

Extracting Information on Pneumonia in Infants Using Natural Language Processing of Radiology Reports

Eneida A Mendonca	Janet Haas	Lyudmila Shagina	Elaine Larson	Carol Friedman
Biomedical Informatics Columbia University New York, NY, USA em264@columbia.edu	Infection Control New-York Presbyterian Hospital New York, NY, USA jah9012@nyp.org	Biomedical Informatics Columbia University New York, NY, USA ls303@columbia.edu	Pharmaceutical and Therapeutic Research Columbia University New York, NY, USA el123@columbia.edu	Biomedical Informatics Columbia University New York, NY, USA cf9@columbia.edu

Abstract

Natural language processing (NLP) is critical for improvement of the healthcare process because it has the potential to encode the vast amount of clinical data in textual patient reports. Many clinical applications require coded data to function appropriately, such as decision support and quality assurance applications. However, in order to be applicable in the clinical domain, performance of the NLP systems must be adequate. A valuable clinical application is the detection of infectious diseases, such as surveillance of healthcare-associated pneumonia in newborns (e.g. neonates) because it produces significant rates of morbidity and mortality, and manual surveillance of respiratory infection in these patients is a challenge. Studies have already demonstrated that automated surveillance using NLP tools is a useful adjunct to manual clinical management, and is an effective tool for infection control practitioners. This paper presents a study aimed at evaluating the feasibility of an NLP-based electronic clinical monitoring system to identify healthcare-associated pneumonia in neo-

nates. We estimated sensitivity, specificity, and positive predictive value by comparing the detection with clinicians' judgments and our results demonstrated that the automated method was indeed feasible. Sensitivity (recall) was 87.5%, and specificity (true negative rates) was 94.1%.

1 Introduction

Several studies have demonstrated the value of natural language processing (NLP) technology for a variety of healthcare applications. For example, NLP techniques have been used to analyze and structure narrative patient reports in order to provide data for applications, such as automated encoding, decision support, patient management, quality assurance, outcomes analysis, and clinical research (Baud et al., 1995; Fiszman et al., 2000; Friedman et al., 1994; Friedman et al., 1999b; Gundersen et al., 1996; Haug, Ranum, and Frederick, 1990; Sager et al., 1995). Additionally, data mining and knowledge discovery techniques have been used to automate the development of rules that detect clinical conditions by interpreting data generated from the natural language processing output of narrative reports (Wilcox and Hripcsak, 2000). NLP is potentially an invaluable tool for

healthcare because it enables access to a rich and varied source of clinical data. However, adequate performance is critical for practical clinical applications as well as timeliness.

One type of infectious disease that is important to monitor is healthcare-associated pneumonia in preterm and full-term neonates because it remains a significant cause of morbidity and mortality in that population (Whitsett et al., 1999). The incidence of pneumonia in Neonatal Intensive Care Units can be high as 10% in the United States (Gaynes et al., 1996), with mortality varying from 5-20% in cases of acquired pneumonia (Gaynes et al., 1996; Zangwill, Schuchat, and Wenger, 1920). Healthcare-associated pneumonia is an infection that is acquired during hospitalization, or in emergency departments and outpatient clinics, and it is neither present nor incubated at the time of admission.

The diagnosis of healthcare-associated pneumonia in neonates is extremely challenging, since neonates often do not exhibit ‘typical’ signs and symptoms of this infection (Bonten, 1999; Cordero et al., 2002; Cordero et al., 2000; Craven and Steger, 1998; Flanagan, 1999). In most cases, the final diagnosis is confirmed only by microbiologic culture, but it is difficult to obtain adequate specimens in neonates because of the invasive nature of this procedure (Heyland et al., 1999). Additionally, culture results are not timely (Cordero et al., 2002) because results are produced after 2 days, whereas results of radiology reports are usually obtained within 2 hours.

