scope of this research project, clearly such a study would be beneficial. By all indications the incident reporting system is not being used in accordance with its design capabilities. Studying the system's usability in terms of the user interface could determine if the system is a barrier to error reporting at the Medical Center. Such an analysis would have benefits beyond this specific Medical Center since the web based incident reporting system is a commercially available system used at more than 600 hospitals and clinics worldwide (rL soultions, 2009). In the event that the user interface design study determines that the interface and system functionality is not a barrier to incident reporting, it is recommended that other research studies assessing the patient safety culture at the Medical Center may be useful. For example, an online patient safety culture survey adapted from the "*Hospital Survey on Patient Safety Culture"* sponsored by the Quality Interagency Coordination Task Force and funded by the Agency for Healthcare Research and Quality(AHRQ), was conducted at the Medical Center during the data collection phase of this research. The survey was designed to measure overall perceptions of patient safety as well as ten dimensions of culture pertaining to patient safety (Stone, 2007). An extension of this survey could help assess the underlying cultural, organizational, or administrative barriers to acknowledging and reporting errors.

To summarize the results in terms of the proposed research questions, applying the clustering data mining technique to the Medical Center's medication administration error data did not to conclusively affirm that:

1) it is possible to find a subset of input attributes that differentiate medication administration errors that occur within a distinct step of the overall medication administration process (Research question 1 not affirmed).
2) it is possible to find a subset of input attributes that are associated with known medication administration error event types which are not documented in the incident reporting dataset (Research question 2 not affirmed).
3) it is possible to predict the occurrence of medication administration error types that are not currently being document in the incident reporting system (Research question 3 not affirmed).

Despite the inability to answer the proposed research questions, the results adequately address the first objective of this research which was "to determine if data mining can provide medication error situational awareness beyond what is currently provided from incident report analysis." This research demonstrated that data mining is a useful proactive tool for evaluating all available information surrounding medication administration errors. The insights obtained from data mining analysis point to distinct opportunities for improving the error reporting and error analysis procedures for the Medical Center. The second objective which was "to determine if the results from data mining analysis can identify areas in the medication administration process that require redesign to improve patient safety" remains unsatisfied. Although many process redesign suggestions can be made given the depth of analysis that this research effort achieved, such recommendations are based on the qualitative findings and not on the quantitative results of data mining.

The initial limitation of this research design was that the outcomes and results obtained could not be generalized for other healthcare organizations since this study was conducted using

the specific environment of one Medical Center. However the difficulty for conducting data mining analysis on data from various independent databases in a healthcare setting can be widely generalized.

Recommendations

At the core of this research is the desire to create a method for analyzing medication administration error data in the context of a healthcare organization's unique clinical processes. This research has the potential to deliver several clinical, cost savings, and technological benefits. Some of the direct and logically achievable benefits through the extension of this research are listed below. However, the most important outcome from this research will be a toolset which healthcare organizations can use to enhance their continuous improvement initiatives that directly impact patient safety.

Clinical Implications:

- This study demonstrated how data mining can be used to identify subtle trends, classify and describe the error event conditions beyond what is provided by incident reporting.
- Although not proven conclusively from the results of this data mining analysis, with improved descriptive data about the error environment, clustering algorithms should be able to predict attributes of unreported medication administration errors.
- This study provided a method for determining the areas within the medication administration process that require further analysis and provide the greatest opportunity for error reduction.
- This study supports clinical process redesign to improve patient safety.

Healthcare Cost Impact:

- The ability to identify sub-processes within the medication administration process that contribute the majority of the errors establishes a process improvement prioritization which reduces the cost associated with conducting a complete top-down or bottom-up analysis requiring direct observation of nurses during the entire medication administration process.
- Reduction in resource costs associated with the extra time a nurse spends caring for and monitoring a patient that has been involved in a medication administration error.

Technology Investment Justification:

- The ability to identify the key processes that influence the majority of medication administration errors gives IT decision makers an opportunity to enhance clinical support systems to assist the nurse's role in medication administration and provide a continuous error monitoring capability.
- The outcome of this research can also lead to the development of functional specifications for improvements in the current clinical support systems that assist nurses in the medication administration process.

Finally, the lessons learned from this research can contribute to research and the evaluation process for commercially available data warehouse/data mining solutions for healthcare organizations relying on multiple systems that are not fully interoperable. Data warehouses provide a single consistent point of access to organizational data, transcending departmental divisions. They are a place where old data is published in a way that can be used to inform business decisions. "The movement toward data warehousing is recognition of the fact that the fragmented information that an organization uses to support day-to-day operations at a

department level can have immense strategic value when brought together."(Ian H. Witten, 2000)

Proposed Publications

From this manuscript the researcher will create three separate articles for publication with an overarching purpose of sharing the methodology and process for using data mining to analyze clinical data obtained from disparate non-integrated IT systems. The first article will consist of a literature review covering the following topic areas: use of data mining to analyze medical error in the healthcare industry, medication administration error reporting systems, and methods for quantifying medication administration errors including the human factors that contribute to medication administration errors for nurses. The second article will be practitioner focused, describe this research method, and identify how it can assist nurses' efforts to use computation tools such as data mining within the context of their patient safety improvement programs. The third and final article will be intended for healthcare informatics professionals and discuss technology application in healthcare.

References

Agency for Healthcare Research and Quality. (2001, July). Making health care safer: A critical analysis of patient safety practices. *Evidence Report/Technology Assessment, No. 43. AHRQ Publication No. 01-E058* Retrieved July, 2009, from http://www.ahrq.gov/clinic/ptsafety/

Apté, C., & Weiss, S. (1997). Data mining with decision trees and decision rules. *Future Generation Computer Systems, 13*(2-3), 197-210.

Barker, K. N., Flynn, E. A., & Pepper, G. A. (2002). Observation method of detecting medication errors. *AJHP, 59*(23), 2314-2316.

Barker, K. N., Flynn, E. A., Pepper, G. A., Bates, D. W., & Mikeal, R. L. (2002). Medication errors observed in 36 health care facilities. *Archives Of Internal Medicine, 162*(16), 1897-1903.

Barker, K. N., & McConnell, W. E. (1962). The problems of detecting medication errors in hospitals. *AJHP, 19*.

Bates, D. W., Teich, J. M., Lee, J., Seger, D., Kuperman, G. J., Ma'Luf, N., et al. (1999). The impact of computerized physician order entry on medication error prevention. *Journal Of The American Medical Informatics Association, 6*(4), 313-321.

Blendon, R. J., Schoen, C., Donelan, K., Osborn, R., DesRoches, C. M., Scoles, K., et al. (2001). Physicians' views on quality of care: A five-country comparison. *Health Aff, 20*(3), 233-243.

Bruskin-Goldring Research. (1999, February). A study of medication errors and specimen collection errors., from http://www.bd.com/contentmanager/b_article.asp?Item_ID=21202&ContentType_ID=1&BusinessCode=20001&d=home&s=error&dTitle=&dc=&dcTitle=

Chaudhry, B. M. D., Wang, J. M. D., Wu, S. P., Maglione, M. M. P. P., Mojica, W. M. D., Roth, E. M. A., et al. Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *[Miscellaneous Article]*: Annals of Internal Medicine May 16, 2006;144(10):742-752.

Cullen, D. J., Bates, D. W., Small, S. D., Cooper, J. B., Nemeskal, A. R., & Leape, L. L. (1995). The incident reporting system does not detect adverse drug events: A problem for quality improvement. *Jt Comm J Qual Improv, 21*(10), 541-548.

Flynn, E. A., Barker, K. N., Pepper, G. A., Bates, D. W., & Mikeal, R. L. (2002a). Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. *AJHP, 59*(5), 436-446.

Flynn, E. A., Barker, K. N., Pepper, G. A., Bates, D. W., & Mikeal, R. L. (2002b). Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. *American Journal Of Health-System Pharmacy, 59*(5), 436-446.

Gilbert, J. E. (2006). Applications quest: Computing diversity. *Communications of the ACM, 49*(3), 99-104.

Holden, R. J., & Karsh, B.-T. (2007). A review of medical error reporting system design considerations and a proposed cross-level systems research framework. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 49*(2), 257-276.

Ian H. Witten, E. F. (2000). Data mining: Practical machine learning tools and techniques with java implementations. San Francisco: Morgan Kaufmann Publishers.

Institute of Medicine. (2000). To err is human: Building a safer health system. Washington, DC: National Academy Press.

Institute of Medicine. (2001). Crossing the quality chasm: A new health system for the 21st century. Washington DC: National Academy Press.

Institute of Medicine. (2004a). Keeping patients safe: Transforming the work environment of nurses. Washington, DC: National Academy Press.

Institute of Medicine. (2004b). Patient safety: Achieving a new standard for care. Washington, DC: National Academy Press.

John Gantz, D. R. (2009). As the economy contracts, the digital universe expands. In I. IDC Research (Ed.), *IDC - Multimedia Whitepaper* (Vol. 2009). Framingham: IDC Go-to-Market Services.

Krizek, T. J. (2000). Surgical error: Ethical issues of adverse events. *Arch Surg, 135*(11), 1359-1366.

Leape, L. L. M., & Berwick, D. M. M. (2005). Five years after to err is human: What have we learned? *JAMA, 293*(19), 2384-2390.

Manasse Hr, Jr. (1989). Medication use in an imperfect world: Drug misadventuring as an issue of public policy, Part 1. *Am J Health Syst Pharm, 46*(5), 929-944.

Osei-Bryson, K.-M., & Rayward-Smith, V. J. Data mining and operational research: Techniques and applications. *J Oper Res Soc, 60*(8), 1043-1044.

Pepper, G. A. (1995). Errors in drug administration by nurses. *American Journal Of Health-System Pharmacy, 52*(4), 390-395.

Pharmacopeia, U. S. (2004). The medication use process. Retrieved 7/22, 2009, from http://www.usp.org/pdf/EN/patientSafety/medicationUseProcess.pdf

Reason, J. (1990). Human error. Cambridge: Cambridge University Press.

rL soultions. (2009). Clients. In R. T. Inc. (Ed.): RadicaLogic Technologies Inc.

Santell, J. P., Hicks, R. W., McMeekin, J., & Cousins, D. D. (2003). Medication errors: Experience of the United States Pharmacopeia (USP) MEDMARX reporting system. *J Clin Pharmacol, 43*(7), 760-767.

Stone, J., Ellison, K.J. & Dubois, E.J. (2007, May). An accident waiting to happen. Paper presented at the National Association of Orthopaedic Nurses 27th Annual Congress: A New Perspective: Energizing the Future, San Jose, CA.

Taylor, J. A., Brownstein, D., Christakis, D. A., Blackburn, S., Strandjord, T. P., Klein, E. J., et al. (2004). Use of incident reports by physicians and nurses to document medical errors in pediatric patients. *Pediatrics, 114*(3), 729-735.

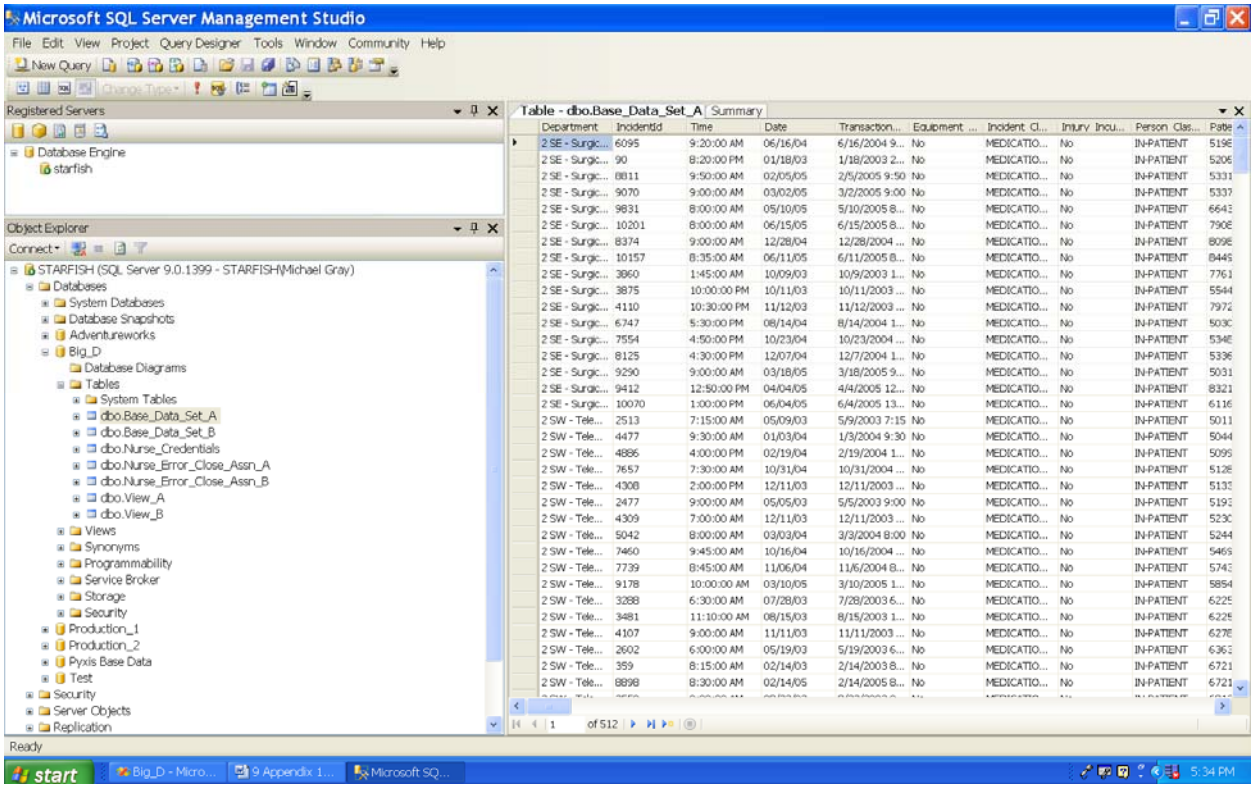
Appendix 1

Risk Monitor Pro Data Fields (Attributes)

Risk Monitor Pro - Application Data Fields	
Department	Med Dose/Rate/Conc (Ordered)
Incident Id	Med Dosage Form (Ordered)
Time	Med Admin Route (Ordered)
Date	Med Strength (Ordered)
Equipment Involved	Med Product Name (Administered)
Incident Classification	Med Generic Name (Administered)
Injury Incurred	Med Dose/Rate/Conc (Administered)
Person Classification	Med Admin Route (Administered)
Last Name	Med Strength (Administered)
1st name	Patient Received Medication
Patient #	ID/Documentation/Consent
Sex	Actions
DOB	Actions Taken to Prevent Recurrence
Person Age	Incident Severity Level
Room #	Injury Degree
Reported By Organization	Injury Nature
Reported By	Location of Injury on Body
Reported By Date	Injury Body Part
Reported By Time	Xray Date
Site	Xray Site
Specific Location	Xray Result
Entered by	Has Blood Test
Entered Date	Equipment Manufactor
Entered Time	Equipment Serial No.
Witness Name	Equipment Out of Service
Witness Address	Equipment Secured
Witness Phone	Notification Type
Specific Incident Type	Notification Date
Med Product Name(Ordered)	Notification Time
Med Generic Name(Ordered)	Factors
	Description

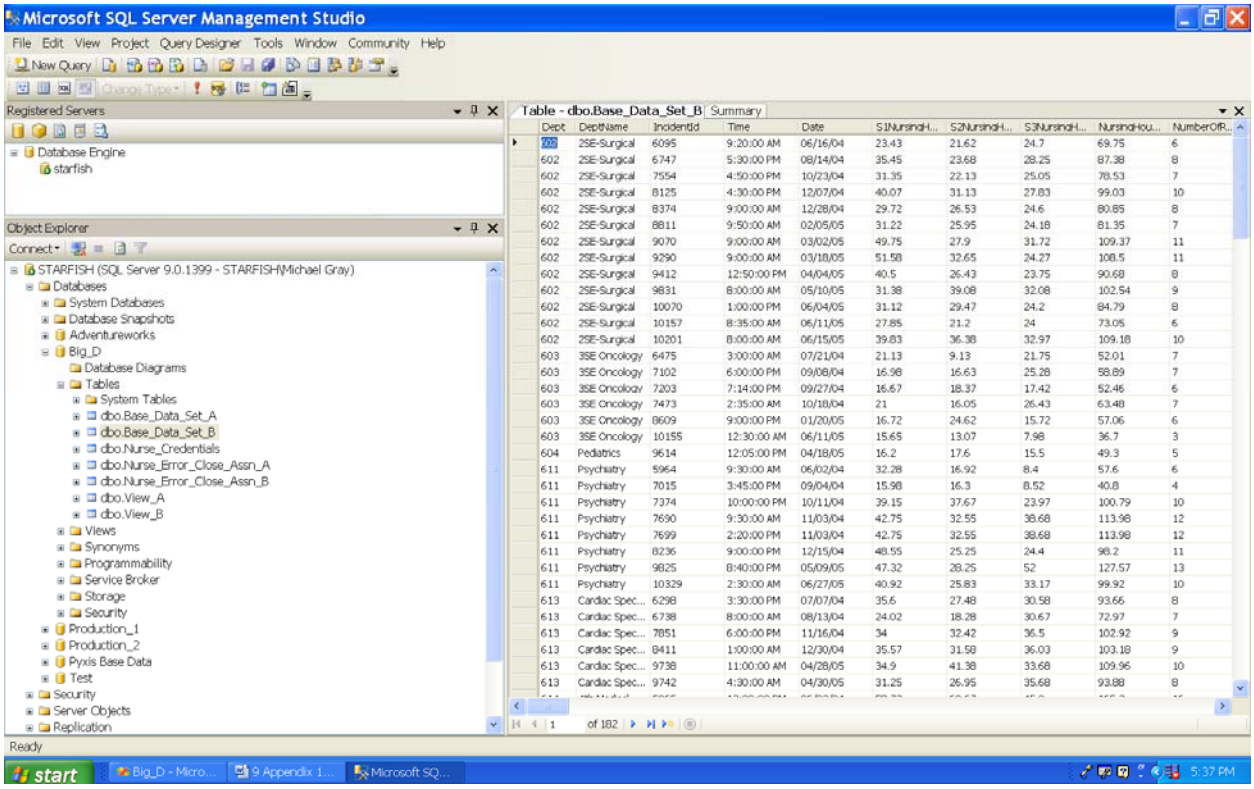
Base Dataset A

512 Rows



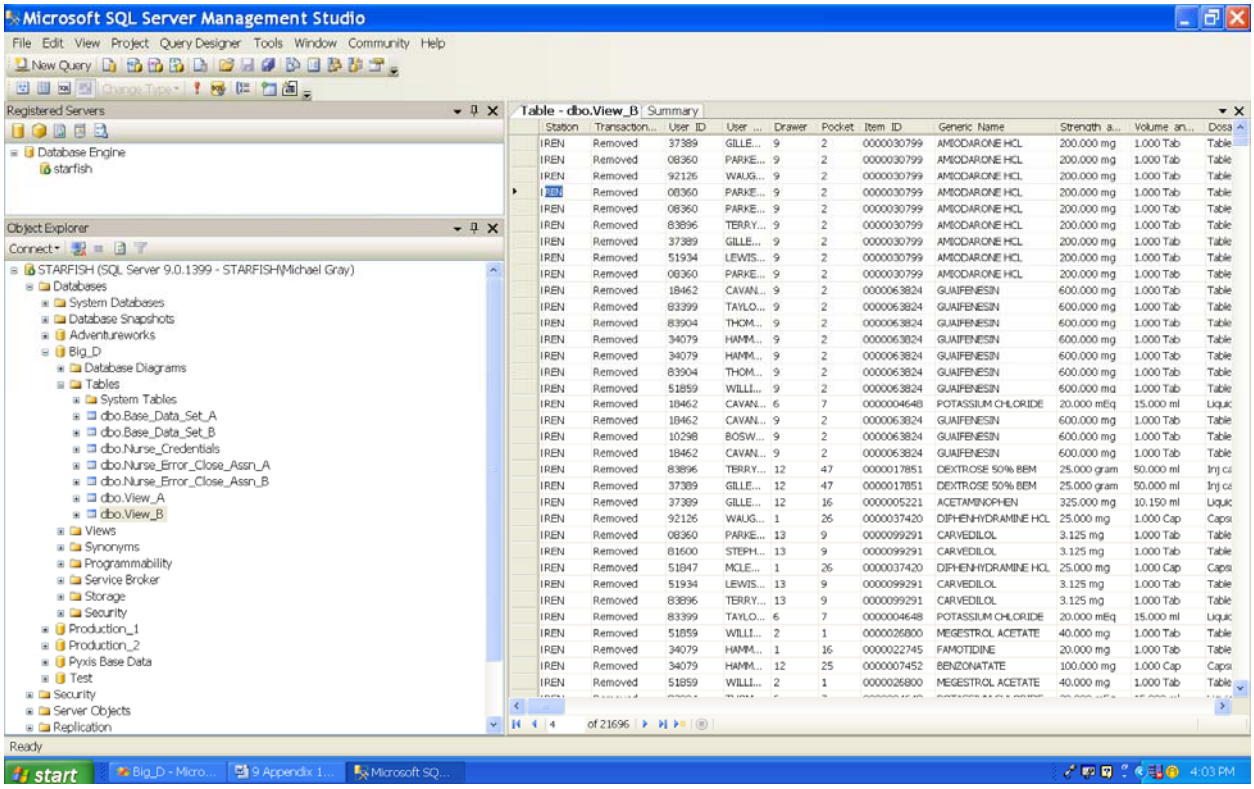
Base Dataset B

182 Rows



Final Base Dataset Table View

21,696 Rows



View B Attributes Table

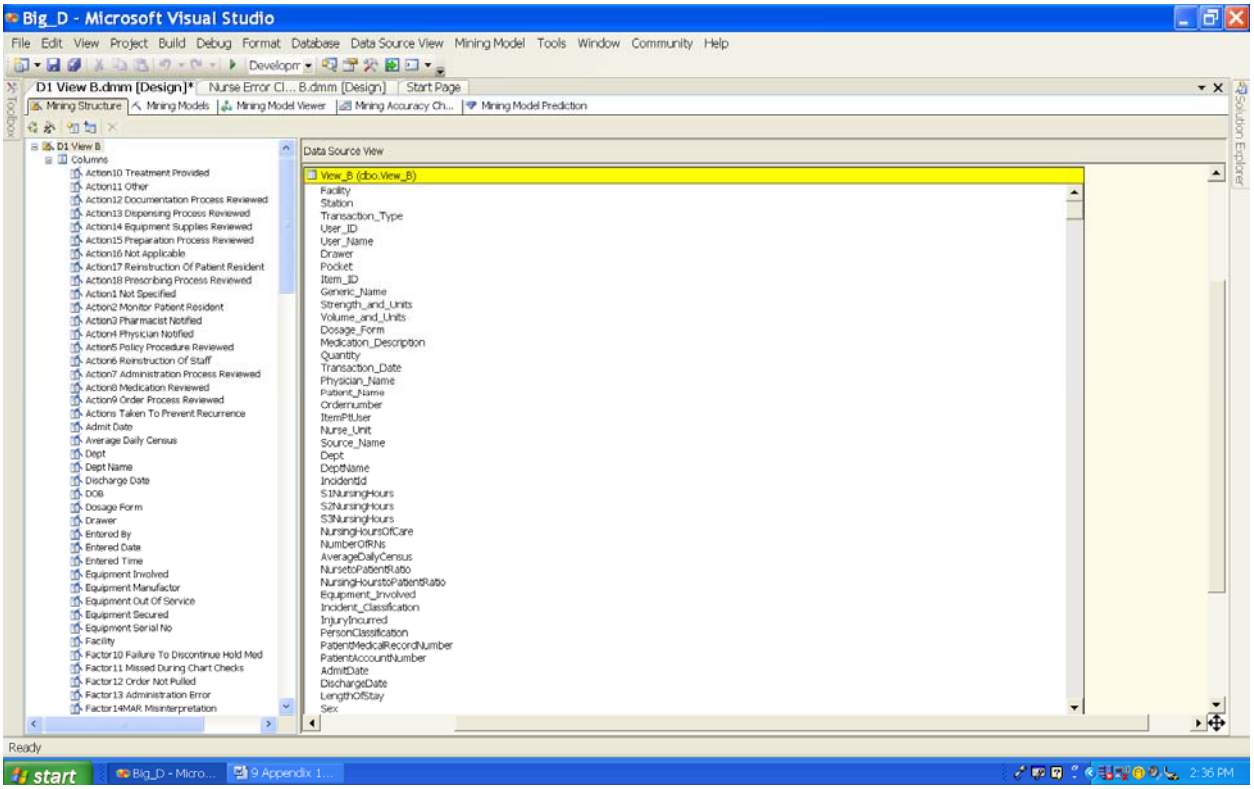
View_B Data Fields (Attributes)	
Action10TreatmentProvided	IDDocumentationConsent18DocumentsUnsigned
Action11Other	IDDocumentationConsent19ConsentIssue
Action12DocumentationProcessReviewed	IDDocumentationConsent1NotSpecified
Action13DispensingProcessReviewed	IDDocumentationConsent2WrongPatient
Action14EquipmentSuppliesReviewed	IDDocumentationConsent3DocumentsMissing
Action15PreparationProcessReviewed	IDDocumentationConsent4PolicyProcedureNotFollowed
Action16NotApplicable	IDDocumentationConsent5TranscriptionError
Action17ReinstructionOfPatientResident	IDDocumentationConsent6DocumentsDelayed
Action18PrescribingProcessReviewed	IDDocumentationConsent7Inappropriate
Action1NotSpecified	IDDocumentationConsent8Illegible
Action2MonitorPatientResident	IDDocumentationConsent9Absent
Action3PharmacistNotified	Incident_Classification
Action4PhysicianNotified	IncidentId

