

## **Big data: a concept in nursing discipline**

**Raed Ali Jaradat\***

riaradat@mutah.edu.jo

### **Abstract**

**Background:** Big data is an important concept for the caring disciplines, such as nursing, and medicine, the rapid development of technology has made massive volumes of data available and achievable anytime and anywhere around the world.

**Methods:** The Morse et al. (1996) criteria for concept maturity were used as a framework. A literature review of the big data concept was undertaken to evaluate the logical, epistemological, linguistic, and pragmatic principles. Big data literature was retrieved by searching the Cumulative Index to Nursing and Allied Health Literature (CINAHL), MEDLINE, PubMed, and Google Scholar. Search terms used were big data, big data analysis, big data and nursing, and data science.

**Conclusion:** Big Data has recently become a popular term among researchers and IT professionals as well as health professionals. Even though the concept meaning is still unclear, but it holds great promise to enhance healthcare quality. Despite that, the concept of big data has been attracting the attention of scholars in the nursing discipline; the lack of a unified definition of the concept limits the use of big data in nursing research. With all its complexity, the concept becoming increasingly more prevalent and affects how nurses learn, practice, conduct research, and develop policy. Consistently defining the concept can help the nurses to prevent, diagnose, and treat health-related issues and enhance healthcare quality

**Keywords:** Big data, Concept analysis, Nursing Discipline, Data Science

---

\* College of Nursing, Mutah University.

Received: 9/8/2023.

Accepted: 29/8/2023.

© All rights reserved to Mutah University, Karak, Hashemite Kingdom of Jordan, 2023.

## مفهوم البيانات الضخمة في مهنة التمريض

رائد علي جرادات\*

### ملخص

المقدمة: تعتبر البيانات الضخمة مفهوماً مهماً لتخصصات الرعاية ، مثل التمريض والطب ، وقد أتاح التطور السريع للتكنولوجيا كميات هائلة من البيانات المتاحة وقابلة للتحقيق في أي وقت وفي أي مكان حول العالم. وليس هناك تعريف محدد لهذا المفهوم في التمريض والهدف من هذه الدراسة تحديد المعنى لمفهوم البيانات الضخمة في مجال التمريض

الطريقة: تم استخدام معايير مورس وزملائه نضج المفاهيم كإطار عمل. تم إجراء مراجعة الأدبيات لمفهوم البيانات الضخمة لتقييم المبادئ المنطقية والمعرفية واللغوية والبراغماتية. تم استرداد أدبيات البيانات الضخمة من خلال البحث في الفهرس التراكمي للتمريض والأدب الصحي المساعد (CINAHL) و MEDLINE و PubMed و Google Scholar. كانت مصطلحات البحث المستخدمة هي البيانات الضخمة، وتحليل البيانات الضخمة، والبيانات الضخمة والتمريض، وعلوم البيانات. وخلصت الدراسة الى انه اصبحت البيانات الضخمة مؤخرًا مصطلحًا شائعًا بين الباحثين ومتخصصي تكنولوجيا المعلومات وكذلك المتخصصين في مجال الصحة. على الرغم من أن معنى المفهوم لا يزال غير واضح، إلا أنه يحمل وعودًا كبيرة لتحسين جودة الرعاية الصحية. على الرغم من أن مفهوم البيانات الضخمة قد جذب انتباه العلماء في تخصص التمريض، إلا أن عدم وجود تعريف موحد للمفهوم يحد من استخدام البيانات الضخمة في أبحاث التمريض. مع كل تعقيدها، أصبح المفهوم أكثر انتشارًا بشكل متزايد، ويؤثر على طريقة تعلم العاملين في مجال التمريض وممارستها وإجراء البحوث وتطوير السياسة. ويمكن أن يساعد تحديد المفهوم بطريقة متسقة على تشخيص وعلاج المشكلات المتعلقة بالصحة وتعزيز جودة الرعاية الصحية

الكلمات المفتاحية: البيانات الضخمة، تحليل المفهوم، مهنة التمريض، علم البيانات

\* كلية التمريض، جامعة مؤتة.

تاريخ تقديم البحث: 2023/8/9.

تاريخ قبول البحث: 2023/8/29 م.

© جميع حقوق النشر محفوظة لجامعة مؤتة، الكرك، المملكة الأردنية الهاشمية، 2023 م.

## **Introduction:**

Big data is an important concept for the caring disciplines, such as nursing, and medicine, the rapid development of technology has made massive volumes of data available and achievable anytime and anywhere around the world. Computing technology has further advanced in recent years, and the accessibility and popularity of social media and smartphones have led to new sources for collecting research data. These technological developments have caused the world to progress to an "era of data"(Wong et al., 2016).

