Published in final edited form as:

J Cardiovasc Nurs. 2017 ; 32(1): 4–6. doi:10.1097/JCN.0000000000000384.

Big Data: Contributions, Limitations, and Implications for Cardiovascular Nurses

Kelly T. Gleason, RN, BSN [PhD student] and Cheryl R. Dennison Himmelfarb, RN, ANP, PhD, FAHA, FPCNA, FAAN [Associate Professor]

Department of Acute and Chronic Care, Johns Hopkins University School of Nursing, 525 N. Wolfe Street / Rm. 527, Baltimore, MD 21205-2110

Big Data is a catch phrase applied to large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information.^{1,2} Applied to healthcare, Big Data can be defined as combining and analyzing large amounts of data to identify associations and make predictions that can inform improvements in quality of healthcare delivery and patient outcomes.^{1,3} Healthcare data sources include electronic health records, machine generated information garnered from devices such as cardiac monitors and ActiGraphs, social media including Facebook status updates and Twitter Posts, and genome data. The availability of healthcare data is growing exponentially as electronic health records, the use of wearable devices, social media and internet use, and genomic information continue to expand.³ Examining data from real-world clinical care to evaluate therapies is a powerful alternative to costly and often impractical randomized controlled trials. While Big Data analytics have the potential to significantly contribute to cardiovascular nursing science and practice, it is important to understand the limitations and how nurses can contribute to ensuring Big Data findings are effectively employed to improve patient care.

The Potential of Big Data to Build Cardiovascular Nursing Science

Big Data analytical applications have great potential to improve quality of cardiovascular care.^{3,4} For example, the Big Data technique of data mining allowed the development of an in-hospital outcome prediction model for acute coronary syndrome patients undergoing percutaneous coronary intervention.⁵ This comparison of claims data for over 14,000 individuals determined that individuals who received advanced biomarker and disease state management services for cardiovascular and cardiometabolic conditions were less likely to experience a myocardial infarction with no significant increase in cost of care.⁶ Registry studies that integrate electronic health record data, medical claims, and survey data that includes patient-reported outcome measures have the potential to further our understanding of key outcomes from real-world clinical care. For example the Atrial Fibrillation: Focus on Effective Clinical Treatment Strategies (AFFECTS) registry provided important insight into community treatment patterns of atrial fibrillation that would not have been possible in a

P: 443-287-4174, F: 410-955-7463, cdennis4@jhu.edu.

Disclosure: Kelly T. Gleason received support from predoctoral fellowship in Interdisciplinary Training in Cardiovascular Health Research, T32 NR012704, and Predoctoral Clinical Research Training Program, TL1 TR001078.

controlled clinical trial.⁷ A data analysis from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF) found that atrial fibrillation symptoms were associated with a higher risk of hospitalization.⁸ Big Data opens doors to research involving historically understudied populations, including racial and ethnic minorities. An examination of medical claims from 460,417 individuals with atrial fibrillation determined that adding ethnicity to the CHA2DS2-VASc (congestive heart failure, hypertension, age 75, diabetes, previous stroke, vascular disease, age 65 to 74, and female sex) score significantly improved stroke prediction.⁹ Through big data analytics, nurse scientists have access to a vast amount of patient-related health medical data to answer clinical care questions.¹

The Limitations of Applying Findings from Big Data

The promise of Big Data is accompanied by complexity and multiple limitations that are important to understand when interpreting findings from studies that employed Big Data. Any cardiovascular nurse who has worked on a monitored inpatient unit can easily imagine the quality issues of data garnered from sensory devices and electronic health records. The false alarms from cardiac monitors, differences in documentation standards not just across hospitals but across units of the same hospital, an adverse event that was inadequately charted as the day became increasingly hectic, and variance in terms used in free text documentation are all potential contributors to low quality data. While ICD-10 codes are standardized, they do not tell the full story of the patient's diagnosis and why the patient received a particular treatment. Internet and mobile app data allows access to a broader population; however, both the quality and the high variability in data types present substantial challenges. One example is using activity data generated from a smartphone's health app: while some individuals carry their smartphone everywhere, others leave the phone behind for workouts and walks; thus, any conclusions drawn from that data are likely incomplete.

