Data Mining Nursing Care Plans of End of Life Patients: A Study to Improve Healthcare Decision Making

F. Almasalha[1a], **D. Xu**[1a], **G. M. Keenan**[2a], **A. Khokhar**[1a,*], **Y. Yao**[1a], **Yu-C. Chen**[1a], **A. Johnson**[3a], **R. Ansari**[1a], and **D. J. Wilkie**[2a]

[1]University of Illinois at Chicago, 851 South Morgan Street, Chicago. IL 60607 USA

[2]University of Illinois at Chicago, 845 South Damen Avenue, Chicago, IL 60612-7350 USA

[3]University of Illinois at Chicago, 2032 ERF, MC 152, Chicago, IL 60607, USA

Abstract

Pain management of end of life patients (EOL) (n=596 episodes) is examined using statistical and data mining processes of the HANDS database of care plans coded with NANDA-I, NOC, and NIC (NNN) terminologies. HANDS episode data (episode =care plans updated at every handoff on a patient while staying on a hospital unit) were gathered in 8 units located in 4 different health care facilities (total episodes = 40,747; EOL episodes = 1,425) over two years. Results show the multiple discoveries such as EOL patients with hospital stays (< 72 hrs.) are less likely (p<0.005) to meet the pain relief goals compared to EOL patients with longer hospital stays. The study demonstrates a major benefit of systematically integrating NNN into electronic health records.

Keywords

Data mining; plan of care; electronic health record; end of life hospital care; pain

1. Introduction

The electronic health record (EHR) became a priority of the U.S. federal government as a result of the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. Through the Act health providers will be given incentives to adopt and use EHRs with all patients by 2015 (Managed Care Outlook, 2010). The aim of HITECH is to improve the quality of patient care by enhancing the efficiency of healthcare delivery systems. With the anticipated increase in use of EHRs the delivery systems are likely to amass and archive an unprecedented amount of health-related information in a relatively short period of time. However, the mere deployment of an EHR will have only limited benefits unless novel applications are developed that extract information nuggets hidden in the captured data and use it as feedback or a knowledge base to continuously improve the quality of health care. This fact is borne out in numerous studies in which investigators have reported major gaps in the availability and usefulness of information in the EHRs and patient record systems to the front line users (Allen, 1998; Hardey, Payne, & Coleman, 2000; Karkkainen, Bondas, & Eriksson, 2005; Keenan & Yakel, 2005; Keenan, Yakel, Dunn

Lopez, Tschannen, & Ford, In Review). A key requirement to extract maximum benefits is that the information be collected and stored in a standardized format. If similar data are stored in different ways it is nearly impossible to employ automated methods for directly using the data for benchmarking purposes or meaningful knowledge discovery. In this paper we present an example of benefits that can be achieved through mining of data gathered from multiple organizations in which the collection and database storage have been standardized.

The data we explore in this paper were nursing care documentation as a part of a three year study of the HANDS plan of care system (Keenan, Tschannen, & Wesley, 2008; Keenan, Yakel, Yao, Xu, Szalacha, Tschannen…Wilkie, in press). Our data source, the HANDS system, is innovative in its use of methods that ensure the generalizability, feasibility, acceptability, and utility to the user at the point of care under real-time conditions. The analyses conducted particularly focus on pain management of patients at or near the end of life (EOL).

Our specific aim is to demonstrate the effectiveness and relative ease of applying data mining techniques when a standardized database is the source utilized in (i) searching for relationships among care plan indicators for meeting target outcomes, (ii) uncovering the relationships among the indicators to cluster them into groups with shared characteristics for discovering cues to improved patient care, and (iii) classifying data into categories with known structures so that the extracted information nuggets can be incorporated to improve decision making at the point of care. Data mining, also known as knowledge discovery, provides a powerful approach for extracting the care related patterns and hidden relationships that become meaningful information for benchmarking and best practices (Cheung, Moody, & Cockram, 2002; Fayyad, 1996). Data mining in the health care field has focused on data from different patient populations (e.g., cancer (Stephan et al., 2002); diabetes (Breault, Goodall, & Fos, 2002); data regarding care and treatments (e.g., nursing (Keenan et al., 2008; Keenan et al, in press), pharmacy (Rudman et al., 2002); techniques used to mine the data (e.g., Bayesian networks (Goodwin, VanDyne, Lin, & Talbert, 2003); neural networks (Lee & Abbott, 2003); and other knowledge management systems (Hsia & Lin, 2006; Kraft, Desouza, & Androwich, 2003; Mullinsa et al., 2006).

There are a number of examples in the literature (Embley, Tao, & Liddle, 2005; Goodwin et al., 2003; Hanauer, Rhodes, & Chinnaiyan, 2009; Heckerman, 1997; Hsia & Lin, 2006; Kraft, Desouza, & Androwich, 2003; Lee & Abbott, 2003; Mullinsa et al., 2006, Trifiro et al., 2009) of tools and their uses in mining health care data and in some cases using it for decision making. For example, Goodwin et al. (2003) explore the use of data mining techniques to build and represent nursing knowledge and relate it to the data present in the patients' records. The major point underscored by these authors is the enormous difficulty encountered in mining data for which the content, collection methods, and storage have not been standardized. In their study of prenatal patients, only 28% of patient records (19,970) from a total of 71,753 patient records were usable after extensive cleaning procedures. Vast amounts of computer processing time consumed multiple machines for extensive periods of time for the cleaning process. In spite of extensive cleaning, the authors report that "noisy" data remained and was thought to mask the effect of potentially important variables (Goodwin et al., 2003).

Other investigators report a variety of approaches to data mining that reveal challenges and insights. A set of knowledge management functions with enabling IT tools are needed to produce nursing knowledge management systems that adequately detect the interplay among different nursing practices and related functions (Hsia and Lin, 2006). Using the HealthMiner System with three unsupervised "Rule Discovery" methods (CliniMiner,

Predictive Analysis, and Pattern Discovery) and 667,000 in-patient and outpatient digital records from one academic center, Mullins et al. (2006) discovered patterns utilizing the data elements of clinical problems, laboratory tests, treatment groups and medications where these were found in the records. Mullins et al. (2006) note the importance of the quality of data to the process. If the data are not entered in a way that makes it easily analyzable (no standardization of data entry process, content, and storage) then the depth of discovery is correspondingly limited. Data mining and statistical analyses (logistic regression and predictive modeling) on a dataset gathered on patients cared for by 15 home health care organizations suggest the enormous potential of data mining when the content, processes of data entry, and storage of electronic health record data are standardized (Westra, Dey, et al., 2011; Westra, Savik, et al., 2011). Using mining of correlations present among nursing diagnoses, outcomes, and interventions, Duan, Street, and Lu (2008) propose care plan recommendations that are ranked for listing based on support and confidence measures generally used in association rule mining. These findings also include a novel measure that anticipates which selections might improve the quality of future rankings (Duan, Street, & Lu, 2008).

