

# Retrospective Derivation and Validation of a Search Algorithm to Identify Emergent Endotracheal Intubations in the Intensive Care Unit

N.J. Smischney<sup>1,2</sup>; V.M. Velagapudi<sup>1</sup>; J.A. Onigkeit<sup>1</sup>; B.W. Pickering<sup>1,2</sup>; V. Herasevich<sup>1,2</sup>; R. Kashyap<sup>1,2</sup>

<sup>1</sup>Department of Anesthesiology, Mayo Clinic, Rochester, Minnesota; <sup>2</sup>Multidisciplinary Epidemiology and Translational Research in Intensive Care (METRIC), Mayo Clinic, Rochester, Minnesota

## Keywords

Electronic medical record, emergent, endotracheal intubation, intensive care unit, search algorithm

## Summary

**Background:** The development and validation of automated electronic medical record (EMR) search strategies are important in identifying emergent endotracheal intubations in the intensive care unit (ICU).

**Objective:** To develop and validate an automated search algorithm (strategy) for emergent endotracheal intubation in the critically ill patient.

**Methods:** The EMR search algorithm was created through sequential steps with keywords applied to an institutional EMR database. The search strategy was derived retrospectively through a secondary analysis of a 450-patient subset from the 2,684 patients admitted to either a medical or surgical ICU from January 1, 2010, through December 31, 2011. This search algorithm was validated against an additional 450 randomly selected patients. Sensitivity, specificity, and negative and positive predictive values of the automated search algorithm were compared with a manual medical record review (the reference standard) for data extraction of emergent endotracheal intubations.

**Results:** In the derivation subset, the automated electronic note search strategy achieved a sensitivity of 74% (95% CI, 69%-79%) and a specificity of 98% (95% CI, 92%-100%). With refinements in the search algorithm, sensitivity increased to 95% (95% CI, 91%-97%) and specificity decreased to 96% (95% CI, 92%-98%) in this subset. After validation of the algorithm through a separate patient subset, the final reported sensitivity and specificity were 95% (95% CI, 86%-99%) and 100% (95% CI, 98%-100%).

**Conclusions:** Use of electronic search algorithms allows for correct extraction of emergent endotracheal intubations in the ICU, with high degrees of sensitivity and specificity. Such search algorithms are a reliable alternative to manual chart review for identification of emergent endotracheal intubations.

## Correspondence to:

Nathan J. Smischney, MD  
Department of Anesthesiology  
Mayo Clinic  
200 First St SW, Rochester, MN 55905  
Email: smischney.nathan@mayo.edu

## Appl Clin Inform 2013; 4: 419–427

DOI: 10.4338/ACI-2013-05-RA-0033

received: May 20, 2013

accepted: August 16, 2013

published: September 4, 2013

**Citation:** Smischney NJ, Velagapudi VM, Onigkeit JA, Pickering BW, Herasevich V, Kashyap R. Retrospective derivation and validation of a search algorithm to identify emergent endotracheal intubations in the intensive care unit. *Appl Clin Inf* 2013; 4: 419–427  
<http://dx.doi.org/10.4338/ACI-2013-05-RA-0033>

## Introduction

Critically ill patients receive frequent interventions for diagnostic and therapeutic purposes. These interventions can be associated with complications. One common and life-saving intervention is endotracheal intubation (ETI). The ETIs that occur in the critical care setting are by definition emergent, with a few exceptions (e.g., ETI performed electively to facilitate a procedure-endoscopy). Unlike the elective ETI that occurs in the operating room, emergent ETI performed in a critically ill patient is often associated with hemodynamic instability (1-3). Studies performed in the intensive care unit (ICU) have addressed complications related to emergency ETIs in the ICU, but these studies primarily looked at airway complications (1, 4, 5). Therefore, relatively little data have been assessed for risk factors of hemodynamic instability during ETI in the ICU.

The risk factors associated with hemodynamic instability surrounding ETI in the ICU are largely unknown. This scarcity of knowledge stems from a lack of scientific investigation and the obstacle of defining the time period when an emergent ETI took place in the critically ill patient – information that usually is buried within the medical record. This obstacle is further compounded by inadequate search strategies or improper filtering with the use of the electronic medical record (EMR) in modern medical practice. Therefore, in order to address these potential risk factors, two elements must be defined: identifying an emergent ETI and at what time did it take place, designated as *time zero*.

