2014

Opportunities for Business Intelligence and Big Data Analytics in Evidence Based Medicine

Omar F. El-Gayar
Dakota State University

Prem Timsina
Dakota State University

Surendra Sarnikar
Dakota State University

2014 47th Hawaii International Conference on System Science

Opportunities for Business Intelligence and Big Data Analytics In Evidence Based Medicine

Omar El-Gayar
Dakota State University, USA
Omar.El-Gayar@dsu.edu

Prem Timsina
Dakota State University, USA
ptimsina@pluto.dsu.edu

Abstract

Evidence based medicine (EBM) is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. Each year, a significant number of research studies (potentially serving as evidence) are reported in the literature at an ever-increasing rate outpacing the translation of research findings into practice. Coupled with the proliferation of electronic health records, and consumer health information, researchers and practitioners are challenged to leverage the full potential of EBM. In this paper we present a research agenda for leveraging business intelligence and big data analytics in evidence based medicine, and illustrate how analytics can be used to support EBM.

1. Introduction

The United States spends more than $2.3 trillion per year in healthcare and is the second largest nation (just below Marshall Island) in healthcare spending as a percentage of GDP [1]. However, such spending has not translated into quality of care. As many as 98,000 people die annually in hospital because of medical errors [2]. Moreover, there is a significant gap between the healthcare we could have and the health care that is currently available in the United States [3, 4]. In response to this situation, The Institute of Medicine proposes several recommendations to increase the quality of care. One of the highly sought areas is the usage of business intelligence and big data analytics (BI&A) techniques to collect, analyze, curate, and present evidence at the point of care, i.e., the practice of evidence based medicine (EBM). For the purpose of this study, we refer to EBM as the "conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patient" [5]. EBM means affirming the individual clinical expertise with the best available external clinical evidence. The best external clinical evidence can be drawn from the literature, and/or through the practice based knowledge. Each year, a significant number of research studies (potentially serving as evidence) are reported in the literature at an ever-increasing rate outpacing the translation of these research findings into practice. Coupled with the proliferation electronic health records, and consumer health information, researchers and practitioners are challenged to leverage the full potential of EBM.

The proliferation of potential evidence, patient data, and health consumer information represent opportunities for the innovative applications of analytics techniques to assist with the translation of data (in a variety of format depending on the source) to consumable knowledge. In this study, we refer to analytics as the "systematic usage of data and related business insights developed through the applied analytical disciplines (statistical, contextual, quantitative, predictive, and cognitive) to deliver the fact-based decision making for management, measurement, learning and planning. Analytics can be predictive, descriptive, or prescriptive" [6].

The objective of this study is to explore the opportunities for leveraging business intelligence and big data analytics for evidence-based medicine. This is accomplished by conducting a systematic review of peer-reviewed journal by searching PubMed, Web of Science, IEEE, ACM, ABI/Inform and EBSCOhost using the keywords—(Evidence Based Medicine) AND (Computer OR Information Technology OR Information System OR Big Data OR Analytics OR Data Mining) to identify current practices and applications of analytics in EBM. We also performed a Google scholar search to scan important articles related to big data analytics in EBM. We included 69 articles in final analysis. This research draws from that systematic review, and highlights existing gaps and opportunities for addressing these gaps by applying analytics techniques along the different steps involved in EBM. The analysis is designed to respond to the requirements and limitation of existing approaches in EBM.

The paper is organized as follows: the next section further describes business intelligence and big data analytics in the context of this study. Next, we illustrate the process of Evidence Based Medicine,

978-1-4799-2504-9/14 $31.00 © 2014 IEEE
DOI 10.1109/HICSS.2014.100

followed by a research agenda for leveraging business intelligence and big data analytics in EBM. The paper concludes with a set of examples of research projects emanating from the research agenda and a brief discussion of challenges and prospects of BI&A in EBM.

2. Evidence Based Medicine (EBM)

Evidence based medicine (EBM) is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients [5]. EBM is the intersection of individual clinical expertise, external evidence, and value to the patient. Evidence can be generated from both extant medical literature and practice-based evidence. Literature sources include randomized-controlled trials (RCT), systematic reviews, clinical guidelines, cohort studies, Quasi-Experimental studies, descriptive studies, and expert opinions [7]. On the other hand, practice-based evidences are generated from the day-to-day data collected in the hospital through treating the patients (electronic health record). Additional sources of practice-based evidence include claims data, insurance, and other administrative hospital data.