Today, healthcare organizations use multiple, complex information technology systems that must be integrated in a meaningful way to manage clinical, and financial issues (Zhu et al., 2019). Morse and her colleagues (1996) suggested that concept analysis begins with assessing a concept's level of maturity. When a concept appears to be mature within a large body of literature yet is associated with inconsistent and competing assumptions or variables, then quantitative and/or qualitative research may be used effectively in concept clarification. They also added clarifying a concept requires a detailed familiarization of the literature to recognize values and assumptions common to the concept under review. Based on these values and assumptions, attributes can then be identified and synthesized.

## **Definition of Big Data**

The concept of "Big Data" is quite vague and there is no agreement on its definition among researchers of different disciplines. Understanding big data is a priority for nurses as the profession for the prevention, diagnosis, and treatment of health-related issues (Wong et al., 2016). Big data is a collection of large amounts of data that is unmanageable using traditional software and digital sources (Nashwan, 2017). Moreover, big data is described not by essentially the volume of the data, but instead by the collection and speed of data (Bhadini & Jothimani, 2016).

The most popular definition of big data is the 4Vs, because of its clear structure and the concept originated from data scientists (Wong et al., 2016). Volume refers to the size of the dataset measured by the unit byte, velocity refers to the speed at which new data is generated, variety refers to the number of data types, sources, or attributes, and veracity refers to the accuracy of a database.

---

## Big Data in Nursing Discipline

The healthcare industry is beginning to see just how beneficial big data can be to patients, doctors, and nurses. For example, a new nurse leader rounding technology makes use of big data to improve the level of care available by hospital nurses (Zhu et al., 2019).

The main components of a healthcare system are the health professionals (physicians or nurses), health facilities (clinics, and hospitals for delivering medicines and other diagnosis or treatment technologies), and a financing institution assistant the previous two (Wong et al., 2016). Working in healthcare organizations with complex information technology with multiple clinical, financial, and rights systems that must be integrated sets the stage for the big data challenge. The ability to integrate different data and analyze it to better understand outcomes is not a small task. Having access to the right information at the right time to support clinical decisions is essential in planning and providing the right care for patients (Clancy, 2020; Topaz & Pruinelli, 2017).

In the nursing discipline, a suitable definition of big data is essential to advance nursing research and practice. Moreover, all clinical practices need the information to make critical and clinical decisions, other care providers improve care quality, improve outcomes, and reduce healthcare costs, insurance, and cost data (Nashwan, 2017; HIMSS, 2015).

Understanding big data is essential for nurses, as the profession aims to provide the best possible care to patients. Capturing healthcare data in an organized way helps build the foundation for accurate, reliable information regarding a patient across various systems and settings of care (Carter-Templeton et al., 2021; Sensmeier, 2016). Using data elements consistently allows information to be collected and reused for multiple purposes including; outcomes measurement, practice level improvements, population health, research, and decision support (ONC, 2014).

Nurses and patients can benefit from the application of data science methods to the massive amount of data now emerging from the healthcare systems. Large healthcare delivery systems are using their vast databases to predict probable outcomes depending on the treatment a patient chooses (Brennan & Bakken, 2015). Moreover, the use of big data technologies can help nurses and other healthcare providers improve care quality, optimize outcomes, and reduce healthcare costs (HIMSS, 2015)

## **Methodology**

The evaluation of the maturity level of big data as a concept in the nursing discipline was guided by the four parameters recommended by Morse and her colleagues (1996). Epistemological, Logical, Pragmatical, and Linguistical. Big data literature was retrieved by searching the Cumulative Index to Nursing and Allied Health Literature (CINAHL), MEDLINE, PubMed, and Google Scholar. Search terms used were big data, big data analysis, big data and nursing, and data science.

## **Results**

### **First: The epistemological principle**

The first principle refers to how the concept has been appropriately clearly defined and well differentiated from other concepts. It is difficult to reach a complete and clear definition of the concept because it is dynamic and changeable over time, and influenced by time, and the advancement of technology (Wong et al., 2016).

Because there are at least 32 such definitions available from different sources, the definition of big data has been described as a “moving target” (De Mauro et al., 2015). The most popular and widespread of these definitions is the 4Vs because of its clear structure and because the concept originated from a data scientist (Wong et al., 2016). According to this definition, Big Data should fulfill four characteristics that are represented by four keywords starting with the letter “V”. These characteristics are “Volume”, “Velocity”, “Variety,” and “Veracity”.