View_B Data Fields (Attributes)	
Action5PolicyProcedureReviewed	IncidentSeverityLevel
Action6ReinstructionOfStaff	InjuryBodyPart
Action7AdministrationProcessReviewed	InjuryDegree
Action8MedicationReviewed	InjuryIncurred
Action9OrderProcessReviewed	InjuryNature
Actions_Taken_to_Prevent_Recurrence	Item_ID
AdmitDate	ItemPtUser
AverageDailyCensus	LengthOfStay
Dept	Location_of_Injury_on_Body
DeptName	MedAdminRouteAdministered
DischargeDate	MedAdminRouteOrdered
DOB	MedDosageFormOrdered
Dosage_Form	MedDoseRateConcAdministered
Drawer	MedDoseRateConcOrdered
Entered_By	MedGenericNameAdministered
Entered_Date	MedGenericNameOrdered
Entered_Time	Medication_Description
Equipment_Involved	MedProductNameAdministered
Equipment_Manufactor	MedProductNameOrdered
Equipment_Out_of_Service	MedStrengthAdministered
Equipment_Secured	MedStrengthOrdered
Equipment_Serial_No	NotificationDate
Facility	NotificationTime
Factor10FailureToDiscontinueHoldMed	NotificationType1NotSpecified
Factor11MissedDuringChartChecks	NotificationType2Manager
Factor12OrderNotPulled	NotificationType3NextOfKin
Factor13AdministrationError	NotificationType4Physician
Factor14MARMisinterpretation	NotificationType5Pharmacist
Factor15MisinterpretationOfOrder	NotificationType6Other
Factor16MedicationOnHold	NotificationType7Supervisor
Factor17PharmacyOrderProcessingError	NotificationType8Administrator
Factor18TranscriptionError	NotificationType9Director
Factor19AllergyReactionUnknown	NumberOfRNs
Factor1NotSpecified	Nurse_Unit
Factor20IncorrectPreparation	NursetoPatientRatio
Factor21OrderError	NursingHoursOfCare
Factor22DispensingError	NursingHourstoPatientRatio
Factor23MisinterpretationOfLabel	Ordernumber
Factor24RateDoseCalculationError	Patient_Name

View_B Data Fields (Attributes)	
Factor25IllegibleHandwriting	PatientAccountNumber
Factor26MedicationUnavailable	PatientMedicalRecordNumber
Factor27AdministrationDelay	PatientReceivedMedication
Factor28MARUnclear	PersonAge
Factor29OrderProcessingDelay	PersonClassification
Factor2AdministrationNotRecordedSignedOff	Physician_Name
Factor30AllergyNotNoted	Pocket
Factor31MedicationDiscontinued	Quantity
Factor32PatientResidentsOwnMedication	Reported_By
Factor33EquipmentSuppliesFaulty	Reported_By_Date
Factor34PumpInfusionSettings	Reported_By_Time
Factor35IncorrectImproperLabel	ReportedByOrganization
Factor36DispensingDelay	RoomNumber
Factor37PrescribingError	S1NursingHours
Factor38OrderProcessingError	S2NursingHours
Factor3NotApplicable	S3NursingHours
Factor4InterferenceByPatientResident	Sex
Factor5Other	Site
Factor6PatientResidentIdentification	Source_Name
Factor7IncorrectAdministration	SpecificIncidentType
Factor8PolicyProcedureIssue	SpecificLocation
Factor9DocumentationError	Station
Generic_Name	Strength_and_Units
Has_Blood_Test	Transaction_Date
IDDocumentationConsent10MRNwrong	Transaction_Type
IDDocumentationConsent11MedicalClearanceNotDocumented	User_ID
IDDocumentationConsent12WrongName	User_Name
IDDocumentationConsent13PatientMedicatedBeforeSigning	Volume_and_Units
IDDocumentationConsent14IncorrectRequisition	Witness_Address
IDDocumentationConsent15IncompleteRequisition	Witness_Name
IDDocumentationConsent16NotesUnsigned	XrayDate
IDDocumentationConsent17Altered	XrayResult
	XraySite

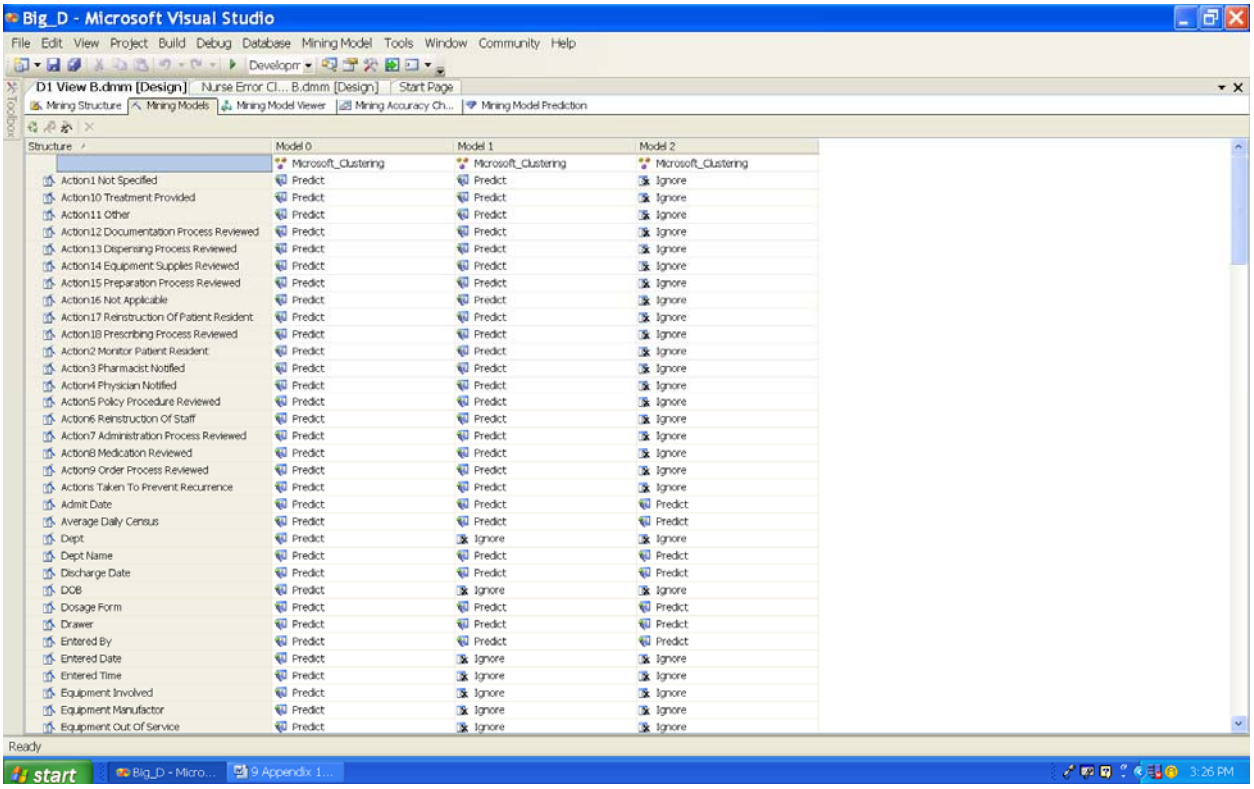
Appendix 2

Data Source View – The “View B” Mining Structure



Microsoft® SQL Server™ 2005 Analysis Services

Mining Models 0, 1, and 2.



Attribute Detail for View B Dataset Mining Models 0, 1, and 2

View_B	Mining Model 0	Mining Model 1	Mining Model 2
Action10TreatmentProvided	I & P	I & P	X*
Action11Other	I & P	I & P	X*
Action12DocumentationProcessReviewed	I & P	I & P	X*
Action13DispensingProcessReviewed	I & P	I & P	X*
Action14EquipmentSuppliesReviewed	I & P	I & P	X*
Action15PreparationProcessReviewed	I & P	I & P	X*
Action16NotApplicable	I & P	I & P	X*
Action17ReinstructionOfPatientResident	I & P	I & P	X*
Action18PrescribingProcessReviewed	I & P	I & P	X*
Action1NotSpecified	I & P	I & P	X*
Action2MonitorPatientResident	I & P	I & P	X*
Action3PharmacistNotified	I & P	I & P	X*
Action4PhysicianNotified	I & P	I & P	X*
Action5PolicyProcedureReviewed	I & P	I & P	X*
Action6ReinstructionOfStaff	I & P	I & P	X*
Action7AdministrationProcessReviewed	I & P	I & P	X*
Action8MedicationReviewed	I & P	I & P	X*
Action9OrderProcessReviewed	I & P	I & P	X*
Actions_Taken_to_Prevent_Recurrence	I & P	I & P	X*
AdmitDate	I & P	I & P	I & P
AverageDailyCensus	I & P	I & P	I & P
Dept	I & P	X*	X*
DeptName	I & P	I & P	I & P
DischargeDate	I & P	I & P	I & P
DOB	I & P	X*	X*
Dosage_Form	I & P	I & P	I & P
Drawer	I & P	I & P	I & P
Entered_By	I & P	I & P	I & P
Entered_Date	I & P	X*	X*
Entered_Time	I & P	X*	X*
Equipment_Involved	I & P	X*	X*
Equipment_Manufactor	I & P	X*	X*
Equipment_Out_of_Service	I & P	X*	X*
Equipment_Secured	I & P	X*	X*
Equipment_Serial_No	I & P	X*	X*
Facility	I & P	X*	X*
Factor10FailureToDiscontinueHoldMed	I & P	I & P	X*
Factor11MissedDuringChartChecks	I & P	I & P	X*
Factor12OrderNotPulled	I & P	I & P	X*
Factor13AdministrationError	I & P	I & P	X*
Factor14MARMisinterpretation	I & P	I & P	X*

I - Input attribute for clustering, P - Attribute was used for prediction of cluster membership,
X - Attribute was not used for cluster development. * Rationale for exclusion prior to clustering

View_B	Mining Model 0	Mining Model 1	Mining Model 2
Factor15MisinterpretationOfOrder	I & P	I & P	X*
Factor16MedicationOnHold	I & P	I & P	X*
Factor17PharmacyOrderProcessingError	I & P	I & P	X*
Factor18TranscriptionError	I & P	I & P	X*
Factor19AllergyReactionUnknown	I & P	I & P	X*
Factor1NotSpecified	I & P	I & P	X*
Factor20IncorrectPreparation	I & P	I & P	X*
Factor21OrderError	I & P	I & P	X*
Factor22DispensingError	I & P	I & P	X*
Factor23MisinterpretationOfLabel	I & P	I & P	X*
Factor24RateDoseCalculationError	I & P	I & P	X*
Factor25IllegibleHandwriting	I & P	I & P	X*
Factor26MedicationUnavailable	I & P	I & P	X*
Factor27AdministrationDelay	I & P	I & P	X*
Factor28MARUnclear	I & P	I & P	X*
Factor29OrderProcessingDelay	I & P	I & P	X*
Factor2AdministrationNotRecordedSignedOff	I & P	I & P	X*
Factor30AllergyNotNoted	I & P	I & P	X*
Factor31MedicationDiscontinued	I & P	I & P	X*
Factor32PatientResidentsOwnMedication	I & P	I & P	X*
Factor33EquipmentSuppliesFaulty	I & P	I & P	X*
Factor34PumpInfusionSettings	I & P	I & P	X*
Factor35IncorrectImproperLabel	I & P	I & P	X*
Factor36DispensingDelay	I & P	I & P	X*
Factor37PrescribingError	I & P	I & P	X*
Factor38OrderProcessingError	I & P	I & P	X*
Factor3NotApplicable	I & P	I & P	X*
Factor4InterferenceByPatientResident	I & P	I & P	X*
Factor5Other	I & P	I & P	X*
Factor6PatientResidentIdentification	I & P	I & P	X*
Factor7IncorrectAdministration	I & P	I & P	X*
Factor8PolicyProcedureIssue	I & P	I & P	X*
Factor9DocumentationError	I & P	I & P	X*
Generic_Name	I & P	I & P	I & P
Has_Blood_Test	I & P	X*	X*
IDDocumentationConsent10MRNwrong	I & P	I & P	X*
IDDocumentationConsent11MedicalClearanceNotDocumented	I & P	I & P	X*
IDDocumentationConsent12WrongName	I & P	I & P	X*
IDDocumentationConsent13PatientMedicatedBeforeSigning	I & P	I & P	X*
IDDocumentationConsent14IncorrectRequisition	I & P	I & P	X*

I - Input attribute for clustering, P - Attribute was used for prediction of cluster membership,
X - Attribute was not used for cluster development. * Rationale for exclusion prior to clustering

View_B	Mining Model 0	Mining Model 1	Mining Model 2
IDDocumentationConsent15IncompleteRequisition	I & P	I & P	X*
IDDocumentationConsent16NotesUnsigned	I & P	I & P	X*
IDDocumentationConsent17Altered	I & P	I & P	X*
IDDocumentationConsent18DocumentsUnsigned	I & P	I & P	X*
IDDocumentationConsent19ConsentIssue	I & P	I & P	X*
IDDocumentationConsent1NotSpecified	I & P	I & P	X*
IDDocumentationConsent2WrongPatient	I & P	I & P	X*
IDDocumentationConsent3DocumentsMissing	I & P	I & P	X*
IDDocumentationConsent4PolicyProcedureNotFollowed	I & P	I & P	X*
IDDocumentationConsent5TranscriptionError	I & P	I & P	X*
IDDocumentationConsent6DocumentsDelayed	I & P	I & P	X*
IDDocumentationConsent7Inappropriate	I & P	I & P	X*
IDDocumentationConsent8Illegible	I & P	I & P	X*
IDDocumentationConsent9Absent	I & P	I & P	X*
Incident_Classification	I & P	X*	X*
IncidentId	KEY	KEY	KEY
IncidentSeverityLevel	I & P	I & P	I & P
InjuryBodyPart	I & P	X*	X*
InjuryDegree	I & P	I & P	X*
InjuryIncurred	I & P	I & P	I & P
InjuryNature	I & P	I & P	X*
Item_ID	I & P	I & P	I & P
ItemPtUser	I & P	X*	X*
LengthOfStay	I & P	I & P	I & P
Location_of_Injury_on_Body	I & P	X*	X*
MedAdminRouteAdministered	I & P	I & P	I & P
MedAdminRouteOrdered	I & P	I & P	I & P
MedDosageFormOrdered	I & P	I & P	I & P
MedDoseRateConcAdministered	I & P	I & P	I & P
MedDoseRateConcOrdered	I & P	I & P	I & P
MedGenericNameAdministered	I & P	I & P	X*
MedGenericNameOrdered	I & P	I & P	X*
Medication_Description	I & P	I & P	I & P
MedProductNameAdministered	I & P	I & P	X*
MedProductNameOrdered	I & P	I & P	I & P
MedStrengthAdministered	I & P	I & P	I & P
MedStrengthOrdered	I & P	I & P	I & P
NotificationDate	I & P	X*	X*
NotificationTime	I & P	X*	X*
NotificationType1NotSpecified	I & P	I & P	X*
NotificationType2Manager	I & P	I & P	X*

I - Input attribute for clustering, P - Attribute was used for prediction of cluster membership,
X - Attribute was not used for cluster development. * Rationale for exclusion prior to clustering

View_B	Mining Model 0	Mining Model 1	Mining Model 2
NotificationType3NextOfKin	I & P	I & P	X*
NotificationType4Physician	I & P	I & P	X*
NotificationType5Pharmacist	I & P	I & P	X*
NotificationType6Other	I & P	I & P	X*
NotificationType7Supervisor	I & P	I & P	X*
NotificationType8Administrator	I & P	I & P	X*
NotificationType9Director	I & P	I & P	X*
NumberOfRNs	I & P	I & P	I & P
Nurse_Unit	I & P	I & P	I & P
NursetoPatientRatio	I & P	I & P	I & P
NursingHoursOfCare	I & P	I & P	I & P
NursingHourstoPatientRatio	I & P	I & P	I & P
Ordernumber	I & P	I & P	I & P
Patient_Name	I & P	X*	X*
PatientAccountNumber	I & P	I & P	I & P
PatientMedicalRecordNumber	I & P	I & P	I & P
PatientReceivedMedication	I & P	I & P	I & P
PersonAge	I & P	I & P	I & P
PersonClassification	I & P	X*	X*
Physician_Name	I & P	I & P	I & P
Pocket	I & P	I & P	I & P
Quantity	I & P	I & P	I & P
Reported_By	I & P	I & P	I & P
Reported_By_Date	I & P	X*	X*
Reported_By_Time	I & P	X*	X*
ReportedByOrganization	I & P	X*	X*
RoomNumber	I & P	I & P	I & P
S1NursingHours	I & P	I & P	I & P
S2NursingHours	I & P	I & P	I & P
S3NursingHours	I & P	I & P	I & P
Sex	I & P	I & P	I & P
Site	I & P	X*	X*
Source_Name	I & P	X*	X*
SpecificIncidentType	I & P	I & P	I & P
SpecificLocation	I & P	I & P	I & P
Station	I & P	I & P	I & P
Strength_and_Units	I & P	I & P	I & P
Transaction_Date	I & P	I & P	I & P
Transaction_Type	I & P	X*	X*
User_ID	I & P	I & P	I & P
User_Name	I & P	I & P	I & P
Volume_and_Units	I & P	I & P	I & P
Witness_Address	I & P	X*	X*

I - Input attribute for clustering, P - Attribute was used for prediction of cluster membership,
X - Attribute was not used for cluster development. * Rationale for exclusion prior to clustering

View_B	Mining Model 0	Mining Model 1	Mining Model 2
Witness_Name	I & P	X*	X*
XrayDate	I & P	X*	X*
XrayResult	I & P	X*	X*
XraySite	I & P	X*	X*

I - Input attribute for clustering, P - Attribute was used for prediction of cluster membership,
X - Attribute was not used for cluster development. * Rationale for exclusion prior to clustering

Attributes Excluded from Mining Model 1 and Mining Model 2

Attribute Excluded	Reason	Rationale
Facility	Not significant	Since this analysis includes only one healthcare system, Facility = MC without exception.
Transaction_Type:	Not significant	This field describes the action reported by the Pyxis machine when a Medication is removed from the pocket. All Medications from the Pyxis machine will have Transaction Type = Removed
Patient_Name:	Redundant	The Patient Account number uniquely identifies a patient in the combined data sets.
ItemPtUser:	Redundant	A combination of three other fields in the Pyxis database - Medication Item Number - Patient Record Number - Pyxis Machine User ID (Nurse Pyxis ID)
Source_Name	Not significant	The Pyxis database records "console" as the source when a Medication is removed from the machine. All Medications from the Pyxis machine will have Source Name = Console
Dept:	Redundant	The “Dept Name” attribute identifies the Department or Unit where the error or near miss occurred. “Dept” is a replicated attribute.
Equipment_Involved	Not significant	No incident reports used this attribute. The attribute was set to its default value of “Not Specified”.
Incident_Classification:	Not significant	All values for Incident Classification equal “MEDICATION/IV/BLOOD”.
PersonClassification:	Not significant	Almost all values for “Person Classification” were “IN-PATIENT” with only a few values listing “RESIDENT.” Both values assume direct control of the care environment by the Medical Center. Therefore, clustering about this data point is not considered significant.
DOB:	Redundant	The “Person Age” attribute serves the main objective which is to determine if the patient’s age has any relationship to the other error related data.
Attribute Excluded	Reason	Rationale
ReportedByOrganization	Redundant	The same attribute as Department and Department Name attributes.
Reported_By_Date	Not significant	Risk Monitor Pro uses the system Date and Time when the reporting person logs into the system and records it as the Reported By Date and Reported By Time. This is not the Date and Time when the error occurred. Therefore this attribute will not be considered for clustering.

Attribute Excluded	Reason	Rationale
Reported_By_Time	Not significant	Risk Monitor Pro uses the system Date and Time when the reporting person logs into the system and records it as the Reported By Date and Reported By Time. This is not the Date and Time when the error occurred. Therefore this attribute will not be considered for clustering.
Site	Not significant	The "Site" attribute is not at the level of the organization that will be useful for this analysis. Department Name and Department Number provides the location and organization where the error occurred and the reporting source department.
Entered_Date	Not significant	Risk Monitor Pro uses the system Date and Time when the reporting person completes an entry and records it as the Entered Date and Entered Time to be the date and time when the record was committed to the Self Reporting System database. This does not reflect the Date and Time when the error occurred. For the purposes of clustering this variable will not be considered.
Entered_Time	Not significant	Risk Monitor Pro uses the system Date and Time when the reporting person completes an entry and records it as the Entered Date and Entered Time to be the date and time when the record was committed to the Self Reporting System database. This does not reflect the Date and Time when the error occurred. For the purposes of clustering this variable will not be considered.
Witness_Name	Not significant	Risk Monitor Pro defaults to Null or Not Specified for this field if the reporting person does not add the information when reporting an error. Inspection of the data recorded indicates, this field most often contains the default value. Therefore it was excluded from clustering.
Witness_Address	Not significant	Risk Monitor Pro defaults to Null or Not Specified for this field if the reporting person does not add the information when reporting an error. Inspection of the data recorded indicates, this field most often contains the default value. Therefore it was excluded from clustering.
Location_of_Injury_on_Body	Not significant	All values for "Location Of Injury On Body" = Null. Therefore it will not be considered in for clustering.
InjuryBodyPart	Not significant	All values for "Injury Body Part" = Null. Therefore it will not be considered in for clustering.
XrayDate:	Not significant	All values for "XrayDate" = Null. Therefore it will not be considered in for clustering.
XraySite	Not significant	All values for "XraySite" = Null. Therefore it will not be considered in for clustering.
XrayResult	Not significant	All values for "XrayResult" = Null. Therefore it will not be considered in for clustering.
Has_Blood_Test	Not significant	This attribute is not a determinant for clustering in a way that adds meaning to the analysis. "Has Blood Test" is not directly linked to any other part of the Medication Administration process as documented at the time of data gathering. Therefore it is excluded from clustering.
Equipment_Manufactor	Not significant	All entries for this attribute are Null. In other words this field defaults to "Not Specified" when someone does not change this option in the Risk Monitor Pro.

Attribute Excluded	Reason	Rationale
Equipment_Serial_No	Not significant	All entries for this attribute are Null. In other words this field defaults to "Not Specified" when someone does not change this option in the Risk Monitor Pro.
Equipment_Out_of_Service	Not significant	All entries for this attribute are Null. In other words this field defaults to "Not Specified" when someone does not change this option in the Risk Monitor Pro.
Equipment_Secured	Not significant	All entries for this attribute are Null. In other words this field defaults to "Not Specified" when someone does not change this option in the Risk Monitor Pro.
NotificationDate	Not significant	Risk Monitor Pro defaults to Null or Not Specified for this field if the reporting person does not add the information when reporting an error. Inspection of the data recorded indicates, this field most often contains the default value. Therefore it was excluded from clustering.
NotificationTime	Not significant	Risk Monitor Pro defaults to Null or Not Specified for this field if the reporting person does not add the information when reporting an error. Inspection of the data recorded indicates, this field most often contains the default value. Therefore it was excluded from clustering.

Additional Attributes Excluded from Mining Model 2

Attribute Excluded	Rationale
Action10TreatmentProvided	Not significant – determined from Model 1 clustering results
Action11Other	Not significant – determined from Model 1 clustering results
Action12DocumentationProcessReviewed	Not significant – determined from Model 1 clustering results
Action13DispensingProcessReviewed	Not significant – determined from Model 1 clustering results
Action14EquipmentSuppliesReviewed	Not significant – determined from Model 1 clustering results
Action15PreparationProcessReviewed	Not significant – determined from Model 1 clustering results
Action16NotApplicable	Not significant – determined from Model 1 clustering results
Action17ReinstructionOfPatientResident	Not significant – determined from Model 1 clustering results
Action18PrescribingProcessReviewed	Not significant – determined from Model 1 clustering results
Action1NotSpecified	Not significant – determined from Model 1 clustering results
Action2MonitorPatientResident	Not significant – determined from Model 1 clustering results
Action3PharmacistNotified	Not significant – determined from Model 1 clustering results
Action4PhysicianNotified	Not significant – determined from Model 1 clustering results
Action5PolicyProcedureReviewed	Not significant – determined from Model 1 clustering results

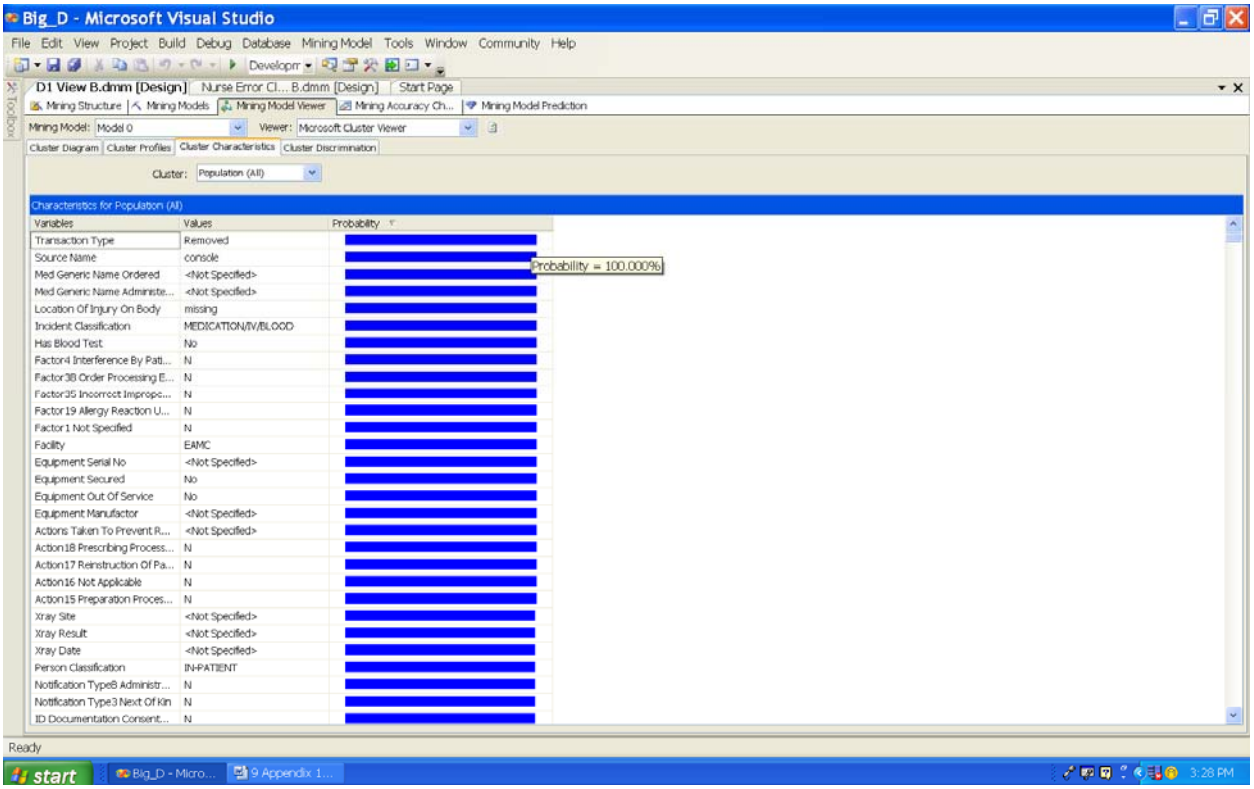
Attribute Excluded	Rationale
Action6ReinstructionOfStaff	Not significant – determined from Model 1 clustering results
Action7AdministrationProcessReviewed	Not significant – determined from Model 1 clustering results
Action8MedicationReviewed	Not significant – determined from Model 1 clustering results
Action9OrderProcessReviewed	Not significant – determined from Model 1 clustering results
Actions_Taken_to_Prevent_Recurrence	Not significant – determined from Model 1 clustering results
Factor10FailureToDiscontinueHoldMed	Not significant – determined from Model 1 clustering results
Factor11MissedDuringChartChecks	Not significant – determined from Model 1 clustering results
Factor12OrderNotPulled	Not significant – determined from Model 1 clustering results
Factor13AdministrationError	Not significant – determined from Model 1 clustering results
Factor14MARMisinterpretation	Not significant – determined from Model 1 clustering results
Factor15MisinterpretationOfOrder	Not significant – determined from Model 1 clustering results
Factor16MedicationOnHold	Not significant – determined from Model 1 clustering results
Factor17PharmacyOrderProcessingError	Not significant – determined from Model 1 clustering results
Factor18TranscriptionError	Not significant – determined from Model 1 clustering results
Factor19AllergyReactionUnknown	Not significant – determined from Model 1 clustering results
Factor1NotSpecified	Not significant – determined from Model 1 clustering results
Factor20IncorrectPreparation	Not significant – determined from Model 1 clustering results
Factor21OrderError	Not significant – determined from Model 1 clustering results
Factor22DispensingError	Not significant – determined from Model 1 clustering results
Factor23MisinterpretationOfLabel	Not significant – determined from Model 1 clustering results
Factor24RateDoseCalculationError	Not significant – determined from Model 1 clustering results
Factor25IllegibleHandwriting	Not significant – determined from Model 1 clustering results
Factor26MedicationUnavailable	Not significant – determined from Model 1 clustering results
Factor27AdministrationDelay	Not significant – determined from Model 1 clustering results
Factor28MARUnclear	Not significant – determined from Model 1 clustering results
Factor29OrderProcessingDelay	Not significant – determined from Model 1 clustering results
Factor2AdministrationNotRecordedSignedOff	Not significant – determined from Model 1 clustering results