2.1. EBM and healthcare cost

According to Bloomberg, "Among advanced economies, the U.S. spends the most on health care on a relative cost basis with the worst outcome", ranking 46th among Bloomberg list of the most efficient health care countries," with an efficiency score of 30.8 [8]. This is far behind the comparable developed countries like Hong Kong (efficiency= 92.6) and Singapore (efficiency= 81.9) [8]. Moreover, the National Health Expenditure's (NHE) fact sheet projects the condition to persistently degrade in the future [9]. Over 2011-2021 NHE projected (a) healthcare cost is predicted to grow an average of 5.7% per year (b) health share of GDP to increased from 17.9 % to 19.6% (c) Medicare Spending to grow in average of 6.1% per year (d) Medicaid spending is estimated to grow an average of 8.2 % per year. Most importantly, in 2007, 62.1% of bankruptcies filers claimed high medical expenses; and a study done in 2013 showed that 25% of senior citizen who filed bankruptcy indicates the healthcare cost as primary reason [10]. In essence, healthcare cost is real problem in United States and if something is not done, healthcare is likely to hamper the overall country's economy and the quality of life of its citizens.

In that regard, there is a significant interest in EBM by various stakeholders (e.g. federal organizations, researchers and practitioners) as a mechanism to reduce cost and improve quality of care.. In principal, EBM will assist in identifying what works for individual patients and in identifying the most suitable intervention. Moreover, there are also evidence that the practice of EBM will results in cost saving. For example, Neubauer et al. showed that use of EBM for patients with non-small-cell lung cancer (NSCLC) resulted in cost saving of 35% over 12 months [11]. Kolodziej [12] further demonstrates various ways for how EBM can result in reduced healthcare cost including: 1) EBM results in a decrease in overall therapy. That is, physician can confidently recommend the most effective therapy as the first line of treatment, which eliminates many unfruitful or detrimental interventions; 2) EBM can reduce the cost by prescribing less expensive Drugs or/and Intervention. For example, if evidence points to two probable therapies with similar benefits and risks, but largely varies in cost; than, physician recommends less expensive therapies; 3) EBM can reduce the number of hospital visit by allowing physicians to minimize the adverse side effect of drugs or therapies. Other important cost saving through the EBM includes: cost saving by preventing or predicting the diseases condition, cost saving by efficient chronic care, and cost saving by reducing the adverse drug effects etc.

2.2. The EBM process

Generally, evidence based medicine consists of 7 steps (Table 1) and as described below [4, 13].

Step 1—identifying the Patient Condition: The first step in EBM is to adequately understand patient's clinical problem. An example of patient's condition may be—48 years old, 200 pounds, white male, with high blood pressure (BP), A1C=7% that is newly diagnosed with Type-2 diabetes. Other relevant information includes: allergy condition; kidney, foot and eye health; cardio-vascular diseases etc. The patient's condition can be understood by the past and present diagnosis, laboratory, and administrative data.

Step 2—Formulating EBM Question: After the patient's condition is identified, this step proceeds with synthesizing those condition into clinical questions. An example of a clinical question may be: Is nutritional therapy as affective as Oral Medication to the 48 years old T2 Diabetes patients with overweight, and high BP? Then, clinical questions are converted into EBM question. One of the ways to structure the EBM question is through the Population-Intervention-Comparison-Outcome (PICO) criteria. PICO is getting significant attention as a means to compare the economic, social and clinical outcome of different intervention alternatives [14, 15]. Population refers to the demographic and clinical information of patient; intervention refers to the possible course of action; while comparison refers to comparing between alternative interventions, or between "intervention" and "no intervention"; finally, outcome means output which we want to access (clinical, economic or social). An example of EBM question may be:

Population—Age: 48; Wt: 200pounds; BP: high; A1C:7%; Allergy: NSAIDs
Intervention—Nutritional Therapy
Comparison—Oral Medication
Outcome—Glucose Control

In general, the medical question on hand can refer to diagnosis, prognosis, treatment, iatrogenic harm, quality of care, or health economics [16].

Step 3—Evidence Gathering: This step focuses on extracting the information related to the EBM question. This step also includes preliminary evaluation of relevant information (individual article, or individual piece of information from electronic health record) before it is used for further analysis. For example, if an article does not satisfy the minimum criteria, then it should be discarded. The information is analyzed, synthesized, and translated into consumable module that solve the EBM question. The sources of evidence could be electronic health record (Practice-Based Evidence) and/or research studies (Literature-Based Evidence).

Step 4—Evidence Evaluation: The practice guidelines that are generated in previous step are further curated in this step. Medical field is information critical domain—correct information assist to make correct decision and save life; however, faulty information may invite death or complicates the patient's condition. As a result, clinicians are reluctant to trust a computer program that display evidence without clear information about reliability and provenance of the information [3, 17, 18]. Researcher and practitioners must adequately evaluate the evidence and maintain confidence in evidence. Evidence evaluation can be done in three steps: grade the individual articles, or individual piece of information; grade the overall evidence (after aggregating the individual piece of information); and calculate the statistical significance of overall information. The output of evidence evaluation is the grade (A-D or similar) and statistical significance (p-value) of evidence.

