Critique of Samuels, et al., 2015

Julianne Ammirati

February 06, 2017

Big Data Analytics as a Future Method for Nurse Researchers

Beginning with the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009, health care facilities nationwide began adopting electronic health record (EHR) systems at a pace not previously matched. The implementation of these EHRs provided researchers with an abundance of digitized health information, but Samuels et al. illuminate the limitations that the use of EHR data poses for researchers and offer explanations for issues in the field that will most likely be addressed as the science of "Big Data" improves (2015). Though the authors delve mostly into the challenges presented by digitized health records, they make it clear that these challenges should be addressed and resolved to allow researchers the access and utilization of EHRs, as the opportunities this data present outweigh the challenges, in many cases.

As technology develops, it is vital for nursing researchers to expand their interest into new realms of innovation in healthcare, and this article helps explain the value of the new abundance of digitized data for research; individuals no longer are required to abstract data from illegible, handwritten patient records. This new branch of nursing research, while both intimidating and promising, has various limitations, not least of which is the fact that many nursing researchers are not typically informatics focused. This article serves to introduce the field of big data while promoting interest and still tapering that excitement due to the restraints within the organization, analysis, and interpretation of big data. This balanced approach by the authors seems to echo the primary purpose of this article: The challenges of big data are not necessarily solvable, but are rather the result of the advantages of big data, and as such, solutions to these challenges will need to optimize a balance between competing interests, such as accessibility and security or nuanced versus clear-cut.

Limitations and Advantages of Big Data

As previously mentioned, the authors introduce the primary limitations and advantages of using EHR data for research purposes, and convincingly argue that these are issues, while not opining about potential ways forward in this field. The primary limitations highlighted are the various structural differences of both the

data itself and the EHR, human error in use/utilization during data input, and information security. These challenges are varied in their sources, as they have both human and technological causes. Issues in data structure are often differences that arise organically through the various data types that are housed within an EHR. For example, there is no simple way to automate the analysis of both text fields and MRIs dicom files. Additionally, there are structural issues that arise because units and hospitals often utilize different proprietary EHR software, which categorizes data in different types of fields with different field structure, which often limits the capacity for comparison across EHR systems.

Of all the prominently identified issues, I most relate to the difficulty EHRs have balancing ease of entry with too many limitations in field structure. There are nuanced differences which exist in the language used by nursing staff in long-form patient notes, and many EHRs try to capture information via drop-down menus or structured fields that only allow specified inputs. The incongruity that exists between the information the nurse wants to input and the information the EHR is willing to accept leads to work-arounds, inadequate charting, and blank fields. The authors convincingly demonstrate that this issue is a problem without an easy solution, and this issue is one that translates to many research methods: How can you maximizing the usability of the data while still capturing the most appropriate and accurate information. As a long-time research coordinator, this was an issue I have struggled with in smaller databases as well, since team members are more likely to document information in more intuitive ways (IE: yes-no check-boxes) and the output from these fields was more easily tracked, but these field formats did not always allow the study team to track all important aspects of the study.

While each of these challenges adds a level of complexity to the use and analysis of big data, they are acceptable because of the sheer scope of big data and the novel information that researchers can gain from this field. The authors prominently mention that 90% of all data an information available up until the writing of this article had been accumulated in the five years prior to manuscript preparation. This information, from Dell EMC (EMC2), is a strong argument that big data analytics are not just a potential method but a necessity to gain meaningful knowledge from the massive amounts of data that are collected. The use of data mining allows researchers to find patterns in data sets that would otherwise contain a wealth of information to difficult to parse.

Striking a Balance in Big Data

While the limitations are concerning, Samuels et al. support their main argument that the trade-offs of working with big data are just collateral in order to use EHR data at full capacity. The process of finding

signal in the noise requires a delicate balance between issues that are indivisible from EHR data entry, storage, and security. While the authors do not propose solutions to limitations, this paper discusses the use of data mining and big data analysis as undeniable future methodologies to be utilized by nursing researchers in order to capitalize on the volume, velocity, and variety of data being collected within the healthcare system.

Markdown file used for pdf production is available at https://github.com/JulianneA/N741_Homework2