Hansell, Int J Nurs Clin Pract 2017, 4: 220
https://doi.org/10.15344/2394-4978/2017/220

International Journal of Nursing & Clinical Practices

Methodology Open Access

Advances in Nursing Research Methodology: Big Data Analytics the Future

Phyllis Shanley Hansell

College of Nursing, Seton Hall University, South Orange, New Jersey 07079, USA

Abstract

Big data is an advanced research methodology available to nurse researchers as part of a collaborative intra-professional team effort. In the Digital age large data sets are generated at a rate that defies definition. These data sets represent found data that exceeds the capacity of normal computer systems and is typically collected for another purpose, that are then used to answer research questions and test hypotheses. The team typically includes at least one or more health professionals who work on a collaborative team with data experts that include a statistician, software programmer and additional information technology experts. These methods are critical to the future advancement of nursing science and healthcare, and need to be integrated into the curriculum of Nursing PhD programs. The management of big data requires special ethical consideration by the IRB to insure that the project is either exempt or excluded and free from human subjects' concerns. Big data methodology is important to and the advancement of nursing science and healthcare in the Digital Age.

Publication History:

Received: December 2, 2016
Accepted: January 28, 2017
Published: January 30, 2017

Keywords:

Big data, Methodology, Nursing research, National Institutes of Healt

Introduction

Through the latter half of the twentieth century nursing research methodology has advanced in ways that would not only impress Florence Nightingale, but would more than likely amaze her. Nightingale's fame was gained largely from her keen observational skills, analytical prowess and her extraordinary ability to graphically present results and to expound on the health issues of the soldiers of the Crimean War. As nursing science has advanced so too have the methodologies that nurse researchers use. In the not too distant past there the notion of high quality research was the randomized clinical trial to test the effectiveness of a nursing intervention. The science of the nursing profession now incorporates an expansive array of methodologies that have evolved in every increasing complexity that enable the investigator to ask research questions and test hypotheses from a variety of perspectives. The coming of age of Nursing Science in the twenty-first century as positioned the profession to on the cutting edge of health care innovation. Following Nightingale's lead to manage and manipulate data, the nurse researcher enters into the multidisciplinary arena of big data that enables the investigative team to join together to tackle big data sets whereby data analytics are deployed to answer important research questions and to test hypotheses [1].

Definition of Big Data

In the era of the twenty-first century Digital Age we are immersed in and surrounded by big data. According to Brennan and Bakken [1] the term encompasses large amounts of data that are generated by sensors, novel research techniques and ubiquitous information technologies. Typically, big data is understood as data sets too large to be managed and analyzed by the usual computer systems. Exactly what constitutes big data is constantly is a state of flux due to the ever increasing capacity of computer systems [2]. Big data is characterized by the "Three Vs" to which a fourth was recently added: volume; variety; velocity and value [3]. Mayer-Schonberger and Cukier [2] point out for us that the sum is more valuable than its parts which is suggestive of a holistic approach that is evident in most nursing theories.

The Data Scientist and the Nurse Researcher

The data scientist is defined by [4] as "a person who takes raw material (data in this case) and uses skill, knowledge and vision to craft it into something of unique value." They further explain that "the data scientist must have the abillity to connect scenarios to life as Nightingale did by using both data and visualization techniques that is nothing beyond story telling.

Mayor-Schonberger [2] see the profession of the data scientist as one with the skills of the statistician, software programmer, infographics designer and story teller. The role of the data scientist represents a unique type of expertise. However, in reality the data scientist is a member of the investigation team with a unique specialization who works collaboratively with the other members of the team. It is within the context of the team that the nurse investigator and the data scientist work as collaborative equal partners to answer the research questions or hypotheses posed by the nurse and other health professionals who are members of the team. Big data analytics is never accomplished as a lone endeavor, but rather always requires a team with mutually shared goals that are achieved through the collective collaborative efforts of the research investigation team. Overall the inter-professional approach to big data possesses the potential to identify and solve healthcare issues and problems that were never before possible.

Purpose of Big Data

The unique purpose of big data analytics is to examine large sets of data and to uncover hidden patterns, new correlations, market trends, preferences and patient care outcomes. Some of the potential benefits of big data include: cost savings; inclusion of entire populations with robust samples thus eliminating potential sample bias. The big data arrive at a velocity that is not under the control of the investigator and possess a level of imprecision that is not found in traditional inquiry [1]. Big data is typically found data that was collected for another purpose. Although it may represent a primary analysis of the data set it is more represents a type of secondary data analysis, big data is most important as it can be used to answer endless research questions or to test multiple hypotheses.

***Corresponding Author:** Prof. Phyllis Shanley Hansell, Graduate Nursing Department, College of Nursing, Seton Hall University, South Orange, New Jersey 07079, USA ,Tel: +19732021940; E-mail: hanselph@shu.edu

Citation: Hansell PS (2017) Advances in Nursing Research Methodology: Big Data Analytics the Future. Int J Nurs Clin Pract 4: 220. doi: https://doi.org/10.15344/2394-4978/2017/220

Nursing and Big Data

In Nursing Science efforts have been made to achieve theory driven science which uniquely provides the necessary frameworks that can guide analytical exploration towards promising unraveling of unique phenomena that will leverage new insights into patient care and nursing practice [1]. In terms big data relevance to nursing science one need not look any further than the 2010 update of the American Nurses Association's Social Policy Statement (2010) which specifies the role of the nurses in the generation and application of technology to healthcare outcomes and planning for health policy and regulation that is responsive to consumer needs and provided the best resource in the provision of health for all [5]. According to Brenner and Bakken [1] it is essential that nurses apply their expertise in multiple areas of relevance to both big data and data science, including the definition of important questions, extending data sources, applying data miming and modeling methods while modeling methods, and addressing the ethical legal and social implications (ELSI).