Step 5—Convert the Critically Appraised Evidence into Consumable Unit (practicable to use at the point of care): Here, the focus is on converting critically appraised evidence into consumable guidelines preferably in machine readable format, e.g., evidence coded into the UMLS, or evidence coded in the Asbru language [19].

Step 6—Evidence Presentation and Use: Regarding the evidence presentation, literature points four key factors [20, 21]: present the EBM at the point of care; present at least 2 alternatives course of actions; show reliability and citation along with the evidence, fit information in a single screen of a computer.

Step 7—Evaluate the result of putting evidence into practice: This step assesses if stakeholders are actually getting benefit by practicing EBM. Not only this, it also helps to know which evidence is working, and what are the improvements needed in the knowledge base of EBM. When technology is used to deliver EBM, this step becomes easier by automating the evaluation task.

TABLE 1: The EBM Process

Steps	Examples of EBM-related tasks
1- Identification of Patient Condition	• Identify overall patient's condition by analyzing past and present health records. This may include demographics, allergies, symptoms, diagnosis, tests, economic status, past condition, family history for hereditary and other related information.
2- Formulate the EBM Question	• Synthesize the clinical question into EBM question. One of the most common ways of generating the EBM question is through the PICO criteria (Population, Intervention, Comparison and Outcome).
3- Evidence Generation and Analysis	• Generate evidence by combining the information from the research studies (literature, RCT, Clinical Guidelines), and practice-based evidence (Hospital Information Systems/ Electronic Health record). Here, practice-based evidence refers to evidence generated from the day-to-day practice in clinical setting.
4- Evidence Evaluation	• Grade individual pieces of information (individual article, or individual information from practice based evidence). • Calculate the statistical significance of collective evidence.
5- Converting Evidence into Consumable Unit	• Convert critically appraised evidence into easily accessible & usable format.
6- Evidence Presentation and Use	• Present at least 2 alternatives course of actions at the point of care. • Show reliability and citation along with the evidence.
7- Evaluation of Implementing EBM into Practice.	• Track the success/failure of each piece of information. • Update the statistical significant/rating of evidence accordingly. • Calculate the long-term effect of using EBM.

3. Business Intelligence and Analytics (BI&A)

With the proliferation of data, business intelligence and big data analytics has emerged as an important study for both practitioners and researchers, regardless of domain. Chen et al. [22] defines BI&A as the "techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions". Examples of emerging issues in BI&A include Big Data Analytics, Text Analytics, Web Analytics, Network Analytics, and Mobile Analytics [22]. Some of the important applications of analytics in the literature in the context of evidence based medicine are (a) Automate the article selection procedure in systematic reviews [23, 24] (b) Automatically structure the abstract of article in PICO (population, intervention, comparison and outcome) criteria [25] (c) mine the free-text document of clinician notes [26] (d) mine the electronic health record and identify the patient's condition [27, 28] (e) convert the clinical guidelines into computer executable format [29-31] (f) grade the evidence and calculate the statistical significance[32, 33] (g) Mine the health records and identify the optimum practice guidelines [34] (h) and, identify the association rules in the electronic health records [35]. These application of data mining while valuable, still lags behind other domains, such as e-commerce. This may be attributed to the notion that healthcare researchers are relatively new to the field or has rarely taken advantage of scalable computational platform or analytical methods [36]. In this study, we demonstrate that the potential exist for researchers and practitioners to employ BI&A in healthcare, specifically the practice of EBM and as a mechanism to provide a new venue to reduce the cost and improve the quality of healthcare.

4. Business Intelligence and Big Data Analytics (BI&A) in EBM

This section illustrates how BI&A can be applied to support various Evidence Based Medicine processes described in Table 1. Each sub-section describes pertinent examples from the literature, identifies research gaps, and lists examples of research issues for leveraging BI&A in EBM as outlined in Table 2.

4.1. Identifying the patient's condition

Extant literature suggests that EBM should be fully integrated with electronic health record (EHR) systems [21, 37]. Currently, only, indexing and matching of keyword's of EHR with pertinent literature is done (for example, matching diseases condition with Meta-thesaurus of article) [28]. For example, Demner-Fushman et al. [27] developed a prototype system that explores the opportunity to automatically extract patient's problems from the team's notes and query evidence resources available in literature. They used the National Library of Medicine (NLM) MetaMap service to identify relevant Unified Medical Language System (UMLS) Metathesaurus concepts in the patient's electronic medical record (EMR); then, used those terms to query resources. The problem with their system is—identifying only the UMLS terms in the patient's EMR does not provide a complete picture of patient's condition, e.g., demographic information, duration of problems, etc. Indeed, if a researcher wants to realize adoptable EBM, more work should be done—EBM systems must automatically mine patient's past and present health record and assist in identifying the overall clinical condition. Examples include demographics, allergies, symptoms, diagnosis, tests, economic status, past condition and family history for hereditary). In essence, there is a limited capability for automatically extracting patients' information from EHR for identifying a patient's condition [27, 28].