Surveillance require routine collection and analysis of relevant data, which must be promptly distributed to the appropriate health care providers, who then must use the data to take action and further prevent morbidity and mortality (Thacker, et al., 1986). Data provided by surveillance tools can be used for several purposes: (a) to identify the natural incidence of particular events, (b) to detect situations that require epidemiologic control measures, (c) to guide actions, allocation of resources, and interventions (Gaynes et al., 1996). Surveillance tools provide baseline information on trends and geographic distribution of conditions. An important aspect is the ability to detect an outbreak at the stage when intervention may affect the expected course of events (AHRQ, 2002). In order to facilitate infectious disease surveillance, several measures have been developed at the national

level. The Centers for Disease Control and Prevention (CDC), for example, has implemented measures to improve data collection and sharing for surveillance purposes. The National Nosocomial Infectious Surveillance System (NNISS) (Richards et al., 2001) is concerned with data standards and sharing on healthcare associated infections. The National Electronic Disease Surveillance System (NEDSS) focuses on standards and the response to biothreats.

2 Background

At the New York Presbyterian Hospital, a general NLP system in the clinical domain, called MedLEE (Medical Language Extraction and Encoding System) (Friedman et al., 1994), is routinely used to parse and encode clinical reports. It has been satisfactorily evaluated for clinical applications that require encoded data that is found in discharge summaries and radiology reports (Friedman et al., 1999b) (Friedman and Hripcsak, 1998; Friedman et al., 1999a). Hripcsak et al showed that, for particular clinical conditions found in chest radiographs, which included pneumonia, the performance of MedLEE was the same as that of physicians, and was significantly superior to that of lay persons and alternative automated methods (Hripcsak et al., 1995). In another study to evaluate a clinical guideline and an automated computer protocol for detection and isolation of patients with tuberculosis, Knirsch et al (Knirsch et al., 1998) demonstrated that automated surveillance is a useful adjunct to clinical management and an effective tool for infection control practitioners. That detection system monitored radiology reports encoded by MedLEE for evidence of radiographic abnormalities suggestive of tuberculosis along with other data in the patient repository that was already coded, such as the patient’s hospital location (for isolation status), laboratory and pharmacy data for immunological compromised status. Most importantly, the system detected patients who should be isolated that were not detected using the normal protocol (i.e. manual detection). MedLEE has also been extended to process pathology reports, echocardiograms, and electrocardiograms, but evaluations of performance in these areas have not yet been undertaken because evaluation is very costly in terms of time and personnel.

3 Methods

3.1 Overview of NLP System

MedLEE is composed of several different modules where each module processes and transforms the text in accordance with a particular aspect of language until a final structured output form is obtained. The structured output consists of primary units of clinical information (i.e. findings, procedures, and medications), along with corresponding modifiers (e.g. body locations, degree, certainty). Figure 1 shows an example of a simplified version of structured output that is generated as a result of processing the sentence *there is evidence of severe pulmonary congestion with question mild consolidation changes*.

finding: congestion body_location: lung certainty: high degree: high
finding: changes certainty: moderate degree: low descriptor: consolidation

Figure 1 – Sample output in simplified form for the sentence *there is evidence of severe pulmonary congestion with question mild consolidation changes*.

The output that is generated represents two primary clinical findings, **congestion** and **changes**. The first finding has a body location modifier **lung**, stemming from *pulmonary*, a certainty modifier **high**, stemming from *evidence of*, and a **degree** modifier **high**, stemming from *severe*. In the second finding, the **certainty** modifier **moderate**, corresponds to *question*, the **degree** modifier **low** corresponds to *mild*, and the **descriptor** corresponds to *consolidation*. Values for **degree** and **certainty** modifiers were automatically mapped to a small set of values in order to facilitate subsequent retrieval. The actual form of output generated by MedLEE is XML, but Figure 1 shows a compatible and more readable form.