Attribute Excluded	Rationale
Factor30AllergyNotNoted	Not significant – determined from Model 1 clustering results
Factor31MedicationDiscontinued	Not significant – determined from Model 1 clustering results
Factor32PatientResidentsOwnMedication	Not significant – determined from Model 1 clustering results
Factor33EquipmentSuppliesFaulty	Not significant – determined from Model 1 clustering results
Factor34PumpInfusionSettings	Not significant – determined from Model 1 clustering results
Factor35IncorrectImproperLabel	Not significant – determined from Model 1 clustering results
Factor36DispensingDelay	Not significant – determined from Model 1 clustering results
Factor37PrescribingError	Not significant – determined from Model 1 clustering results
Factor38OrderProcessingError	Not significant – determined from Model 1 clustering results
Factor3NotApplicable	Not significant – determined from Model 1 clustering results
Factor4InterferenceByPatientResident	Not significant – determined from Model 1 clustering results
Factor5Other	Not significant – determined from Model 1 clustering results
Factor6PatientResidentIdentification	Not significant – determined from Model 1 clustering results
Factor7IncorrectAdministration	Not significant – determined from Model 1 clustering results
Factor8PolicyProcedureIssue	Not significant – determined from Model 1 clustering results
Factor9DocumentationError	Not significant – determined from Model 1 clustering results
IDDocumentationConsent10MRNwrong	Not significant – determined from Model 1 clustering results
IDDocumentationConsent11MedicalClearanceNotDocumented	Not significant – determined from Model 1 clustering results
IDDocumentationConsent12WrongName	Not significant – determined from Model 1 clustering results
IDDocumentationConsent13PatientMedicatedBeforeSigning	Not significant – determined from Model 1 clustering results
IDDocumentationConsent14IncorrectRequisition	Not significant – determined from Model 1 clustering results
IDDocumentationConsent15IncompleteRequisition	Not significant – determined from Model 1 clustering results
IDDocumentationConsent16NotesUnsigned	Not significant – determined from Model 1 clustering results
IDDocumentationConsent17Altered	Not significant – determined from Model 1 clustering results
IDDocumentationConsent18DocumentsUnsigned	Not significant – determined from Model 1 clustering results
IDDocumentationConsent19ConsentIssue	Not significant – determined from Model 1 clustering results
IDDocumentationConsent1NotSpecified	Not significant – determined from Model 1 clustering results

Attribute Excluded	Rationale
IDDocumentationConsent2WrongPatient	Not significant – determined from Model 1 clustering results
IDDocumentationConsent3DocumentsMissing	Not significant – determined from Model 1 clustering results
IDDocumentationConsent4PolicyProcedureNotFollowed	Not significant – determined from Model 1 clustering results
IDDocumentationConsent5TranscriptionError	Not significant – determined from Model 1 clustering results
IDDocumentationConsent6DocumentsDelayed	Not significant – determined from Model 1 clustering results
IDDocumentationConsent7Inappropriate	Not significant – determined from Model 1 clustering results
IDDocumentationConsent8Illegible	Not significant – determined from Model 1 clustering results
IDDocumentationConsent9Absent	Not significant – determined from Model 1 clustering results
InjuryDegree	Not significant – determined from Model 1 clustering results
InjuryNature	Not significant – determined from Model 1 clustering results
MedGenericNameAdministered	Redundant - Other attributes exist for this data type.
MedGenericNameOrdered	Redundant - Other attributes exist for this data type.
MedProductNameAdministered	Redundant - Other attributes exist for this data type.
NotificationType1NotSpecified	Not significant – determined from Model 1 clustering results
NotificationType2Manager	Not significant – determined from Model 1 clustering results
NotificationType3NextOfKin	Not significant – determined from Model 1 clustering results
NotificationType4Physician	Not significant – determined from Model 1 clustering results
NotificationType5Pharmacist	Not significant – determined from Model 1 clustering results
NotificationType6Other	Not significant – determined from Model 1 clustering results
NotificationType7Supervisor	Not significant – determined from Model 1 clustering results
NotificationType8Administrator	Not significant – determined from Model 1 clustering results
NotificationType9Director	Not significant – determined from Model 1 clustering results

Appendix 3

Cluster Characteristics – Model 0



D1 View B **Model 0** **Population Profile** **Size: 147**		
Variables	**Values**	**Probability**
Transaction Type	Removed	100.00%
Source Name	console	100.00%
Med Generic Name Ordered	<Not Specified>	100.00%
Med Generic Name Administered	<Not Specified>	100.00%
Location Of Injury On Body	missing	100.00%
Incident Classification	MEDICATION/IV/BLOOD	100.00%
Has Blood Test	No	100.00%
Factor4 Interference By Patient Resident	N	100.00%

D1 View B
Model 0
Population Profile
Size: 147

Variables	Values	Probability
Factor38 Order Processing Error	N	100.00%
Factor35 Incorrect Improper Label	N	100.00%
Factor19 Allergy Reaction Unknown	N	100.00%
Factor1 Not Specified	N	100.00%
Facility	MC	100.00%
Equipment Serial No	<Not Specified>	100.00%
Equipment Secured	No	100.00%
Equipment Out Of Service	No	100.00%
Equipment Manufactor	<Not Specified>	100.00%
Actions Taken To Prevent Recurrence	<Not Specified>	100.00%
Action18 Prescribing Process Reviewed	N	100.00%
Action17 Reinstruction Of Patient Resident	N	100.00%
Action16 Not Applicable	N	100.00%
Action15 Preparation Process Reviewed	N	100.00%
Xray Site	<Not Specified>	99.32%
Xray Result	<Not Specified>	99.32%
Xray Date	<Not Specified>	99.32%
Person Classification	IN-PATIENT	99.32%
Notification Type8 Administrator	N	99.32%
Notification Type3 Next Of Kin	N	99.32%
ID Documentation Consent3 Documents Missing	N	99.32%
ID Documentation Consent19 Consent Issue	N	99.32%
ID Documentation Consent17 Altered	N	99.32%
ID Documentation Consent16 Notes Unsigned	N	99.32%
ID Documentation Consent13 Patient Medicated Before Signing	N	99.32%
ID Documentation Consent12 Wrong Name	N	99.32%
Factor37 Prescribing Error	N	99.32%
Factor34 Pump Infusion Settings	N	99.32%
Factor30 Allergy Not Noted	N	99.32%
Factor28MAR Unclear	N	99.32%
Factor23 Misinterpretation Of Label	N	99.32%
Factor16 Medication On Hold	N	99.32%
Factor10 Failure To Discontinue Hold Med	N	99.32%
Equipment Involved	No	99.32%
Notification Type9 Director	N	98.64%
Injury Body Part	<Not Specified>	98.64%
ID Documentation Consent8 Illegible	N	98.64%
ID Documentation Consent7 Inappropriate	N	98.64%
ID Documentation Consent15 Incomplete Requisition	N	98.64%
ID Documentation Consent14 Incorrect Requisition	N	98.64%
Factor33 Equipment Supplies Faulty	N	98.64%
Factor32 Patient Residents Own Medication	N	98.64%

D1 View B
Model 0
Population Profile
Size: 147

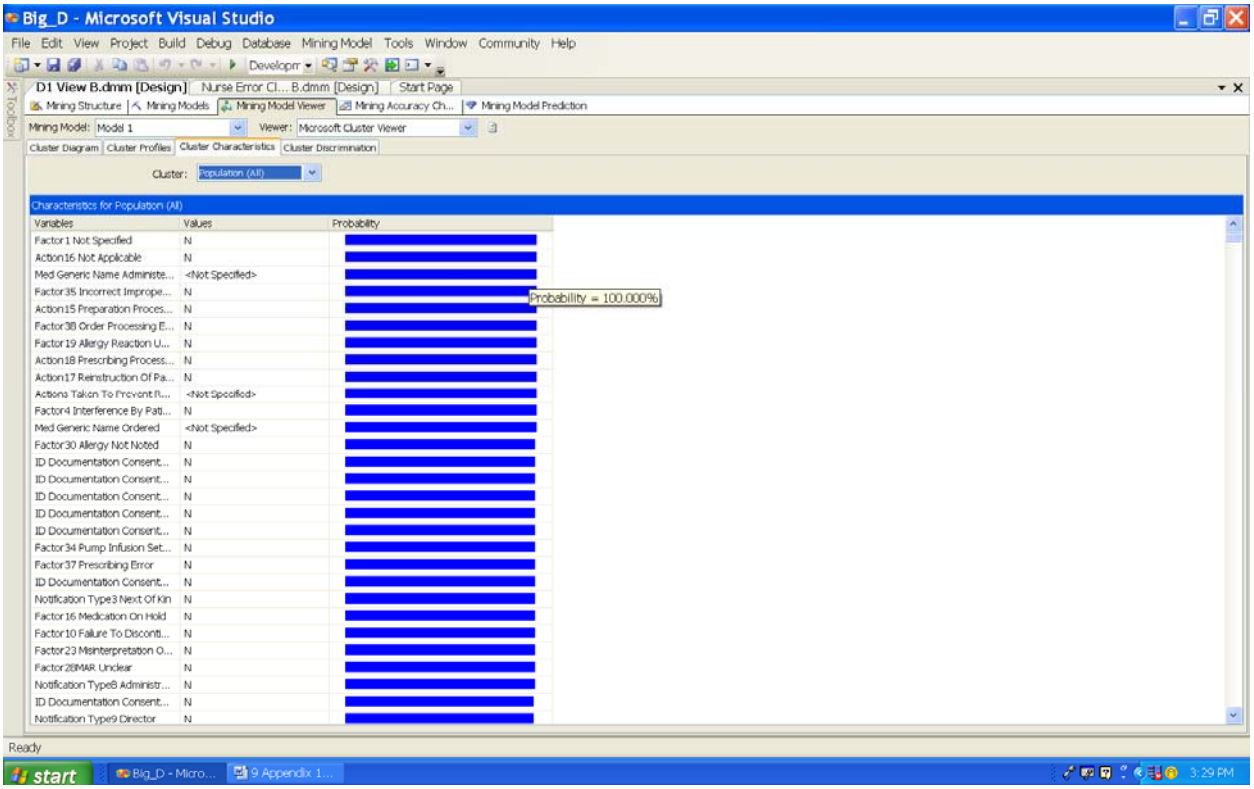
Variables	Values	Probability
Factor29 Order Processing Delay	N	98.64%
Factor26 Medication Unavailable	N	98.64%
Factor25 Illegible Handwriting	N	98.64%
Factor12 Order Not Pulled	N	98.64%
Action14 Equipment Supplies Reviewed	N	98.64%
ID Documentation Consent18 Documents Unsigned	N	97.96%
ID Documentation Consent11 Medical Clearance Not Documented	N	97.96%
Factor21 Order Error	N	97.96%
Injury Nature	<Not Specified>	97.28%
Injury Incurred	No	97.28%
Injury Degree	<Not Specified>	97.28%
ID Documentation Consent10MR Nwrong	N	97.28%
Factor6 Patient Resident Identification	N	97.28%
Factor31 Medication Discontinued	N	97.28%
Factor24 Rate Dose Calculation Error	N	97.28%
ID Documentation Consent9 Absent	N	96.60%
Factor9 Documentation Error	N	96.60%
Factor36 Dispensing Delay	N	96.60%
Factor2 Administration Not Recorded Signed Off	N	96.60%
Action12 Documentation Process Reviewed	N	96.60%
Factor27 Administration Delay	N	95.92%
Factor14MAR Misinterpretation	N	95.92%
Action9 Order Process Reviewed	N	95.92%
Action13 Dispensing Process Reviewed	N	95.92%
ID Documentation Consent6 Documents Delayed	N	95.24%
Factor8 Policy Procedure Issue	N	95.24%
Factor3 Not Applicable	N	95.24%
Action11 Other	N	95.24%
Notification Type6 Other	N	94.56%
Factor22 Dispensing Error	N	93.88%
Factor17 Pharmacy Order Processing Error	N	93.88%
Action5 Policy Procedure Reviewed	N	93.20%
Factor20 Incorrect Preparation	N	92.52%
Factor15 Misinterpretation Of Order	N	92.52%
Action6 Reinstruction Of Staff	N	92.52%
Notification Type5 Pharmacist	N	91.84%
ID Documentation Consent2 Wrong Patient	N	91.16%
Factor11 Missed During Chart Checks	N	91.16%
Action10 Treatment Provided	N	91.16%
ID Documentation Consent4 Policy Procedure Not Followed	N	90.48%
Factor5 Other	N	90.48%

D1 View B
Model 0
Population Profile
Size: 147

Variables	Values	Probability
Factor18 Transcription Error	N	90.48%
Action7 Administration Process Reviewed	N	90.48%
Quantity	1	89.80%
Notification Type7 Supervisor	N	89.80%
ID Documentation Consent5 Transcription Error	N	88.44%
Notification Type2 Manager	N	87.76%
Factor7 Incorrect Administration	N	87.76%
Action8 Medication Reviewed	N	87.76%
Action3 Pharmacist Notified	N	87.76%
Action2 Monitor Patient Resident	N	86.40%
Witness Address	<Not Specified>	85.71%
Factor13 Administration Error	N	83.67%
Site	MC-Main Hospital	81.63%
Notification Type4 Physician	N	79.59%
Med Strength Administered	<Not Specified>	77.55%
Strength And Units	missing	74.83%
Volume And Units	missing	72.79%
Dosage Form	missing	72.79%
Action4 Physician Notified	N	68.71%
ID Documentation Consent1 Not Specified	Y	68.03%
Reported By Organization	MC- Main Hospital	67.35%
Witness Name	<Not Specified>	66.67%
Med Strength Ordered	<Not Specified>	66.67%
Incident Severity Level	Severity Level 1	62.59%
Sex	F	59.18%
Med Admin Route Administered	<Not Specified>	56.46%
Notification Time	<Not Specified>	54.42%
Notification Type1 Not Specified	Y	53.74%
Notification Date	<Not Specified>	53.74%
Action1 Not Specified	N	52.38%
Med Dose Rate Conc Administered	<Not Specified>	51.02%
Patient Received Medication	Yes	50.34%

Appendix 4

Cluster Characteristics – Model 1



D1 View B **Model 1** **Population Profile** **Size: 147**		
Variables	**Values**	**Probability**
Factor1 Not Specified	N	100.00%
Action16 Not Applicable	N	100.00%
Med Generic Name Administered	<Not Specified>	100.00%
Factor35 Incorrect Improper Label	N	100.00%
Action15 Preparation Process Reviewed	N	100.00%
Factor38 Order Processing Error	N	100.00%
Factor19 Allergy Reaction Unknown	N	100.00%
Action18 Prescribing Process Reviewed	N	100.00%

D1 View B
Model 1
Population Profile
Size: 147

Variables	Values	Probability
Action17 Reinstruction Of Patient Resident	N	100.00%
Actions Taken To Prevent Recurrence	<Not Specified>	100.00%
Factor4 Interference By Patient Resident	N	100.00%
Med Generic Name Ordered	<Not Specified>	100.00%
Factor30 Allergy Not Noted	N	99.32%
ID Documentation Consent17 Altered	N	99.32%
ID Documentation Consent16 Notes Unsigned	N	99.32%
ID Documentation Consent19 Consent Issue	N	99.32%
ID Documentation Consent12 Wrong Name	N	99.32%
ID Documentation Consent13 Patient Medicated Before Signing	N	99.32%
Factor34 Pump Infusion Settings	N	99.32%
Factor37 Prescribing Error	N	99.32%
ID Documentation Consent3 Documents Missing	N	99.32%
Notification Type3 Next Of Kin	N	99.32%
Factor16 Medication On Hold	N	99.32%
Factor10 Failure To Discontinue Hold Med	N	99.32%
Factor23 Misinterpretation Of Label	N	99.32%
Factor28MAR Unclear	N	99.32%
Notification Type8 Administrator	N	99.32%
ID Documentation Consent7 Inappropriate	N	98.64%
Notification Type9 Director	N	98.64%
ID Documentation Consent14 Incorrect Requisition	N	98.64%
ID Documentation Consent15 Incomplete Requisition	N	98.64%
Factor12 Order Not Pulled	N	98.64%
Action14 Equipment Supplies Reviewed	N	98.64%
Factor32 Patient Residents Own Medication	N	98.64%
ID Documentation Consent8 Illegible	N	98.64%
Factor25 Illegible Handwriting	N	98.64%
Factor29 Order Processing Delay	N	98.64%
Factor26 Medication Unavailable	N	98.64%
Factor33 Equipment Supplies Faulty	N	98.64%
ID Documentation Consent11 Medical Clearance Not Documented	N	97.96%
Factor21 Order Error	N	97.96%
ID Documentation Consent18 Documents Unsigned	N	97.96%
Factor24 Rate Dose Calculation Error	N	97.28%
Injury Incurred	No	97.28%
Factor6 Patient Resident Identification	N	97.28%
Factor31 Medication Discontinued	N	97.28%
Injury Nature	<Not Specified>	97.28%
ID Documentation Consent10MR Nwrong	N	97.28%
Injury Degree	<Not Specified>	97.28%

D1 View B
Model 1
Population Profile
Size: 147

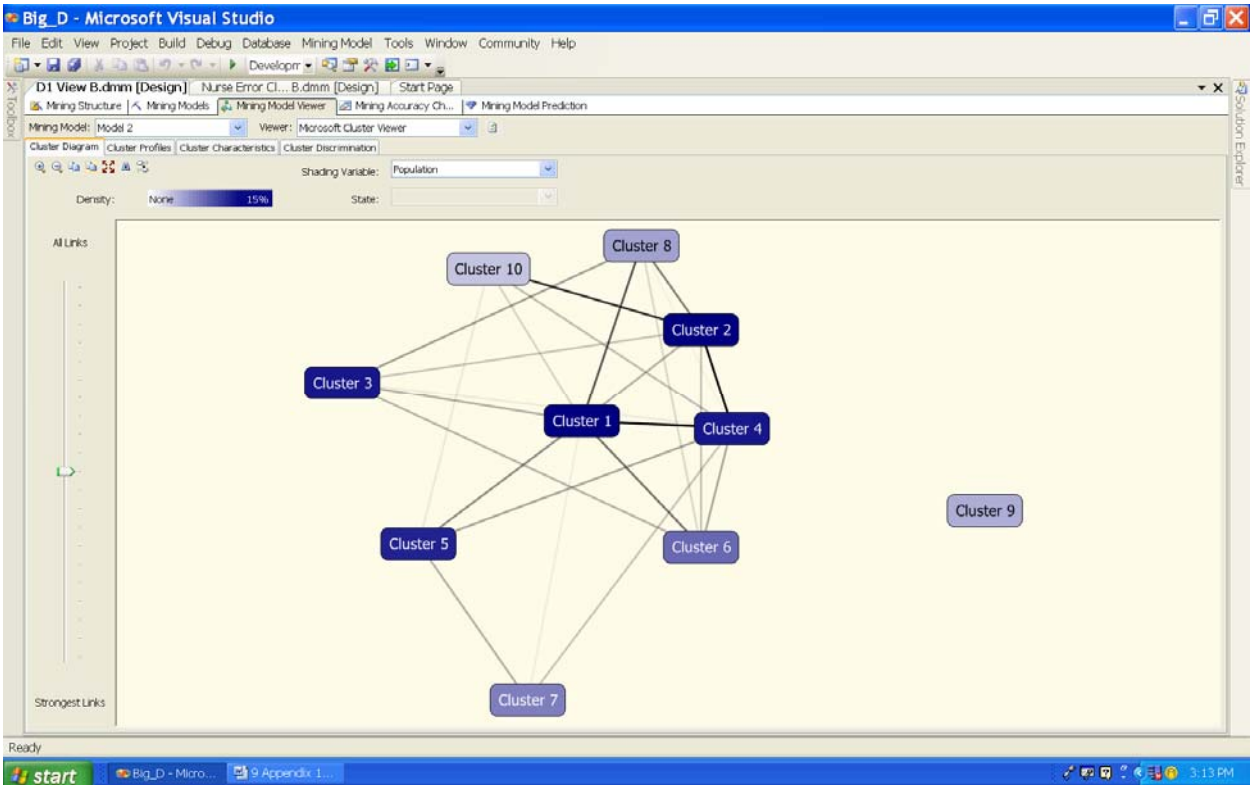
Variables	Values	Probability
ID Documentation Consent9 Absent	N	96.60%
Action12 Documentation Process Reviewed	N	96.60%
Factor2 Administration Not Recorded Signed Off	N	96.60%
Factor36 Dispensing Delay	N	96.60%
Factor9 Documentation Error	N	96.60%
Factor14MAR Misinterpretation	N	95.92%
Action13 Dispensing Process Reviewed	N	95.92%
Factor27 Administration Delay	N	95.92%
Action9 Order Process Reviewed	N	95.92%
Factor8 Policy Procedure Issue	N	95.24%
Action11 Other	N	95.24%
ID Documentation Consent6 Documents Delayed	N	95.24%
Factor3 Not Applicable	N	95.24%
Notification Type6 Other	N	94.56%
Factor17 Pharmacy Order Processing Error	N	93.88%
Factor22 Dispensing Error	N	93.88%
Action5 Policy Procedure Reviewed	N	93.20%
Factor15 Misinterpretation Of Order	N	92.52%
Factor20 Incorrect Preparation	N	92.52%
Action6 Reinstruction Of Staff	N	92.52%
Notification Type5 Pharmacist	N	91.84%
ID Documentation Consent2 Wrong Patient	N	91.16%
Action10 Treatment Provided	N	91.16%
Factor11 Missed During Chart Checks	N	91.16%
Factor5 Other	N	90.48%
ID Documentation Consent4 Policy Procedure Not Followed	N	90.48%
Action7 Administration Process Reviewed	N	90.48%
Factor18 Transcription Error	N	90.48%
Notification Type7 Supervisor	N	89.80%
Quantity	1	89.80%
ID Documentation Consent5 Transcription Error	N	88.44%
Notification Type2 Manager	N	87.76%
Factor7 Incorrect Administration	N	87.76%
Action8 Medication Reviewed	N	87.76%
Action3 Pharmacist Notified	N	87.76%
Action2 Monitor Patient Resident	N	86.40%
Factor13 Administration Error	N	83.67%
Notification Type4 Physician	N	79.59%
Med Strength Administered	<Not Specified>	77.55%
Strength And Units	missing	74.83%
Dosage Form	missing	72.79%
Volume And Units	missing	72.79%

D1 View B
Model 1
Population Profile
Size: 147

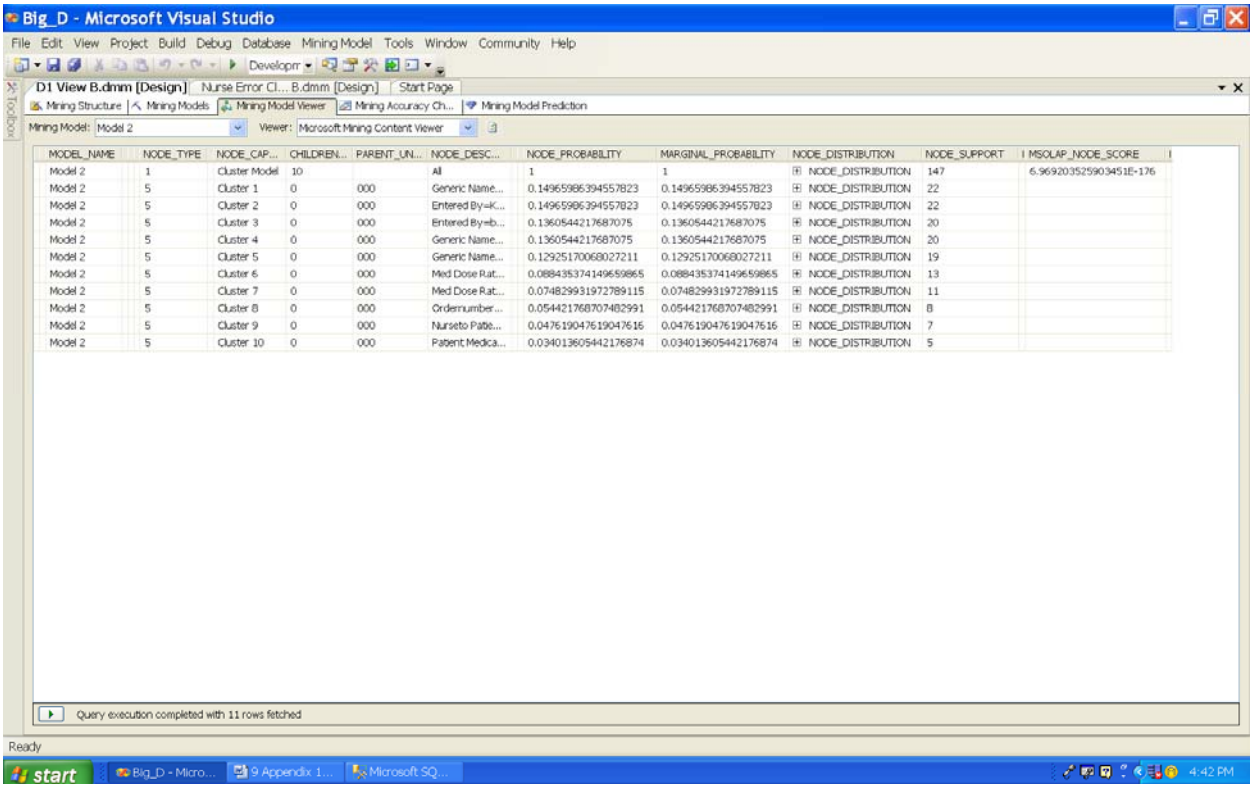
Variables	Values	Probability
Action4 Physician Notified	N	68.71%
ID Documentation Consent1 Not Specified	Y	68.03%
Med Strength Ordered	<Not Specified>	66.67%
Incident Severity Level	Severity Level 1	62.59%
Sex	F	59.18%
Med Admin Route Administered	<Not Specified>	56.46%
Notification Type1 Not Specified	Y	53.74%
Action1 Not Specified	N	52.38%
Med Dose Rate Conc Administered	<Not Specified>	51.02%
Patient Received Medication	Yes	50.34%