In order to automatically identify patient's condition, a system should use a combination of analytics techniques. Applying data analytics will be facilitated if the electronic health record is in a standardized format, e.g., CDA, SNOMED, and LONIC. Otherwise, ad-hoc processes may be required. Data mining and Text analytics techniques such as information retrieval and information extraction can be applied to EHR. Ad-hoc techniques to mine the unstructured physician notes can also be employed.

4.2. Formulating the EBM question

Currently, clinicians have to manually translate patient's condition into an EBM question following the Population-Intervention-Comparison-Outcome (PICO) criteria. However, this is very time consuming and inefficient. For example, the average hospital visit in United States in primary care setting takes 20.8 minutes; with less than 10 minutes for more than 20% of the patients. [38]. During a visit, it is almost impossible for the physician to conduct the additional work needed to adequately formulate the clinical question, to generate the search query, and to execute those queries against an evidence-base. There must be a better way!

Fortunately, with recent developments in technology, patient's condition can be automatically synthesized into PICO criteria. One of the key techniques in synthesizing clinical question into PICO criteria is Topic Modeling—it uncovers the hidden thematic structure in document collection. Topic-modeling assist in creating new methods to browse, search and summarize large archives of text into structured format (based on topic) [39].

After the clinical questions are structured into the PICO criteria, it is possible to automatically create the search query; moreover, search query must be slightly customized based on the underlying technology (Vendor of EMR system, database system, analytics tools and other) used. The system also allows clinician to alter the clinical questions, or generate the clinical question manually.

4.3. Evidence Generation and Analysis

There is a need for being able to generate usable information from existing literature as well as the ever-increasing repository of information in electronic health records. For example, as of May 2013, over 145 thousand clinical trials are registered in "clinicaltrial.gov" alone; moreover, there is a huge number of clinical guidelines, cohort studies, expert opinions, and other evidence sources. On the other hand, as the healthcare is moving into electronic format, large amounts of data is being produced and stored in EHR systems on a daily basis. In that regard, IBM Watson [40] use proprietary DeepQA technology to generate evidence by retrospective analysis of existing data as well as data from perspective studies. IBM partnered with Wellpoint and Memorial Sloan Kettering Cancer Center (MSKCC) to create commercial applications for Watson in Healthcare. At MSKCC, Watson will be trained about the Oncology so that it can help with diagnostic and decisions in cancer. IBM Watson is incredible example of application of BI&A for EBM; However, it is still in the research phase, and many more research and business issues should be figure it out before it is delivered into the market. Another example from literature is study by Cohen et al [41]. They proposed text mining based pipelining framework that supported the creation and updating of evidence reports that provide assistance for the literature collection, collation, and triage steps of the systematic review process. Their system has four components: (a) meta-search engine; (b) classifier—classifies article based in RCT, cohort studies etc. (c) aggregator, which groups the articles publication's relatedness (d) the fourth component ranks the classified and grouped articles based on likelihood of inclusion in systematic review. The proposed approach automates some aspects of the systematic review process. Other examples include [18, 42-46]; nonetheless, these studies do not yet realize the full potential enabled by an ever-increasing evidence base..

Big data analytics enables all the data (medical literature, electronic health record, clinical notes, x-ray and other imaging data, insurance and claims data, and more) to be leveraged to produce translate data to relevant information for EBM support [47]. A meta-search Engine provides the mechanism to search all the available resources (e.g., PubMed and Cochrane) by translating user queries into the respective query languages of each search engine [48]. Also, Enterprise Search System provides the capabilities for searching and retrieving any number of unstructured and structured data sources within a organization by single query [49]. The data analysis can then proceed with the application of various analytics technologies — semantic web technology, text analytics, statistical machine learning and others. Semantic web technology is particularly important in EBM because it bridge the vocabularies gap between the different ontology, and help to address the variety aspects of medical literature, clinical notes, and electronic health records [50]. Text analytics technologies that are particularly important for EBM data analytics are [22, 47] (a) Concept Linking: finding related document based on shared concept (b) Entity Extraction: extracting the name, location, diagnosis, allergy, dates, medication, and other (c) Topic Tracking: tracking the interested information in raw data (d) Document Categorization: Categorizing documents based on pre-specified criteria (e) Clustering: Grouping documents into clusters (f) Question Answering: finding the answer specific questions. There are also various other business intelligence and big data analytics techniques that are useful for the EBM data analysis (Table 2).

Examples of on how these techniques can be helpful to analyze the EBM data include the use of document categorization to automatically classify RCT as relevant or not-relevant during systematic review [51]. Entity extraction can then be used to extract the criteria for the EBM question [52] (population, intervention, comparison, outcome) Another example is the use of entity extraction to extract important concepts in a RCT (demographic, intervention, outcomes, conclusion, funding source) and grade the trials based on the Agency for Healthcare Research and Quality (AHRQ) criteria using regression or statistical machine learning [52, 53]. Topic linking can be used to link all the patients in an EHR who have a particular combination of related conditions, e.g., gout and have allergy of NSAIDs.