Below is a brief overview of the system. More detailed descriptions were previously published (Friedman et al., 1994). When MedLEE was originally developed, it was intended to be used in conjunction with decision support applications, where high precision was critical. Therefore, it was initially designed to maximize precision and required a complete parse. However, subsequent clinical applications required high recall, and we discovered that flexibility was critical. Currently, MedLEE attempts to find a complete parse and only resorts to partial parsing when a full parse cannot be obtained. When generating the structured output, the method that was used to obtain the parse is saved along with the structured output so that the user can filter in or out findings accordingly.

Preprocessor - The preprocessor recognizes sentence boundaries, and also performs lexical lookup in order to recognize and categorize words, phrases, and abbreviations, and to specify their target forms. The lexicon was manually developed using clinical experts because of the need for high precision. In a study we used the UMLS (Unified Medical Language System) (Lindberg, Humphreys, and McCray, 1993), a controlled vocabulary developed and maintained by the National Library of Medicine, to automatically generate a lexicon. This lexicon was subsequently used by MedLEE instead of the MedLEE lexicon to process a set of reports. Results showed a significant loss of precision (from 93% to 86%) and recall (from 81% to 60%) when using the UMLS lexicon (Friedman, et al., 2001). Terms with ambiguous senses may be disambiguated in this stage based on contextual information. The preprocessor can also handle tagged text so that lexical definitions can be specified in the text, bypassing the need for lexical lookup for cases where the text is already tagged. This feature is particularly useful for handling local terminology (such as the names of local facilities), as well as for resolving domain specific ambiguities.

Parser - The parser uses a grammar and lexicon to identify and interpret the structure of the sentence, and to generate an intermediate structure based on grammar specifications. The grammar is a set of rules based on semantic and syntactic co-occurrence patterns. Development of manual rules

are costly, and we are currently investigating stochastic methods to help extend the grammar automatically.

Composer - The composer is needed to compose multi-word phrases that appear separately in the input sentence to facilitate retrieval later on. For example, the discontinuous words *spleen* and *enlarged* in *spleen appears enlarged* would be mapped to a phrase **enlarged spleen** so that a subsequent retrieval could look for that phrase rather than the individual components.

Encoder - The encoder maps the target terms in the intermediate structure to a standard clinical vocabulary (i.e. *enlarged spleen* is mapped to the preferred vocabulary concept *splenomegaly*) in the UMLS.

Chunker - The chunker increases sensitivity by using alternative strategies to break up and structure the text if the initial parsing effort fails.

3.2 Design of Feasibility Study

A two-year crossover design study was conducted independently of this NLP effort (03/01/2001-01/31/2002, 03/01/2002-01/31/2003) in two neonatal intensive care units (NICU) in New York City to study the impact of hand hygiene products on healthcare acquired infection:

- NICU-A: a 40-bed care unit, which cares for acutely ill neonates, including those requiring surgery for complex congenital anomalies and extra corporeal membrane oxygenation
- NICU-B: a 50-bed unit associated with a large infertility treatment practice

A trained infection control practitioner (ICP), using the CDC National Nosocomial Infection Surveillance System (NNIS) definitions, performed the surveillance for infections in both units. Cases were reviewed manually, including analysis of computerized radiology, pathology and microbiology reports as well as chart reviews and interviews with patient care providers. The diagnosis of infection was validated with the physician coinvestigator from each unit.

As part of this study, we evaluated the feasibility of using the NLP system (MedLEE) to auto-

matically identify potential cases of healthcare-associated pneumonia in neonates. The NLP system was not changed, but medical logic rules that accessed the NLP output had to be developed. The rules were developed by a medical expert based on modifications to a previous rule to detect pneumonia in adults (Hripcsak et al., 1995). Modifications were made in accordance with the CDC NNIS definition of healthcare-associated pneumonia in neonates. The final rule was then adapted to function properly with the output generated by MedLEE. For example, the rule looks for 38 different findings or modifier-finding combinations, such as *pneumatocele* and *persistent opacity*, and then filters out findings that are not applicable because they occur with certain modifiers (e.g. *no*, *rule out*, *cannot evaluate*, *resolved*, a total of 62 modifier). Therefore the automated monitoring system consists of two components: a) the MedLEE NLP system, and b) medical rules that access the output generated by MedLEE. In this first phase, the medical expert defined the rules broadly, to identify reports consistent with pneumonia (and not only healthcare-associated pneumonia) with the intention of continuing the effort if performance in identifying all forms of pneumonia was satisfactory. This means that the automated system could not differentiate between pneumonia and healthcare-associated pneumonia at this point. There were no probabilities associated with findings or combination of findings. The second phase of the study will use the results present in this work to refine the rules in order to differentiate between healthcare-associated and other types of pneumonia.