In summary, the current incentives for speeding adoption of EHRs are certain to result in the ever increasing amount and availability of clinical data in an electronic format that is a potential gold mine for knowledge discovery. The examples of tools and mining of health data above highlight the enormous potential of data mining large databases generated by EHRs to discover knowledge nuggets that can l improve health care decision making. Most of the work to date is limited in breadth and depth. The techniques investigated have been domain specific and are not amenable to generalization. Most of these techniques are specific to the format and nature of the underlying data sets. This limitation is in big part because of the poor quality of data collected with EHRs due to lack of standardization of the collection process, content, and format of the data. As a result, extensive data cleaning is required to extract sufficient usable content from multiple systems whose collection, content, and storage systems differ. This preparation work is costly, particularly when the algorithms cannot be universally applied to similar data that are captured and stored in so many different ways. Consequently, data mining is frequently cost prohibitive for poor quality data. In this paper we demonstrate the use of multiple data mining techniques on patient care plans from EOL patients available in an EHR system to show the enormous potential for efficient and effective knowledge discovery when the process of collection, content, and format of the data are standardized. We next show how the discovery of good practices could be used in a feedback mechanism to improve health care delivery of patients with similar symptoms and illnesses.

This paper presents a study that involved the use of data-mining techniques applied to nursing care plan data collected on 569 EOL patients. EOL care takes a disproportionate share of health care expenditures with 30% of Medicare expenditures going to the 5% of beneficiaries who die each year (Zhang et al., 2009). About one-third of the Medicare expenditures in the last year of life are spent in the last month, which means that Americans spend about $300 million for care during that last month, usually for life sustaining care (e.g., mechanical ventilators, resuscitation; Zhang et al., 2009). Costs for care of veterans at the EOL are similar (Yu, 2006). Unless interventions are used more effectively, especially for patients hospitalized in the last month of life, these costs are expected to rise substantially as the baby boomers age and require EOL care (Smith et al., 2003). From a health care expenditure perspective, there is an urgent need to provide cost effective care at EOL that also facilitates dignity and comfort for the dying.

2. Material and Methods

The data analyzed in this study were derived from a primary data set gathered on 33,451 medical surgical patients (40,747 unique episodes) during a three-year study of hospitalized patients in 8 medical-surgical units of four Midwestern medical centers (R01 HS01505402, 2005–2008). The data consisted of all nursing care plans entered into the HANDS system during each patient's hospital stay. The HANDS system is an electronic application and database that nurses use to enter and track patients' clinical diagnoses, outcomes, and interventions within and across episodes of care. An episode (continuous patient stay in a single hospital unit) consists of the plans of care that nurses documented at every formal handoff (admission, shift-change update, or discharge) in the electronic HANDS tool for a period of two years. The main attributes available in the database for each plan of care include patient demographics, nurse demographics, all nursing diagnoses, associated outcomes and their ratings at the time of each handoff, and interventions coded with ANA-approved standardized terminologies. Nursing diagnoses are coded with NANDA-I (NANDA International, 2003) terms, outcomes are coded using terms and rating scales from the Nursing Outcomes Classification (NOC) (Moorhead, Johnson, & Maas, 2004), and interventions are coded with terms from the Nursing Intervention Classification (NIC) (Bulechek, Butcher, & Dochterman, 2007). An expected outcome rating (goal at discharge from the unit) is recorded on the plan of care to which each NOC is first entered. The patient's discharge disposition is entered into the database at the end of each unit episode. An abstract view of different entities in the HANDS database and their interdependencies are captured in a high level ER diagram shown in Figure 1.

In the primary study, the investigators used representative convenience sampling to select units that would support generalizability of findings. Four unique hospitals were first approached and agreed to participate; 2 large community hospitals (LCH), 1 university hospital (UH), and 1 small community hospital (SCH). The hospitals were located in one Midwest state and housed from 85 to 865 in-patient beds (See Table 1). Next, 8 units were recruited from these hospitals to participate for either a 12-month (4 units) or 24-month (4 units) study period. There were 3 primary inclusion criteria; unit willingness to participate in the study, stable staffing as perceived by unit and upper administration, and agreement that all nurses employed on a unit would complete the required training session and use the HANDS system for the entire study duration.

2.1 The EOL Dataset

The analyses for the current study were conducted on a subset of EOL episodes derived from the primary dataset. We first identified all of the episodes of care from the primary data set in which any of the following indicators were present: 1) NOC outcome: Comfortable Death; 2) NOC outcome: Dignified Life Closure; 3) NIC intervention: Dying Care; 4) Discharged to hospice medical facility; 5) Discharged to hospice home care; or 6) Expired There were a total of 1,425 episodes (1,394 unique patients; one patient had 3 episodes and 29 had 2 episodes) that met these criteria. In this, our EOL subsample, the patients' average age was 77.5 (median=83, SD=12.8, ranged from 22–105). Since race and gender were not required fields (appeared elsewhere on a patient's record), not all episodes include this information. Of the 413 episodes with gender information, 55% were females and 45% were males. Of the 187 with race information, 87% were white, 11% were black, 1% were Asians, and 1% were other (0% Native Indian or Alaskan; 0% Hispanics). From the 1,425 episodes, we next isolated those episodes in which at least one NANDA-I pain diagnosis (Chronic Pain or Acute Pain) was recorded. A total of 569 EOL episodes met this criteria and were the source of data analyzed in this study.

An episode in the HANDS database consists of a set of POCs, one submitted at every formal handoff (typically end of shift) to reflect the care provided during the shift and the status of outcomes at shift change. The first plan of care for an episode (admission plan) is entered at the first handoff and includes all of the pertinent nursing diagnoses, outcomes, and interventions (coded with the corresponding NANDA-I, NOC, or NIC (NNN) label and code) that were addressed during the admission shift. Each subsequent plan (updates/discharge) builds on the previous and includes all changes (additions, deletions, resolutions to NNNs) that occurred during the time period since the last plan submission. The status of all NOC outcomes on a POC are rated (or re-rated) at every handoff. An expected NOC rating (goal at discharge from a unit) is entered only once, on the POC to which it was first added. The entry of the NOC status and expected outcome ratings make it possible to monitor progress toward outcome goals during an episode and to assess goal attainment (met/not met) at discharge from a unit. All NOC outcomes have a rating scale that ranges from 1 to 5 with unique anchors for each NOC. A "1" is the worst possible rating and 5 is the best possible rating. A value of 5 reflects the score that a healthy person of the same age, sex, and cognitive ability would score on the specified outcome. The reliability and validity of NOC outcomes were established in previous studies (Keenan, Barkauskas, et al., 2003; Keenan, Stocker, et al., 2003; Keenan et al., in press).