4.4. Evidence Evaluation

Despite of strong guidelines and recommendation, there is no adequate mechanism for evaluating the evidence. For example, West at al. [7] proposed three criteria for rating evidence: (a) quality—aggregate rating of individual studies (b) quantity—number of studies, sample size, and effect magnitude (c) consistency—for any topic, the extent to which different studies reported similar findings. While the study demonstrates the viability of the criteria in evaluating relevant evidence, it is not scalable given the sheer volume of medical knowledge/evidence. There is a need to automate the application of these criteria to the the literature and practice based information..

BI&A techniques may be used to rate the individual articles as well as the overall evidence. The individual piece of information can be rated based on the study question, population, randomization, blinding, intervention, and more [33]. For example, a researcher can use text analytics and natural language processing to identify if the articles satisfy the criteria given by AHRQ and rate the individual articles accordingly [7].

4.5. Convert Critically Appraised Evidence into Consumable Unit

Recommendation and guidelines for computer-based EBM highly encouraged generating the final evidence in computer interpretable format [17, 21, 37]. Surprisingly, we found none of the EBM systems in literature satisfying this requirement. There are a few studies that discuss techniques for the conversion of clinical guidelines into computer interpretable format [29-31]. For an example, Essaihi et al. [31] evaluated if limited set of action types (such as Prescribe, Perform, Test etc.) could be used to model the clinical guidelines (represented in if...then rule) and used for the decision purpose. In another example, Sim et al [30] propose an ontology for the randomized control trial. These techniques help to understand the fundamental ideas; however, there is yet a need to be able to analyze incomplete, highly unstructured guidelines; and accurately convert into computer readable format. One approach is to use a combination of advanced analytics techniques (entity extraction, topic modeling, clustering, semantic web and statistical machine learning), and synthesize the practice guidelines into standard Medical Ontology. The exact identification of biomedical and clinical terms, and matching to standard ontology is very critical in this step [54]. In essence, term identification, consist of three steps [54, 55]: 1) recognize the text string as possible term; 2) classify the term (diseases, drug, body part, physiological function); and 3) map the term to a standardized medical ontology. For an example, one can use a combination of analytics techniques to mine the practice guidelines generated by computer, and those practice guidelines can be represented into UMLS format. According to NLM, "UMLS integrates and distributes key terminology, classification and coding standards, and associated resources to promote creation of more effective and interoperable biomedical information systems and services, including electronic health records" [56]. Moreover, future research can further explore adequacy of various techniques in accurately modeling incomplete practice guidelines into computer readable format.

4.6. Evidence Presentation

The final task of EBM systems, namely EBM presentation, is to ensure that the right information gets to the right person, at the right place, at the right time, and in the right format [18]. Any information delivered should have metadata related to provenance, reliability, and quality of the delivered information. It is also important that the information presented should be concise, and if possible, should fit in a single screen [20, 21]. However, majority of the EBM systems in the literature have not considered all these factors.

EBM data visualization techniques present the economic, clinical, and social outcome of different treatment alternatives in the form of text. Physicians do not have time to read extensive documentation. Better way must be found! Thinking out of the box, and moving further from the traditional healthcare data visualization; it is time for researchers and practitioners to adopt recent developments in areas such as visual analytics, e.g., SAS Visual Analytics [57]. Visual Analytics present big data in an interactive graph and chart in a way that is not overwhelming. For an example, in spite of presenting the economic, clinical, and diseases progression of different treatment alternatives in the form of text in three different screens; visual analytics can present all the information in single screen in the form of chart, or graph [57].

4.7. Evaluation of putting evidence into practice

Almost every literature suggests "evaluation of putting evidence into practice" is a crucial step of evidence-based medicine. However, existing literature have performed only limited task of evaluation. For example, Bigus et al. [18] stated that EBM for Cardiac condition (Cardiac Decision Support) optimize the physician time for getting relevant information. They have not done the micro-level evaluation. Other studies have also done similar kinds of evaluation. However, the potential is there for applying analytics to track the outcome (success or failure) of particular evidence, and update the confidence of the evidence accordingly.

System continuously tracks whether the use of EBM is providing any benefit to patients. The key idea—when clinician follows the evidence suggested by EBM system, and the outcome is as expected or better than expected; then, statistical significant or rating of evidence increases. Accordingly, if the outcome is worse than the expected, than statistical significant or rating of evidence decreases.