All chest radiograph reports of neonates admitted to NICU-A were processed using the automated monitoring system. To better assess true performance, no corrections were made to the reports despite misspellings and even the inclusion of other types of reports in the same electronic files as the chest radiograph reports. For instance, it is not uncommon to have a combined chest-abdomen radiograph in a neonate.

4 Results

During the 2 years of the study, from the total of 1,688 neonates admitted to the NICU-A, 1,277 neonates had 7,928 chest radiographs. Based on the experts' evaluation, only 7 neonates had

healthcare-associated pneumonia at least one point during the hospital stay. Cases were definitively confirmed by cultures. These patients had a total of 168 chest radiographs, but only 13, which were associated with the 7 patients, were positive because they contracted pneumonia at some point after their admission.

The automated system found the presence of pneumonia in 125 chest radiographs that were associated with 82 patients, including 6 of the 7 patients identified by the experts. The missed case was a neonate with cardiac problems, and the chest radiograph did not show findings of healthcare-associated pneumonia. A pulmonary biopsy performed subsequently showed findings which were consistent with healthcare-associated pneumonia.

For healthcare-associated pneumonia, the sensitivity (recall) of the automated system was 85.7%, while specificity (false positive rate) was 94.1%, and the positive predictive value (precision) was only 7.32%.

One of the authors (EAM), who is a board certified pediatric intensive care physician, manually analyzed the false positive cases (e.g. errors in precision), and found that several of the false positive cases actually had radiographic findings corresponding to pneumonia. Other errors require expert review of the entire patient charts to determine whether or not healthcare-associated pneumonia was present.

The expert reviewer (EAM) also encountered several occurrences of a missed abbreviation (“BPD”). Another common error was the misspelling of terms.

5 Discussion

Natural language processing has the potential to extract valuable data from narrative reports. The significance is that a vast amount of NLP structured data could then be exploited by automated tools, such as decision support systems. Automated alerts (Dexter et al., 2001; Hripcsak et al., 1990; Kuperman et al., 1999; Rind et al., 1994) require coded clinical data to do an intelligent analysis of patient status or condition. An automated tool, which notifies appropriate personnel about patients with a particular condition or infection facilitates timely and adequate response, including treatment, medication prophylaxis, and isolation.

Conditions such as healthcare-associated pneumonia carry significant rates of morbidity and mortality. Surveillance of respiratory infection in these patients is a challenge, and especially in neonates admitted to neonatal intensive care units. Isolated positive cultures alone do not distinguish between bacterial colonization and respiratory infection. Surveillance based on radiology and laboratory findings can be valuable as a complement to daily manual chart review and clinical rounds.

An NLP system cannot be used in a clinical environment without an infrastructure to support its use. At the NYPH, a clinical event monitor (Hripcsak et al., 1996) based on Arden Syntax for Medical Logic Modules – MLM (Hripcsak et al., 1990; Hripcsak et al., 1994) provides clinical decision support. When a clinical event occurs (such as uploading of a radiograph reports), appropriate medical logic modules are triggered based on the type of event. However, in order to be used by the monitoring system, narrative data must be coded. We envision the integration and use of this automated NLP system to facilitate surveillance of healthcare-associated pneumonia in a real clinical environment. An additional issue is that the data from the NLP system has to be represented in a way that can be manipulated by the clinical information system, and easily retrieved by the medical rules. Therefore it is not enough to evaluate an NLP system in isolation of a clinical application. The NLP system may perform very well in isolation, but the rules that access the data may be very complex. They may involve complex inferencing, or may be difficult to write because of the representation generated by the NLP system.

