**IT601- Research Topics**

**Predictive Analytics in healthcare: Promise and potential**

**Introduction**

The healthcare industry historically has generated large amounts of data, driven by record keeping, compliance & regulatory requirements, and patient care (Raghupathi, 2010). While most data are stored in hard copy form, the current trend is toward rapid digitization of these large amounts of data. Driven by mandatory requirements and the potential to improve the quality of healthcare delivery meanwhile reducing the costs, these massive quantities of data (known as ‘big data’) hold the promise of supporting a wide range of medical and healthcare functions, including among others clinical decision support, disease surveillance, and population health management (Burghard, 2012). Predictive analytics is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data. The goal is to go beyond knowing what has happened to providing a best assessment of what will happen in the future (SAS, 2017). It makes it possible to harness the power of big data to improve the health of patients and lower the cost of health care (Cohen, Amarasingham, Shah, Xie, & Lo, 2014). Predictive analytics encompasses a variety of statistical techniques from predictive modeling, machine learning, and data mining that analyze current and historical facts to make predictions about future, or otherwise unknown events (Encyclopedia WTF). In medicine, the convergence of meaningful use of electronic medical records, ICD-10 diagnostic coding, data warehouses, and integrated healthcare systems are bringing such predictive analytics to the bedside and clinics in order to improve the health of the nation. The U.S. is investing a significant amount of resources into the informational technology infrastructure with the intent of harnessing such big data to help predict, diagnose, and treat medical conditions and thereby improve population health (Ustun, Westover, Rudin, & Bianchi, 2016).

In healthcare, the term big data typically refers to large quantities of electronic health record, administrative claims, and clinical trial data as well as data collected from smartphone applications, wearable devices, social media, and personal genomics services (Hernandez & Zhang, 2017). Predictive analytics refers to innovative methods of analysis developed to overcome challenges associated with big data, including a variety of statistical techniques ranging from predictive modeling to machine learning to data mining (Hernandez et al., 2017). Predictive analytics using big data have been applied successfully in several areas of medication management, such as in the identification of complex patients or those at highest risk for medication noncompliance or adverse effects (Hernandez et al., 2017). Other industries have successfully used predictive analytics to tailor service delivery in real time. Familiar examples include Amazon’s product recommendation system for online shopping based on an individual’s prior purchases, and American Airlines’ ticket pricing system based on prior customer purchasing trends. Sports teams like the Oakland Athletics have relied heavily on analytics to select player rosters, outperforming expectations despite having a much smaller payroll than other teams. These organizations use large amounts of data and sophisticated machine learning algorithms to meet consumer and organizational needs. In health care, predictive analytics offers an automated means to forecast future health outcomes for individuals or populations based on algorithms derived from historical patient data. Some smartphone apps have successfully applied predictive analytics to influence health care: Ginger.io, for example, uses analytics based on cell phone data to identify patients at risk for depression crises, cueing physicians and caregivers to intervene (Madan, Bebrian, Lazer, & Pentland, 2010). As more electronic health data become available, some health systems have begun to develop predictive models around clinical issues, such as acute intensive care unit decompensation and hospital readmissions (Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014). In 2009, the Secretary-General of the United Nations (UN), Ban Ki-moon started the UN Global Pulse (UNGP*)* initiative, with the explicit goal of harnessing big data technology for human development (Pulse UG, 2012). The Global Pulse program is aimed at forming a network of innovation centers, called the *Pulse Labs*, all over the world. Ideally, these Pulse labs will bring together people from different fields of life together to make use of the free and open source computing methods/software toolkits to analyze data to help the development and humanitarian operations especially in the developing countries. In (Kirkpatrick, 2013), Kirkpatrick, the director of the UN Global Pulse innovation initiative, presents the case for deploying big data techniques and analytics in the field of human development. It is highlighted that data—especially from mobile phone and social media—can be utilized in fighting hunger, disaster and poverty. This report talks about “data philanthropy” where the companies, whose businesses revolve around data, can collaborate with the UN in predicting imminent humanitarian crises and help take possible steps to avoid situations that can lead to disasters. In the developing countries, the farmers are often less informed about the soil conditions, extreme changes in the weather patterns, plantation, topography and access to markets (CFA, 2015; Kshetri, 2014). Data collected from different sensors, satellite imagery and field experts can be analyzed and predictive models can be formed. Based on these models the most relevant information can then be sent using cellular network to individual farmers.