TABLE 2: Examples of specific Business Intelligence and Big Data Analytics research projects in EBM

EBM process	Examples of research projects	Relevant BI&A techniques
Identify the Patient's Condition	Explore and evaluate mechanisms to identify a patient's condition by analyzing electronic health records and pertinent literature.	Information Retrieval, Text Analytics
Convert EBM question into search query	Develop and evaluate approaches to translate the patient's condition into a an EBM question (PICO criteria).	Text Analytics; Topic Modeling
Gather, consolidates, rate, and analyze the raw information	Develop meta-search engines that are able to search and consolidate the results from pertinent literature, and enterprise search systems using EBM questions expressed as PICO criteria.	Meta-search Engine; Enterprise Search System; Information Extraction; Text Analytics; Topic Modeling; document representation, Statistical Machine Learning; Statistical NLP regression, classification, and association analysis
Evaluate and calculate the statistical significance	Explore and evaluate methods for assessing the quality and relevance of the results.	Entity extraction; Topic Tracking; Document Categorization; Clustering; Concept Linking; Question Answering; Topic Modeling; document representation, Statistical Machine Learning; Statistical NLP regression, classification, and association analysis
Convert evidence into machine readable format	Develop and evaluate approaches to support the conversion and curation of evidence into computer interpretable format. Use a combination of advanced analytics techniques (entity extraction, topic modeling, clustering, semantic web and statistical machine learning), and synthesize the evidence into standard Medical Ontology.	Asbru, UMLS, entity extraction, topic modeling, clustering, semantic web and statistical machine learning
Present the evidence in the appropriate format, time, and place	Explore, develop, and evaluate the use of visual analytics for the Evidence Presentation.	Visual Analytics
Evaluate the outcome of putting Evidence into Practice	Track the success or failure of applying particular evidence, then update statistical significant accordingly. Calculate the long term success—cost, physician and patient satisfaction, average hospital visit time of patient	Statistical machine learning, regression, data mining

5. Conclusion

The study explores opportunities for leveraging business intelligence and big data analytics in evidence based medicine. The study first describes the steps involved in evidence based medicine and then proceeds to identify current needs and discusses the potential for business intelligence and Big Data analytics in addressing these gaps. The study presents some of the emerging research areas relating to the use of Big Data Analytics for EBM. In so doing, this study provides a research agenda for health informatics researchers and data scientists to address issues of pressing needs, namely, reducing the cost and improving the cost of healthcare by broadening the practice of evidence based medicine through the applications of business intelligence big data analytics.

6. References

[1]http://data.worldbank.org/indicator/SH.XPD.TOTL.ZS?order=wbapi_data_value_2010+wbapi_data_value+wbapi_data_value-last&sort=desc, accessed March 08, 2013,

[2] Institute of Medicine, "Shaping the Future for Health", Institute of Medicine, 1999, pp. 1-8.

[3] Bates, D., Cohen, M., Leape , L., Overhage, J., Shabot, M., and Sheridan, T., "Reducing the Frequency of Errors in Medicine Using Information Technology", J Am Med Inform Assoc, 2001, pp. 299–308.

[4] Wells, L., "Role of Information Technology in Evidence Based Medicine : Advantages and Limitations", 2007, pp. 1-10.

[5] Sackett, D.L., William Rosenberg, J.A., Gray, R., Brian Haynes, and Richardson, W.S., "Evidence Based Medicine: What It Is and What It Isn't", Bmj, 1996, pp. 71-72.

[6] Ibm, "The Value of Analytics in Healthcare: From Insights to Outcomes", in (Editor, 'ed.'^'eds.'): Book The Value of Analytics in Healthcare: From Insights to Outcomes, 2012

[7] West, S., King, V., Carey, T.S., Lohr, K.N., Mckoy, N., Sutton, S.F., and Lux, L., "Systems to Rate the Strength of Scientific Evidence", Agency for Healthcare Research and Quality, 2002,

[8]http://www.bloomberg.com/visual-data/best-and-worst/most-efficient-health-care-countries, accessed Sep 3, 2013,

[9]http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet.html, accessed Aug 30, 2013,

[10] Kelley, A.S., Mcgarry, K., Fahle, S., Marshall, S.M., Du, Q., and Skinner, J.S., "Out-of-Pocket Spending in the Last Five Years of Life", J Gen Intern Med, 28(2), 2013, pp. 304-309.

[11] Neubauer, M.A., Hoverman, J.R., Kolodziej, M., Reisman, L., Gruschkus, S.K., Hoang, S., Alva, A.A., Mcarthur, M., Forsyth, M., Rothermel, T., and Beveridge, R.A., "Cost Effectiveness of Evidence-Based Treatment Guidelines for the Treatment of Non-Small-Cell Lung Cancer in the Community Setting", J Oncol Pract, 6(1), 2010, pp. 12-18.

[12]http://www.cancernetwork.com/practice/content/article/10165/1821731, accessed Aug 30, 2013,

[13] Hurwitz, S.R., Tornetta, P., and Wright, J.G., "How to Read the Literature to Change Your Practice: An Evidence-Based Medicine Approach", The Journal of Bone & Joint Surgery, 88A(8), 2006, pp. 1873-1879.