For healthcare-associated pneumonia, sensitivity (recall) and specificity (rate of true negatives) were appropriate for the clinical application (87.7% and 94.1% respectively), but the positive predictive value (precision) was low (7.32%), as expected in this phase. Low precision was primarily due to the broad rule that was used to detect pneumonia, and was not due to the NLP system itself. This rule now needs to be refined to detect only healthcare-associated pneumonia, and distinguish among radiograph findings moderately or highly suggestive of healthcare-associated pneumonia. That would require substantial effort involving manual chart review by an expert. Additional data from other sources, such as laboratory results, should also be combined with radio-

graph findings to add precision to the automated system. This will be done in the future as well as an evaluation. The data from NICU-B was reserved as a test set for this purpose.

The MedLEE system was not adapted in any way for this effort. Additionally, the rules were based on expert knowledge but there was no training of the rules because of the sparseness of the data. One type of NLP error was caused by a missed abbreviation *BPD*. A straightforward solution would be to include the abbreviation in the lexicon, but, this will create problems because of the ambiguous nature of the abbreviation. *BPD* has multiple meanings, including *broncopulmonary dysplasia*, *borderline personality disorder*, *biparietal diameter*, *bipolar disorder*, and *bilio-pancreatic diversion*, among others. This is not surprising since abbreviations are known to be highly ambiguous (Aronson and Rindflesch, 1994; Nadkarni, Chen, and Brandt, 2001), and are widespread in clinical text. In chest radiographs of neonates, *BPD* generally denotes *broncopulmonary dysplasia*, a condition that predisposes the patient to respiratory infection. In other types of radiology reports, for instance abdominal echography, *BPD* generally means *biparietal diameter*, a measure of the gestation age. Word sense disambiguation is a difficult problem, which is widely discussed in the computational linguistics literature. A review of methods for word sense disambiguation is presented by Ide and colleagues (Ide and Veronis, 1998). In the clinical setting, an important part of the solution will involve identifying the particular domain and use of special purpose domain-specific disambiguators that tag ambiguous abbreviations and specify their appropriate sense prior to parsing, based on the domain and other contextual information. Defining the appropriate domain granularity will be important, but may be a difficult task because the granularity may vary with the abbreviation. For example, in the case of radiographic reports, possibly the domain should involve all chest x-rays or only chest x-rays of neonates, or the specific type of reports.

In this study, we wanted to first evaluate the feasibility of automated surveillance based on NLP in a real clinical situation. The situation that presented itself was important but only involved a small population of positive cases. The results that were obtained are not meant to be definitive but to expose the issues associated with the use of an

automated system that uses NLP in a real environment. This study established a relationship with clinicians who need this technology. It is this collaboration, which is critical for furthering use of and validation of NLP in the clinical domain. In this study, for instance, upon reviewing our results, the infection control practitioner felt she may have missed some cases when following her typical manual surveillance, and would welcome the assistance of an automated system, even if it generated a manageable amount of false positives (false alerts). We do not know what that amount should be, but estimate that an amount in the range of a few false positives per week would be acceptable. In that case, the 82 false positives, accounting for 2 years of cases, would be very acceptable. This would need further studying.

Routine surveillance of infectious diseases in hospitals is generally accomplished by manual review of charts and clinical rounds by the ICPs. In case of suspected infection, the data are collected using surveillance protocols that target inpatients at high risk of infection. The CDC NNIS definition for healthcare-associated pneumonia is a 2-page written protocol with two different criteria. It is well known that interpretation of guidelines and protocols vary among health care providers, even within the same institution. A recent study on surveillance of ventilator-associated pneumonia (VAP) in very-low-weight infants retrospectively compare VAP surveillance diagnoses made by the hospital ICPs with those made by a panel of experts with the same clinical, laboratory, and radiologic data corroborates the variation among experts (Cordero, et al., 2000). An accurate NLP system, which codes reports consistently, should improve data collection for surveillance.