2.2 Data Mining Methods

After approval from the Institutional Review Board at the University of Illinois at Chicago, the principal investigator of the primary study provided a copy of the database that had all HIPAA (Health Insurance Portability and Accountability Act) identifiers encrypted and no decryption tool was available to investigators of the study reported here. To achieve the study aim, we employed association mining, clustering, and classification techniques to glean the hidden information. Association Rule Mining identifies hidden correlations that may be present among different attributes of data. It is exhaustive in terms of computation complexity and therefore needs to be applied carefully on date sets that have been cleaned and rid of errors. Clustering techniques allow determination of groups that have similar characteristics, and classification techniques involves training using samples and helps in identifying the membership of any given datum to one of the known classes.

These techniques have shown remarkable ability to extract hidden patterns and associations present among different variables in historical databases (Stephan et al., 2002). However, these techniques need to be applied carefully to the underlying datasets. Although generic at some abstract level, these techniques involve many technical steps that need to be tailored to the underlying data and their complexity varies with the characteristics of the data. We used the public domain data mining software "Rapid Miner" in our experiments. Given the availability of a very large sample size (1,425 EOL episodes), we used t-tests or chi-square tests and a two-tailed alpha level of .05 to determine the statistical significance of results.

3. Results

Historically pain is one of the most prominent symptoms for EOL patients, but the first striking fact our study revealed was that out of 1,425 EOL episodes, only 41.3% included a Pain NANDA-I in their plans of care. Furthermore, we observed that the percentage of pain diagnosis among EOL patients was significantly lower in LCH1. Note that even in the intensive care unit of LCH1, the percentage of dying patients with a pain diagnosis was only 42.3%. The difference between LCH1 and the other 3 hospitals was statistically significant ($Z(1)=-12.5$, p-value< 0.0005).

We next clustered episodes based on the pain outcome in the last care plan of every episode while considering different variables including: discharge status, age, gender, and length of

stay. For clustering of data with one or two attributes, we used sorting and group-by operations within the SQL framework. For clustering of data with a larger set of attributes we used the k-means clustering algorithm.

Clustering of episodes based on just the final pain outcome revealed that in more than 55% of the EOL pain episodes patients were discharged to a hospice with pain worse than expected. We found this trend to be uniform across all hospitals. Episodes were also clustered based on their final pain outcome and discharged status. Figure 2 shows (in percentages) that among EOL patients, those with "Expired" status had a significantly higher percentage of episodes with pain that did not meet the expected outcome. This trend was statistically significant ($X^2(1)$=7.49, p<.0006).Further analysis of these clusters revealed that episodes with short length of stay (less than 72 hours) had a significantly higher ratio of episodes that did not meet the expectations. This comparison was also statistically significant ($X^2(1)$=60.43, p < 0.001) (Figure 3).

We also studied the role of different NIC interventions that were applied to meet the expected pain related NOC outcomes at discharge. Using the results from the descriptive analysis, we identified the top six most frequently occurring NIC interventions that were applied to manage pain across all hospitals. These combinations covered 96% of the data. Then we used the k-means clustering method to cluster the final outcome based on these six NIC interventions. As shown in Figure 4, we first note that in most of the hospitals/units the number of different NIC combinations applied to manage pain were limited to just a few, a very small fraction of possible combinations. Second, the NIC intervention combination Pain Management, Medication Management, and Positioning, that produced more positive results than negative, was found to have been applied in only 81 episodes out of a 569 episodes on the study. When statistically compared against the final pain outcome of all the other NIC combinations, the outcome of this NIC combination was significantly different from the others ($X^2(1)$=34.46,p<.001).

Finally, we investigated the feasibility of utilizing the NOC rating trajectory data to predict whether a patient would meet the expected pain related outcome ratings at the end of an episode. Our analysis revealed that an improvement in a patient's outcomes from hour 12 to hour 24 indicates an above average chance in meeting the expected outcome before discharge or death. We found that 56.4% of the patients, whose pain related outcomes improved during this period, met the expected outcome ratings at the end of the episode. Among patients whose pain outcomes did not improve during this period, only 40.3% met the expected outcome. Z-Test ($Z(1)$=−2.45) yields a p-value of 0.007, indicating a significant gap between these two groups.

Using this statistical analysis as a cue, we clustered the patients based on their initial pain condition (pain related NOC outcome ratings at the end of the first plan of care, i.e., first handoff), expected NOC outcome ratings, and the rating of these same NOC outcomes at the end of first 24 hours. In order to reduce noise in the data, we discretized the 5 pain levels into the following three categories: severe (pain rating 1/2), medium (pain rating 3/4), and no pain (pain rating 5). We clustered patients into nine static clusters using the k-means algorithm. Figure 5 depicts the outcome of this clustering. For example, 30% of the patients who came in with medium pain, after first 24 hours did not reach the expected outcome of 'no pain' and in fact their pain level increased.

In order to investigate interactions among NANDA-Is pain diagnoses, and NOC outcomes linked to these, we applied the association rule mining method on a data set that contained the final pain outcome, and the top five most frequent NANDA-I diagnoses among the episodes. To better understand the results, with the input from our domain experts, we

grouped all NANDA-I diagnoses related to cardiac or pulmonary problems into a single category of diagnoses and referred to it as "cardiopulmonary concern."

Our association rule mining revealed several interesting results. One rule that was of particular interest revealed that (Acute or Chronic Pain) and worse than expected pain outcome rating implies patients had a high likelihood of cardiopulmonary concern as one of the NANDA-Is. The rule was supported by 30% of the EOL pain episodes and can be stated with a confidence value of 84%. Its interpretation is as follows: among all the episodes that had acute pain and were discharged from the hospital to a hospice or died with worse pain, 84% of them also had other NANDA-Is related to cardiopulmonary functions. We computed the p-value for this rule using Null-Hypothesis testing and it is < 0.001. Clinically and statistically significant rules such as above will be included in our proposed user interface (visualization screens) to serve as a critical reminder to the health care staff whenever a new patient arrives with conditions similar to this rule.

4. Discussion

From the data analysis perspective, this article presents the usefulness of multiple data mining tools for gleaning information from historical health care datasets and then using the information to improve decision making. However key to the successful use of diverse mining tools is the fact that the underlying data set was based on standardized language in the form of well established nursing terminologies—NANDA-Is, NICs, NOCs and capture of this information into a standardized database architecture. Such standardization greatly reduced noise present in the data. In the absence of standardized data sets, pre-processing overheads will limit the scope and applicability of the data mining techniques we studied. We observed that the standardization made preprocessing steps common to a diverse set of techniques. The data mining techniques pursued in this study can be easily extended to larger data sets and data sets that have similar standardized attributes.

From the clinical analysis perspective, the insights discovered through mining the dataset in this study revealed enormously important trends that have been previously invisible to clinicians. The results provide powerful evidence to support decision making that can dramatically improve the care of dying patients.