**Predictive Analytics in Healthcare**

Big data analytics in healthcare is bringing a huge cultural change in the way conventional medical diagnosis and treatment operates. Big data can revolutionize medical diagnosis by integrating data gathered from various medical records of a patient, as well as real time wearable sensors, to analyze and diagnose the patient’s current health status and provide an early warning sign if the health of a patient is on a dangerous track. Doing this helps in taking preventing measurements to diagnose and treat a potentially harmful disease during early stages. In terms of making treatment more efficient and convenient, it is possible for a person having a smart phone to access medical service providers via a healthcare app (Press TA, 2015) to obtain quick and more personalized response from the convenience of one’s home. The use of predictive analytics has accelerated in numerous industries in the past decade, with the emergence of real-time electronic data sets so large and complex that traditional data-processing tools have proved inadequate. With the advent of the EHR, it has become possible to apply predictive analytics to health care. The use of predictive analytics in health care leverages decades of work in statistics, computer science, and clinical decision support. In this emerging era of big data, predictive analytics models can use a variety of current or historical information such as claims, clinical, social, and genomic data to make predictions about the future. The early use of predictive analytics models in medicine has focused on identifying patients at high or low risk for serious complications or adverse clinical events, preventing those adverse events, and optimally allocating scarce clinical resources. The most common example is identifying patients at high risk of hospital readmission (Cohen et al., 2014). Kaiser Permanente of Northern California (KPNC), an integrated health service organization, has used predictive analytics to reduce antibiotic overuse in neonates. KPNC used maternal health data from more than 600000 livebirths to determine the probability of early-onset neonatal sepsis in non-premature infants prior to birth. These data were integrated with objective clinical data from the new born at birth to assess the probability of sepsis by categorizing newborns as at low, medium, or high risk of sepsis. KPNC obstetricians and neonatologists then used this score to determine whether to administer antibiotics (Escobar et al., 2014). After implementation of this algorithm, use of systemic antibiotics in the neonatal period among newborns of 34 weeks or more gestation was estimated to decrease by 33% to 60%, and up to an estimated 250000 newborns nationally could potentially be spared antibiotics at birth annually (Escobar et al., 2014). Post discharge Care Hospital readmissions represent an important driver of spending, with all-cause 30-day readmissions costing the US health system more than $41 billion annually, and thus are a major quality indicator for health systems (Hines, Barret, Jiang, & Steiner, 2014). Parkland Health and Hospital System used an algorithm based on 29 clinical, social, behavioral, and utilization factors available within 24 hours of admission to predict risk of readmission for patients with heart failure (Amarasingham et al., 2013). In a prospective study, 228 patients with heart failure deemed at high risk of 30-day readmission received targeted evidence-based interventions including (1) detailed patient education by a multidisciplinary team including a pharmacist, nutritionist, and case manager; (2) follow-up telephone calls within 48 hours to ensure medication adherence; (3) outpatient heart failure specialist appointments within 7 days; and (4) a primary care appointment scheduled according to the urgency of non-cardiac issues. Compared with 834 patients enrolled in the study prior to intervention, there was a 26%relative reduction in risk-adjusted odds of readmission among 913 patients with heart failure enrolled in the post intervention period (26% vs 21% 30-day readmission rates) (Amarasingham et al., 2013).

**Concerns**

Some health professionals have raised concerns about the application of predictive analytics, not the least of which is the perceived diminution of the role of the physician in managing clinical uncertainty (Sniderman, D’Agostino, & Pencina, 2015). Other concerns include protection of patient privacy, diminishment of patient preferences, and inadequate medical training ([Amarasingham,](https://www.ncbi.nlm.nih.gov/pubmed/?term=Amarasingham%20R%5BAuthor%5D&cauthor=true&cauthor_uid=25006140) [Patzer,](https://www.ncbi.nlm.nih.gov/pubmed/?term=Patzer%20RE%5BAuthor%5D&cauthor=true&cauthor_uid=25006140) [Huesch,](https://www.ncbi.nlm.nih.gov/pubmed/?term=Huesch%20M%5BAuthor%5D&cauthor=true&cauthor_uid=25006140) [Nguyen,](https://www.ncbi.nlm.nih.gov/pubmed/?term=Nguyen%20NQ%5BAuthor%5D&cauthor=true&cauthor_uid=25006140) & [Xie,](https://www.ncbi.nlm.nih.gov/pubmed/?term=Xie%20B%5BAuthor%5D&cauthor=true&cauthor_uid=25006140) 2014). Health professionals had similar hesitations more than a decade ago when considering implementing EHRs. However, algorithms routinely outperform practitioners’ clinical intuition without decision support. Algorithms also may enhance the quality of interaction between physicians and patients— for example, machine learning algorithms based on retrospective data can provide survival projections that may help inform discussions regarding end-of-life care for patients with advanced cancer. However, physicians will still need to exercise clinical judgment, and with appropriate training can combine new insights learned from predictive analytics alongside patient preferences to make higher-value treatment decisions.