6 Conclusion

Surveillance of infectious disease is critical for health care but manual methods are costly, inconsistent, and error prone. An automated system using natural language processing would be an invaluable tool that could be used to improve surveillance, including emerging infectious diseases and biothreats. We performed a feasibility study in conjunction with an infectious disease control study to detect the presence of healthcare-associated pneumonia in neonates. The results showed that an automated system consisting of

NLP and clinical rules could be used for automated surveillance. Further work will include refinement of the rules, further evaluation, integration with the clinical environment, and identification of other surveillance applications.

Acknowledgment

This work was supported in part by grants LM06274 from the National Library of Medicine, 1 R01 NR05197-01A1 from the National Institute of Nursing Research, and by a gift from the Sulzberger Foundation.

References

- AHRQ. Bioterrorism preparedness and response: use of information technologies and decision support systems. *Evid Rep Technol Assess (Summ)* 2002; (59):1-8.
- Aronson AR, Rindfleshch TBA. Exploring a large thesaurus for information retrieval. *Proc. RIAO* 1994; 197-216.
- Baud RH, Rassinoux AM, Wagner JC *et al.* Representing clinical narratives using conceptual graphs. *Methods Inf Med* 1995; 34(1-2):176-86.
- Bonten MJ. Controversies on diagnosis and prevention of ventilator-associated pneumonia. *Diagn Microbiol Infect Dis* 1999; 34(3):199-204.
- Cordero L, Ayers LW, Miller RR, Seguin JH, Coley BD. Surveillance of ventilator-associated pneumonia in very-low-birth-weight infants. *Am J Infect Control* 2002; 30(1):32-9.
- Cordero L, Sananes M, Coley B, Hogan M, Gelman M, Ayers LW. Ventilator-associated pneumonia in very low-birth-weight infants at the time of nosocomial bloodstream infection and during airway colonization with *Pseudomonas aeruginosa*. *Am J Infect Control* 2000; 28(5):333-9.
- Craven DE, Steger KA. Ventilator-associated bacterial pneumonia: challenges in diagnosis, treatment, and prevention. *New Horiz* 1998; 6(2 Suppl):S30-45.
- Dexter PR, Perkins S, Overhage JM, Maharry K, Kohler RB, McDonald CJ. A computerized reminder system to increase the use of preventive care for hospitalized patients. *N Engl J Med* 2001; 345(13):965-70.
- Fiszman M, Chapman WW, Aronsky D, Evans RS, Haug PJ. Automatic detection of acute bacterial pneumonia from chest X-ray reports. *J Am Med Inform Assoc* 2000; 7(6):593-604.
- Flanagan PG. Diagnosis of ventilator-associated pneumonia. *J Hosp Infect* 1999; 41(2):87-99.
- Friedman C, Alderson PO, Austin JH, Cimino JJ, Johnson SB. A general natural-language text processor for clinical radiology. *Journal of the American Medical Informatics Association* 1994; 1(2):161-74.
- Friedman C, Hripcsak G. Evaluating natural language processors in the clinical domain. *Methods of Information in Medicine* 1998; 37:311-575.
- Friedman C, Hripcsak G, Shagina L, Liu H. Representing information in patient reports using natural language processing and the extensible markup language. *J Am Med Inform Assoc* 1999a; 6(1):76-87.
- Friedman C, Knirsch C, Shagina L, Hripcsak G. Automating a severity score guideline for community-acquired pneumonia employing medical language processing of discharge summaries. *Proc AMIA Symp* 1999b; 256-60.
- Friedman C, Liu H, Shagina L, Johnson S, Hripcsak G. Evaluating the UMLS as a source of lexical knowledge for medical language processing. *Proc AMIA Symp* 2001; 189-93.
- Gaynes RP, Edwards JR, Jarvis WR, Culver DH, Tolson JS, Martone WJ. Nosocomial infections among neonates in high-risk nurseries in the United States. *National Nosocomial Infections Surveillance System. Pediatrics* 1996; 98(3 Pt 1):357-61.
- Gundersen ML, Haug PJ, Pryor TA *et al.* Development and evaluation of a computerized admission diagnoses encoding system. *Computers and Biomedical Research* 1996; 29(5):351-72.
- Haug PJ, Ranum DL, Frederick PR. Computerized extraction of coded findings from free-text radiologic reports. *Work in progress. Radiology* 1990; 174(2):543-8.
- Heyland DK, Cook DJ, Griffith L, Keenan SP, Brun-Buisson C. The attributable morbidity and mortality of ventilator-associated pneumonia in the critically ill patient. *The Canadian Critical Trials Group. Am J Respir Crit Care Med* 1999; 159(4 Pt 1):1249-56.
- Hripcsak G, Clayton PD, Jenders RA, Cimino JJ, Johnson SB. Design of a clinical event monitor. *Comput Biomed Res* 1996; 29(3):194-221.
- Hripcsak G, Clayton PD, Pryor TA, Haug PJ, Wigertz O, van der Lei J. The Arden Syntax for medical logic modules. Miller RA. *Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care*. Washington, D.C.: IEEE Computer Press, 1990: 200-4.