As such findings from this study provide only a peek into the wealth of information that can be garnered from data gathered in well designed documentation systems that support practice and generate high quality data amenable to mining. Figure 6 shows an example of a user interface to a system that incorporates the discovered knowledge into future decision-making. The proposed system provides alerts when it discovers, through classification and training, symptoms matching with the historical data in the database, and suggests best practices when feasible. For example, the figure shows the record of a patient whose pain management is not responding well to existing treatments, and the pain trend is pointing towards a known category. Based on this it indicates a red arrow, and clicking on that arrow provides the description of NICs that have worked well in similar circumstances. Note that the decision-making is still in the hands of care provider. The system works just like an intelligent assistant aimed at minimizing errors and improving health care delivery

A more elaborate study that considers different indicators and combinations of other NANDA-Is in EOL patients and how they interact with NOCs and NICs is expected to reveal an abundance of details of indicator relationships that will be extremely useful to the health care providers, administrators, etc. Although researchers repeatedly find that the vast majority of dying patients state a preference to die at home, most die in the hospital (Murray, Fiset, Young, & Kryworuchko, 2009). For more than 15 years since the large

SUPPORT study focused on improving outcomes in the dying (Teno et al., 2000; Teno et al., 1997), medical investigators have tested interventions to improve the quality of dying and to assure that the place of death is consistent with patients' wishes but success has been elusive (Curtis et al., 2011). Large studies of nursing interventions to improve outcomes of dying patients in hospitals, however, are not available. One reason for this research gap is lack of EHRs that adequately represent nursing care of hospitalized patients with standardized language. Based on a decade of development and testing (Keenan et al., In press), the HANDS system with its technical infrastructure and integration of standardized language, provides a tested method for gathering data that can advance practice-based knowledge (Keenan et al., In press) for all types of patients and settings.

Too often, the EHR contains insufficient information to represent nursing knowledge (Goodwin et al., 2003) and yet nursing care has a dramatic impact on the cost and quality of care. Using standardized language to represent nursing in documentation and storing it in a standardized database structure provides a rich source of data for evaluating and improving nurse practices. The HANDS documentation method is standardized and linked in a complex relational SQL database. In this article we provide one example, EOL patient care, of how capturing nursing care in a electronic documentation system can become a source of powerful evidence to transform nursing and health care practices. In conclusion, a major benefit of utilizing standardized languages in documentation systems that gather and store data in standardized databases is that the generation of data that are amenable to easy analysis with informatics tools that can cost-effectively be utilized to transform health care practice. We tapped only a fraction of the insights that are possible to uncover with data mining tools. Additional research is needed to maximize discovery of meaningful information about EOL and all types of patient care from mining actual practice based data. Also, improved data visualization techniques and user interfaces need to be investigated to transparently present gleaned information at decision making portals.

Acknowledgments

References

Allen D. Record-keeping and routine nursing practice: The view from the wards. Journal of Advanced Nursing. 1998; 27(6):1223–1230. [PubMed: 9663874]

Breault JL, Goodall CR, Fos PJ. Data mining a diabetic data warehouse. Artificial Intelligence in Medicine. 2002; 26(1–2):37–54. [PubMed: 12234716]

Bulechek, GM.; Butcher, HK.; Dochterman, JM. Nursing interventions classification (NIC). St. Louis, MO: Mosby; 2007.

Cheung RB, Moody LE, Cockram C. Data mining strategies for shaping nursing and health policy agendas. Policy Politics Nursing Practice. 2002; 3(3):248–260.

Curtis JR, Nielsen EL, Treece PD, Downey L, Dotolo D, Shannon SE, Engelberg RA. Effect of a quality-improvement intervention on end-of-life care in the intensive care unit: a randomized trial. American Journal of Respiratory and Critical Care Medicine. 2011; 183(3):348–355. [PubMed: 20833820]

Duan, L.; Street, W.; LU, D. A nursing care plan recommender system using a data mining approach. In: Li, J.; Aleman, D.; Sikora, R., editors. Paper presented at the Proceedings of the 3rd INFORMS Workshop on Data Mining and Health Informatics; Washington, DC. 2008.

Embley DW, Tao C, Liddle SW. Automating the extraction of data from HTML tables with unknown structure. Data & Knowledge Engineering. 2005; 54(1):3–28.

Fayyad, UM. Advances in knowledge discovery and data mining. Menlo Park, CA: AAAI Press; 1996.

Goodwin L, VanDyne M, Lin S, Talbert S. Data mining issues and opportunities for building nursing knowledge. Journal of Biomedical Informatics. 2003; 36(4–5):379–388. [PubMed: 14643734]

Hanauer DA, Rhodes DR, Chinnaiyan AM. Exploring clinical associations using '-omics' based enrichment analyses. PLoS ONE. 2009; 4(4):e5203. [PubMed: 19365550]

Hardey M, Payne S, Coleman P. 'Scraps': Hidden nursing information and its influence on the delivery of care. Journal of Advanced Nursing. 2000; 32(1):208–214. [PubMed: 10886453]

Heckerman D. Bayesian networks for data mining. Data Mining and Knowledge Discovery. 1997; 1(1):79–119.

Hsia T, Lin L. A framework for designing nursing knowledge management systems. Interdisciplinary Journal of Information, Knowledge, and Management. 2006; 1:13–22.

Karkkainen O, Bondas T, Eriksson K. Documentation of individualized patient care: A qualitative metasynthesis. Nursing Ethics. 2005; 12(2):123–132. [PubMed: 15791782]

Keenan G, Barkauskas V, Johnson M, Maas M, Moorhead S, Reed D. Establishing the validity, reliability, and sensitivity of NOC in adult care nurse practitioner clinics. Outcomes Management. 2003; 7(7):74–83. [PubMed: 12715604]

Keenan G, Stocker J, Barkauskas V, Johnson M, Maas M, Moorhead S, Reed D. Assessing the reliability, validity, and sensitivity of nursing outcomes classification in home care settings. Journal of Nursing Measurement. 2003; 11(2):135–155. [PubMed: 15274522]

Keenan G, Tschannen GD, Wesley M. Standardized nursing terminologies can transform practice. Journal of Nursing Administration. 2008; 38(3):103–106. [PubMed: 18327054]

Keenan, G.; Yakel, E. Promoting safe nursing care by bringing visibility to the disciplinary aspects of interdisciplinary care; AMIA Annual Symposium Proceedings; 2005. p. 385-389.

Keenan G, Yakel E, Dunn Lopez K, Tschannen D, Ford YB. Challenges to nurses' efforts of retrieving, documenting and communicating patient care. Journal of the American Medical Informatics Association. (In Review).

Keenan G, Yakel E, Yao YW, Xu D, Szalacha L, Tschannen D, Wilkie DJ. Maintaining a consistent big picture: Meaningful use of a web-based POC EHR system. International Journal of Nursing Knowledge. (In press).