Increasingly, automated search strategies are used to identify certain elements in a patient's EMR (6, 7). Recently, Singh et al (7) developed and validated automated electronic search strategies to extract Charlson comorbidities from EMRs. These investigators found that by using the electronic search strategies, they were able to achieve a sensitivity and specificity that were greater than 90%.

The primary objective of the present study was to develop and validate an automated electronic search strategy for emergent ETIs in the ICU. Identifying emergent ETIs is a necessary first step before establishing time zero, which is defined as when mechanical ventilation took place in the critically ill patient. Our secondary aim was to compare the sensitivity, specificity, and positive and negative predictive values of our electronic search strategy with a comprehensive manual review of the medical record (the reference standard), detecting emergent ETIs from the EMR.

## Methods

The study was approved by the Mayo Clinic Institutional Review Board for the use of existing medical records of patients who gave prior research authorization.

## Study Population

The derivation and validation subsets were obtained retrospectively from two critical care units at Mayo Clinic in Rochester, Minnesota. These subsets were a heterogeneous population of medical and surgical ICU patients admitted from January 1, 2010, through December 31, 2011. The data included only first-time ICU admission. Critically ill patients who have been admitted repeatedly to the ICU have significant alterations in their physiology that may hinder identification of possible risk factors leading to hemodynamic instability after emergent ETI for future study. Furthermore, although patients in other locations, such as the emergency department, may have complications from emergent ETI, they are potentially different with regard to their pathophysiology. Therefore, the present study focused on the critically ill population.

The total retrospective cohort contained 6,714 consecutive patients (► Figure 1). This number was reduced to 2,689 patients who received mechanical ventilation on first ICU admission during the study period, excluding 101 patients who did not provide prior research authorization. From this cohort, five patients were excluded because of age restriction (i.e., younger than 18 years). No other exclusion criteria were used. A subset of 450 randomly selected patients was used for derivation in 2010. The automated search algorithm was further validated against 450 randomly selected patients from the cohort in 2011, with use of statistical software (JMP version 9.0; SAS Institute Inc). The selection of 450 patients in both subsets was chosen to limit manual annotation burden while ensuring

a robust sample size for the two subsets. The selection of two different years was chosen to add more scientific rigor to the final search algorithm.

## Manual Data Extraction Strategies

No implicit gold standard is available for the identification of emergent ETI, but manual data extraction is the traditional method for assessing data in clinical research. The medical records of the derivation and validation groups were reviewed manually and independently by two critical care clinicians (N.J.S. and V.M.V.). The data were collected from procedure notes marked as *intubation* and then reviewed to assess the emergent vs. nonemergent status of the ETI note, which was listed as yes or no. In addition, all emergent procedure notes, whether related to ETI or not, were reviewed during each patient's ICU stay. The emergent ETI notes were used for data extraction because they provided the only documentation in the patient's EMR that indicates the emergent nature of the ETI. The accuracy or timing, or both, of emergent ETI notes have not been validated previously. However, there is no other form of documentation regarding the emergent status of ICU intubation. At our institution, these notes are required in documentation of ETI in the ICU.

Additional factors that represent ETI, such as mechanical ventilation parameters, were not used because these may be present for patients receiving noninvasive ventilation, as well as patients who require ventilatory support through a tracheotomy. The research team performing manual data extraction was not aware of the automated electronic note search strategy results.

## Automated Electronic Note Search Strategy

Mayo Clinic is unique regarding the numerous databases available for clinical research. The present retrospective study used data from the Mayo Clinic Life Sciences System (8), an exhaustive database with abundant patient information. Centralized for all Mayo Clinic hospital data, this database contains such patient information as demographic characteristics, diagnoses, laboratory test results, flow sheet data, and clinical and pathologic information gathered from various resources in the institution. This database has been validated previously and is reliable (9, 10). We used a Web-based commercial software tool set (Data Discovery and Query Builder [DDQB]; International Business Machines Corp) for data access. Within DDQB, a medical record can be searched for diagnosis codes (free text terms), laboratory test results, or procedure codes. DDQB is based on Boolean logic to create free text searches (10). With it, a researcher can search quickly for a unique entity by using a text search strategy.