[14] Ashrafian, H., Sevdalis, N., and Athanasiou, T., "Evidence-Based Surgery: Creating the Culture", Surg Clin North Am, 86(2006, pp. 91-100, ix.

[15] Bauer, J., Spackman, S., Chiappelli, F., Prolo, P., and Stevenson, R., "Making Clinical Decisions Using a Clinical Practice Guideline", J Calif Dent Assoc, 34(2006, pp. 519-528.

[16] Rosenberg, W., and Donald, A., "Evidence Based Medicine: An Approach to Clinical Problem-Solving", Bmj, 310(1995, pp. 1122-1126.

[17] Bates, D.W., "Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-Based Medicine a Reality", Journal of the American Medical Informatics Association Volume, 10(6), 2003, pp. 523-530.

[18] Bigus Et Al., "Information Technology for Healthcare Transformation", Ibm Journal of Research and Development, 55(2011,

[19] Miksch, S., and Seyfang, A., "Continual Planning with Time-Oriented, Skeletal Plans, In: W. Horn (Ed.), Proceedings of Ecai, Ios Press, Amsterdam", 2000,

[20] Bates, D.W., Kuperman, G.J., Wang, S., Gandhi, T., Kittler, A., Volk, L., Spurr, C., Khorasani, R., Tanasijevic, M., Middleton, B., Horasani, R.a.K., Anasijevic, M.I.T., Iddleton, B.L.M., and Ms, C., "Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-Based Medicine a Reality", J Am Med Inform Assoc, 10(2003, pp. 523-530.

[21] Lester, W.T., Zai, A.H., Grant, R.W., and Chueh, H.C., "Designing Healthcare Information Technology to Catalyse Change in Clinical Care", Inform Prim Care, 16(2008, pp. 9-19.

[22] Chen, H., Chiang, R.H.L., and Storey, V.C., "Business Intelligence and Anlaytics: From Big Data to Big Impact", MIS Quarterly, 36(4), 2012, pp. 1165-1188.

[23] Abernethy, A.P., Etheredge, L.M., Ganz, P.A., Wallace, P., German, R.R., Neti, C., Bach, P.B., and Murphy, S.B., "Rapid-Learning System for Cancer Care", J Clin Oncol, 28(2010, pp. 4268-4274.

[24] Frunza, O., Inkpen, D., and Matwin, S., "Building Systematic Reviews Using Automatic Text Classification Techniques", 2010, pp. 303-311.

[25] Kim, S.N., Martinez, D., Cavedon, L., and Yencken, L., "Automatic Classification of Sentences to Support Evidence Based Medicine", BMC Bioinformatics, 12 Suppl 2(2011, pp. S5.

[26] Chu, S., and Cesnik, B., "Knowledge Representation and Retrieval Using Conceptual Graphs and Free Text Document Self-Organisation Techniques", International Journal of Medical Informatics, 62(2001, pp. 121-133.

[27] Demner-Fushman, D., Seckman, C., Fisher, C., Hauser, S.E., Clayton, J., and Thoma, G.R., "A Prototype System to Support Evidence-Based Practice", AMIA Annu Symp Proc, 2008, pp. 151-155.

[28] O'sullivan, D.M., Wilk, S.A., Michalowski, W.J., and Farion, K.J., "Automatic Indexing and Retrieval of Encounter-Specific Evidence for Point-of-Care Support", Journal of Biomedical Informatics, 43(2010, pp. 623-631.

[29] Marshall, C., Rosenbrand, K., and Ollenschlaeger, G., "Transforming Written Guidelines into Electronic Formats--International Perspectives", Stud Health Technol Inform, 101(2004, pp. 123-126.