According to Grady [6] the understanding of big data methods and their application has the potential to advance the understanding of human health, as well as, response to human disease. Understanding and expertise with big data is an essential requisite skill for nurse scientists of today and the future [7]. In concert with the other institutes at the National Institutes of Health (NIH), the National Institute for Nursing Research (NINR) is a contributor along with the other institutes and centers at NIH to The Big Data Knowledge Initiative which includes the development of massive open online data management system, curriculum development and training in data science [6]. Nursing science cannot address or be prepared for the future without a comprehensive understanding of advanced methods such as big data. It is therefore essential that big data analytics be integrated into the curricula of PhD Programs so that the important questions can be answered.

"Big Data" Processes

The steps in big data analytics approximate those of a typical secondary analysis. The steps of the process in a 'big data project include:

1. Perform the literature review.
2. Identify literature gaps.
3. Obtain permission from data owner to use data.
4. Define the research question and/or hypothesis.
5. Clean and organize the data.
6. Determine the sample.
7. Evaluate the original sample.
8. Establish safeguards to protect data.
9. Seek IRB approval for exempt status.
10. Perform data analysis.
11. Disseminate findings.

The major difference that sets big data apart from traditional research is the use of existing found data and the need to implement an intervention and collect data from a defined sample. In some respects, it can be less time consuming and include a much more extensive data set. However, it is important to recognize and be mindful that big data analytics requires a team approach, and can be highly complex in its implementation. Big data presents amazing intra-professional opportunities for nurse scientists to play a significant role in the advancement of nursing science and healthcare. It is a most amazing opportunity for the profession to make a significant difference that will positively affect the healthcare outcomes for many individuals.

Ethics and the Protection of Human Subjects and "Big Data"

Big data represents a major technological advance in research methodology that is important for nurse scientists to avail themselves of. It also represents an epistemic shift wherein data is viewed as infinitely networkable, indefinitely reusable, and completely divorced from the context of collection [8,9]. Although big data presents many analytical advantages that may even accelerate the research process there are important differences that need to be considered especially in relation of the ethics of the conduct of research.

In 1981 the Common Rule which mandates human subjects Institutional Review Board (IRB) review for all federally funded study that includes an independent study ethics review prior to commencing [10] (HHS) which requires most projects researchers' due diligence to identifying and remedy potential physiological, psychological and informational harm to humansubjects in such a manner that human subjects cannot be identified subjects.

The Common Rule has not been usually applied to the core disciplines of big data since the disciplines of mathematics, computing and statistics are understood to be conducting research on systems not human subjects [11]. Clearly a new category in addition to exempt needed to be identified to encompass projects that used big data attached to human subjects. Complaints and concerns relative to the consistence of interpretation and application of the exempt status for big data projects. As a result, HHS has proposed a new category [10] of excluded that would receive no ethical review due to inherently low risk to human subjects of particular interest is the exclusion of:

1. Research involving the collection or study of information that has been acquired solely for non-research purposes.
2. Data for research study was acquired for research studies other than the proposed research study when the sources are publically available.
3. The information is recorded by the investigator in such a manner that human subjects cannot be identified, directly or through identifiers' linked to the subjects, and the investigator does not contact the subjects, and the investigator will not re-identify subjects or otherwise conduct an analysis that would lead to individually identifiable private information (HHS, 101 (b)(2).

Data that are determined to be Excluded based upon the fact that the data has been de-identified [11]. The Exempt category will continue to application with regards to big data to include secondary analysis of data using identifiable information collected for non-research purposes.

The coming of age for big data has resulted in the careful examination of the protection of human subjects included in big data projects. It will ultimately be the responsibility of the affiliated local IRB to examine the source of data to determine if the project warrants either Excluded or Exempt status and the investigators have made every effort to protect the rights of human subjects.

Conclusions

Big data is everywhere and nurse investigators need to be schooled in the how to of big data projects. Big data is a multidisciplinary team effort that includes health professionals, statisticians, software developers, high tech and big data specialists. It is important that PhD students in Nursing have big data analytics integrated into their program of study. Looking to the future big data methods are important to nursing science and health care research. Important questions and hypotheses need to generated by nurses and other members of the health professions to collaborate on big data to identify patterns of health seeking behavior and to seek answers in big data sources including the electronic health care record. The best of nursing research is in our future.

Competing Interests

The author declare no competing interests.

References

1. Brennan PF, Bakken S (2015) Nursing Needs Big Data and Big Data Needs Nursing. J Nurs Scholarsh 47: 477-484.
2. Mayer-Schonberger V, Cukier K (2013) Big Data: A Revolution that will transform hoe we live work and think. London: John Murray.
3. Laney D (2001) 3D data management: controlling data volume, velocity and variety.
4. Mohanty S, Brown W (2013) Big data imperatives, Berkeley, CA: Apress.
5. American Nurses Association (2010) Nursing's Social Policy Statement. Silver Spring Maryland: American Nurses' Association,
6. Grady PA, Gough LL (2015) Nursing Science: Claiming the Future. J Nurs Scholarsh 47: 512-521.
7. National Institutes of Health (2015) NIH big data to knowledge (BD2K).
8. Boyd D, Crawford K (2012) Critical Questions for big data, Information, Communications and Society 15: 662-679.
9. Kitchem R (2014) Big data new epistemologies and paradigm shifts, Big Data and Society 1,1.
10. Department of Health and Human Services (2015) Notice of Proposed rulemaking: Federal Policy for the protection of Human Subjects, Federal Register.
11. Metcalf J (2016) Computing ethics-big data analytics and revision of the common rule. Communications of the ACM 59: 33-33.