Appendix 5

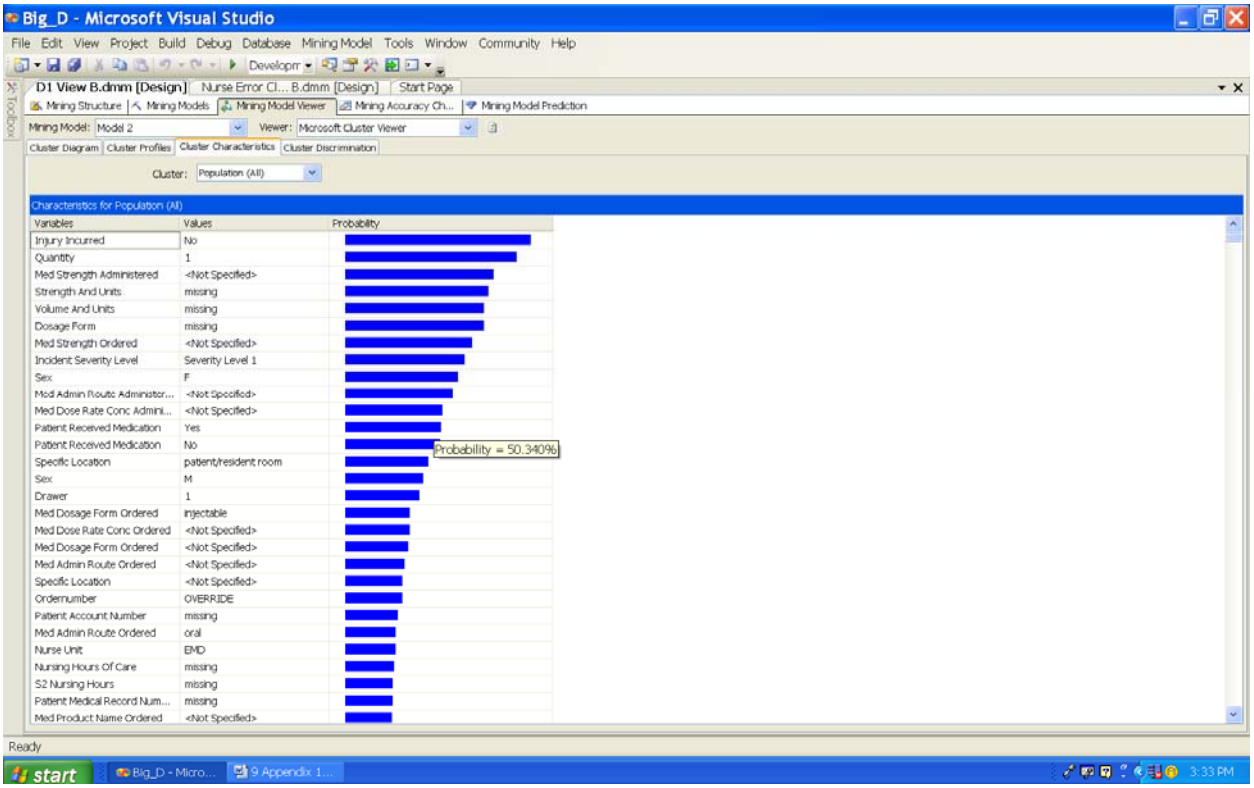
Cluster Diagram – Model 2



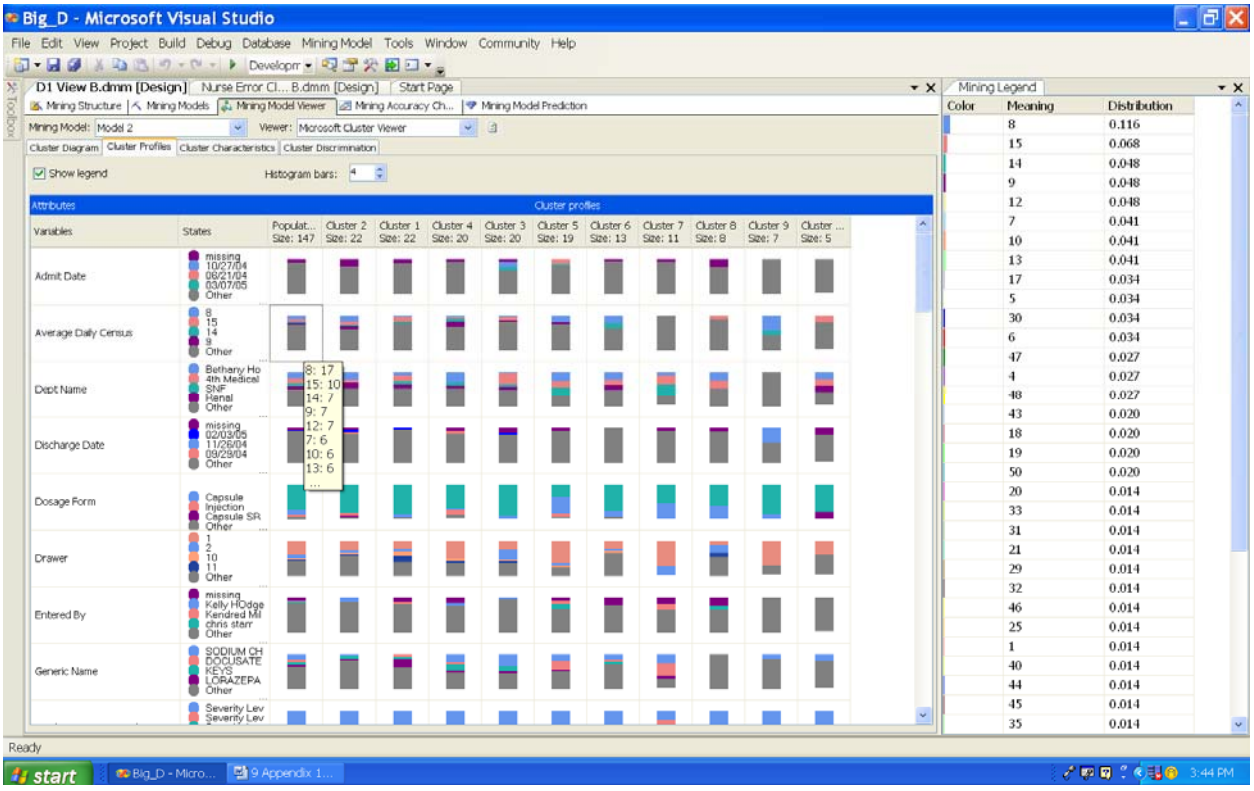
Cluster Summary – Model 2

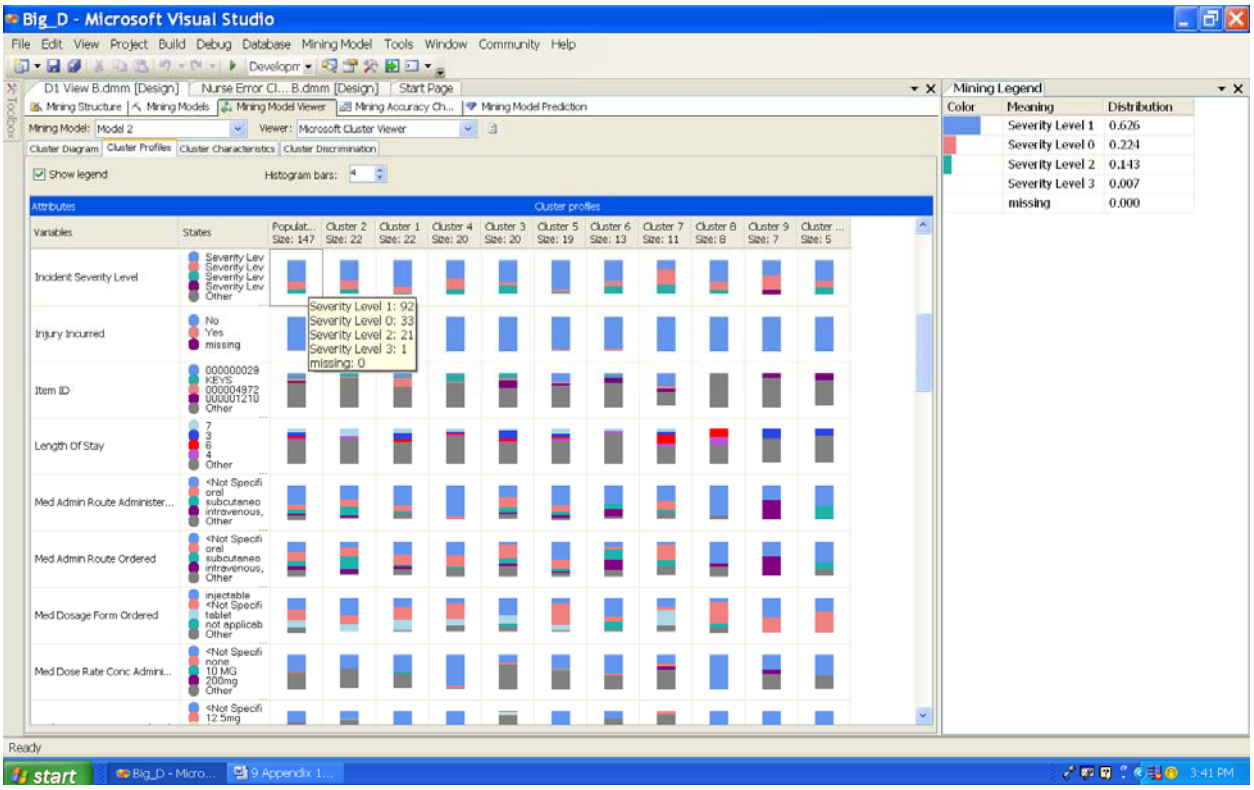


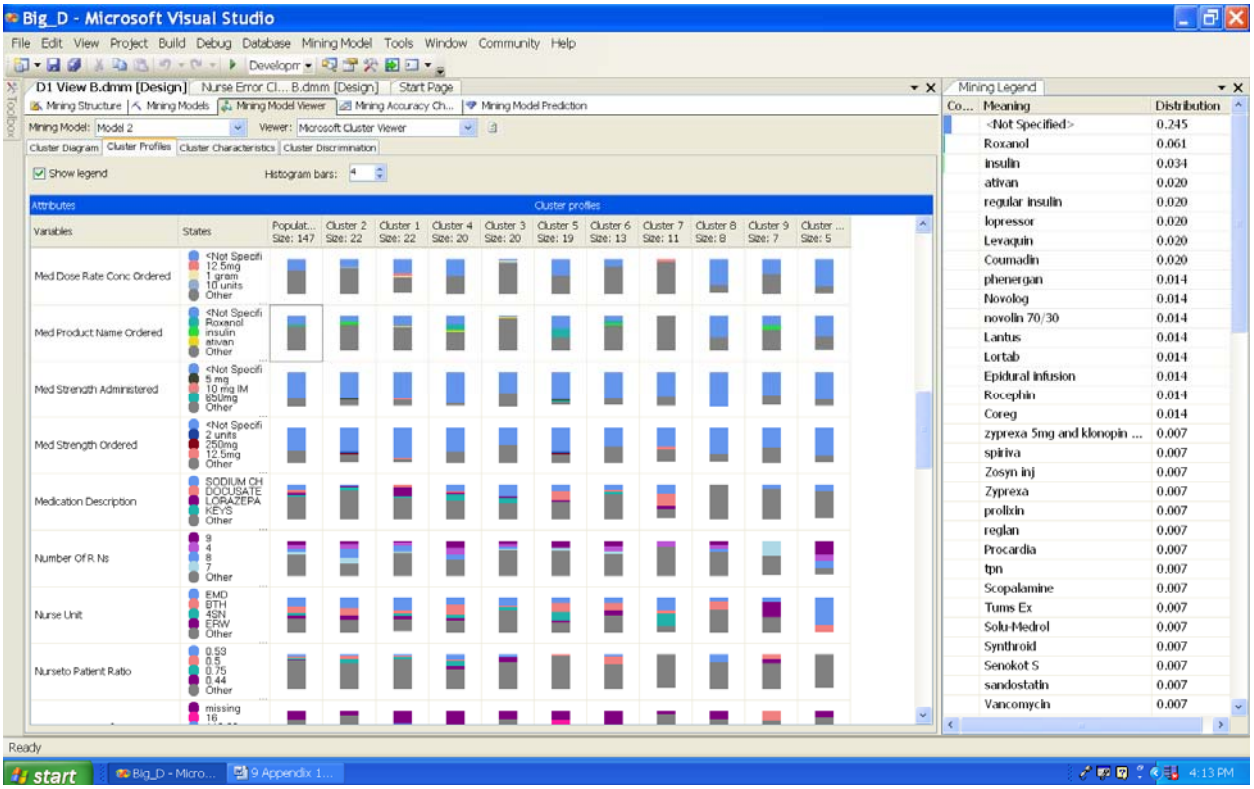
Cluster Results – Model 2

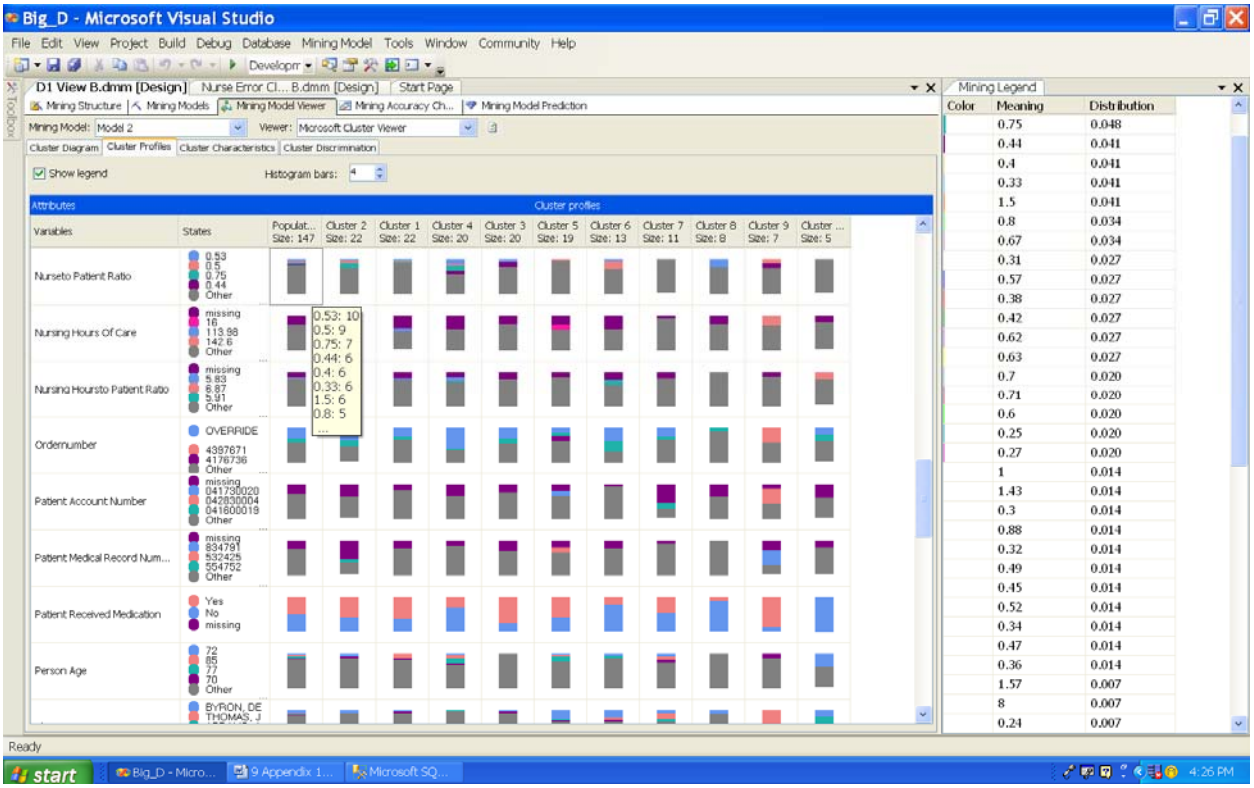


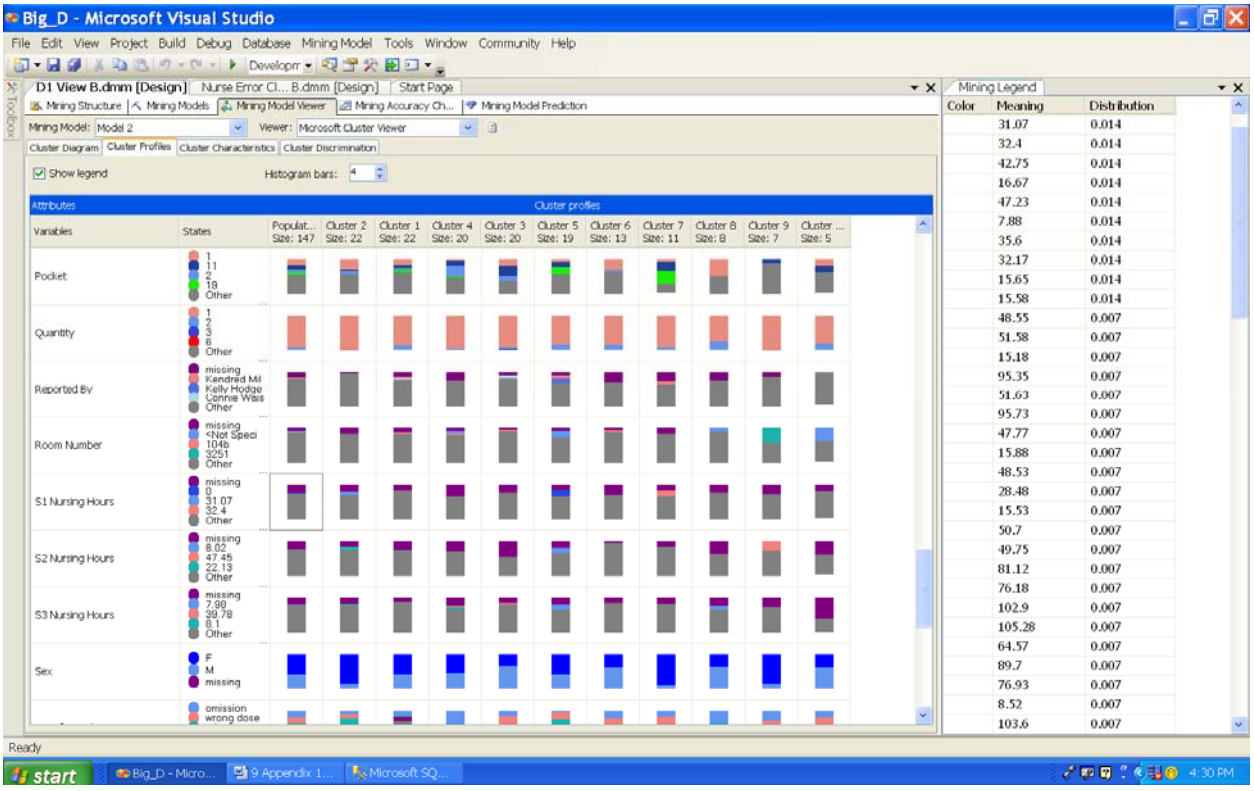
D1 View B **Model 2** **Population Profile** **Size: 147**		
Variables	**Values**	**Probability**
Injury Incurred	No	97.28%
Quantity	1	89.80%
Med Strength Administered	<Not Specified>	77.55%
Strength And Units	missing	74.83%
Volume And Units	missing	72.79%
Dosage Form	missing	72.79%
Med Strength Ordered	<Not Specified>	66.67%
Incident Severity Level	Severity Level 1	62.59%
Sex	F	59.18%
Med Admin Route Administered	<Not Specified>	56.46%
Med Dose Rate Conc Administered	<Not Specified>	51.02%
Patient Received Medication	Yes	50.34%

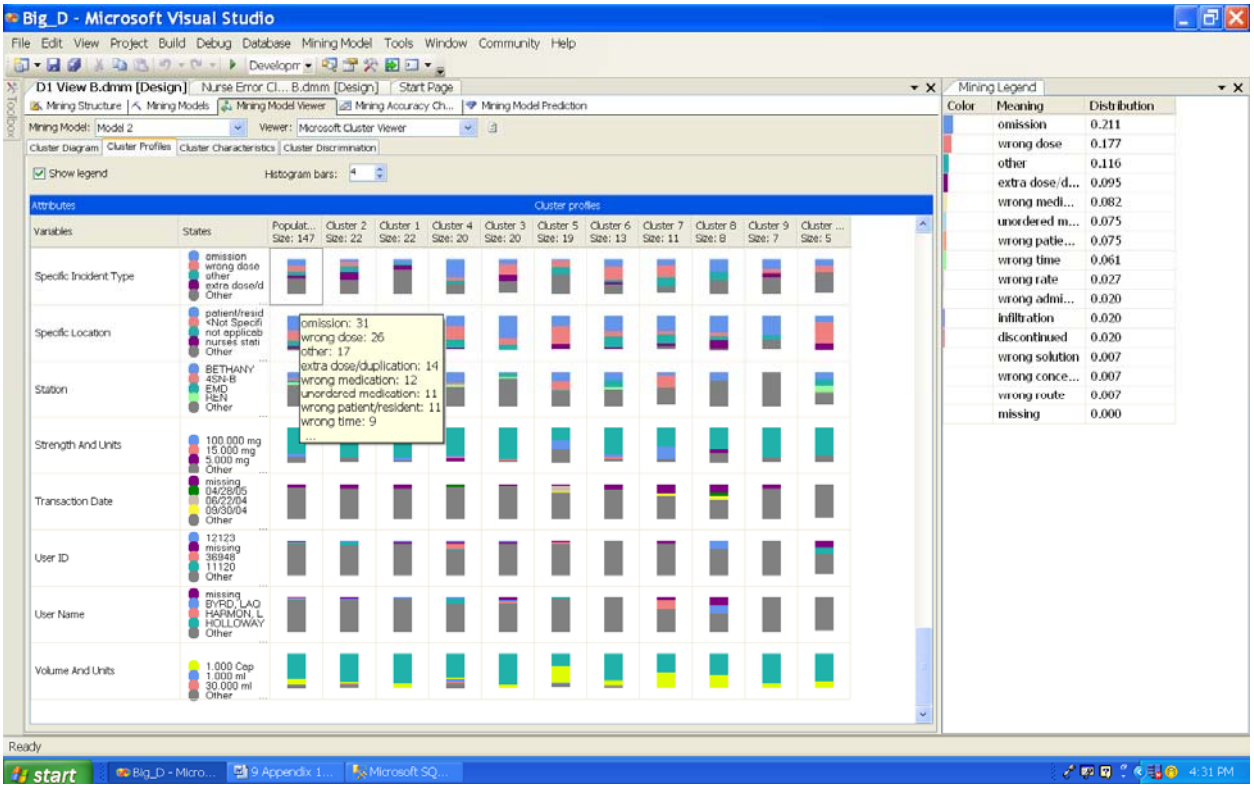

Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Admit Date
Average Daily Census
Dept Name
Discharge Date
Dosage Form
Drawer
Entered By
Generic Name
8: 17
15: 10
14: 7
9: 7
12: 7
7: 6
10: 6
13: 6
Mining Legend
Color
Meaning
Distribution
8 0.116
15 0.068
14 0.048
9 0.048
12 0.048
7 0.041
10 0.041
13 0.041
17 0.034
5 0.034
30 0.034
6 0.034
47 0.027
4 0.027
48 0.027
43 0.020
18 0.020
19 0.020
50 0.020
20 0.014
33 0.014
31 0.014
21 0.014
29 0.014
32 0.014
46 0.014
25 0.014
1 0.014
40 0.014
44 0.014
45 0.014
35 0.014
Ready
start
3:44 PM