To develop the electronic search strategy, we entered the term "Emergency: Yes." Additional criteria consisted of classifying "Event Type on Clinical Note" equal to "Intubation." The electronic search strategy was refined continuously through the addition or edit of terms to enhance sensitivity and specificity to greater than 90% in the derivation subset. For example, "Intubation" when used alone achieved poor performance. The performance improved, however, with the search term "Sedated" added but not with "Sedation." The final search terms used to build the automated electronic search were "Emergency: Yes" and modification of event type, such that "Event Type on Clinical Note" equaled one of the following: "Intubation," "Intubation V2," and "Sedated." The term "Sedated" was selected because ICU patients receiving sedation potentially have mechanical ventilation. The term "Intubation V2" represents version two of the intubation note, labeled in the EMR after the note has been finalized. We restricted the notes to the ICU duration of study. Additional restrictions were placed on location of emergent ETI whereby the included ETIs were only those performed in the ICU, as opposed to ETIs performed in the emergency department or on the medical ward. To validate the automated electronic search, sensitivity and specificity were calculated through comparison to the reference standard of comprehensive manual medical record review (► Figure 1). The automatic search strategy was done by an independent critical care research physician (R.K.).

## Statistical Analyses

We calculated sensitivity and specificity of the automated electronic note search strategy on the basis of comparisons of test results and the reference standard in both the derivation and validation pa-

tient subsets using Clinical Calculator 1 (<http://www.vassarstats.net/clin1.html>). The 95% CIs were calculated with an exact test for proportions. We used JMP statistical software version 9.0 for all data analysis.

## Results

The final retrospective cohort consisted of 2,684 eligible adult patients who were admitted to a medical or surgical ICU at Mayo Clinic in Minnesota during the study period. This retrospective cohort included patients with first-time ICU admission who required mechanical ventilation. In the derivation subset, 41 of the 450 patients had ETIs performed outside the time period of first ICU admission on manual review of their medical record and the ETI was coded as “0” or “nonemergent.”

The automated electronic search strategy achieved a sensitivity of 74% and a specificity of 98% in the derivation subset (▶ Table 1). Disagreement between the automated electronic search strategy and the manual review occurred in 72 patients. On review of these 72 medical charts, we noted several reasons for the findings, including first-time ICU admission dates, procedure note classifications, and procedure note locations. On the basis of these findings, the automated electronic note search strategy was refined further to arrive at our final automated search algorithm with a sensitivity of 95% and a specificity of 96% in the same derivation subset of 450 patients. In the validation subset of 450 randomly selected patients in 2011, the automated electronic search strategy achieved a sensitivity of 95% and a specificity of 100%.

The initial negative predictive value (NPV) and positive predictive value (PPV) in the derivation subset were 53% and 99%, with a prevalence of 0.77. Posttest refinements were made within the same derivation subset. In the validation subset, the automated electronic search strategy achieved a concordance for NPV of 99% (95% CI, 98%-100%) and for PPV of 98% (95% CI, 90%-100%).

## Discussion

The use of electronic search strategies has been increasing in the past decade with the increase in EMR adoption and the ability to combine distributed data sources. Survey findings from the Centers for Disease Control and Prevention reported an increase in EMR use by US office-based physicians from 18% in 2001 to more than 50% in 2011 (11). The push to adopt the EMR has been driven in large part by the US government. Incentive programs developed by the federal government promote adoption of EMRs, including the Health Information Technology for Economic and Clinical Health Act in 2009 (12, 13).

However, with the adoption of EMRs, the amount of information that can be assimilated is enormous and can potentially lead to barriers in clinical research. Development of electronic search algorithms can prove useful for clinical and research purposes. Thus, to identify risk factors associated with hemodynamic instability during emergent ETI, it is necessary to first identify whether an emergent ETI took place. After this is accomplished, establishing the timing of mechanical ventilation can be performed.

The results of the present study indicate that the development and validation of an electronic search algorithm within the EMR for identifying patients who required emergent ETI in the ICU is a reliable alternative to a manual chart review. By using this electronic search algorithm, we achieved both sensitivity and specificity greater than 90%, with an NPV of 99% and a PPV of 98%. This study's findings are in accordance with previously published studies showing that the use of electronic search strategies offers highly valid and reliable data extraction methods (10).