Kraft, M.; Desouza, K.; Androwich, I. Data mining in healthcare information systems: Case study of a Veterans' Administration spinal cord injury population. Paper presented at the Hawaii International Conference on System Sciences, HI; 2003.

Lee SM, Abbott PA. Bayesian networks for knowledge discovery in large datasets: basics for nurse researchers. Journal of Biomedical Informatics. 2003; 36(4–5):389–399. [PubMed: 14643735]

Managed Care Outlook. The road to EHR implementation is paved with incentives and challenges. Managed Care Outlook. 2010; 23(1)

Moorhead, S.; Johnson, M.; Maas, M. Iowa Outcomes Project. Nursing outcomes classification (NOC). St. Louis, MO: Mosby; 2004.

Mullinsa IM, Siadatya MS, Lymana J, Scullya K, Garrettb CT, Millerb WG, Knaus WA. Data mining and clinical data repositories: Insights from a 667,000 patient data set. Computers in Biology and Medicine. 2006; 36:1351–1377. [PubMed: 16375883]

Murray MA, Fiset V, Young S, Kryworuchko J. Where the dying live: a systematic review of determinants of place of end-of-life cancer care. Oncology Nursing Forum. 2009; 36(1):69–77. [PubMed: 19136340]

NANDA International. Nursing diagnoses: Definitions and classifications 2003–2004. Philadelphia, PA: Author; 2003.

Rudman WJ, Brown C, Andrew MD, Hewitt CR, Carpenter WO, Campbell B, Noble SL. The use of data mining tools in identifying medication error near misses and adverse drug events. Topics in Health Information Management. 2002; 23(2):94–103.

Smith T, Coyne P, Cassel B, Penberthy L, Hopson A, Hager M. A high-volume specialist palliative care unit and team may reduce in-hospital end-of-life care costs. Journal of Palliative Medicine. 2003; 6(5):699–705. [PubMed: 14622449]

Stephan C, Cammann H, Semjonow A, Diamandis EP, Wymenga LF, Lein M, Jung K. Multicenter evaluation of an artificial neural network to increase the prostate cancer detection rate and reduce unnecessary biopsies. Clinical Chemistry. 2002; 48(8):1279–1287. [PubMed: 12142385]

Teno JM, Fisher E, Hamel MB, Wu AW, Murphy DJ, Wenger NS, Harrell FE Jr. Decision-making and outcomes of prolonged ICU stays in seriously ill patients. Journal of the American Geriatrics Society. 2000; 48(5 Suppl):S70–S74. [PubMed: 10809459]

Teno JM, Lynn J, Connors AF Jr, Wenger N, Phillips RS, Alzola C, Knaus WA. The illusion of end-of-life resource savings with advance directives. Journal of the American Geriatrics Society. 1997; 45(4):513–518. [PubMed: 9100723]

Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salame G. EU-ADR Group. Data mining on electronic health record databases for signal detection in pharmacovigilance: Which events to monitor? Pharmacoepidemiology and Drug Safety. 2009; 18(12):1176–1184. [PubMed: 19757412]

Westra B, Dey S, Steinbach M, Kumar V, Oancea C, Savik K, Dierich M. Interpretable predictive models for knowledge discovery from home-care electronic health records. Journal of Healthcare Engineering. 2011; 2(1):55–71.

Westra B, Savik K, Oancea C, Chormanski L, Holmes JH, Bliss D. Predicting improvements in urinary and bowel incontinence for home health patients using electronic health record data. Journal of Wound Ostomy and Continence Nursing. 2011; 38(1):77–87.

Keenan G, Yakel E, Keenan G, Tschannen D, Ford Y, Sorokin O. The case for the nurse as logical integrator of interdisciplinary information for the her. Journal of the American Medical Informatics Association ealth Services Research. (In Review).

Yu, W. End of life care: Medical treatments and costs by age, race, and region. 2006. IIR-02-189, from http://www.hsrd.research.va.gov/research/abstracts.cfm?Project_ID=2141693100

Zhang B, Wright AA, Huskamp HA, Nilsson ME, Maciejewski ML, Earle CC, Prigerson HG. Health care costs in the last week of life: Associations with end-of-life conversations. Archives of Internal Medicine. 2009; 169(5):480–488. [PubMed: 19273778]
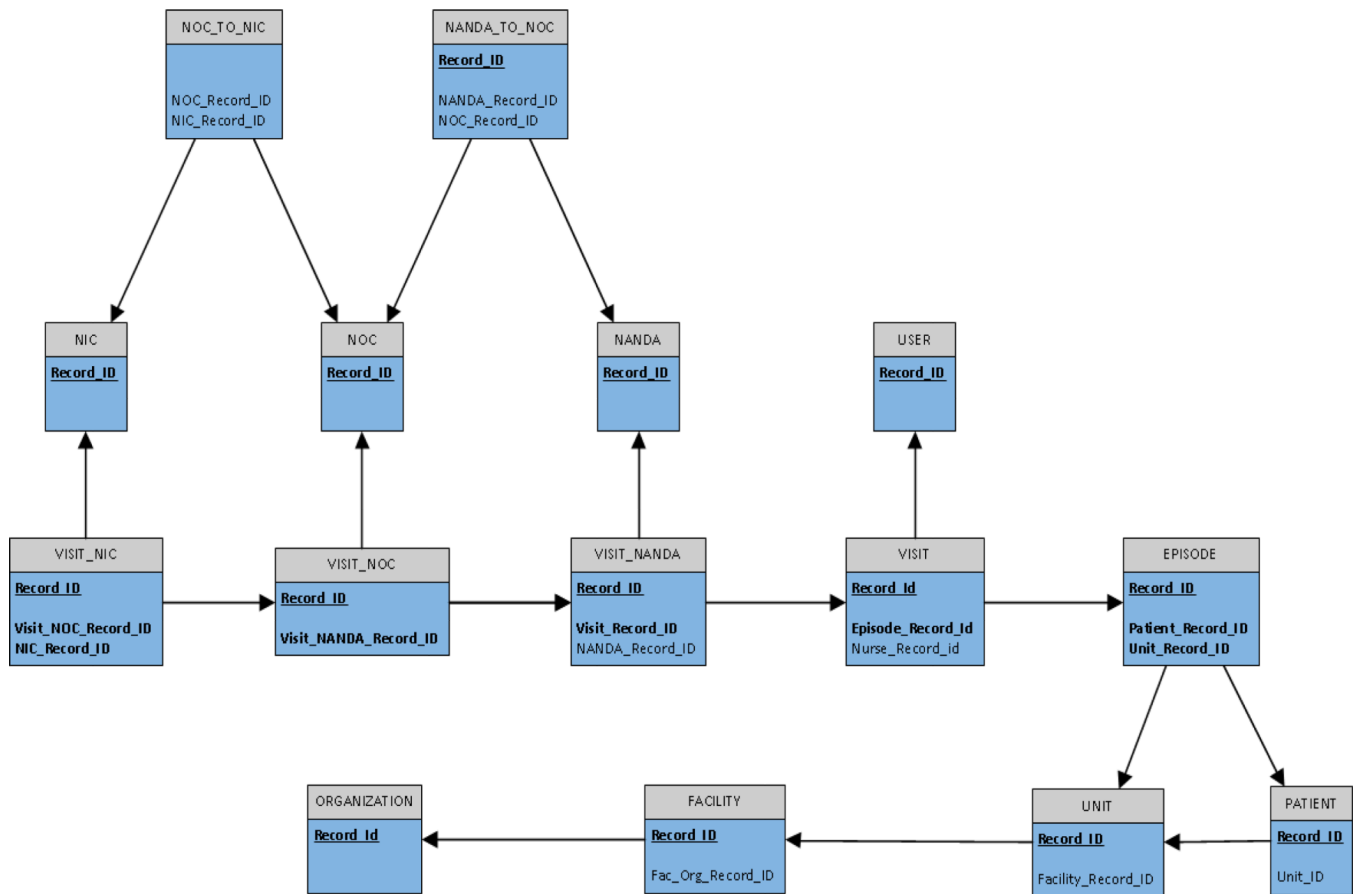
Figure 1. An abstract diagram showing different entities and their inter-dependencies in the HANDS database