**Reference**

1. Raghupathi, W. (2010). Data Mining in Health Care. Healthcare Informatics: Improving Efficiency and Productivity. *Taylor & Francis*, 211-223.
2. Burghard, C. (2012). Big Data and Analytics Key to Accountable Care Success. *IDC Health Insights.*
3. Dembosky, A. (2012, December 12). Data Prescription for Better Healthcare. *Financial Times.* Retrieved from <http://www.ft.com/intl/cms/s/2/55cbca5a-4333-11e2-aa8f-00144feabdc0.html# axzz2W9cuwajK>
4. Feldman, B., Martin, E. M., & Skotnes, T. (2012). Big Data in Healthcare Hype and Hope. *Dr. Bonnie 3600.*
5. Fernandes, L., O’Connor, M., & Weaver, V. (2012). Big data, bigger outcomes. *J AHIMA,* 38-42.
6. Encyclopedia WTF: Predictive Analytics
7. Ustun, B., Westover, M. B., Rudin, C., & Bianchi, M. T. (2016). Clinical prediction models for sleep apnea: The importance of medical history over symptoms.*Journal of Clinical Sleep Medicine: JCSM : Official Publication of the American Academy of Sleep Medicine, 12*(2), 161.
8. Hernandez, I., & Zhang, Y. (2017). Using predictive analytics and big data to optimize pharmaceutical outcomes. *Am J Health System Pharm.*, 74(18), 1494-1500.
9. Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., & Lo, B. (2014). The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care. [*Health Aff (Millwood)*](https://www.ncbi.nlm.nih.gov/pubmed/25006139), 33(7), 1139-1147.
10. Madan, A., Bebrian, M., Lazer, D., & Pentland, A. (2010). Social sensing to model epidemiological behavior change. *Proceedings of UBICOMP, 12th ACM* *Conference on Ubiquitous Computing*, 291-300.
11. Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care. *Health Aff* *(Millwood)*, 33(7), 1123-1131.
12. Pulse U. G. (2012). Big data for development: Challenges & opportunities. *Naciones Unidas, Nueva York, mayo*.
13. Kirkpatrick, R. (2013). Big data for development. Big Data. 1(1), 3–4.
14. Connected Farmer Alliance (2015). Retrieved from <http://www.technoserve.org/our-work/projects/connected-farmer-alliance>.
15. Press TA (2015). ResearchKit: 5 Things to Know About Apple’s Medical Apps. 2015. Retrieved from http://www.nytimes.com/aponline/2015/03/11/technology/ap-us-tec-apple-researchkit.html.
16. Kshetri, N. (2014). The emerging role of big data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society*, 1(2).
17. Escobar, G. J., Puopolo, K. M., Wi, S., Turk, B. J., Kuzniewicz, M. W., Walsh, E. M., . . . Draper, D. (2014). Stratification of Risk of Early-Onset Sepsis in Newborns ≥34 Weeks’ Gestation. *Pediatrics*, 133(1), 30-36.
18. Hines, A. L., Barrett, M. L., Jiang, J., & Steiner, C. A. (2011). Conditions with the Largest Number of Adult Hospital Admissions by Payer. Retrieved from <https://www.hcupus.ahrq>. gov/reports/statbriefs/sb172-Conditions-Readmissions-Payer.jsp.
19. [Amarasingham](https://www.ncbi.nlm.nih.gov/pubmed/?term=Amarasingham%20R%5BAuthor%5D&cauthor=true&cauthor_uid=23904506), R.,  [Patel](https://www.ncbi.nlm.nih.gov/pubmed/?term=Patel%20PC%5BAuthor%5D&cauthor=true&cauthor_uid=23904506), C. P., [Toto](https://www.ncbi.nlm.nih.gov/pubmed/?term=Toto%20K%5BAuthor%5D&cauthor=true&cauthor_uid=23904506), K., [Nelson](https://www.ncbi.nlm.nih.gov/pubmed/?term=Nelson%20LL%5BAuthor%5D&cauthor=true&cauthor_uid=23904506), L. L., [Swanson](https://www.ncbi.nlm.nih.gov/pubmed/?term=Swanson%20TS%5BAuthor%5D&cauthor=true&cauthor_uid=23904506), S. T., [Moore](https://www.ncbi.nlm.nih.gov/pubmed/?term=Moore%20BJ%5BAuthor%5D&cauthor=true&cauthor_uid=23904506), J. B., . . .  [Halm](https://www.ncbi.nlm.nih.gov/pubmed/?term=Halm%20EA%5BAuthor%5D&cauthor=true&cauthor_uid=23904506) A. E. (2013). Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. *BMJ Qual Saf*, 22(12), 998-1005.
20. Sniderman, A. D., D’Agostino, R. B. Sr., & Pencina, M. J. (2015). The role of physicians in the era of predictive analytics. *JAMA*, 314(1), 25-26.
21. [Amarasingham, R](https://www.ncbi.nlm.nih.gov/pubmed/?term=Amarasingham%20R%5BAuthor%5D&cauthor=true&cauthor_uid=25006140)., [Patzer, R. E](https://www.ncbi.nlm.nih.gov/pubmed/?term=Patzer%20RE%5BAuthor%5D&cauthor=true&cauthor_uid=25006140)., [Huesch, M](https://www.ncbi.nlm.nih.gov/pubmed/?term=Huesch%20M%5BAuthor%5D&cauthor=true&cauthor_uid=25006140)., [Nguyen, N. Q](https://www.ncbi.nlm.nih.gov/pubmed/?term=Nguyen%20NQ%5BAuthor%5D&cauthor=true&cauthor_uid=25006140)., & [Xie, B](https://www.ncbi.nlm.nih.gov/pubmed/?term=Xie%20B%5BAuthor%5D&cauthor=true&cauthor_uid=25006140). (2014). Implementing electronic health care predictive analytics. *Health Aff (Millwood)*, 33(7), 1148-1154.