Importantly, Big Data is not only defined by the size of a data set, but can also be characterized by traits including velocity, variety, veracity, exhaustivity, relationality, extensionality, and scalability (Kitchin & McArdle, 2016). Data scientists and researchers have described variety as the greatest challenge of the big data era (Jin et al., 2015; Leventhal, 2014). Others stated that the volume of data is not the only characteristic, four key characteristics define big data: volume, velocity, variety, and value (Oracle, 2013).

Without a clear consensus on the big data issue, scholars in nursing have adopted very different definitions of big data, while some did not attempt to define the term before using it (Brennan & Bakken, 2015). Among the 38 publications related to big data in PubMed Nursing Journals

---

search conducted, 22 explicitly defined big data, nine of the 22 adopted the popular 3Vs or 4Vs definition, while the remaining publications adopted different, less structured definitions (Wong et al., 2016).

In nursing science, the 4Vs definition of big data is adopted; nursing scholars may have to face the problem that some of the studies that claim to have adopted big data have not done so. Moreover, the 4Vs definition may become a barrier to nursing scholars joining the big data research family because of the independent nature of nursing scholars, who are not unique in this regard (Wong et al., 2016). Another characteristic of the data generated today is its increasing variety in type. The diversity of data types is one of the challenges that organizations must tackle to make value out of the extensive informational sets available today (Manyika et al., 2011; Russom, 2011).

### **Second: The logical principle**

In the second principle, the researcher evaluates the relation and integration of the "Big Data" concept with other concepts in the nursing discipline. The logical principle focuses on the boundaries of the concept and overlapping with other concepts.

Big data have been described as related to five dimensions as a concept contributed to "bigness": Volume, Velocity, Veracity, variety, and Value (Gephart et al., 2018). There are many terms related to big data, those terms include: electronic health records (EHRs), (Sensmeier, 2016; Topaz & Pruinelli, 2017), and data science and healthcare informatics (Gephart et al., 2018)

"Digital universe" is a concept that may contribute to confusion about big data concept. The term "digital universe" is quantitatively defined as a massive amount of data created, replicated, and consumed in a single year. For example, the digital universe in 2017 expanded to about 16,000 Exabytes (EB). (IDC) predicted that the digital universe will expand to 40,000 EB by the year 2020, over the past decade, the expanded digital universe constitutes "big data" that has been used by Information Technology (IT), which led to the birth of a new field of science termed "Data Science" (Dash et al., 2019).

Data science utilizes special methods and tools that turn big data into action. Moreover, in healthcare practice, healthcare informatics focuses on hardware, software, and computer technology; this field starting to shift focus from using technology to retrieve data and sophisticated analysis of

data that are informed by data science (Gephart et al., 2018). Dash and his colleagues (2019) present “Data Science” as dealing with various aspects of big data including management and analysis of data to extract deeper insight to improve the system functionality. Therefore, big data are more complex data analytics and focused study that is applied by healthcare professionals and computer scientists (Gephart et al., 2108; Zue et al., 2019).

### **Third: The pragmatic principle**

The third principle refers to how the concept has been appropriately operationalized, and the applicability and usefulness of the big data concept in health discipline.

Nursing has operationalized the concept in terms of strategies to enhance the discipline and recognized the power of big data to advance human health and disease. Nursing and patients can benefit from the application of data science methods to the massive amount of data now emerging from the healthcare systems (Wong et al., 2016). In addition, large healthcare delivery systems are using their vast databases to predict probable outcomes depending on the treatment a patient chooses and big data can reveal patterns shared by thousands of patients (Brennan & Bakken, 2015). Moreover, in healthcare delivery services big data can be used to enhance productivity, for example, management of chronic or long-term conditions is expensive. Use of in-home monitoring devices to measure vital signs, and monitor progress to improve patient health and reduce both office visits and hospital admittance (Oracle, 2013).

With the continuous acceleration of the informatization process, the medical and health field has gradually entered the era of big data. Diverse data can provide new methods and ideas for nursing practices and have application value in many areas, including nursing evaluation, improving the level of nursing practice, monitoring disease, nursing research and clinical decision support, improving care quality, optimizing outcomes, and reducing the cost of health care (Gephart et al.,2018; Wong et al., 2016).

In addition, documentation is an integral part of nurse's work where nurses include relevant information about the patients, and they contribute to big data through their documentation in the Electronic Health Record (EHR), to study patients, reduce healthcare-related complications, and develop policy initiatives (Fox & Hendler, 2011; Shaw, 2014).