In registry trials, treatments are not provided at a set time point; each individual in the study has their own unique time of diagnosis and treatment trajectory. Furthermore, large registries that examine treatment outcomes in a diverse patient group must also take into account multiple factors including the type of hospital (i.e., size, teaching status) patients received care. Further, patients receiving care within a hospital receive care from different providers on different units with different protocols. Patient-reported data, which can measure symptoms, functional status, and quality of life are key to informing patient-centered research. However, patient-reported outcomes are not routinely captured, and when patient-reported outcomes are measured it often is not in parallel with the procedures and medications of interest. Another issue is that Big Data allows us to leverage large sample sizes to generate statistically significant findings that may be clinically irrelevant. The National Institutes of Health Big Data to Knowledge initiative acknowledges insufficient training in the development and use of methods and tools needed to analyze Big Data as a major impediment to maximizing the value of the growing volume and complexity of data available.¹⁰ The data sources that contribute to healthcare Big Data are undeniably complicated, and drawing conclusions about clinical care from Big Data findings without taking into account all potential confounders is unwise.

How Cardiovascular Nurse Scientists Can Contribute to the Use of Big Data

For electronic health data to be effectively used to answer questions important to nursing care, nurses must continue playing a major role in building and improving electronic health records and mobile health technology. The transition to electronic health records allows for an easier analysis of nurse documentation than was previously possible; thus enabling an opportunity to efficiently build evidence from nursing practice. However, nurses must also be a part of implementing documentation standards. One demonstration of this happening is a group who worked across six health systems to align a minimum set of physiologic nursing assessment data with national standardized coding to support the exchange and aggregation of comparable nursing data.¹¹ Considerably more work of this type is needed to facilitate a systematic approach to comparing nursing data. Mobile health technologies allow for clinicians to access data documented by patients for real-time use in clinical care. Cardiovascular nurses have important insight into how mobile health technologies can be used to improve patient care, and will be key to providing patient education about the use of mobile health technologies. Nurses are uniquely poised to include patient-reported outcomes in Big Data and leverage that data to develop an evidence base that better supports patients and providers to improve management for individuals with cardiovascular diseases.

References

1. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst.* 2014; 2:3. [PubMed: 25825667]
2. McAfee A, Brynjolfsson E. Big data: the management revolution. *Harv Bus Rev.* 2012; 90(10):60–66. 68, 128. [PubMed: 23074865]
3. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol.* 2016; 13(6):350–359. [PubMed: 27009423]
4. Bakken S, Reame N. The Promise and Potential Perils of Big Data for Advancing Symptom Management Research in Populations at Risk for Health Disparities. *Annu Rev Nurs Res.* 2016; 34:247–260. [PubMed: 26673385]
5. Sladojevic M, Cankovic M, Cemerlic S, Mihajlovic B, Adic F, Jarakovic M. Data Mining Approach for in-Hospital Treatment Outcome in Patients with Acute Coronary Syndrome. *Med Pregl.* 2015; 68(5–6):157–161. [PubMed: 26234022]
6. Thompson S, Varvel S, Sasinowski M, Burke JP. From Value Assessment to Value Cocreation: Informing Clinical Decision-Making with Medical Claims Data. *Big Data.* 2016; 4(3):141–147. [PubMed: 27642718]
7. Reiffel JA, Kowey PR, Myerburg R, et al. Practice patterns among United States cardiologists for managing adults with atrial fibrillation (from the AFFECTS Registry). *Am J Cardiol.* 2010; 105(8): 1122–1129. [PubMed: 20381664]
8. Freeman JV, Simon DN, Go AS, et al. Association Between Atrial Fibrillation Symptoms, Quality of Life, and Patient Outcomes: Results From the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). *Circ Cardiovasc Qual Outcomes.* 2015; 8(4):393–402. [PubMed: 26058720]
9. Kabra R, Cram P, Girotra S, Vaughan Sarrazin M. Effect of race on outcomes (stroke and death) in patients >65 years with atrial fibrillation. *Am J Cardiol.* 2015; 116(2):230–235. [PubMed: 26004053]
10. Margolis R, Derr L, Dunn M, et al. The National Institutes of Health's Big Data to Knowledge (BD2K) initiative: capitalizing on biomedical big data. *J Am Med Inform Assoc.* 2014; 21(6):957–958. [PubMed: 25008006]

11. Matney SA, Settergren TT, Carrington JM, Richesson RL, Sheide A, Westra BL. Standardizing Physiologic Assessment Data to Enable Big Data Analytics. *West J Nurs Res*. 2016