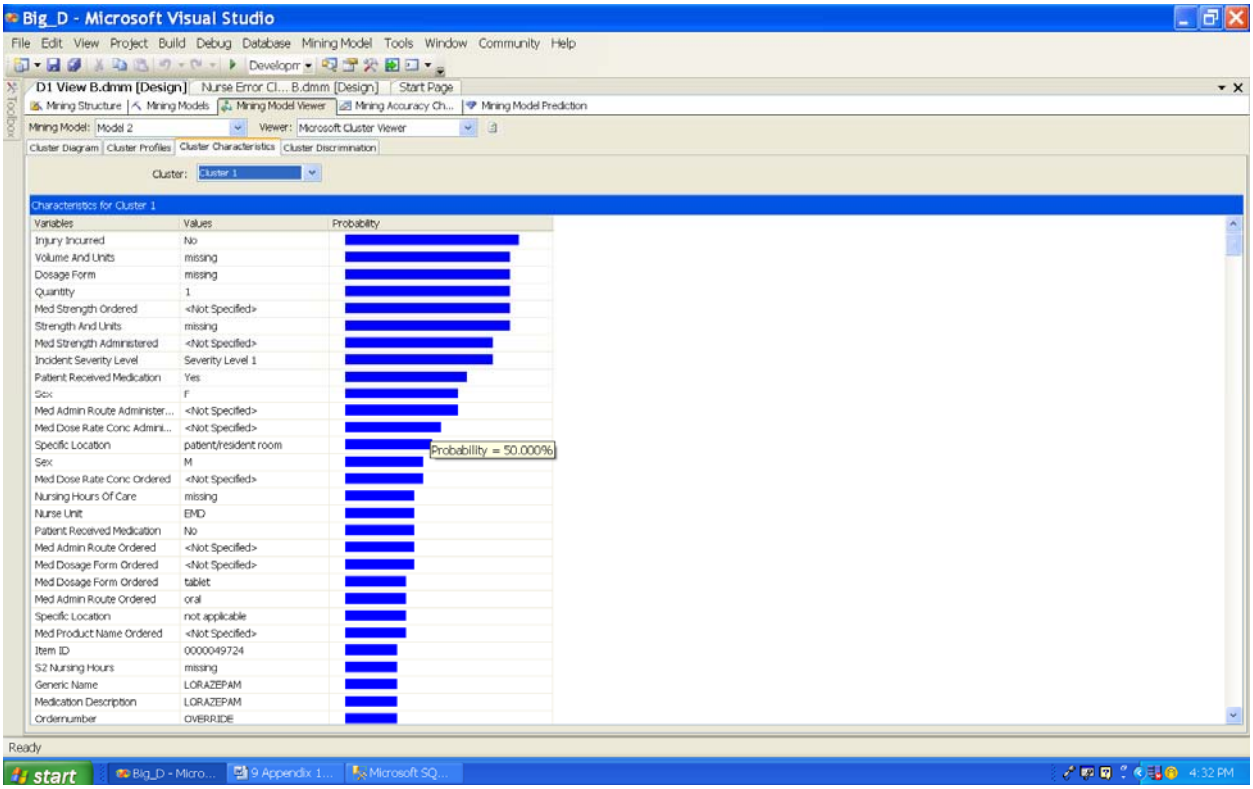

Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Incident Severity Level
Injury Incurred
Item ID
Length Of Stay
Med Admin Route Administer...
Med Admin Route Ordered
Med Dosage Form Ordered
Med Dose Rate Conc Admini...
Severity Level 1: 92
Severity Level 0: 33
Severity Level 2: 21
Severity Level 3: 1
missing: 0
Mining Legend
Color
Meaning
Distribution
Severity Level 1 0.626
Severity Level 0 0.224
Severity Level 2 0.143
Severity Level 3 0.007
missing 0.000
Ready
start
3:41 PM


Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Med Dose Rate Conc Ordered
Med Product Name Ordered
Med Strength Administered
Med Strength Ordered
Medication Description
Number Of R Ns
Nurse Unit
Nurseto Patient Ratio
Mining Legend
Co...
Meaning
Distribution
<Not Specified> 0.245
Roxanol 0.061
insulin 0.034
ativan 0.020
regular insulin 0.020
lopressor 0.020
Levaquin 0.020
Coumadin 0.020
phenergan 0.014
Novolog 0.014
novolin 70/30 0.014
Lantus 0.014
Lortab 0.014
Epidural infusion 0.014
Rocephin 0.014
Coreg 0.014
zyprexa 5mg and klonopin ... 0.007
spiriva 0.007
Zosyn inj 0.007
Zyprexa 0.007
prolixin 0.007
reglan 0.007
Procardia 0.007
tpn 0.007
Scopalamine 0.007
Tums Ex 0.007
Solu-Medrol 0.007
Synthroid 0.007
Senokot S 0.007
sandostatin 0.007
Vancomycin 0.007
Ready
start
4:13 PM

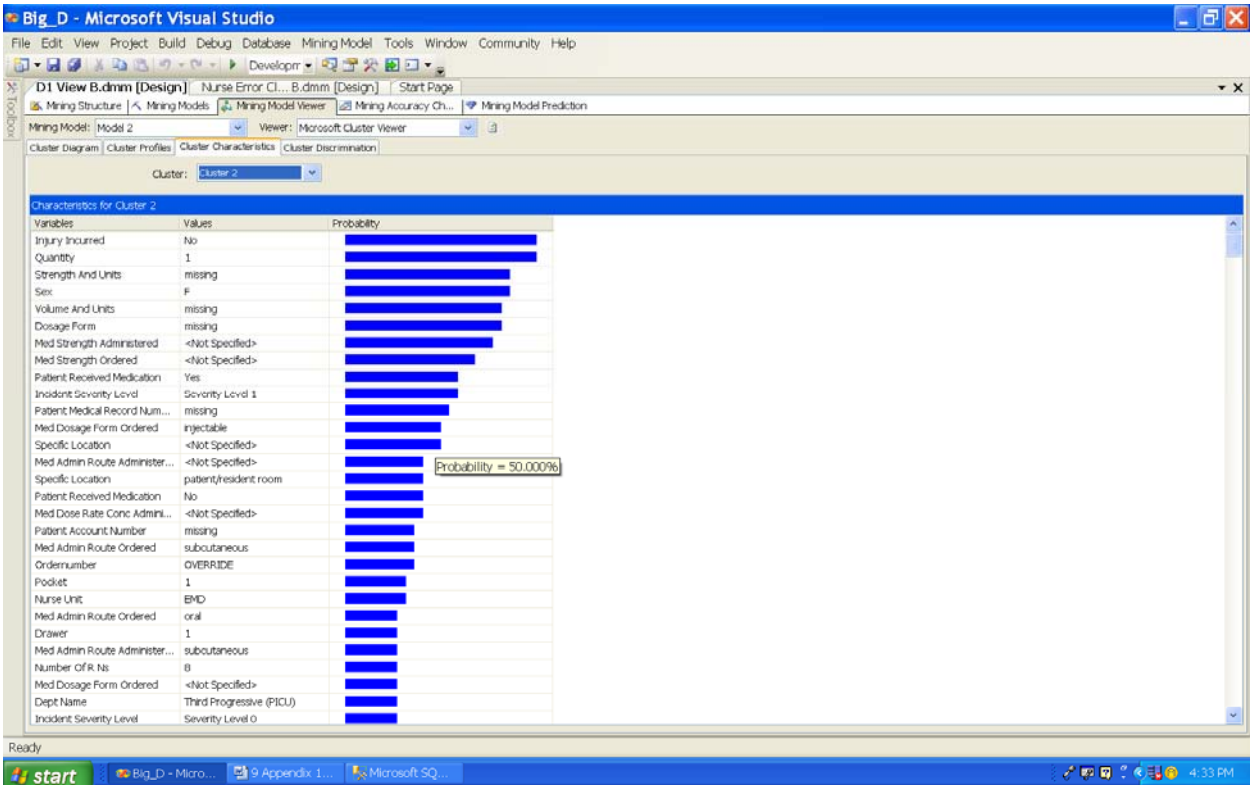

Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Nurseto Patient Ratio
Nursing Hours Of Care
Nursing Hoursto Patient Ratio
Ordernumber
Patient Account Number
Patient Medical Record Num...
Patient Received Medication
Person Age
0.53: 10
0.5: 9
0.75: 7
0.44: 6
0.4: 6
0.33: 6
1.5: 6
0.8: 5
Mining Legend
Color
Meaning
Distribution
0.75 0.048
0.44 0.041
0.4 0.041
0.33 0.041
1.5 0.041
0.8 0.034
0.67 0.034
0.31 0.027
0.57 0.027
0.38 0.027
0.42 0.027
0.62 0.027
0.63 0.027
0.7 0.020
0.71 0.020
0.6 0.020
0.25 0.020
0.27 0.020
1 0.014
1.43 0.014
0.3 0.014
0.88 0.014
0.32 0.014
0.49 0.014
0.45 0.014
0.52 0.014
0.34 0.014
0.47 0.014
0.36 0.014
1.57 0.007
8 0.007
0.24 0.007
Ready
start
4:26 PM


Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Populat... Size: 147
Cluster 2 Size: 22
Cluster 1 Size: 22
Cluster 4 Size: 20
Cluster 3 Size: 20
Cluster 5 Size: 19
Cluster 6 Size: 13
Cluster 7 Size: 11
Cluster 8 Size: 8
Cluster 9 Size: 7
Cluster ... Size: 5
Pocket
Quantity
Reported By
Room Number
S1 Nursing Hours
S2 Nursing Hours
S3 Nursing Hours
Sex
Mining Legend
Color
Meaning
Distribution
31.07 0.014
32.4 0.014
42.75 0.014
16.67 0.014
47.23 0.014
7.88 0.014
35.6 0.014
32.17 0.014
15.65 0.014
15.58 0.014
48.55 0.007
51.58 0.007
15.18 0.007
95.35 0.007
51.63 0.007
95.73 0.007
47.77 0.007
15.88 0.007
48.53 0.007
28.48 0.007
15.53 0.007
50.7 0.007
49.75 0.007
81.12 0.007
76.18 0.007
102.9 0.007
105.28 0.007
64.57 0.007
89.7 0.007
76.93 0.007
8.52 0.007
103.6 0.007
Ready
start
4:30 PM


Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Specific Incident Type
Specific Location
Station
Strength And Units
Transaction Date
User ID
User Name
Volume And Units
omission: 31
wrong dose: 26
other: 17
extra dose/duplication: 14
wrong medication: 12
unordered medication: 11
wrong patient/resident: 11
wrong time: 9
Mining Legend
Color
Meaning
Distribution
omission 0.211
wrong dose 0.177
other 0.116
extra dose/d... 0.095
wrong medi... 0.082
unordered m... 0.075
wrong patie... 0.075
wrong time 0.061
wrong rate 0.027
wrong admi... 0.020
infiltration 0.020
discontinued 0.020
wrong solution 0.007
wrong conce... 0.007
wrong route 0.007
missing 0.000
Ready
start
4:31 PM

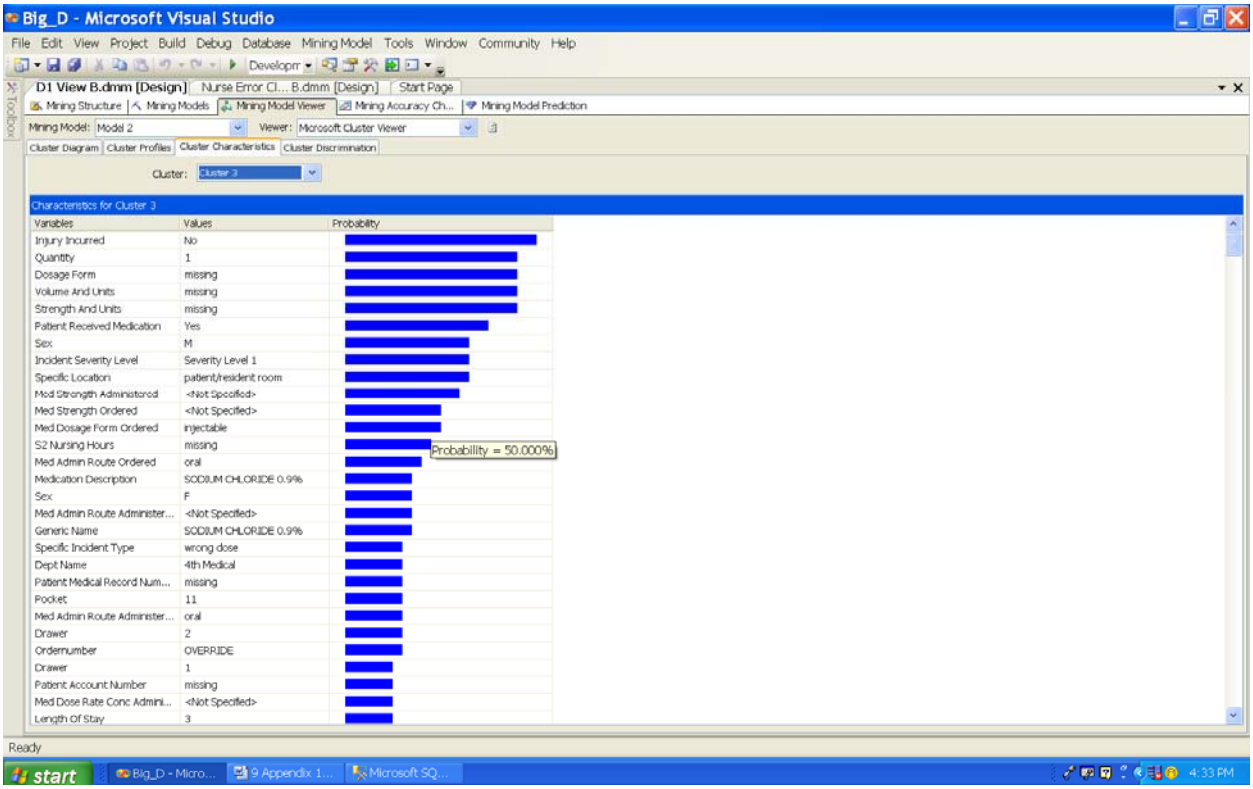


D1 View B **Model 2** **Cluster 1 Profile** **Size: 22**		
Variables	**Values**	**Probability**
Injury Incurred	No	90.91%
Volume And Units	missing	86.36%
Dosage Form	missing	86.36%
Quantity	1	86.36%
Med Strength Ordered	<Not Specified>	86.36%
Strength And Units	missing	86.36%
Med Strength Administered	<Not Specified>	77.27%
Incident Severity Level	Severity Level 1	77.27%
Patient Received Medication	Yes	63.64%
Sex	F	59.09%
Med Admin Route Administered	<Not Specified>	59.09%
Med Dose Rate Conc Administered	<Not Specified>	50.00%

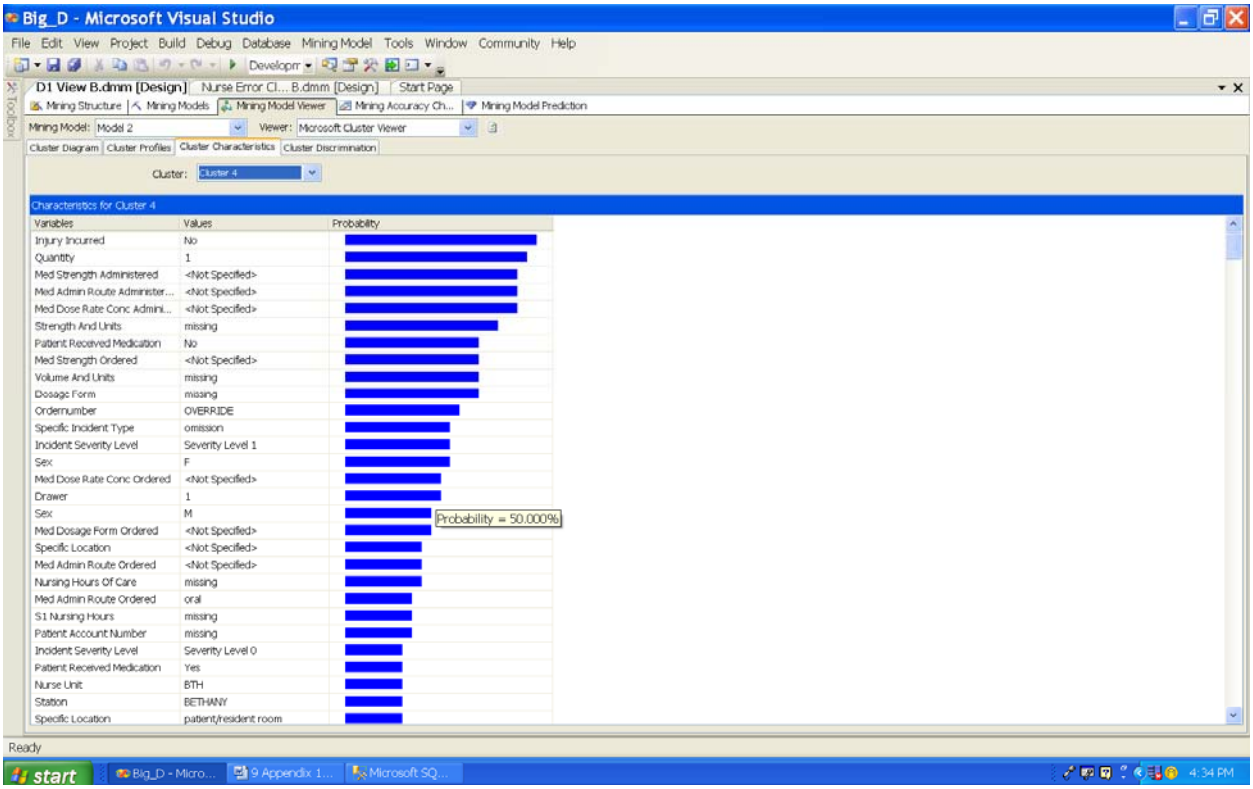


D1 View B
Model 2
Cluster 2 Profile
Size: 22

Variables	Values	Probability
Injury Incurred	No	100.00%
Quantity	1	100.00%
Strength And Units	missing	86.36%
Sex	F	86.36%
Volume And Units	missing	81.82%
Dosage Form	missing	81.82%
Med Strength Administered	<Not Specified>	77.27%
Med Strength Ordered	<Not Specified>	68.18%
Patient Received Medication	Yes	59.09%
Incident Severity Level	Severity Level 1	59.09%
Patient Medical Record Number	missing	54.55%
Med Dosage Form Ordered	injectable	50.00%
Specific Location	<Not Specified>	50.00%

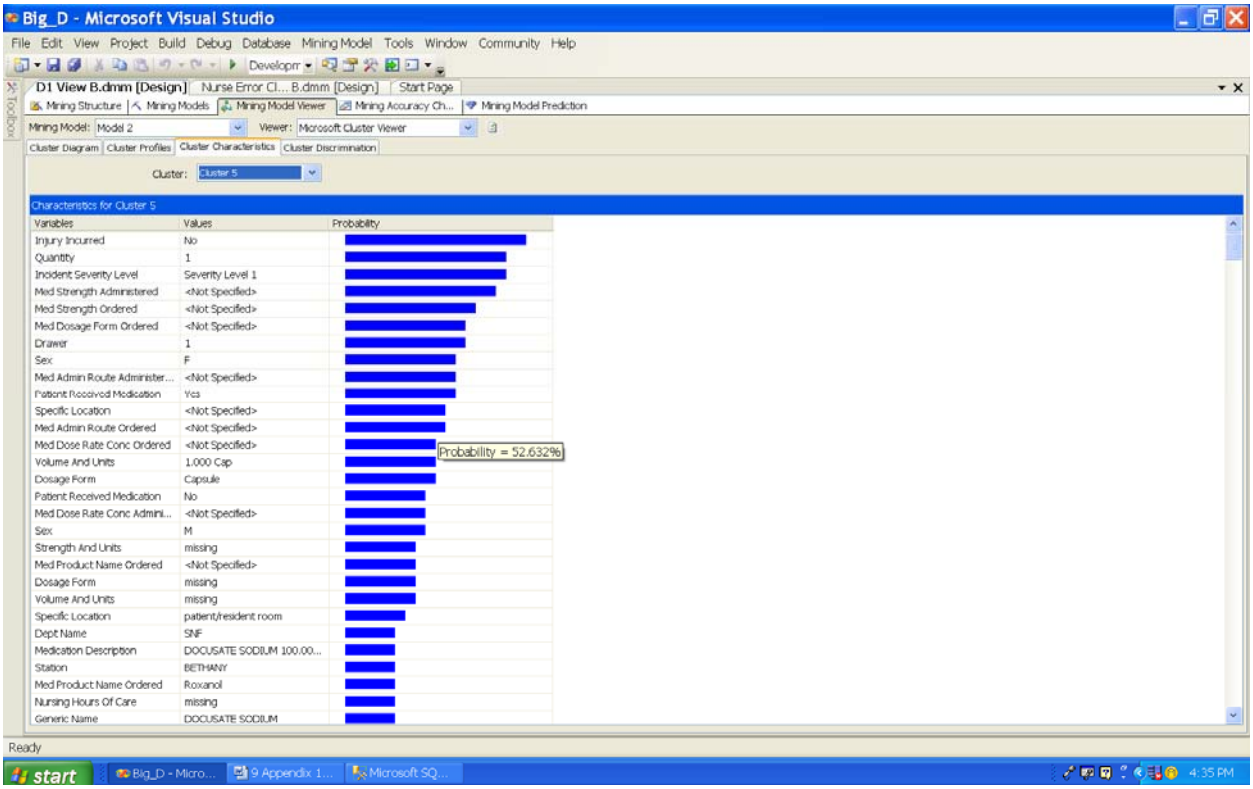


D1 View B **Model 2** **Cluster 3 Profile** **Size: 20**		
Variables	**Values**	**Probability**
Injury Incurred	No	100.00%
Quantity	1	90.00%
Dosage Form	missing	90.00%
Volume And Units	missing	90.00%
Strength And Units	missing	90.00%
Patient Received Medication	Yes	75.00%
Sex	M	65.00%
Incident Severity Level	Severity Level 1	65.00%
Specific Location	patient/resident room	65.00%
Med Strength Administered	<Not Specified>	60.00%
Med Strength Ordered	<Not Specified>	50.00%
Med Dosage Form Ordered	injectable	50.00%



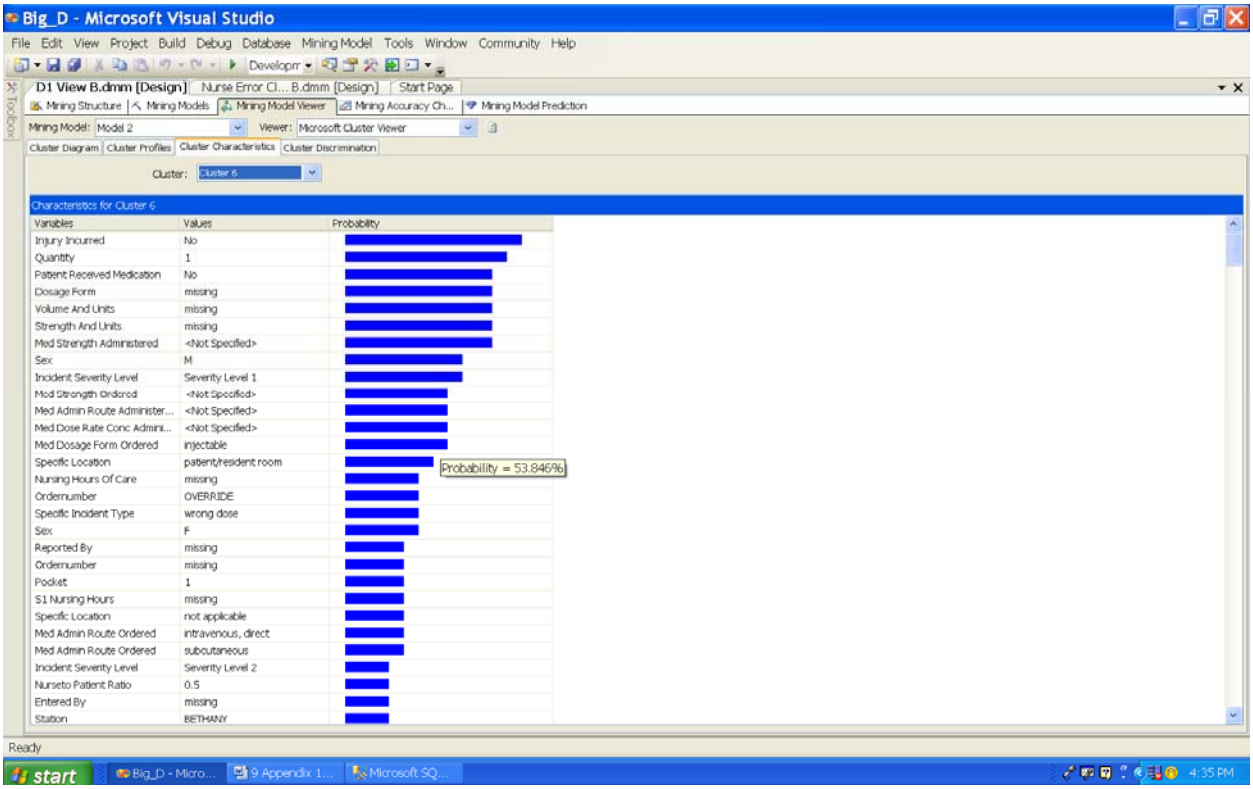
D1 View B
Model 2
Cluster 4 Profile
Size: 20

Variables	Values	Probability
Injury Incurred	No	100.00%
Quantity	1	95.00%
Med Strength Administered	<Not Specified>	90.00%
Med Admin Route Administered	<Not Specified>	90.00%
Med Dose Rate Conc Administered	<Not Specified>	90.00%
Strength And Units	missing	80.00%
Patient Received Medication	No	70.00%
Med Strength Ordered	<Not Specified>	70.00%
Volume And Units	missing	70.00%
Dosage Form	missing	70.00%
Ordernumber	OVERRIDE	60.00%
Specific Incident Type	omission	55.00%
Incident Severity Level	Severity Level 1	55.00%
Sex	F	55.00%
Med Dose Rate Conc Ordered	<Not Specified>	50.00%
Drawer	1	50.00%