Each box shows a unique feature being stored in the database for each patient and/or care provider. Copyright 2012 HANDS Team, used by permission.
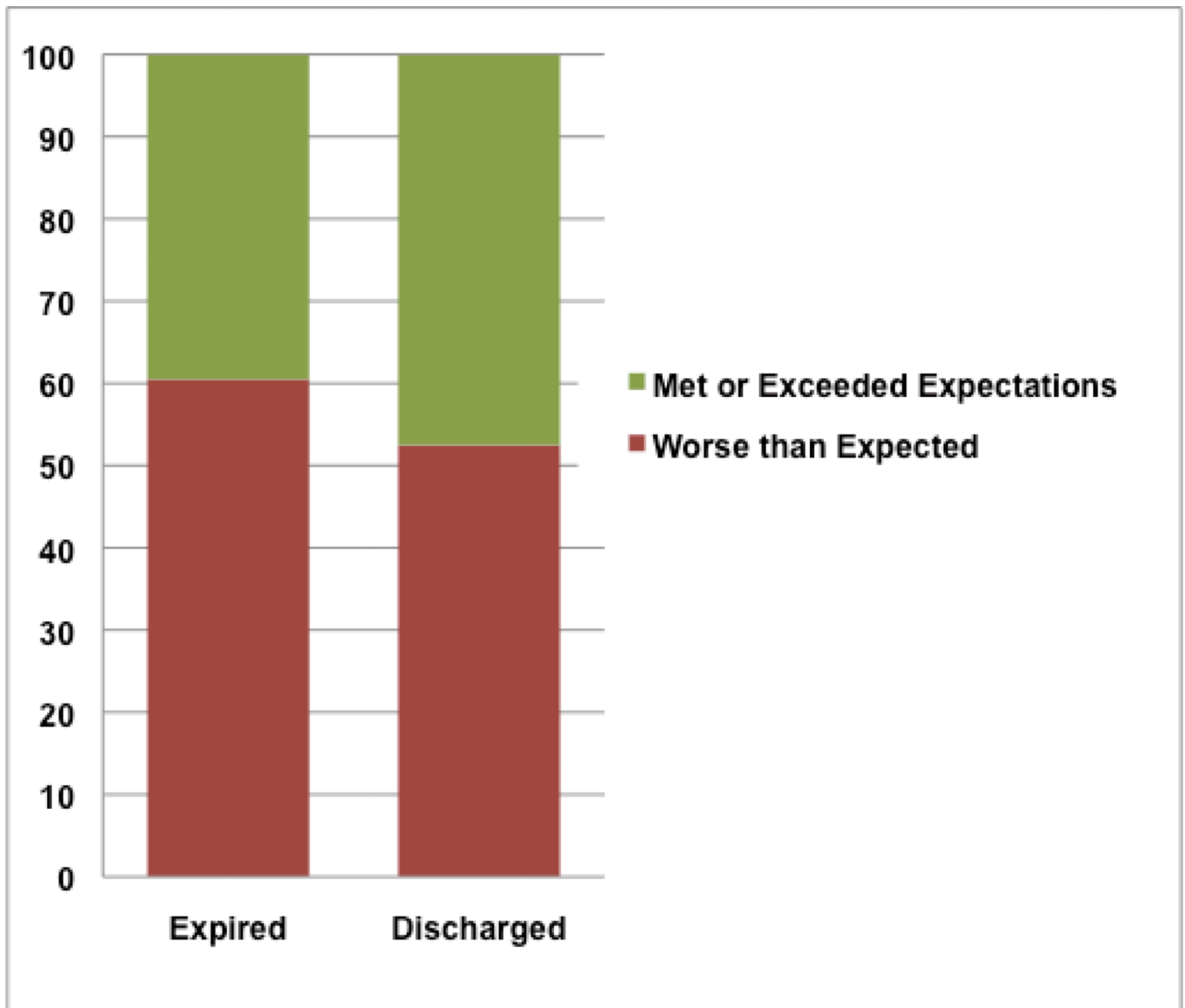
Figure 2. Percentage of EOL patients (all hospitals) meeting the expected pain level ratings at discharge
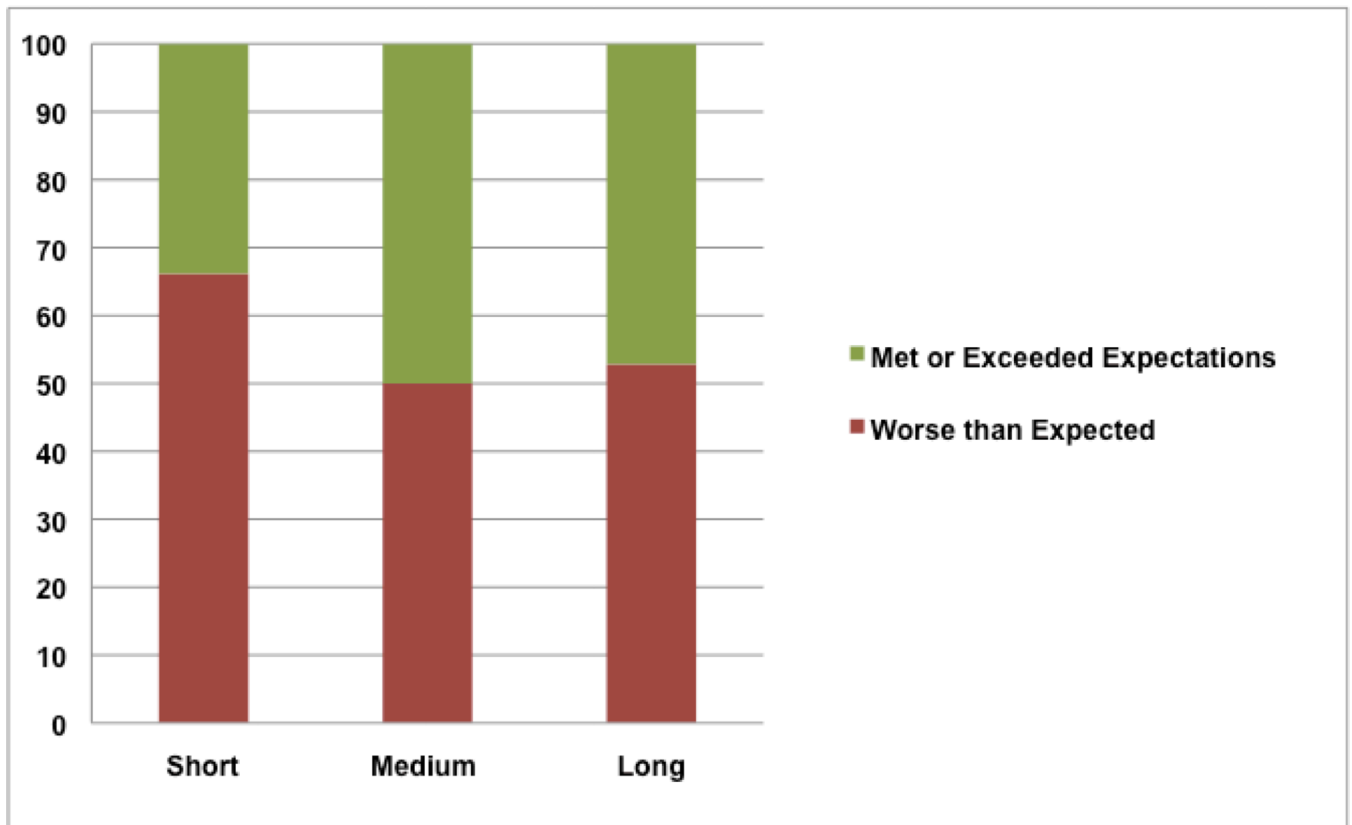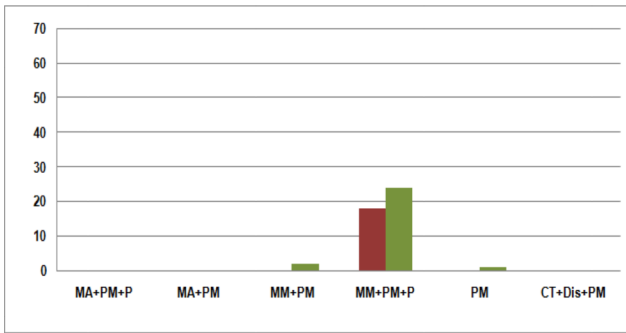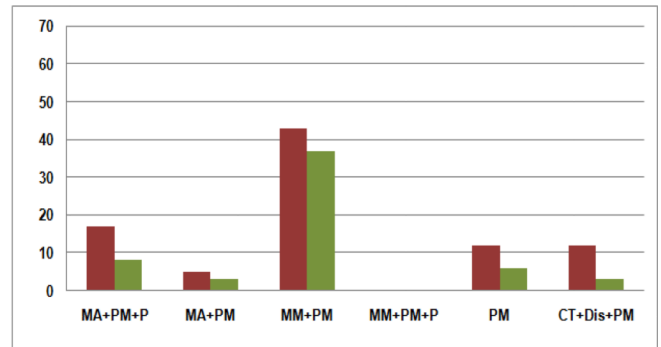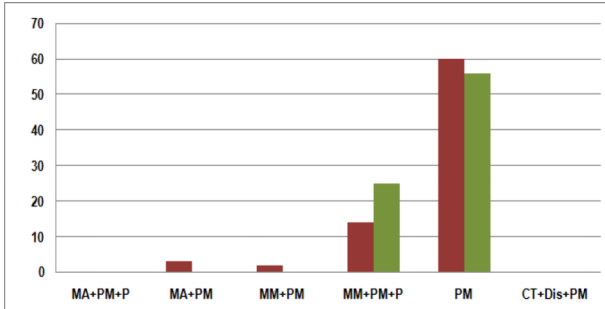More than half of the patients do worse than expected.($X^2(1)$=7.49, p-value: 0.0006).