- Hripcsak G, Friedman C, Alderson PO, DuMouchel W, Johnson SB, Clayton PD. Unlocking clinical data from narrative reports: a study of natural language processing. *Annals of Internal Medicine* 1995; 122(9):681-8.
- Hripcsak G, Ludemann P, Pryor TA, Wigertz OB, Clayton PD. Rationale for the Arden Syntax. *Comput Biomed Res* 1994; 27(4):291-324.
- Ide N, Veronis J. Introduction to the special issue on word sense disambiguation: the state of the art. *Computational Linguistics* 1998; 24:1-40.
- Knirsch CA, Jain NL, Pablos-Mendez A, Friedman C, Hripcsak G. Respiratory isolation of tuberculosis patients using clinical guidelines and an automated clinical decision support system. *Infect Control Hosp Epidemiol* 1998; 19(2):94-100.
- Kuperman GJ, Teich JM, Tanasijevic MJ *et al.* Improving response to critical laboratory results with automation: results of a randomized controlled trial. *J Am Med Inform Assoc* 1999; 6(6):512-22.
- Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods of Information in Medicine* 1993; 32(4):281-91.
- Nadkarni P, Chen R, Brandt C. UMLS concept indexing for production databases: a feasibility study. *J Am Med Inform Assoc* 2001; 8(1):80-91.
- Richards C, Emori TG, Edwards J, Fridkin S, Tolson J, Gaynes R. Characteristics of hospitals and infection control professionals participating in the National Nosocomial Infections Surveillance System 1999. *Am J Infect Control* 2001; 29(6):400-3.
- Rind DM, Safran C, Phillips RS *et al.* Effect of computer-based alerts on the treatment and outcomes of hospitalized patients. *Arch Intern Med* 1994; 154(13):1511-7.
- Sager N, Lyman M, Nhan NT, Tick LJ. Medical language processing: applications to patient data representation and automatic encoding. *Methods Inf Med* 1995; 34(1-2):140-6.
- Thacker SB, Redmond S, Rothenberg RB, Spitz SB, Choi K, White MC. A controlled trial of disease surveillance strategies. *Am J Prev Med* 1986; 2(6):345-50.
- Whitsett JA, Pryhuber GS, Rice WR, Warner BB, Wert SE. Acute respiratory disorders. Avery GB, Fletcher MA, MacDonald MG 5th edition. New York: Lippincott Williams & Wilkins, 1999: 485-508.
- Wilcox A, Hripcsak G. Medical text representations for inductive learning. *Proc AMIA Symp* 2000; 923-7.
- Zangwill KM, Schuchat A, Wenger JD. Group B streptococcal disease in the United States, 1990: report from a multistate active surveillance system. *MMWR CDC Surveill Summ* 1990; 41(6):25-32.