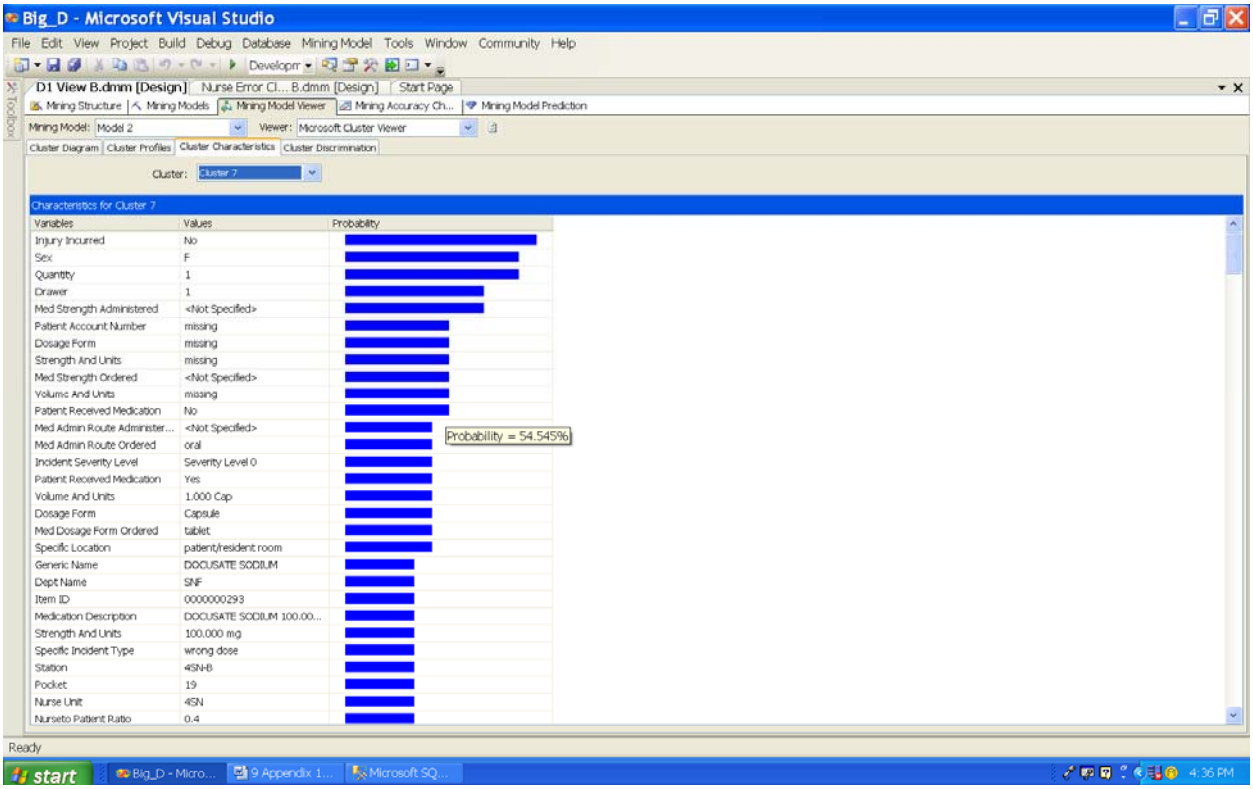
D1 View B
Model 2
Cluster 5 Profile
Size: 19

Variables	Values	Probability
Injury Incurred	No	94.74%
Quantity	1	84.21%
Incident Severity Level	Severity Level 1	84.21%
Med Strength Administered	<Not Specified>	78.95%
Med Strength Ordered	<Not Specified>	68.42%
Med Dosage Form Ordered	<Not Specified>	63.16%
Drawer	1	63.16%
Sex	F	57.90%
Med Admin Route Administered	<Not Specified>	57.90%
Patient Received Medication	Yes	57.90%
Specific Location	<Not Specified>	52.63%
Med Admin Route Ordered	<Not Specified>	52.63%



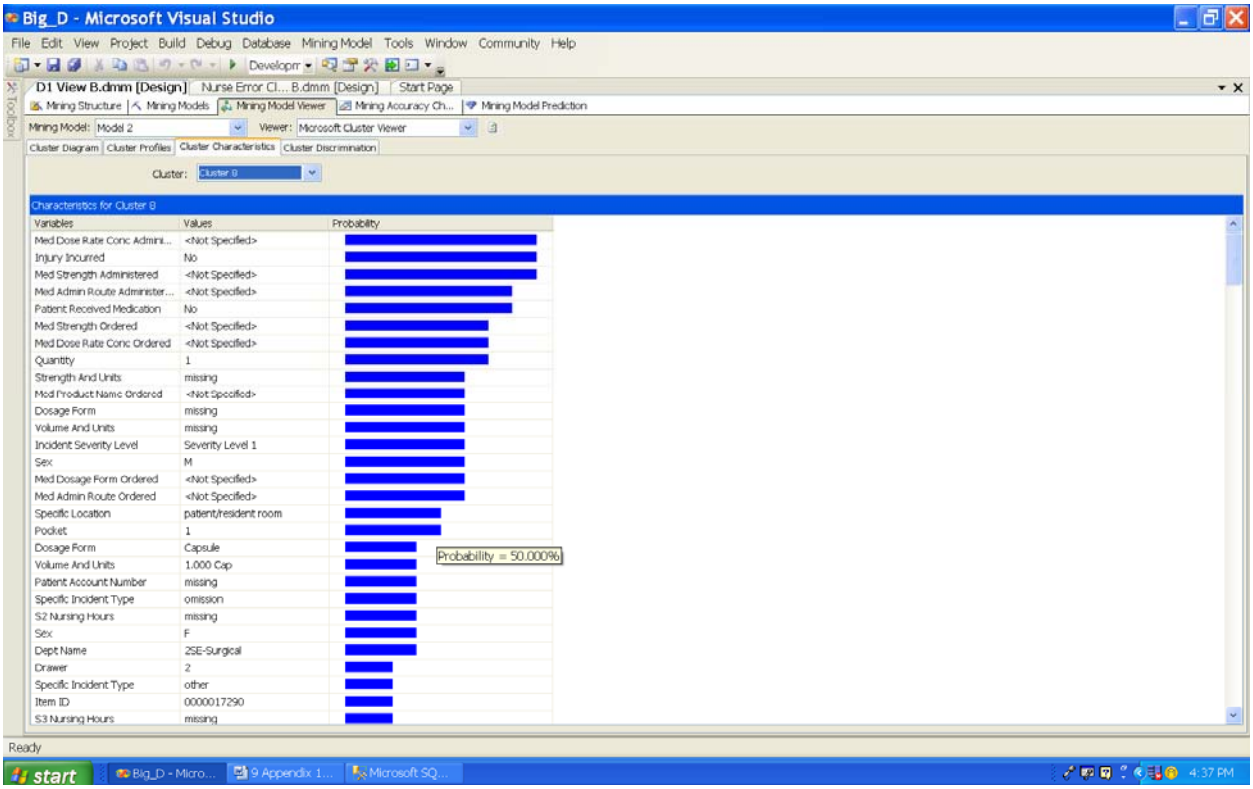
D1 View B
Model 2
Cluster 6 Profile
Size: 13

Variables	Values	Probability
Injury Incurred	No	92.31%
Quantity	1	84.62%
Patient Received Medication	No	76.92%
Dosage Form	missing	76.92%
Volume And Units	missing	76.92%
Strength And Units	missing	76.92%
Med Strength Administered	<Not Specified>	76.92%
Sex	M	61.54%
Incident Severity Level	Severity Level 1	61.54%
Med Strength Ordered	<Not Specified>	53.85%
Med Admin Route Administered	<Not Specified>	53.85%
Med Dose Rate Conc Administered	<Not Specified>	53.85%
Med Dosage Form Ordered	injectable	53.85%

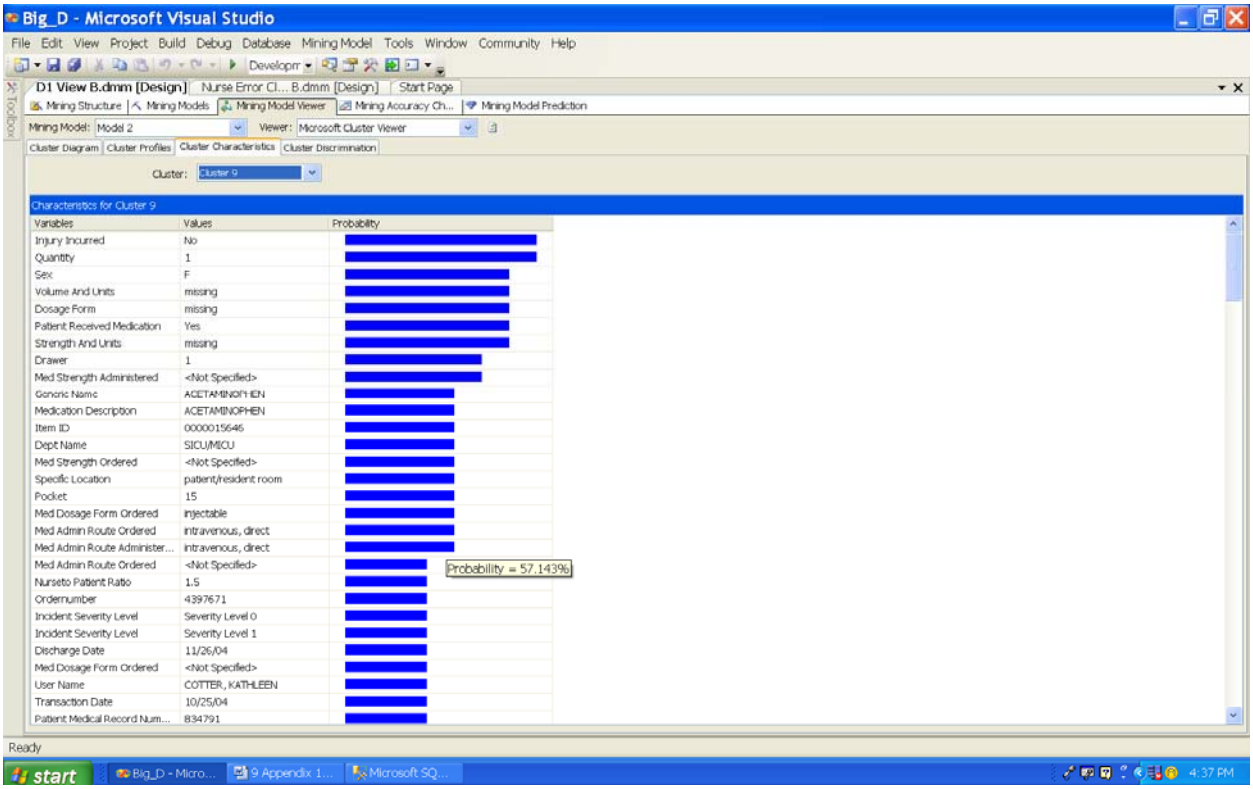


D1 View B
Model 2
Cluster 7 Profile
Size:11

Variables	Values	Probability
Injury Incurred	No	100.00%
Sex	F	90.91%
Quantity	1	90.91%
Drawer	1	72.73%
Med Strength Administered	<Not Specified>	72.73%
Patient Account Number	missing	54.55%
Dosage Form	missing	54.55%
Strength And Units	missing	54.55%
Med Strength Ordered	<Not Specified>	54.55%
Volume And Units	missing	54.55%
Patient Received Medication	No	54.55%

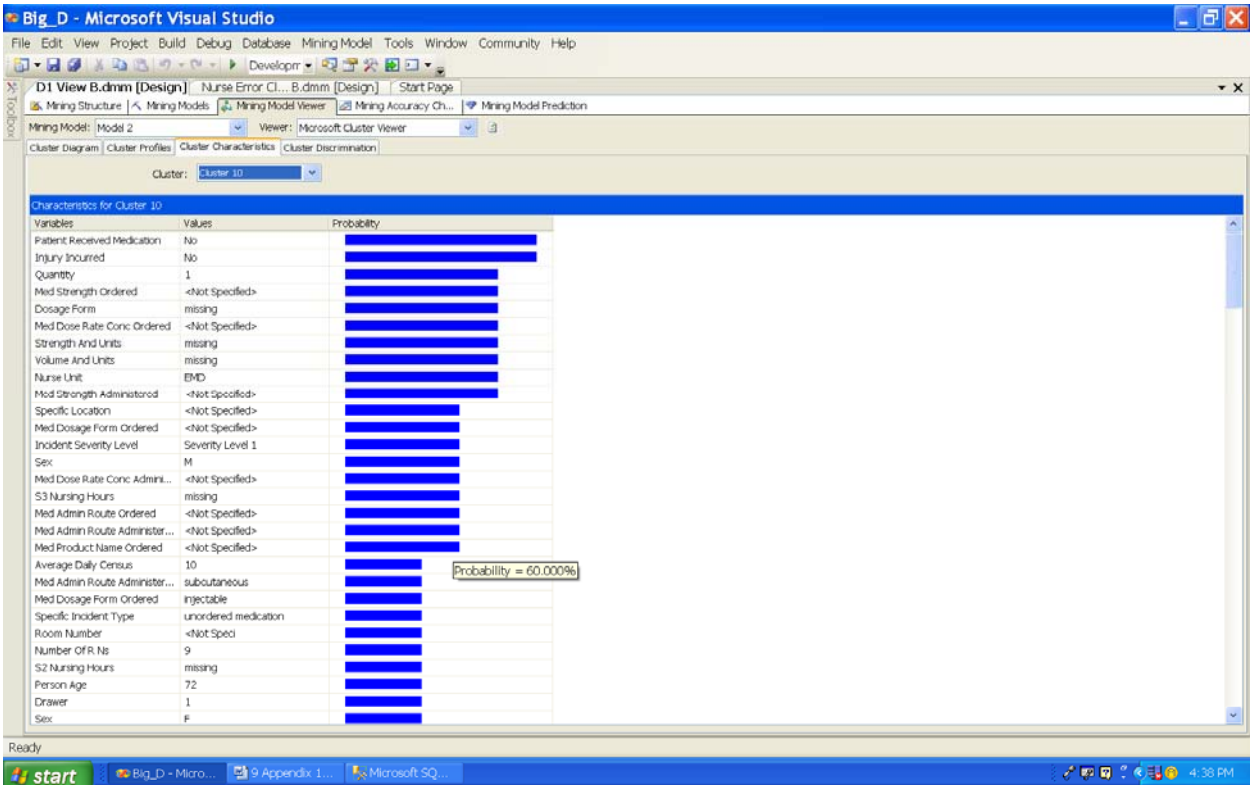


D1 View B **Model 2** **Cluster 8 Profile** **Size: 8**		
Variables	**Values**	**Probability**
Med Dose Rate Conc Administered	<Not Specified>	100.00%
Injury Incurred	No	100.00%
Med Strength Administered	<Not Specified>	100.00%
Med Admin Route Administered	<Not Specified>	87.50%
Patient Received Medication	No	87.50%
Med Strength Ordered	<Not Specified>	75.00%
Med Dose Rate Conc Ordered	<Not Specified>	75.00%
Quantity	1	75.00%
Strength And Units	missing	62.50%
Med Product Name Ordered	<Not Specified>	62.50%
Dosage Form	missing	62.50%
Volume And Units	missing	62.50%
Incident Severity Level	Severity Level 1	62.50%
Sex	M	62.50%
Med Dosage Form Ordered	<Not Specified>	62.50%
Med Admin Route Ordered	<Not Specified>	62.50%
Specific Location	patient/resident room	50.00%
Pocket	1	50.00%



D1 View B
Model 2
Cluster 9 Profile
Size: 7

Variables	Values	Probability
Injury Incurred	No	100.00%
Quantity	1	100.00%
Sex	F	85.71%
Volume And Units	missing	85.71%
Dosage Form	missing	85.71%
Patient Received Medication	Yes	85.71%
Strength And Units	missing	85.71%
Drawer	1	71.43%
Med Strength Administered	<Not Specified>	71.43%
Generic Name	ACETAMINOPHEN	57.14%
Medication Description	ACETAMINOPHEN	57.14%
Item ID	15646	57.14%
Dept Name	SICU/MICU	57.14%
Med Strength Ordered	<Not Specified>	57.14%
Specific Location	patient/resident room	57.14%
Pocket	15	57.14%
Med Dosage Form Ordered	injectable	57.14%
Med Admin Route Ordered	intravenous, direct	57.14%
Med Admin Route Administered	intravenous, direct	57.14%



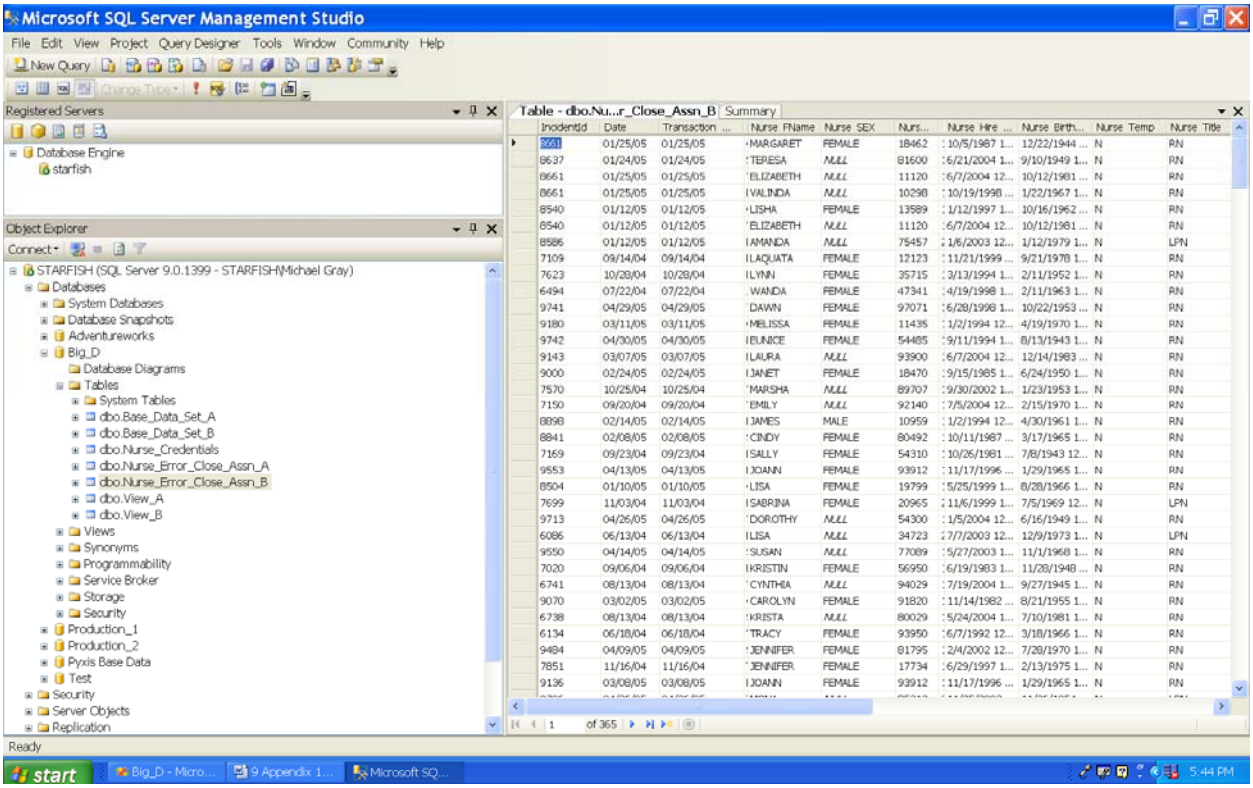
D1 View B
Model 2
Cluster 10 Profile
Size: 5

Variables	Values	Probability
Patient Received Medication	No	100.00%
Injury Incurred	No	100.00%
Quantity	1	80.00%
Med Strength Ordered	<Not Specified>	80.00%
Dosage Form	missing	80.00%
Med Dose Rate Conc Ordered	<Not Specified>	80.00%
Strength And Units	missing	80.00%
Volume And Units	missing	80.00%
Nurse Unit	EMD	80.00%
Med Strength Administered	<Not Specified>	80.00%
Specific Location	<Not Specified>	60.00%
Med Dosage Form Ordered	<Not Specified>	60.00%
Incident Severity Level	Severity Level 1	60.00%
Sex	M	60.00%
Med Dose Rate Conc Administered	<Not Specified>	60.00%
S3 Nursing Hours	missing	60.00%
Med Admin Route Ordered	<Not Specified>	60.00%
Med Admin Route Administered	<Not Specified>	60.00%