Figure 3. Percentage of patients meeting their pain related NOC expected outcome ratings by their length of stay
(short < 72 hours, medium 72 to 199 hours, and long >= 200 hours) across all four hospitals, *p*-value: < .001.
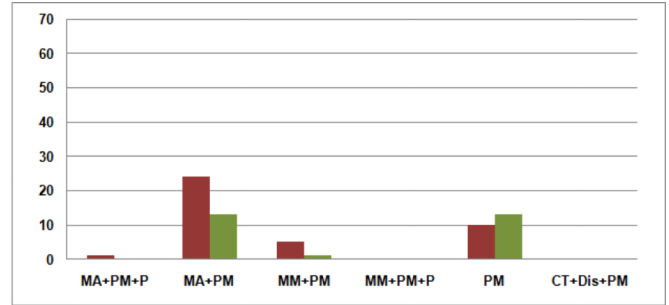
Small Community Hospital (SC): Hands1-1

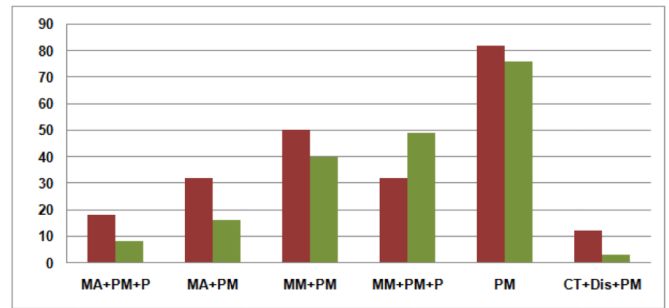Large Community Hospital 1 (LCH1): Hands1-2

Large Community Hospital 2 (LCH2): Hands2

University Hospital (UH): Hands 3

MM: Medication Management
PM: Pain Management
MA: Medication Administration
P: Positioning
CT: Calming Technique
Dis: Distraction

■ Worse then Expected
■ Met or Exceeded Expectations

Across all 4 hospitals, p-value: < 0.001

Figure 4. The impact of different combinations of NIC interventions on the pain related NOC outcome ratings
Shown are the numbers of patients who received the different combinations of NIC interventions by hospital andwhether they met or failed to meet their expected pain level rating (at discharge).In this example, adding Positioning (P) to Medication Management (MM) and Pain Management (PM) resulted in a statistically significant greater likelihood of meeting the expected pain-level rating at discharge or death.