Misclassifications between our electronic note search strategy and manual medical record review occurred in 72 cases. We reviewed these 72 charts and the reasons for the misclassifications. The top three reasons, which together account for 80% of misclassifications, were related to the time frame surrounding first ICU admission, the location of the procedure, and the type of emergent procedure. For example, DDQB initially extracted all the emergent procedure notes occurring in the two ICUs during the specified period. These notes included emergent bronchoscopy or ETI in the emergency department, or both. These misclassifications were not surprising because patients having a bron-

choscopy are usually intubated for this procedure and therefore are registered as being intubated emergently by DDQB. The same reasoning resulted in any ETI occurring within the hospital as being an emergent ETI by DDQB. This review was refined to include the procedure note location (emergency department vs. ICU) and type (bronchoscopy vs. nonbronchoscopy). Further restrictions were placed on the time frame surrounding first ICU admissions (ETI occurring during the first ICU admission of the study period rather than throughout the patient's entire medical record), which resulted in the final reported sensitivity, specificity, NPV, and PPV.

This process was a necessary first step in identifying emergent ETIs in the critical care setting before establishing mechanical ventilation initiation and evaluating the risk factors that may be associated with hemodynamic instability during the emergent procedure. With EMR adoption, a vast amount of information can be assimilated quickly. If these records were analyzed retrospectively, time constraints may become a barrier because of the abundance of information.

Several studies address the complications related to emergent ETIs in the ICU (14–16). However, these investigations are mainly prospective studies with small sample sizes and therefore the information is easier to acquire. For example, to obtain the described data from manual medical record review, the time invested by our research team ranged from one minute to more than five minutes per medical record. The automatic electronic note search strategy was derived from DDQB using keyword phrases within specific note sections. This approach resulted in a tremendous reduction in time commitment compared with manual medical record review. Furthermore, the strategy not only is useful for research purposes, but also is of value in the patient care setting (17). The search strategy distinctly differs from the recently advanced approach of “natural language processing” in the following ways. We performed a free text search for a limited number of keywords. Essentially, this was a simple search to match words. Natural language processing is semantic mapping of uncertain text to controlled terminology (Systematized Nomenclature of Medicine-Clinical Terms [SNOMED CT]) (18). Natural language processing requires dedicated real-time software and hardware that may make the system more complex and less reliable. A direct query was submitted to a database using standard open database connectivity connection. Because DDQB is commercial-based software and the medical record is transitioning from paper to electronic, our approach is applicable to any electronic health system. Therefore, any type of standard or customized software can be used for this purpose. We used JMP.

The search strategy used in this context has several limitations. First, performance of the electronic note search strategy is dependent on the foundation of information from which it is derived. Inconsistencies in the database and text search phrases can lead to inaccurate results and thus limit the applicability of this approach to areas with a similar database. For example, although we used a free text search, we performed this task within a structured note. The procedure note has a limited number of designations in the fields we searched, which increases the specificity score (“Intubation V2” is an example). Therefore, the search algorithm used may be aided by a natural language processing approach to algorithm development. Second, we focused only on procedure notes classified as ETI within the critical care setting. If all the clinical notes were analyzed during the specified time interval, our results may have been different. This timely analysis was not feasible because of the time barrier involved in reviewing the entire medical record during the specified interval. Third, the timing of using any electronic search strategy is limited. Search algorithms are not real-time acquisition tools and depend on when the procedure note is posted. Therefore, they cannot be used in real time. Fourth, the iterative nature of query development requires independent validation of each modification. However, after the new query algorithm is built with optimal results, it could be automated. Fifth, data could have been entered in error or the database could have been corrupted (19). However, this fifth limitation likely accounts for only a small proportion of the database. Sixth, data collection may be reported inaccurately, as with any retrospective study. Seventh, this comprehensive EMR system is unique to Mayo Clinic. However, the way in which the search strategy was performed can be applied to any standard or customized software.

## Conclusions

Emergent ETIs within the critical care setting can be identified correctly through the use of an automated electronic search algorithm. The achieved sensitivity and specificity can approach 100% through refinements in the electronic note search strategy and can serve to expedite clinical research and, ultimately, patient care. The present study reports on the development and validation of an electronic search algorithm regarding emergent ETIs in the ICU. This will be followed by a study describing the development and validation of a search algorithm for the establishment of time zero. Both these electronic search algorithms may then be applied to ultimately determine risk factors associated with hemodynamic instability in the critically ill patient during emergent ETI in the ICU.

### Clinical Relevance Statement

Through the creation of electronic search algorithms, one can accurately and reliably extract data from the electronic medical record. The search strategy created for identifying emergent ETIs in the ICU resulted in a high degree of sensitivity and specificity. The free text search algorithm may bring more structure and standardization in the EMR by information modeling and terminologies, such as SNOMED-CT. Such algorithms can be used with any type of standard or customized software and are a reliable alternative to manual chart review.