[30] Sim, I., Olasov, B., and Carini, S., "An Ontology of Randomized Controlled Trials for Evidence-Based Practice: Content Specification and Evaluation Using the Competency Decomposition Method", Journal of Biomedical Informatics, 37(2004, pp. 108-119.
[31] Essaihi, A., Michel, G., and Shiffman, R.N., "Comprehensive Categorization of Guideline Recommendations: Creating an Action Palette for Implementers", AMIA Annu Symp Proc, 2003, pp. 220-224.
[32] Lehmann, H.P., and Goodman, S.N., "Bayesian Communication: A Clinically Significant Paradigm for Electronic Publication", J Am Med Inform Assoc, 7(2000, pp. 254-266.
[33] Lohr, K.N., "Rating the Strength of Scientific Evidence: Relevance for Quality Improvement Programs", Int J Qual Health Care, 16(2004, pp. 9-18.
[34] Pineau, J., Bellemare, M.G., Rush, A.J., Ghizaru, A., and Murphy, S.A., "Constructing Evidence-Based Treatment Strategies Using Methods from Computer Science", Drug Alcohol Depend, 88 Suppl 2(2007, pp. S52-60.
[35] Masuda, G., "A Framework for Dynamic Evidence Based Medicine Using Data Mining", 2002, pp. 0-5.
[36]http://www.bcr.org/content/big-data-analytics-biomedical-research, accessed Jun 10, 2013, 2012.
[37] Sim, I., Gorman, P., Greenes, R.A., Haynes, R.B., Kaplan, B., Lehmann, H., and Tang, P.C., "Clinical Decision Support Systems for the Practice of Evidence-Based Medicine", J Am Med Inform Assoc, 8(6), 2001, pp. 527-534.
[38] Chen, L.M., Farwell, W.R., and Jha, A.K., "Primary Care Visit Duration and Quality. Does Good Care Take Longer?", Arch Intern Med, 169(20), 2009, pp. 1866-1872.
[39] http://www.cs.princeton.edu/~blei/topicmodeling.html, accessed June 13, 2013, June 13, 2013.
[40] http://lifesciences.ieee.org/publications/newsletter/may-2012/108-ibm-watson-data-analytics-a-big-data-analytics-approach-for-a-learning-healthcare-system, accessed August 25, 2013,
[41] Cohen, A.M., Adams, C.E., Davis, J.M., Yu, C., Yu, P.S., Meng, W., Duggan, L., Mcdonagh, M., and Smalheiser, N.R., "The Essential Role of Systematic Reviews , and the Need for Automated Text Mining Tools", 2010, pp. 376-380.
[42] Best, A., Hiatt, R.A., and Norman, C.D., "Knowledge Integration: Conceptualizing Communications in Cancer Control Systems", Patient Education and Counseling, 71(2008, pp. 319-327.
[43] Latoszek-Berendsen, A., Tange, H., Van Den Herik, H.J., and Hasman, A., "From Clinical Practice Guidelines to Computer-Interpretable Guidelines. A Literature Overview", Methods Inf Med, 49(2010, pp. 550-570.
[44] Lenz, R., Blaser, R., Beyer, M., Heger, O., Biber, C., Baumlein, M., and Schnabel, M., "It Support for Clinical Pathways--Lessons Learned", International Journal of Medical Informatics, 76 Suppl 3(2007, pp. S397-402.
[45] Riano, D., Real, F., Lopez-Vallverdu, J.A., Campana, F., Ercolani, S., Mecocci, P., Annicchiarico, R., and Caltagirone, C., "An Ontology-Based Personalization of Health-Care Knowledge to Support Clinical Decisions for Chronically Ill Patients", Journal of Biomedical Informatics, 45(2012, pp. 429-446.
[46] Stolba, N., Nguyen, T.M., and Tjoa, A.M., "Towards Sustainable Decision-Support System Facilitating Ebm", Conf Proc IEEE Eng Med Biol Soc, 2007(2007, pp. 4355-4358.
[47] http://www.slideshare.net/perficientinc/using-big-data-for-improved-healthcare-operations-and-analytics, accessed June 15, 2013,
[48] Coiera, E., Walther, M., Nguyen, K., and Lovell, N.H., "Architecture for Knowledge-Based and Federated Search of Online Clinical Evidence", Journal of Medical Internet Research, 7(2005, pp. e52.
[49]http://publib.boulder.ibm.com/infocenter/db2luw/v8/index.jsp?topic=/com.ibm.db2.ii.of.doc/admin/iiysaovview.htm, accessed June 13, 2013,
[50] http://www.w3.org/2001/sw/wiki/Bigdata, accessed June 14, 2013,
[51] Aphinyanaphongs, Y., and Aliferis, C.F., "Text Categorization Models for Retrieval of High Quality Articles in Internal Medicine", AMIA Annu Symp Proc, 2003, pp. 31-35.
[52] Takeuchi, K., and Collier, N., "Bio-Medical Entity Extraction Using Support Vector Machines", BioMed '03 Proceedings of the ACL 2003 workshop on Natural language processing in biomedicine, 13(2003, pp. 57-64.
[53] West, S., King, V., Ts, C., and Et Al, "Systems to Rate the Strength of Scientiwc Evidence. Evidence Report, Technology Assessment", AHRQ Publication No. 02-E016. Rockville, MD: Agency for Healthcare Research and Quality, 47(2002,
[54] Krauthammer, M., and Nenadic, G., "Term Identication in the Biomedical Literature", J Biomed Inf, 37(2004, pp. 512-526.
[55] Gooch, P., and Roudsari, A., "A Tool for Enhancing Metamap Performance When Annotating Clinical Guideline Documents with Umls Concepts. ", Paper presented at the IDAMAP Workshop at 13th Conference on Artificial Intelligence in Medicine (AIME'11) Bled, Solvenia, 2011,
[56] http://www.nlm.nih.gov/research/umls/, accessed June 13, 2013,
[57] Sas, "Data Visualization Techniques: From Basics to Big Data with Sas® Visual Analytics", 2013,