Initial Pain Rating / Expected Pain Rating	Severe pain (1 - 2)	Medium pain (3 - 4)	No pain (5)
Severe pain (1 - 2)	2% ↑	3% ↑	
Medium pain (3 - 4)	7% ↑	20% ↑	
	13% ↓	10% ↓	
No pain (5)	0% ↑	6% ↑	3% ↑
	6% ↓	30% ↓	0% ↓

Percentage is over total number of 524 EOL episodes, 24 hours after admission.

Met or Exceeded Expectations
Worse than Expected

Figure 5. Predicting patient's discharge pain rating from rating at 24 hours post admission
Using the patient's pain rating at 24 hours post admission this table shows the distribution of
expected pain-level ratings at discharge or death, and the percentages of patients meeting or
failing to meet their expected ratings. The highlighted cell shows a patient with medium pain
at the 24-hour point who is expected to have no pain by the time of discharge. 36% of
patients fall into this category, but of those only 1 in 6 will meet the pain free goal.
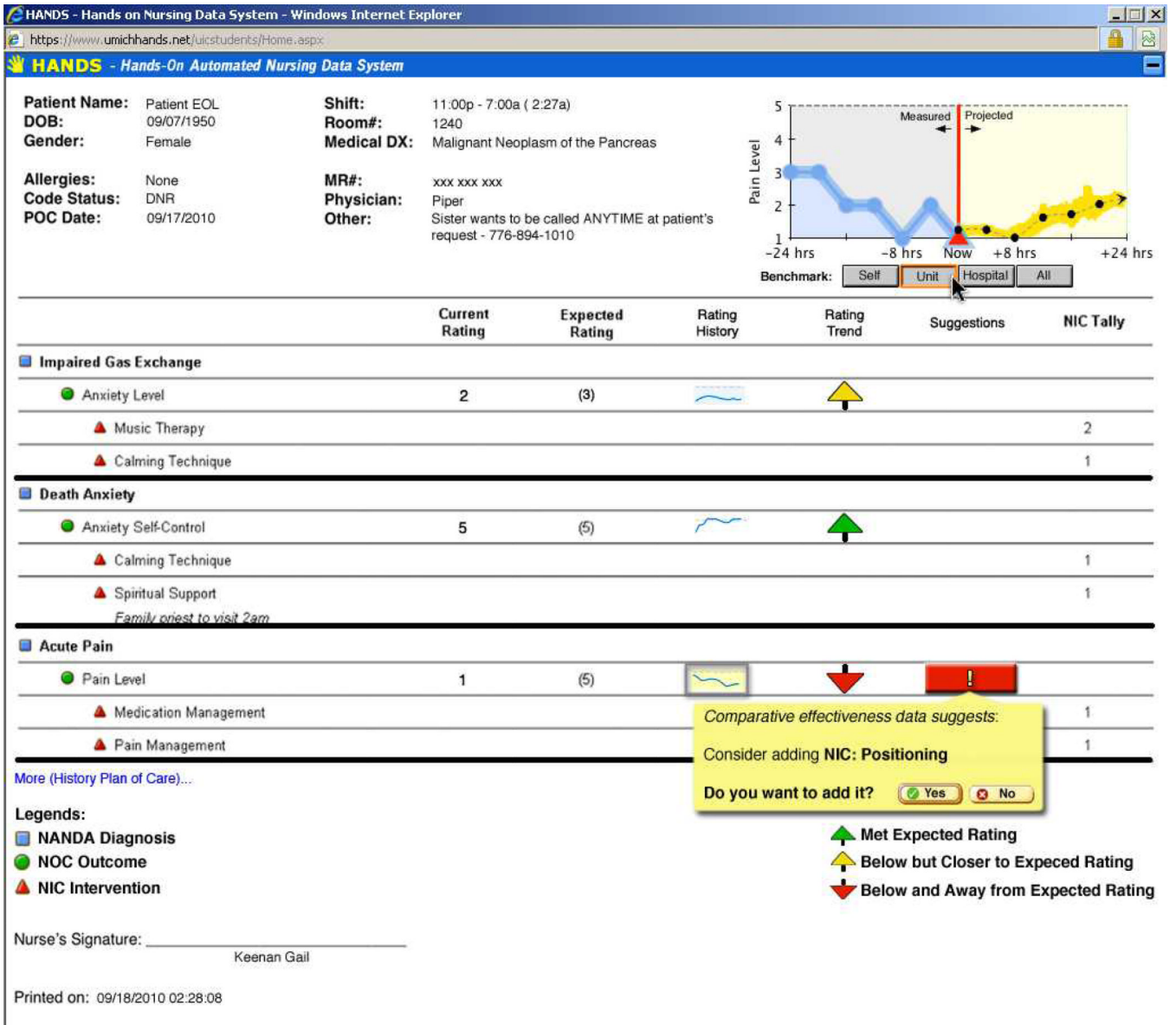Copyright 2012 HANDS Team, used by permission.

Figure 6. User Interface of an example decision-making assistant based on knowledge discovery using data mining of Nursing Care Plans

Examples shows that system has identified a patient with pain-trends that match with the historic data, and tags the patient data records with a red arrow and based on the data mining outcomes suggests interventions that are known to work well for such patients. Copyright 2012 HANDS Team, used by permission.

Table 1

All EOL Episodes and those with NANDA-I Pain diagnoses

Number of EOL patients as a percentage of all patients and the number of EOL patients with a pain diagnosis as a percentage of all EOL patients in the HANDS database from 2005 to 2007. LCH1 = Large community hospital; UH= University hospital; LCH2 = Large community hospital; SCH = Small community hospital, all in the Midwest. Copyright 2012 HANDS Team, used by permission.

Hospital	Unit	Unit Type	Total Number of Episodes of Care	Total Number of EOL Episodes	% EOL Episodes	Number of EOL Episodes with Pain	% Pain of EOL Episodes
LCH1	1	General Medical	5,640	189	3.4 %	72	38.1 %
LCH1	2	Medical ICU	1,228	163	13.3 %	69	42.3 %
LCH1	3	Gerontology	9,565	519	5.4 %	113	21.8 %
UH	4	Cardiac Surgical	6,112	51	0.8 %	33	64.7 %
UH	5	Neuro-Surgical	8,216	97	1.2 %	65	67.0 %
LCH2	6	Medical Gerontology	1,673	116	6.9 %	71	61.2 %
LCH2	7	General Medical	3,432	156	4.5 %	104	66.7 %
SCH	8	Medical Surgical	4,881	134	2.7 %	69	51.5 %
	Total		40,747	1,425	3.5 %	596	41.8 %