---

**Fourth: The linguistic principle**

In the fourth principle, the concept "Big Data" is the evaluation for appropriate and consistent use in the nursing context. Nursing literature uses the concept "big data" to reflect having structured access to capture data that describe building accurate, reliable information regarding the patient across multiple settings of care (Sensmeier, 2016, Topaz & Pruinelli, 2017; Zhu et al., 2019). Furthermore, although the big data concept is relatively new, its related concepts like health care data, EHRs, and data analytics are previously used in nursing literature (Topaz & Pruinelli, 2017).

Consistency and accuracy of data are realized through interpretable systems. Advances in EHRs, and among the first professional groups, nurses were starting standardization of big data terminology, but the huge number of terms makes such standardization difficult (Sensmeier, 2016). To analyze narrative big data, nurses will need support and collaboration of other disciplines, such as computer linguistics and computer science (Topaz & Pruinelli, 2017). Furthermore, when the nursing narrative is evaluated, nursing intervention should be mapped to standardized nursing terminologies (like the International Classification for Nursing Practice, ICNP). This mapping will allow the use of extracted data for automated reasoning (e.g., clinical decisions). For example, when the computer algorithm is standardized to identify patients with pressure ulcers (ICNP code 10015612) in a large pooled clinical note, the automated specific nursing intervention that will be generated by the system (e.g., facilitate adherence to the regime, ICNP code 10036273) (Topaz & Pruinelli, 2017).

**Discussion**

Big data concept was analyzed based on Morse and her colleagues' principles (1996). Even though there is a bulky body of literature addressing the concept of big data, there is no consistency in the definition, and it is still difficult to reach a complete and clear definition for the concepts because it is dynamic and changeable over time and influenced by technological advancement.

The lack of a formal definition has led research to evolve into multiple and inconsistent paths. A convincing definition of a concept is an enabler of its scientific development. As Ronda-Pupo and Guerras-Martin (2012) suggest, the level of consensus shown by a scientific community on a definition of a concept can be used as a measure of the progress of a



discipline. Big Data has instead evolved so quickly and disorderly that such a universally accepted formal statement denoting its meaning does not exist. For Lany (2012), the most popular and widespread definition is the 4Vs because of its clear structure.

In the past several decades, the health natural language processing systems mostly focused on the work with texts form of medical domain, now natural language processing is concerned with creating automated approaches to understanding human language, that combine linguistics computer science, and informatics to create computer algorithms capable to understand generated texts from the EHRs, social media, patient, and family data ((Topaz & Pruinelli, 2017). Nurses should encourage the use of standardized accepted terminologies that can address the documentation needs of the entire nursing care terms and to minimize free-text documentation, regardless of the care setting (Nashwan, 2017). Unified standardized nursing terminology has become a necessity to promote the development of nursing science (Zhu et al., 2019).

Big data is an emerging area in nursing that could lead to the development of the next generation of more sophisticated and compelling research for evidence-based nursing practices. However, for the time being, in nursing research as in many other disciplines, the existing definitions of Big Data may not be suitable (Hilbert & Lopez, 2011). Therefore, further research is necessary, still, the concept requires a continuous and expensive research and development effort to keep up with the pace at which data size increases especially with the growing share of byte-hungry data types such as images, sounds, and videos.

## **Conclusion**

Big Data has recently become a popular term among researchers and IT professionals as well as health professionals. Even though the concept meaning is still unclear, but it holds great promise to enhance healthcare quality. The nursing discipline needs to understand how an organization should start with big data, in terms of handling, organizing, and analyzing. IT departments who are the technical supporters must collaborate with other stakeholders to select the most appropriate information to identify the best software and hardware for big data analytic process to generate the desired outcomes.

---

Regardless, the concept of big data has been attracting the attention of scholars in the nursing discipline; the lack of a unified definition of the concept limits the use of big data in nursing research. With all its complexity, the concept becoming increasingly more prevalent affects how nurses learn, practice, conduct research, and develop policy. In addition, defining the concept consistently can help nurses prevent, diagnose, and treat health-related issues and enhance healthcare quality.

We hope that the analysis and evaluation of "Big Data" concept is valuable for nurse practitioners, educators, researchers, and policymakers to provide a vision for the growth path and help prepare the next generation of nurses to improve patient outcomes.

