BIG DATA USED FOR NURSING research can provide greater precision and granularity for measuring care quality, patient outcomes, and cost of nursing care. However, big data for nursing research is complex and requires collaboration from multiple organizations. A national action plan to make nursing data sharable, comparable, timely, and more relevant to improving health (Delaney, Pruinelli, Alexander, & Westra, 2016) was developed at the 2013 Nursing Knowledge: Big Data Science Initiative hosted by the University of Minnesota. The Nursing Value Workgroup (NVW) formed from the Big Data Science initiative in 2014 plans to measure nursing value and develop new real-time metrics to monitor costs, quality, and effectiveness of nursing care through understanding the individual contribution of a nurse to a patient.

The vision of the NVW is to create a warehouse that hosts data from multiple disparate data sources within multiple organizations for use in research to discover, explore, and test questions related to nursing value. Business and managerial acumen is necessary to take on governance and board development work.

The collective open minds of several diverse partners including academia, practice, and industry allowed for robust discussion, freethinking, and led to development of an innovative governance structure for a nursing value research data warehouse.
Nurse staffing software available today provides data that can potentially measure nursing cost more precisely by recording variables such as individual nurse work time, salary, benefits, and patient acuity. Electronic health record (EHR) data captures the interaction between nurse and patient, which can be translated into time spent per patient. These data can be linked directly to each patient through the nurse-patient assignment. The linkage of nurses and patients is accomplished through unique identifiers; therefore, other data can also be linked such as patient demographics, diagnosis data (e.g., diagnosis-related group), and other hospital-level or clinically relevant information. In addition, data unique to each nurse can be collected such as nurse demographics and experience level. The combination of unique patient and nurse data allows analysis that until now has been nearly impossible (Welton, 2016). One author completed a pilot study using data to explore patient and nurse characteristics contributing to nursing cost per patient using unique patient and nurse data (Jenkins & Welton, 2014). Extracting, organizing, warehousing, and analyzing such data can be a complex undertaking involving many different players and organizations. To ensure validity, reliability, and standardization of these data for research, a data warehouse governance structure is imperative.

## Methods

A collaborative inter-organizational data warehouse requires explicit governance so all parties understand and agree to roles, requirements, processes, and legal implications for use of data for research. A task force of the NVW focused on creating a governance structure for the data warehouse. The consortium model, governing operational board structure, governance charter, and lessons learned of a nursing value research data warehouse (NVRDW) are presented.

### Task Force Members

The original NVRDW governance task force was a subcommittee of the NVW comprised of nine members including representatives from hospitals, academia, and industry. Additional members from primary care and industry lent expertise for thorough planning. Representatives were from diverse settings, both national and international. The task force included informatics experts such as clinical analyst, chief nurse executive strate-
Committee members completed a literature review searching PubMed, CINAHL, and Google Scholar databases using key words governance, data warehouse, plus additional words healthcare, research, and nursing. The task force reviewed 11 recent articles plus data warehouse governance reports from IBM, Patient-Centered Outcomes Research Institute, and Kaiser to gather key concepts and models (Allen et al., 2014; Ballard et al., 2014; Blumer, Giblin, Lemermeyer, & Kwan, 2017; Delaney et al., 2016; Haarbrandt, Wilschko, & Marschollek, 2016; Holmes, 2016; Holmes et al., 2014; Hripcsak et al., 2014; Kim et al., 2014; Paolino et al., 2016; Price, Shea, & Gephart, 2015; Richesson, Horvath, & Rusin-covitch, 2014).

Big data governance is an emerging topic in healthcare and nursing literature given the explosion of use of big data (Delaney et al., 2016; Price et al., 2015). Poor data governance was reported as the greatest barrier to working with big data (Blumer et al., 2017). Clear data governance structure enforces definitions, standards, and processes for a large data warehouse (Richesson et al., 2014). A tenfacet framework used in a systematic review of 39 articles published on data governance (Holmes et al., 2014) provided a model useful when considering components of a governance structure. The ten areas of focus for governance of data with definitions are included in Table 1.

**Key Decisions**

The NVRD W governance task force met via phone or WebEx™ at least biweekly for 12 months. Two co-chairs led the task force.

**Literature Review**

Committee members completed a literature review searching PubMed, CINAHL, and Google Scholar databases using key words governance, data warehouse, plus additional words healthcare, research, and nursing. The task force reviewed 11 recent articles plus data warehouse governance reports from IBM, Patient-Centered Outcomes Research Institute, and Kaiser to gather key concepts and models (Allen et al., 2014; Ballard et al., 2014; Blumer, Giblin, Lemermeyer, & Kwan, 2017; Delaney et al., 2016; Haarbrandt, Wilschko, & Marschollek, 2016; Holmes, 2016; Holmes et al., 2014; Hripcsak et al., 2014; Kim et al., 2014; Paolino et al., 2016; Price, Shea, & Gephart, 2015; Richesson, Horvath, & Rusin-covitch, 2014).