### Conflicts of Interest

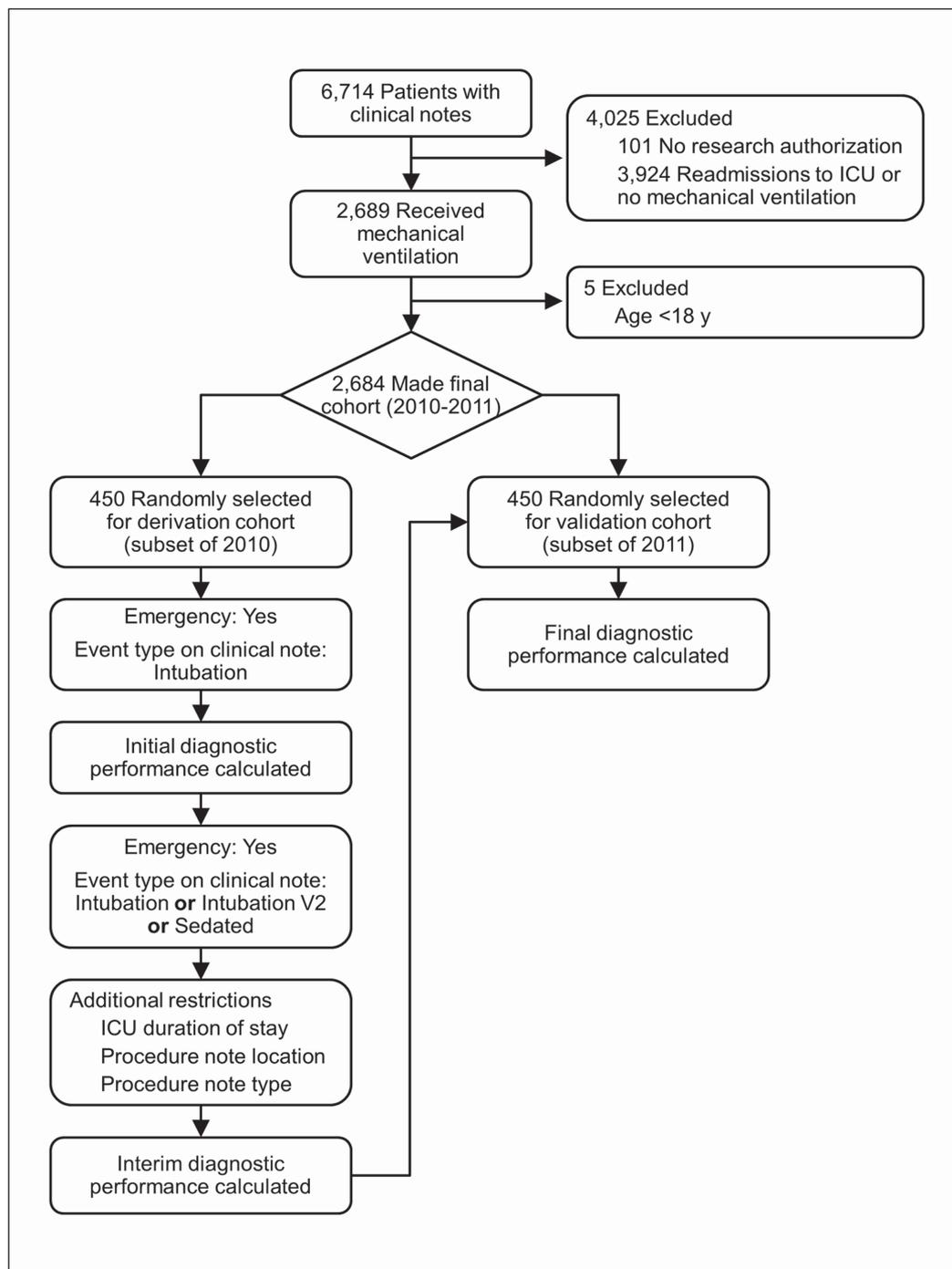
The authors have no conflicts of interest in the research.

### Financial Support and Disclosure

This work was supported by the Division of Critical Care Medicine with no direct financial support.

### Human Subject Research Approval

The study was approved by the Mayo Clinic Institutional Review Board for the use of existing medical records of patients who gave prior research authorization.