## References:

- Bhadani, A. & Jothimani, D. (2016). Big Data: challenges, opportunities , and realities. Ineffective big data management and opportunities for implementation. IGI Global. (pp. 1-24).
- Brennan, P. & Bakken, S. (2015). Nursing needs big data and big data needs nursing. *Journal of Nursing Scholarship*. 47(5). 477 –484.
- Carter-Templeton, H., Nicoll, L., Wrigley, J. & Wyatt, T. (2021). Big data in nursing: A bibliometric analysis. *The Online Journal of Issues in Nursing*, 26(3).
- Clancy, T. (2020). Artificial intelligence and nursing: the future is now. *JONA: The Journal of Nursing Administration*, 50(3), 125-127.
- Dash, S., Shakyawar, S. K., Sharma, M.; & Kaushik, S. (2019). Big data in healthcare: management, analysis, and future prospects. *Journal of Big Data*. 6(54), doi.org/10.1186/s40537-019-0217-0
- De Mauro, A., Greco, M. & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. *AIP Conf. Proc.* 1644, 97–104.
- Fox, P. & Hendler, J. (2011). Changing the equation on scientific data visualization. *Science*, 331, 705-708.
- Gephart, S., Davis, M. & Shes, K. (2018). Perspectives on Policy and the Value of Nursing Science in a Big Data Era. *Nursing Science Quarterly* 3(1) 78-81. doi.org/10.1177/0894318417741122.
- HIMSS Guiding Principles for Big Data in Nursing. (2015). HIMSS CNO-CNIO Vendor Roundtable. Retrieved from [www.himss.org/big10](http://www.himss.org/big10) (accessed 20 April 2020).
- Jin, X., Wah, B. W., Cheng, X. & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*. 2(2), 59-64.
- Kitchin, R., & McArdle, G. (2016). What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*. 3(1), 1-10. doi:10.1177/2053951716631130.
- Lany, D. (2012). 3D Data Management: Controlling Data Volume, Velocity and Variety. Available online: [blogs.gartner.com/doug-laney/files/2012/01/ad949-](http://blogs.gartner.com/doug-laney/files/2012/01/ad949-) (accessed on 20 April 2020).

- 
- Leventhal, R. (2014). Survey: Variety, not volume, the biggest challenge to analyzing big data [Web log post]. Retrieved from <http://www.healthcare-informatics.com/news-item>.
- Manyika, J., Chui, M., Brown, B. & Bughin, J. (2011). Big Data: The next Frontier for Innovation, Competition ,and Productivity.
- Morse, J., Hupcey, J., Mitcham, C. & Lenz, E. (1996). Concept analysis in nursing research: A critical appraisal. *Research and Theory for Nursing Practice*. 10(3), 253–277.
- Nashwan, A. (2017). Big data and the future of nursing. Retrieved from <https://www.researchgate.net/publication/319547432>.
- Office of the National Coordinator (ONC) for Health Information Technology. (2014) Health IT enabled quality improvement a vision to achieve better health and health care.
- [www.healthit.gov/sites/default/files/HITEnabledQualityImprovement-111214](http://www.healthit.gov/sites/default/files/HITEnabledQualityImprovement-111214).
- Oracle (2013). Big Data for the Enterprise. Retrieved from [www.oracle.com/us/products/database/big-data-forenterprise-519135](http://www.oracle.com/us/products/database/big-data-forenterprise-519135).
- Ronda-Pupo, G. & Guerras-Martin, L. (2012). Being aware of the limitations of big data methods. *Strateg. Manag. J.* 33, 162.
- Russom, P. (2011) TDWI Best Pract. Report, Fourth Quarter.
- Shaw, J. (2014). Why “Big Data” is a big deal. *Harvard Magazine*. Retrieved from <http://harvardmagazine.com/why-big-data-is-a-big-deal>.
- Sensmeier, J. (2016). Understanding the impact of big data on nursing Knowledge. Retrieved from [www. Nursingcriticalcare.com](http://www.Nursingcriticalcare.com).
- Topaz, M., & Pruinelli, L. (2017). Big Data and Nursing: Implications for the Future. IMIA and IOS Press. doi:10.3233/978-1-61499-738-2-165.
- Wong, H., Chiang, V., Choi, K. & Loke, A. (2016). The Need for a Definition of Big Data for Nursing Science: A Case Study of Disaster Preparedness *Int. J. Environ. Res. Public Health*. 13, 1015.
- Zhu, R., Han, S., Su, Y., Zhang, C.; Yu, Q, & Duan, Z. (2019). The application of big data and the development of nursing science: A discussion paper. Retrieved from <http://www.elsevier.com/journals/international-journal-of-nursing-sciences/2352-0132>