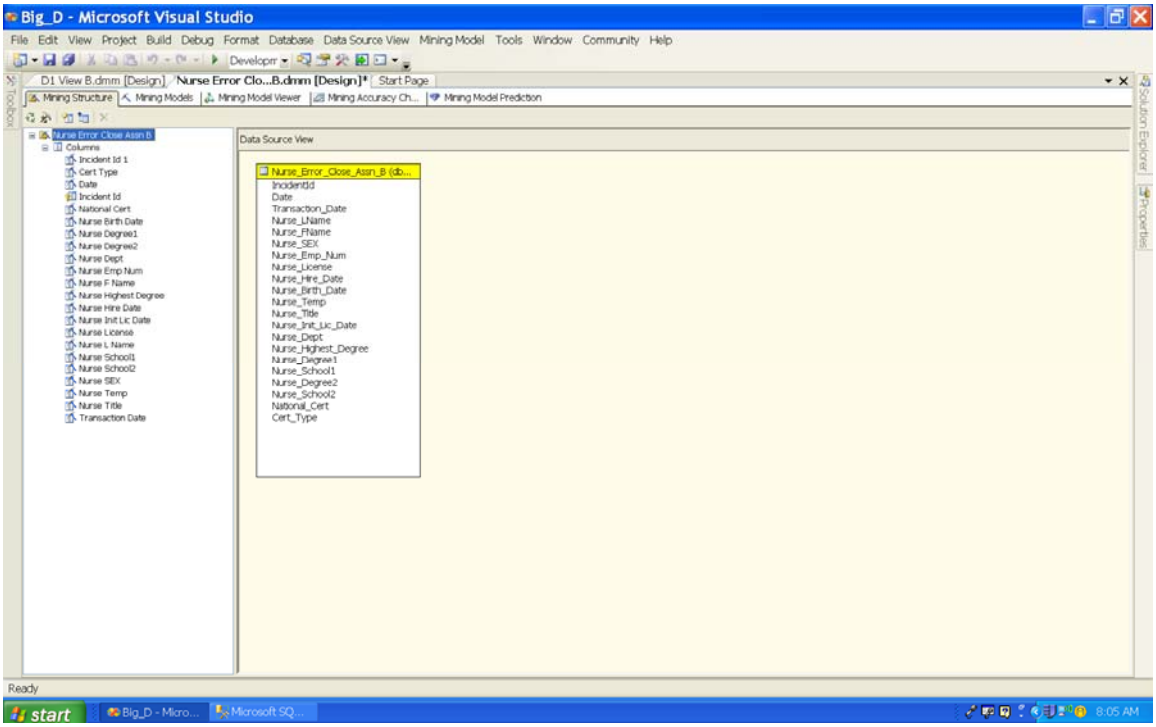
Appendix 6

Error Close Association Table

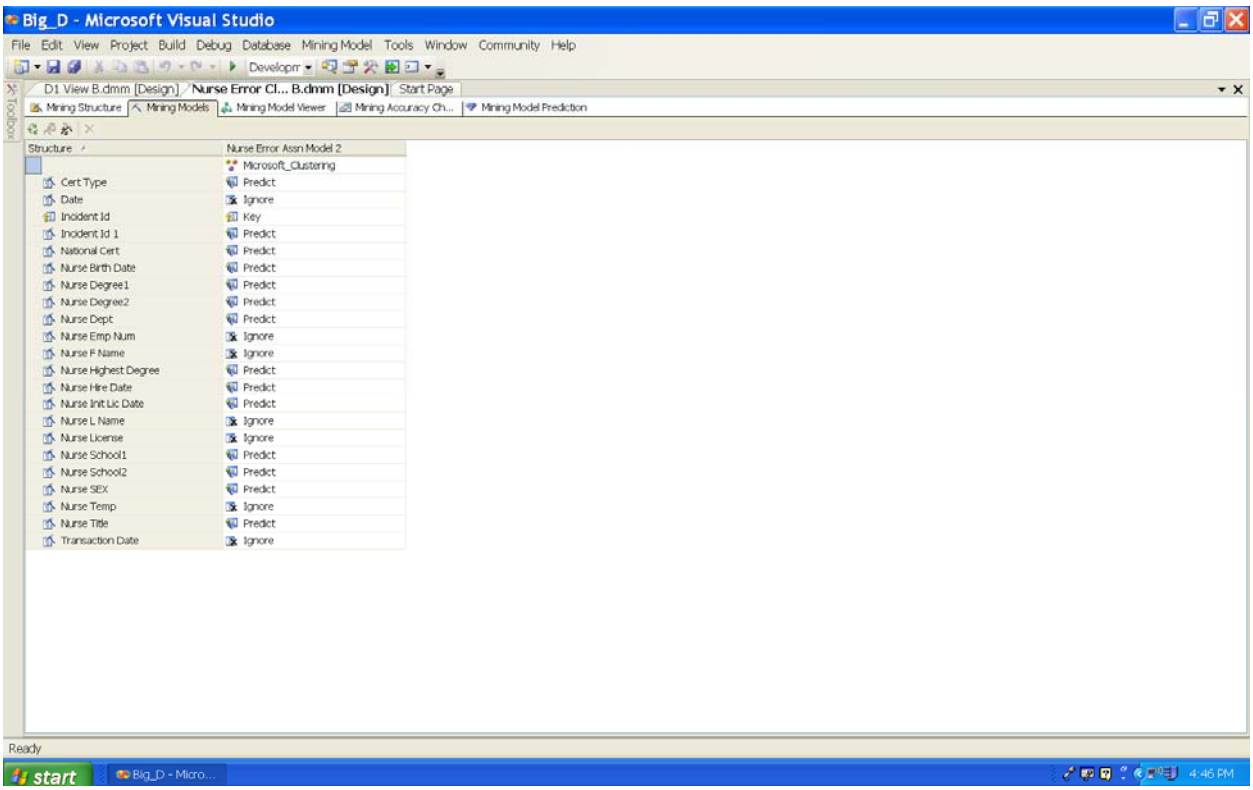
365 Rows



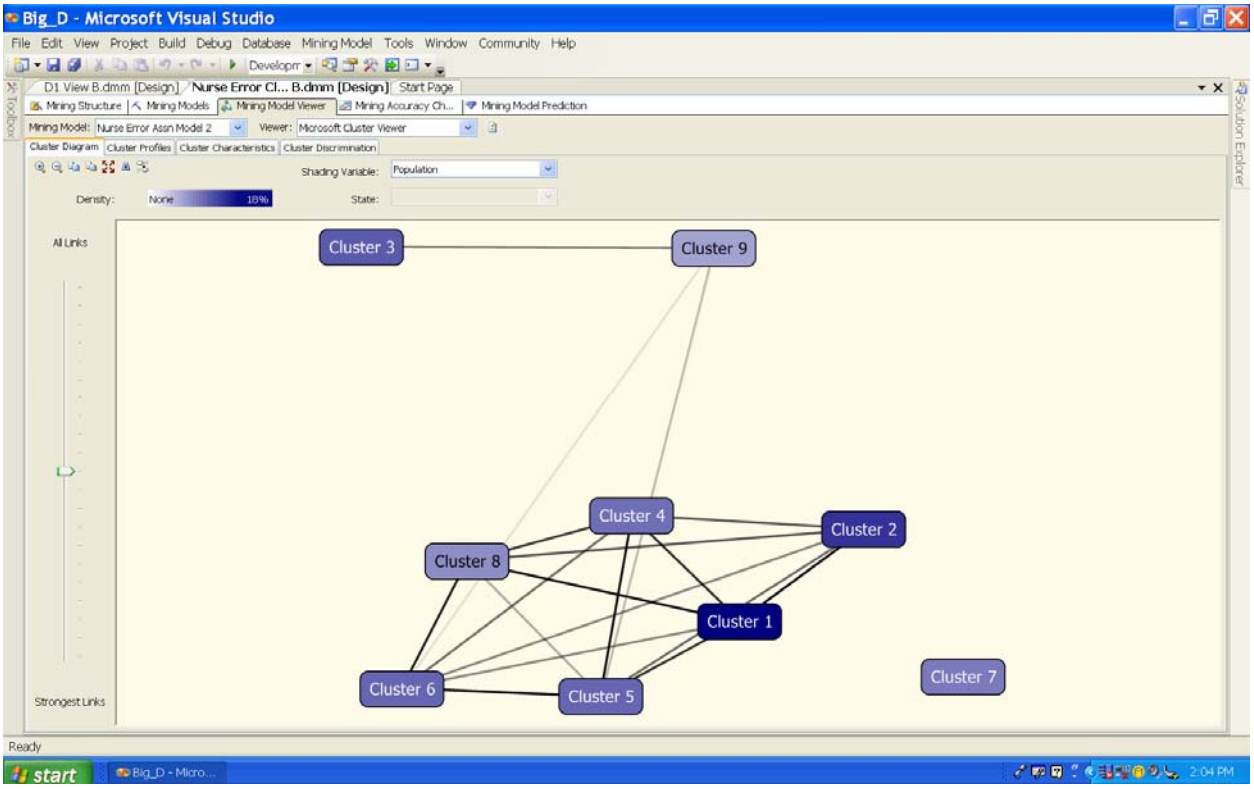
Data Source View – Mining Structure

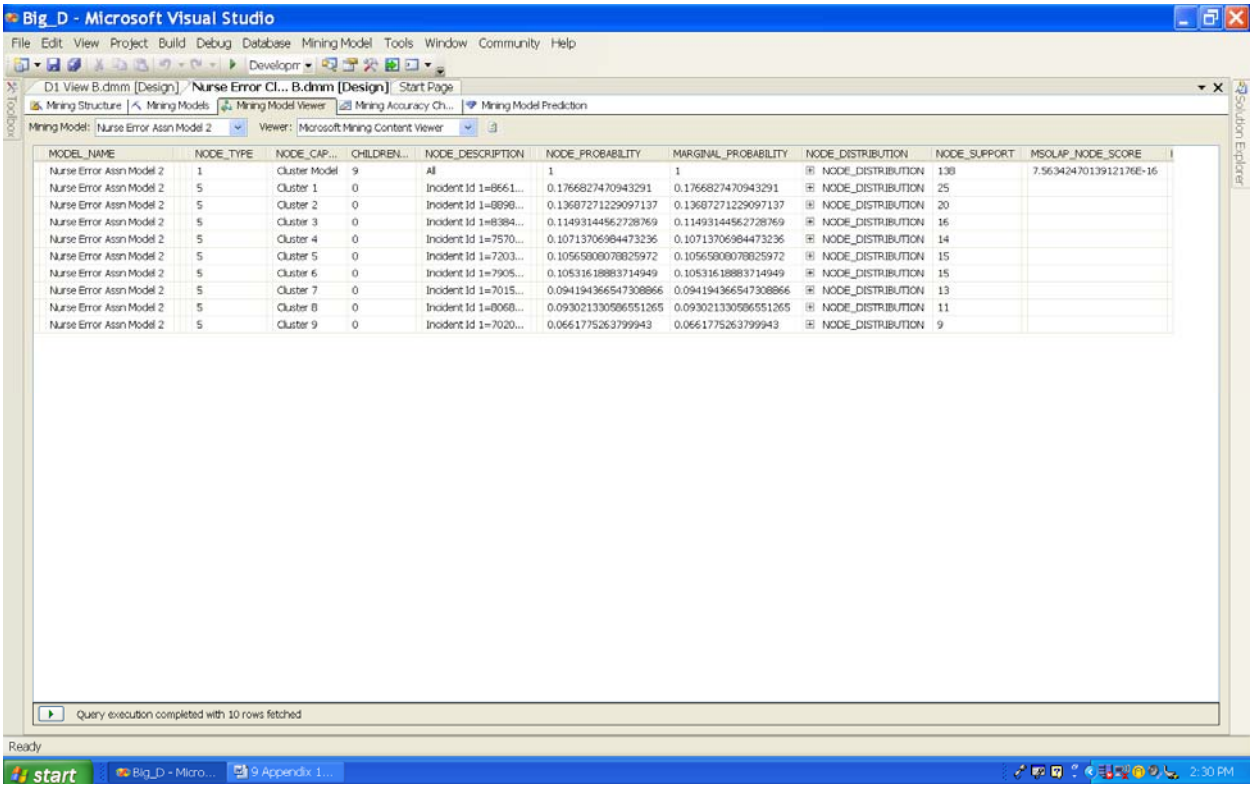


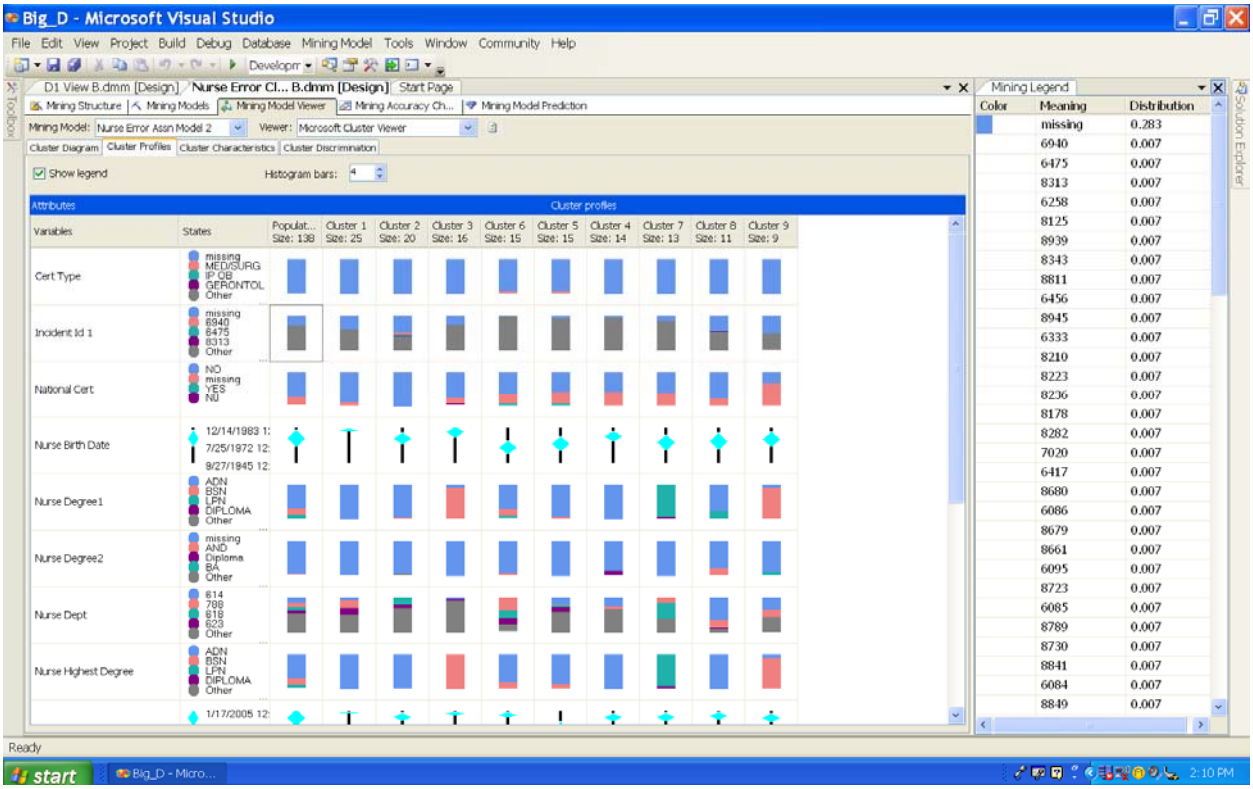
Nurse Error Association Model

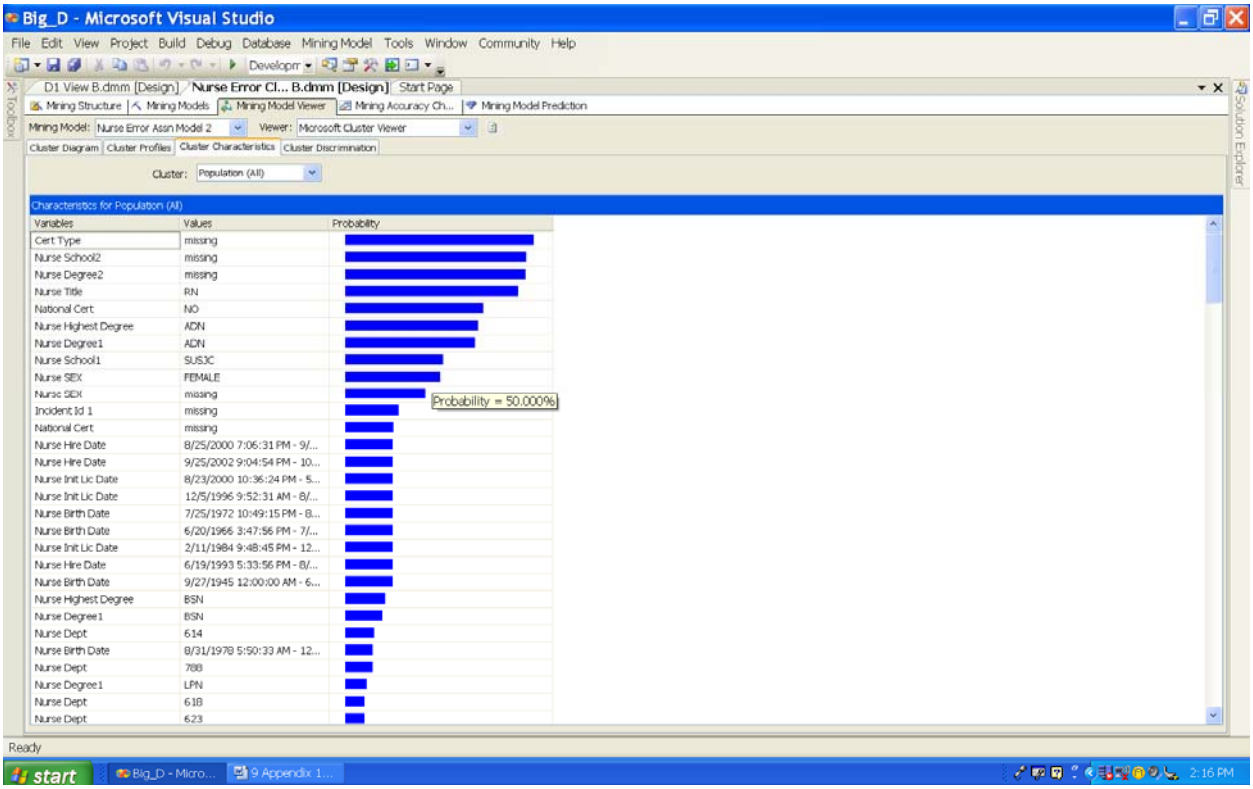


Cluster Diagram

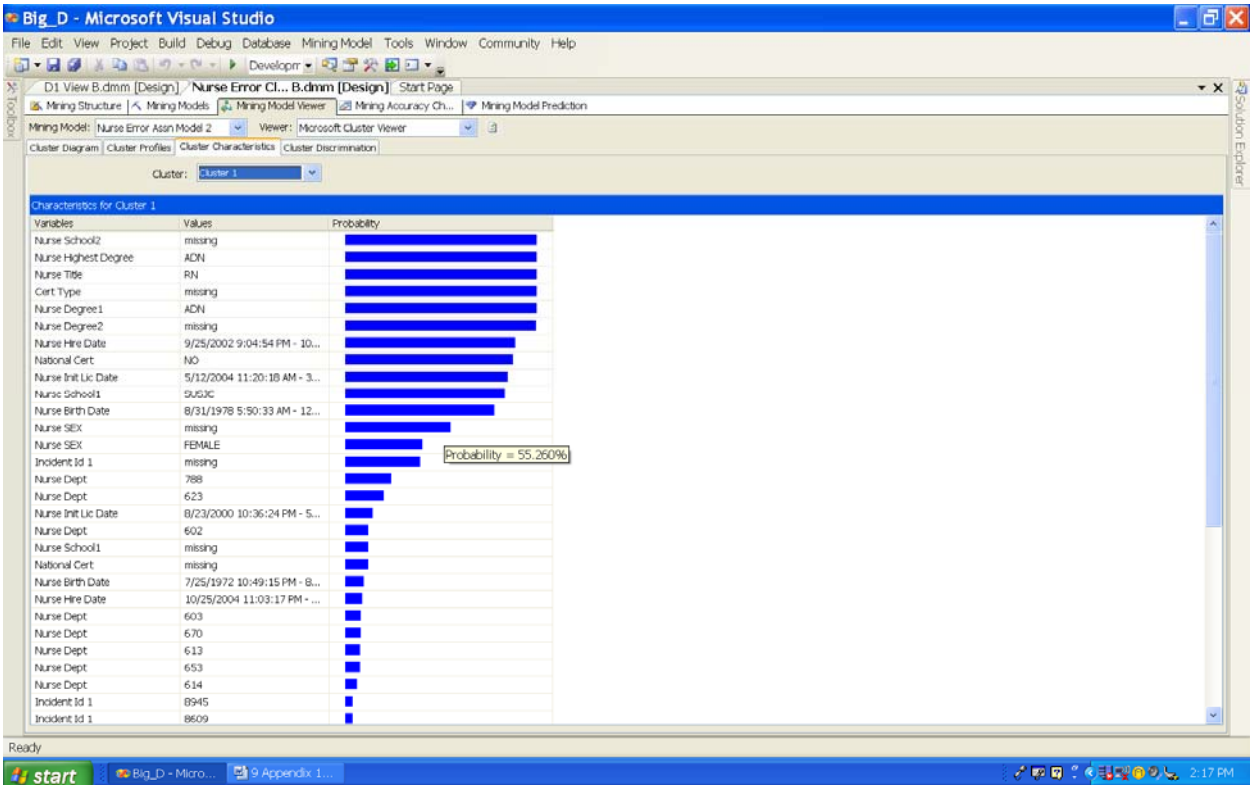





Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Nurse Error Assn Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Popular... Size: 138
Cluster 1 Size: 25
Cluster 2 Size: 20
Cluster 3 Size: 16
Cluster 6 Size: 15
Cluster 5 Size: 15
Cluster 4 Size: 14
Cluster 7 Size: 13
Cluster 8 Size: 11
Cluster 9 Size: 9
Cert Type
missing
MED/SURG
IP OB
GERONTOL
Other
Incident Id 1
missing
6940
6475
8313
Other
National Cert
NO
missing
YES
NU
Nurse Birth Date
12/14/1983 1:
7/25/1972 12:
9/27/1945 12:
Nurse Degree1
ADN
BSN
LPN
DIPLOMA
Other
Nurse Degree2
missing
AND
Diploma
BA
Other
Nurse Dept
614
788
818
823
Other
Nurse Highest Degree
ADN
BSN
LPN
DIPLOMA
Other
1/17/2005 12:
Mining Legend
Color
Meaning
Distribution
missing 0.283
6940 0.007
6475 0.007
8313 0.007
6258 0.007
8125 0.007
8939 0.007
8343 0.007
8811 0.007
6456 0.007
8945 0.007
6333 0.007
8210 0.007
8223 0.007
8236 0.007
8178 0.007
8282 0.007
7020 0.007
6417 0.007
8680 0.007
6086 0.007
8679 0.007
8661 0.007
6095 0.007
8723 0.007
6085 0.007
8789 0.007
8730 0.007
8841 0.007
6084 0.007
8849 0.007
Ready
start
Big_D - Micro...
2:10 PM

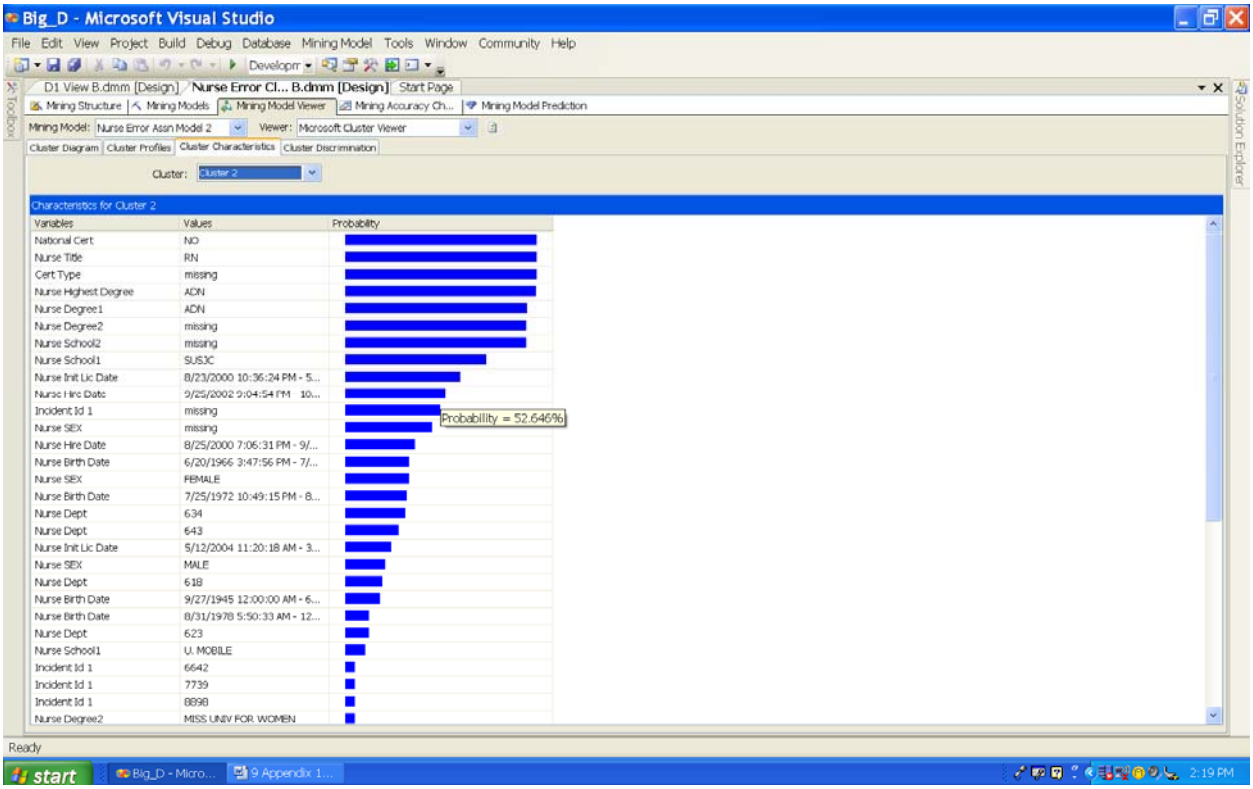


Nurse Error Close Assn. B **Model 2** **Population Profile** **Size: 138**		
Variables	**Values**	**Probability**
Cert Type	missing	98.55%
Nurse School2	missing	94.93%
Nurse Degree2	missing	94.20%
Nurse Title	RN	90.58%
National Cert	NO	72.46%
Nurse Highest Degree	ADN	69.57%
Nurse Degree1	ADN	68.12%
Nurse School1	SUSJC	51.45%
Nurse SEX	FEMALE	50.00%

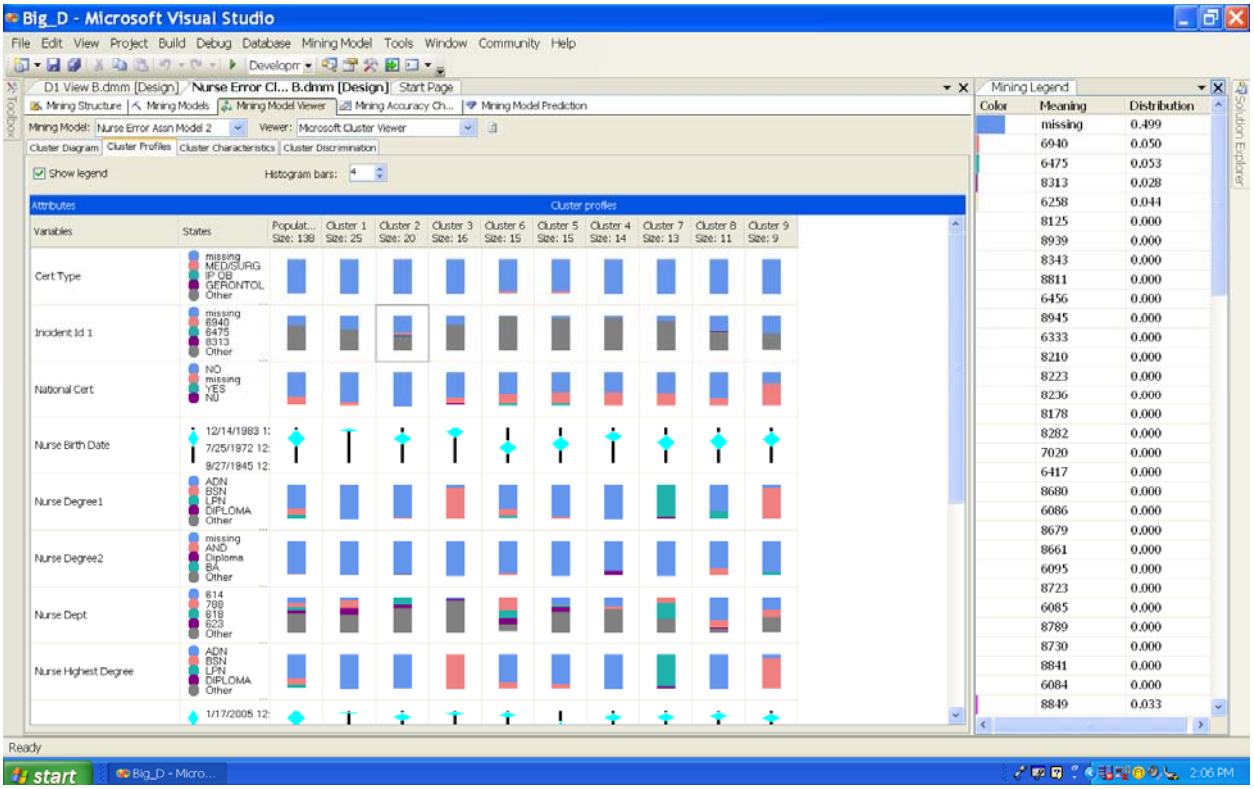


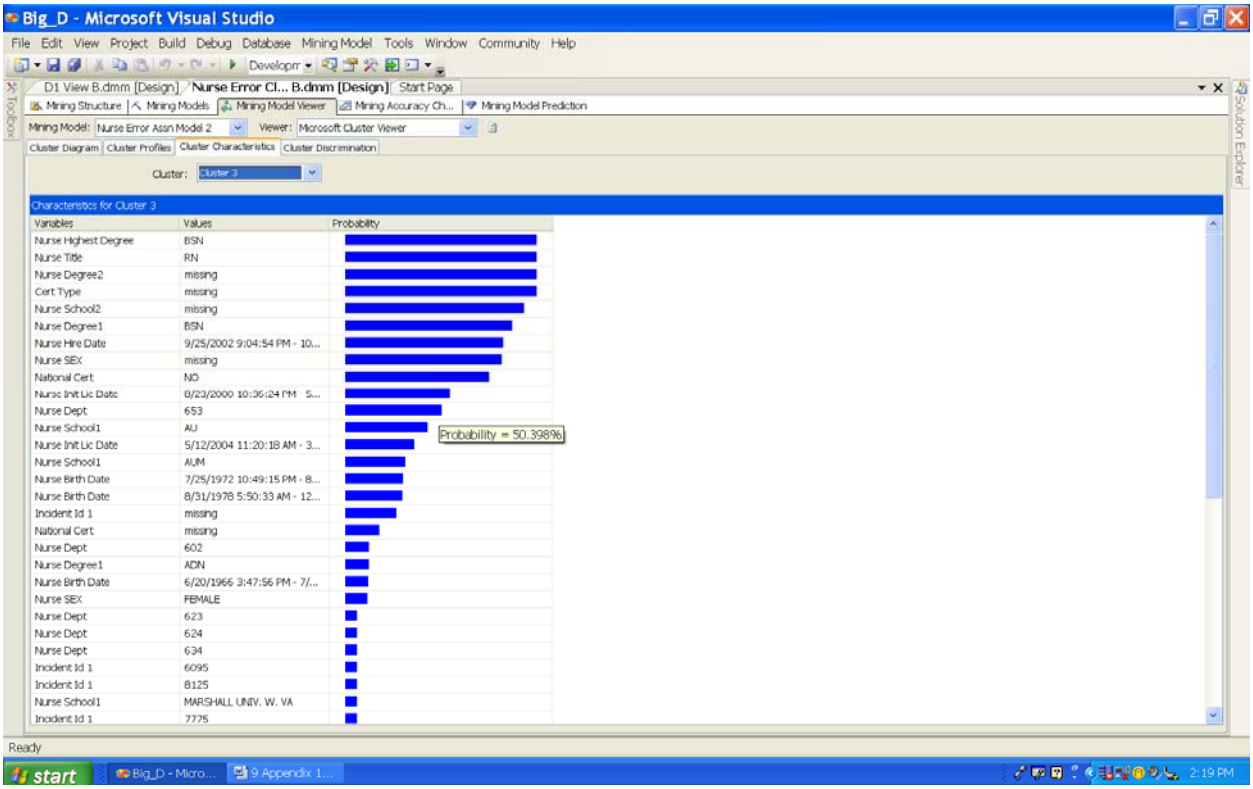
Nurse Error Close Assn. B
Model 2
Cluster 1 Profile
Size: 25