**Fig. 1** Electronic search strategy flow diagram of patients in the Mayo Clinic Life Sciences System from January 2010 through December 2011. ICU indicates intensive care unit.

**Table 1** Derivation and validation of electronic note search algorithm.

	Search Algorithm (Emergency: Yes)	Sensitivity % (95% CI)	Specificity % (95% CI)
<b>Initial signature</b>	event type: "intubation"	74 (69–79)	98 (92–100)
<b>Final signature</b>	event type: "intubation" or "intubation V2" or "sedated"	95 (91–97)	96 (92–98)
<b>Validation</b>	event type: "intubation" or "intubation V2" or "sedated"	95 (86–99)	100 (98–100)

**Table 2** Abbreviations

<b>DDQB</b>	Data Discovery and Query Builder
<b>EMR</b>	electronic medical record
<b>ETI</b>	endotracheal intubation
<b>ICU</b>	intensive care unit
<b>NPV</b>	negative predictive value
<b>PPV</b>	positive predictive value
<b>SNOMED-CT</b>	Systematized Nomenclature of Medicine-Clinical Terms

## References

1. Le Tacon S, et al. Complications of difficult tracheal intubations in a critical care unit. *Ann Fr Anesth Reanim.* 2000; 19(10): 719-724.
2. Stauffer JL, Olson DE, Petty TL. Complications and consequences of endotracheal intubation and tracheotomy: a prospective study of 150 critically ill adult patients. *Am J Med* 1981; 70(1): 65-76.
3. Griesdale DE, et al. Complications of endotracheal intubation in the critically ill. *Intensive Care Med* 2008; 34(10): 1835-1842.
4. Schwartz DE, Matthay MA, Cohen NH. Death and other complications of emergency airway management in critically ill adults: a prospective investigation of 297 tracheal intubations. *Anesthesiology* 1995; 82(2): 367-376.
5. Mort TC. Emergency tracheal intubation: complications associated with repeated laryngoscopic attempts. *Anesth Analg* 2004; 99(2): 607-613.
6. Clifford L, et al. Electronic health record surveillance algorithms facilitate the detection of transfusion-related pulmonary complications. *Transfusion* 2013; 53(6): 1205-1216.
7. Singh B, et al. Derivation and validation of automated electronic search strategies to extract Charlson comorbidities from electronic medical records. *Mayo Clin Proc* 2012; 87(9): 817-824.
8. Chute CG, Beck SA, Fisk TB, Mohr DN. The Enterprise Data Trust at Mayo Clinic: a semantically integrated warehouse of biomedical data. *J Am Med Inform Assoc* 2010; 17(2): 131-135.
9. Herasevich V, et al. Validation of an electronic surveillance system for acute lung injury. *Intensive Care Med* 2009; 35(6): 1018-1023.
10. Alsara A, et al. Derivation and validation of automated electronic search strategies to identify pertinent risk factors for postoperative acute lung injury. *Mayo Clin Proc* 2011; 86(5): 382-388.
11. Hsiao CJ, Hing E, Socey TC, Cai B. Electronic health record systems and intent to apply for meaningful use incentives among office-based physician practices: United States, 2001-2011. *NCHS Data Brief* 2011; (79): 1-8.
12. Bays RA, Kaelin LD. Electronic medical records for the office. *J Vasc Surg* 2010; 51(5): 1302-1308.
13. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. *N Engl J Med* 2010; 363(6): 501-504.
14. Mort TC. The incidence and risk factors for cardiac arrest during emergency tracheal intubation: a justification for incorporating the ASA Guidelines in the remote location. *J Clin Anesth* 2004; 16(7): 508-516.
15. Jaber S, et al. Clinical practice and risk factors for immediate complications of endotracheal intubation in the intensive care unit: a prospective, multiple-center study. *Crit Care Med* 2006; 34(9): 2355-2361.
16. Simpson GD, Ross MJ, McKeown DW, Ray DC. Tracheal intubation in the critically ill: a multi-centre national study of practice and complications. *Br J Anaesth* 2012; 108(5): 792-799.
17. Kutney-Lee A, Kelly D. The effect of hospital electronic health record adoption on nurse-assessed quality of care and patient safety. *J Nurs Adm* 2011; 41(11): 466-472.
18. Murff HJ, et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA* 2011; 306(8): 848-855.
19. Wisniewski MF, et al. Development of a clinical data warehouse for hospital infection control. *J Am Med Inform Assoc* 2003; 10(5): 454-462.