Variables	Values	Probability
Nurse School2	missing	100.00%
Nurse Highest Degree	ADN	100.00%
Nurse Title	RN	100.00%
Cert Type	missing	100.00%
Nurse Degree1	ADN	100.00%
Nurse Degree2	missing	99.98%
Nurse Hire Date	9/25/2002 9:04:54 PM - 10/25/2004 11:03:17 PM	88.89%
National Cert	NO	87.88%
Nurse Init Lic Date	5/12/2004 11:20:18 AM - 3/9/2005 12:00:00 AM	85.09%
Nurse School1	SUSJC	83.62%
Nurse Birth Date	8/31/1978 5:50:33 AM - 12/14/1983 12:00:00 AM	78.06%
Nurse SEX	missing	55.26%



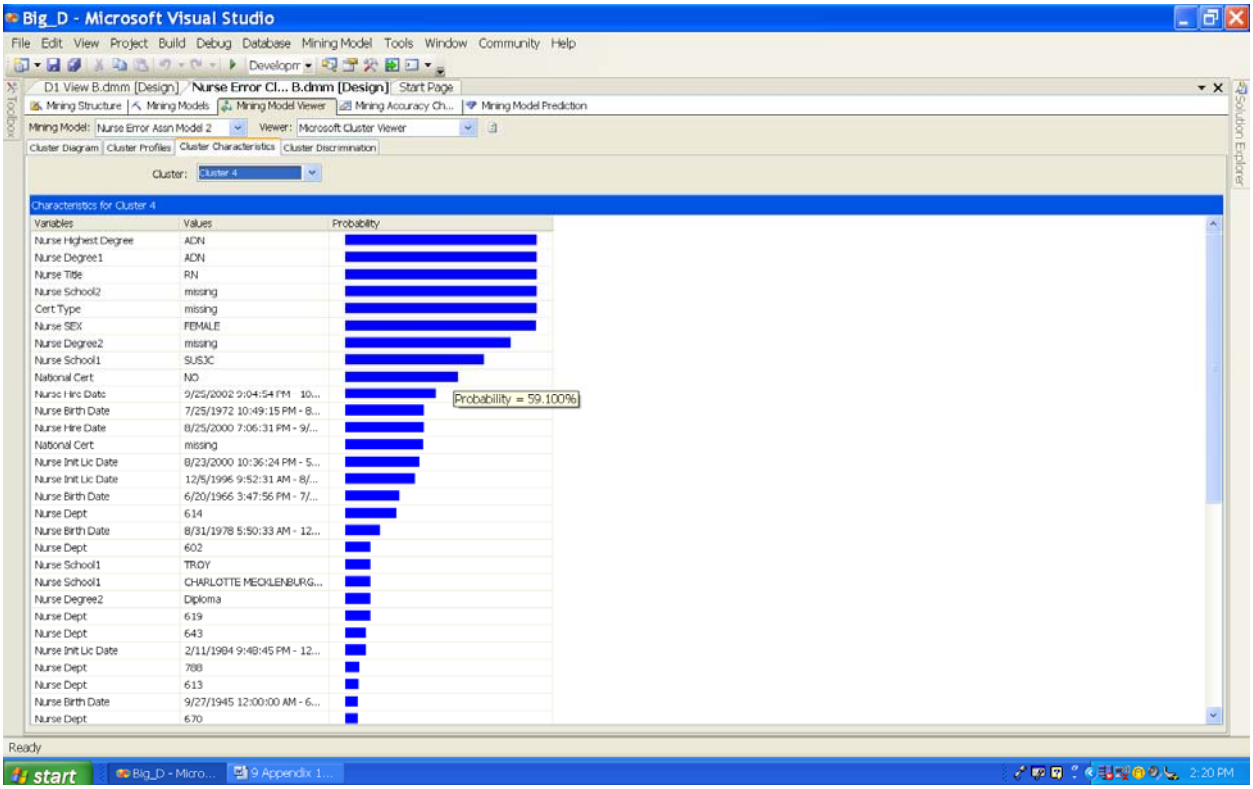
Nurse Error Close Assn. B **Model 2** **Cluster 2 Profile** **Size: 20**		
Variables	**Values**	**Probability**
National Cert	NO	100.00%
Nurse Title	RN	100.00%
Cert Type	missing	100.00%
Nurse Highest Degree	ADN	99.97%
Nurse Degree1	ADN	94.97%
Nurse Degree2	missing	94.71%
Nurse School2	missing	94.71%
Nurse School1	SUSJC	73.72%
Nurse Init Lic Date	8/23/2000 10:36:24 PM - 5/12/2004 11:20:18 AM	60.12%
Nurse Hire Date	9/25/2002 9:04:54 PM - 10/25/2004 11:03:17 PM	52.65%


Big_D - Microsoft Visual Studio
File Edit View Project Build Debug Database Mining Model Tools Window Community Help
D1 View B.dmm [Design]
Nurse Error Cl... B.dmm [Design]
Start Page
Mining Structure
Mining Models
Mining Model Viewer
Mining Accuracy Ch...
Mining Model Prediction
Mining Model: Nurse Error Assn Model 2
Viewer: Microsoft Cluster Viewer
Cluster Diagram
Cluster Profiles
Cluster Characteristics
Cluster Discrimination
Show legend
Histogram bars:
Attributes
Cluster profiles
Variables
States
Popular... Size: 138
Cluster 1 Size: 25
Cluster 2 Size: 20
Cluster 3 Size: 16
Cluster 6 Size: 15
Cluster 5 Size: 15
Cluster 4 Size: 14
Cluster 7 Size: 13
Cluster 8 Size: 11
Cluster 9 Size: 9
Cert Type
missing
MED/SURG
IP OB
GERONTOL
Other
Incident Id 1
missing
6940
6475
8313
Other
National Cert
NO
missing
YES
NU
Nurse Birth Date
12/14/1983 1:
7/25/1972 12:
9/27/1945 12:
Nurse Degree1
ADN
BSN
LPN
DIPLOMA
Other
Nurse Degree2
missing
AND
Diploma
BA
Other
Nurse Dept
614
788
818
823
Other
Nurse Highest Degree
ADN
BSN
LPN
DIPLOMA
Other
1/17/2005 12:
Mining Legend
Color
Meaning
Distribution
missing 0.499
6940 0.050
6475 0.053
8313 0.028
6258 0.044
8125 0.000
8939 0.000
8343 0.000
8811 0.000
6456 0.000
8945 0.000
6333 0.000
8210 0.000
8223 0.000
8236 0.000
8178 0.000
8282 0.000
7020 0.000
6417 0.000
8680 0.000
6086 0.000
8679 0.000
8661 0.000
6095 0.000
8723 0.000
6085 0.000
8789 0.000
8730 0.000
8841 0.000
6084 0.000
8849 0.033
Ready
start
Big_D - Micro...
2:06 PM

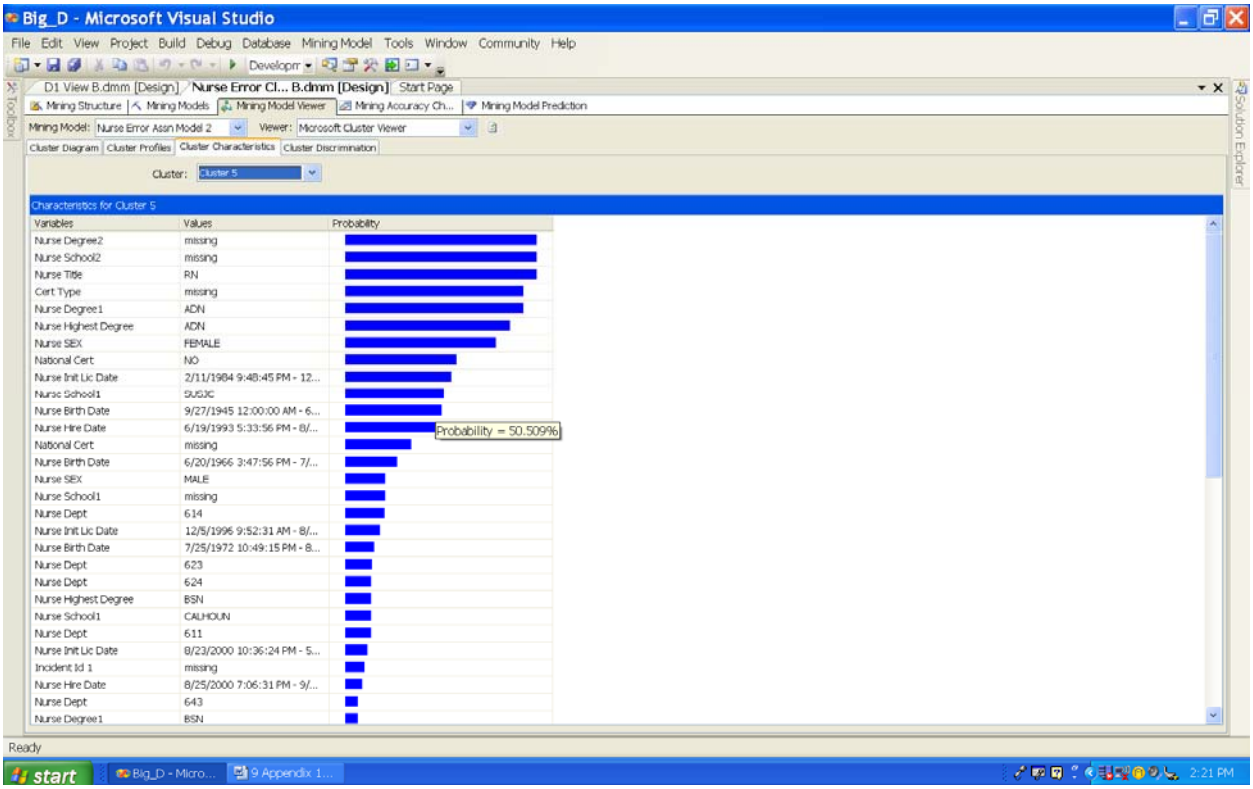


Nurse Error Close Assn. B
Model 2
Cluster 3 Profile
Size: 16

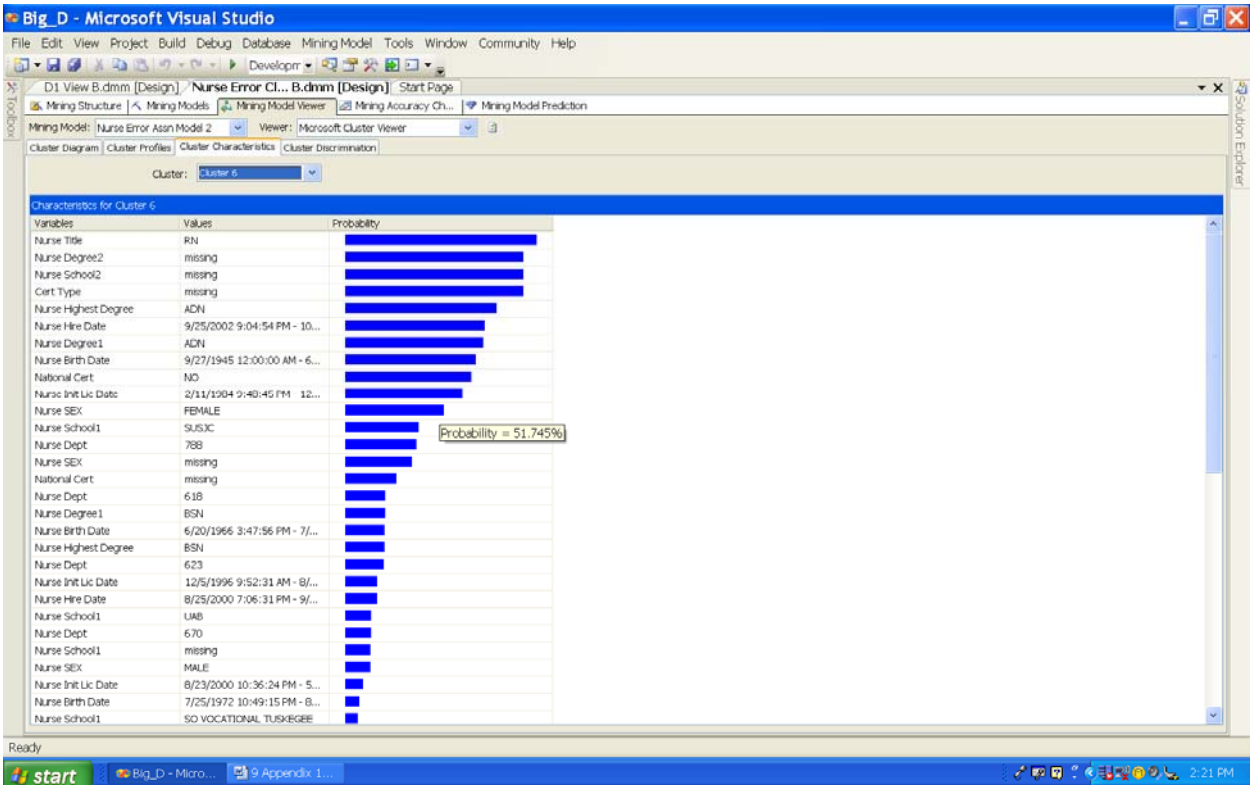
Variables	Values	Probability
Nurse Highest Degree	BSN	100.00%
Nurse Title	RN	100.00%
Nurse Degree2	missing	100.00%
Cert Type	missing	100.00%
Nurse School2	missing	93.70%
Nurse Degree1	BSN	87.43%
Nurse Hire Date	9/25/2002 9:04:54 PM - 10/25/2004 11:03:17 PM	82.87%
Nurse SEX	missing	81.90%
National Cert	NO	75.55%
Nurse Init Lic Date	8/23/2000 10:36:24 PM - 5/12/2004 11:20:18 AM	54.80%
Nurse Dept	653	50.40%



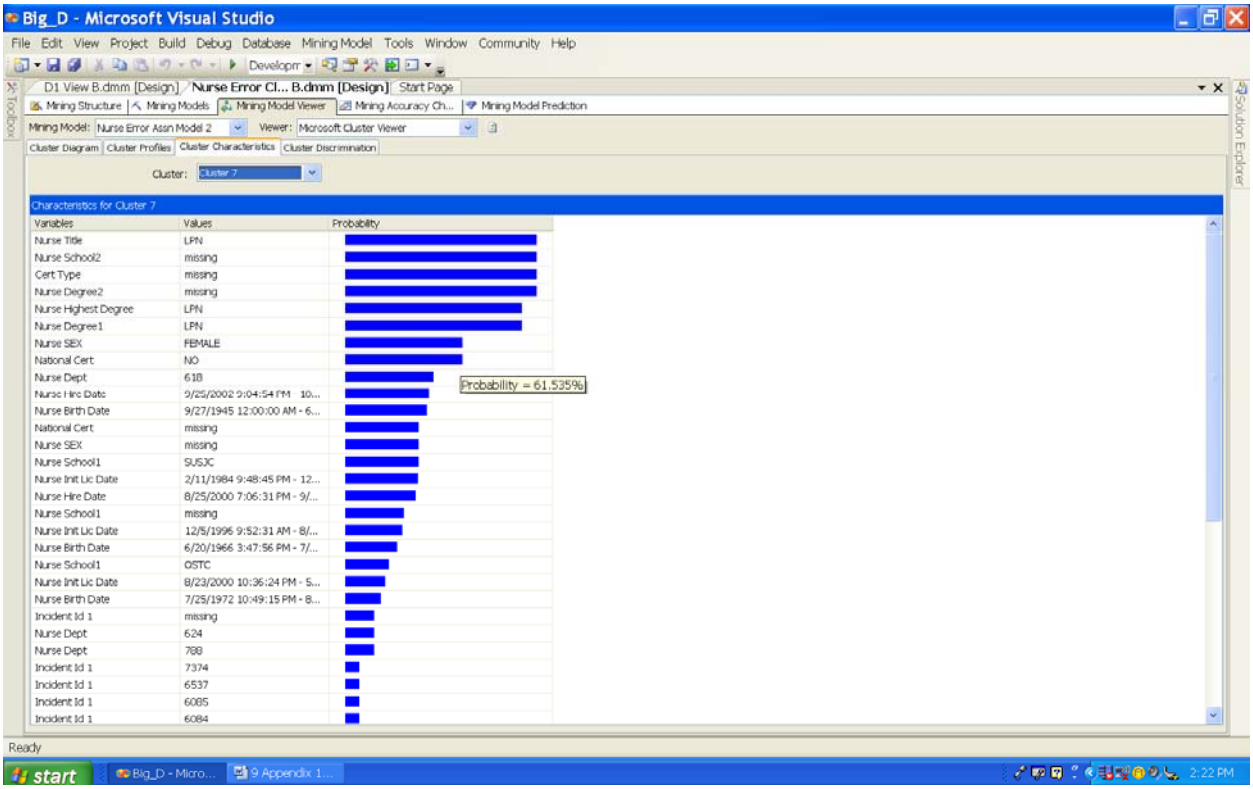
Nurse Error Close Assn. B **Model 2** **Cluster 4 Profile** **Size: 14**		
Variables	**Values**	**Probability**
Nurse Highest Degree	ADN	100.00%
Nurse Degree1	ADN	100.00%
Nurse Title	RN	100.00%
Nurse School2	missing	100.00%
Cert Type	missing	100.00%
Nurse SEX	FEMALE	99.96%
Nurse Degree2	missing	86.51%
Nurse School1	SUSJC	72.80%
National Cert	NO	59.10%



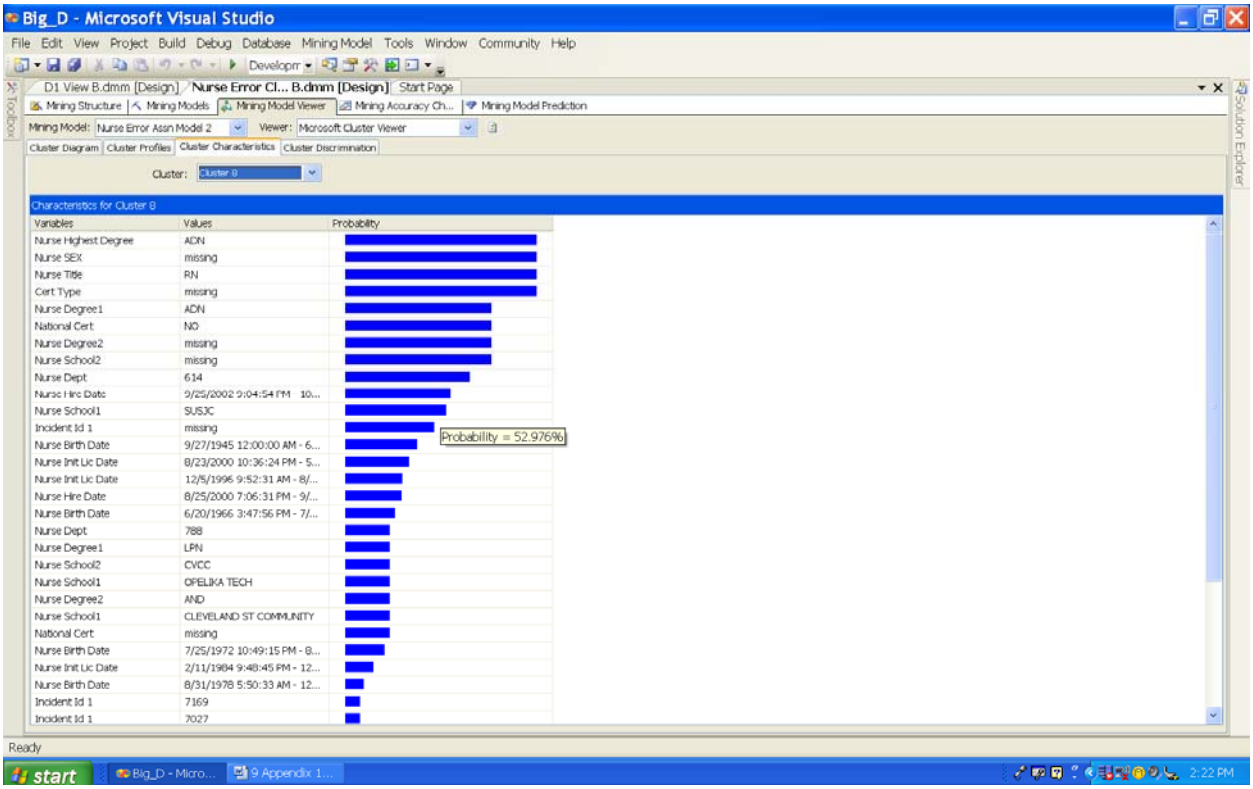
Nurse Error Close Assn. B **Model 2** **Cluster 5 Profile** **Size: 15**		
Variables	**Values**	**Probability**
Nurse Degree2	missing	100.00%
Nurse School2	missing	100.00%
Nurse Title	RN	100.00%
Cert Type	missing	93.14%
Nurse Degree1	ADN	93.14%
Nurse Highest Degree	ADN	86.29%
Nurse SEX	FEMALE	78.86%
National Cert	NO	58.29%
Nurse Init Lic Date	2/11/1984 9:48:45 PM - 12/5/1996 9:52:31 AM	55.59%
Nurse School1	SUSJC	51.58%
Nurse Birth Date	9/27/1945 12:00:00 AM - 6/20/1966 3:47:56 PM	50.51%



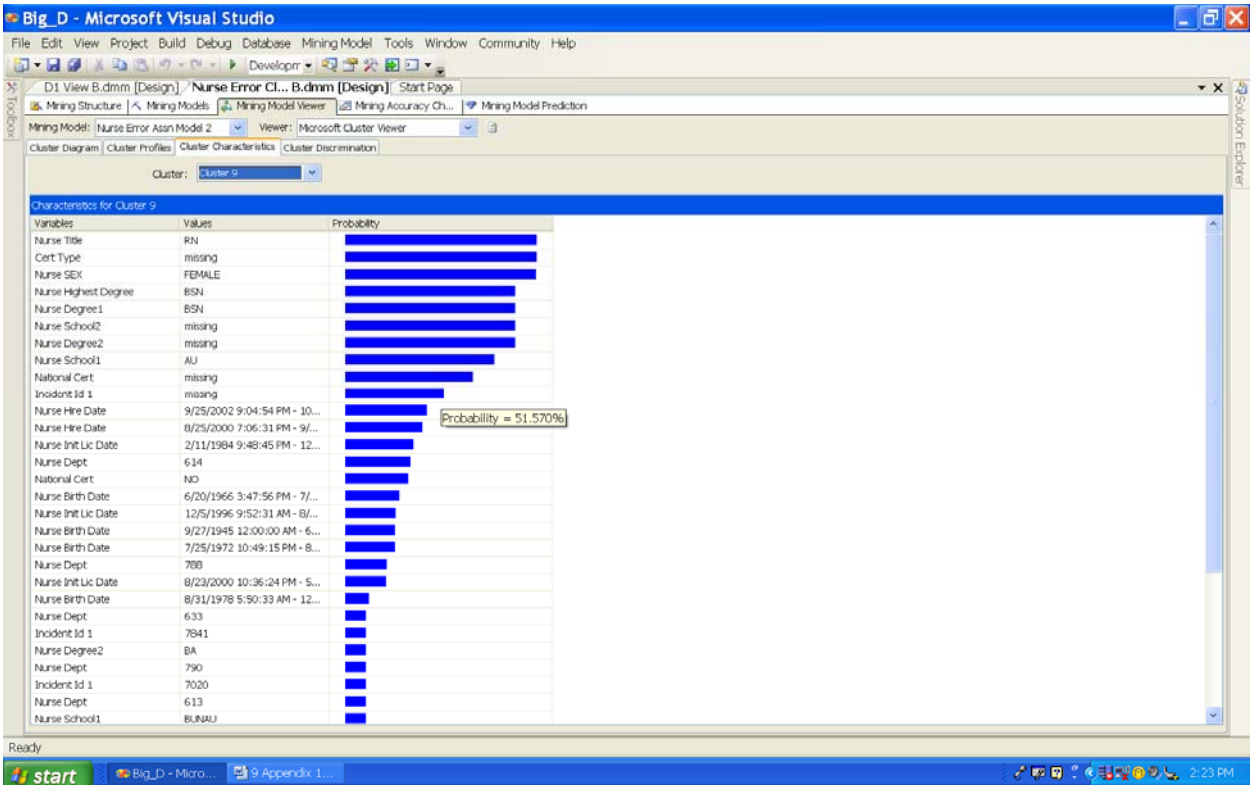
Nurse Error Close Assn. B **Model 2** **Cluster 6 Profile** **Size: 15**		
Variables	**Values**	**Probability**
Nurse Title	RN	100.00%
Nurse Degree2	missing	93.12%
Nurse School2	missing	93.12%
Cert Type	missing	93.12%
Nurse Highest Degree	ADN	79.36%
Nurse Hire Date	9/25/2002 9:04:54 PM - 10/25/2004 11:03:17 PM	73.19%
Nurse Degree1	ADN	72.13%
Nurse Birth Date	9/27/1945 12:00:00 AM - 6/20/1966 3:47:56 PM	68.46%
National Cert	NO	66.16%
Nurse Init Lic Date	2/11/1984 9:48:45 PM - 12/5/1996 9:52:31 AM	61.40%
Nurse SEX	FEMALE	51.75%



Nurse Error Close Assn. B **Model 2** **Cluster 7 Profile** **Size: 13**		
Variables	**Values**	**Probability**
Nurse Title	LPN	100.00%
Nurse School2	missing	100.00%
Cert Type	missing	100.00%
Nurse Degree2	missing	100.00%
Nurse Highest Degree	LPN	92.32%
Nurse Degree1	LPN	92.32%
Nurse SEX	FEMALE	61.54%
National Cert	NO	61.54%



Nurse Error Close Assn. B **Model 2** **Cluster 8 Profile** **Size: 11**		
Variables	**Values**	**Probability**
Nurse Highest Degree	ADN	100.00%
Nurse SEX	missing	100.00%
Nurse Title	RN	100.00%
Cert Type	missing	100.00%
Nurse Degree1	ADN	76.63%
National Cert	NO	76.63%
Nurse Degree2	missing	76.63%
Nurse School2	missing	76.63%
Nurse Dept	614	65.19%
Nurse Hire Date	9/25/2002 9:04:54 PM - 10/25/2004 11:03:17 PM	55.39%
Nurse School1	SUSJC	52.98%



Nurse Error Close Assn. B
Model 2
Cluster 9 Profile
Size: 9

Variables	Values	Probability
Nurse Title	RN	100.00%
Cert Type	missing	100.00%
Nurse SEX	FEMALE	99.88%
Nurse Highest Degree	BSN	89.07%
Nurse Degree1	BSN	89.05%
Nurse School2	missing	89.05%
Nurse Degree2	missing	89.05%
Nurse School1	AU	78.10%
National Cert	missing	67.01%
Incident Id 1	missing	51.57%