Big data governance is an emerging topic in healthcare and nursing literature given the explosion of use of big data (Delaney et al., 2016; Price et al., 2015). Poor data governance was reported as the greatest barrier to working with big data (Blumer et al., 2017). Clear data governance structure enforces definitions, standards, and processes for a large data warehouse (Richesson et al., 2014). A tenfacet framework used in a systematic review of 39 articles published on data governance (Holmes et al., 2014) provided a model useful when considering components of a governance structure. The ten areas of focus for governance of data with definitions are included in Table 1.

**Key Decisions**

The NVRD W governance task force met via phone or WebEx™ at least biweekly for 12 months. Two co-chairs led the task force.

**Literature Review**

Committee members completed a literature review searching PubMed, CINAHL, and Google Scholar databases using key words governance, data warehouse, plus additional words healthcare, research, and nursing. The task force reviewed 11 recent articles plus data warehouse governance reports from IBM, Patient-Centered Outcomes Research Institute, and Kaiser to gather key concepts and models (Allen et al., 2014; Ballard et al., 2014; Blumer, Giblin, Lemermeyer, & Kwan, 2017; Delaney et al., 2016; Haarbrandt, Wilschko, & Marschollek, 2016; Holmes, 2016; Holmes et al., 2014; Hripcsak et al., 2014; Kim et al., 2014; Paolino et al., 2016; Price, Shea, & Gephart, 2015; Richesson, Horvath, & Rusin-covitch, 2014).

Big data governance is an emerging topic in healthcare and nursing literature given the explosion of use of big data (Delaney et al., 2016; Price et al., 2015). Poor data governance was reported as the greatest barrier to working with big data (Blumer et al., 2017). Clear data governance structure enforces definitions, standards, and processes for a large data warehouse (Richesson et al., 2014). A ten-facet framework used in a systematic review of 39 articles published on data governance (Holmes et al., 2014) provided a model useful when considering components of a governance structure. The ten areas of focus for governance of data with definitions are included in Table 1.

**Key Decisions**

The NVRD W governance task...
force made several key decisions to define mission, vision, guiding principles, data warehouse model, partnership, governance and board structure, consortium model, and responsibilities of parties.

Mission, vision, and guiding principles. The first task force meeting resulted in a group consensus mission to create a governance structure for a nursing value data warehouse to facilitate research. The group held true to the original mission throughout its yearlong work. The vision defined in the charter, congruent with the NVW vision (Welton & Harper, 2015), was to develop a pre-eminent repository of nursing and patient data to enable measuring patient-level nursing intensity and cost per patient across the continuum of care. The guiding principles included creating a vendor-agnostic and setting-neutral warehouse structured and defined to support sharable and comparable data. Further delineation of data warehouse, purpose resulted in five goals for facilitating research. The goals included identifying new big data science and micro-costing methods to identify, predict, and ultimately understand and improve outcomes of patient care sensitive to the practice of nursing.

Research warehouse vs. data use agreements. Since the early 1980s, data for nursing research has been collected, stored, processed, and retrieved from data warehouses emanating from data in the clinical record. The ability to integrate data within the warehouse spares users from needing to learn multiple applications and facilitates the development of standardized reports. Data warehouses are common within an organization and becoming more prevalent across organizations.

Another model is the shared data use agreement. This model addresses the issue of need for large storage for big data research. In this model, each organization maintains its own data, but agrees to share for specific uses and users.

The NVRDW governance task force considered both models and decided to create a data warehouse for four reasons: (a) Data being collected are not typically coded in a standard way within or between systems, so the data warehouse serves as the point to scrub, integrate, and normalize the data by creating a common model for retrieval and storage. (b) Funding was not available for the technical and legal expertise to develop a shared network and retrieve data for multiple studies. (c) Control of the data architecture is ensured through creation of a common data model and data governance structure. (d) The familiarity of a warehouse would facilitate nurse researchers working across a variety of technologic solutions. This approach provides flexibility to identify and share data and provides a framework for future funding to sustain the data warehouse.

Public private partnership. The Big Data Science Initiative facilitates collaboration among academia, healthcare systems, government agencies, associations, and small and large businesses. Public/private partnerships (P3) are a contractual agreement between government and non-governmental organizations designed to benefit the public (Meyer, 2012).

A P3 can bring rewards, as each entity brings resources, talent, and perspectives, but risk is inherent as each entity brings different, competing, and sometimes conflicting priorities. Initial and ongoing clarity of the mission, goals, resources, responsibilities, and a timeline are key to success.

The NVRDW governance task force focused on development of the mission, vision, and a set of bylaws. A variety of issues arose. Mid-year, one of the partners identified funding to establish the warehouse, prior to finalizing the governance structure, resulting in a rethinking of the governance. The NVRDW governance group adopted a model described by IBM (Ballard et al., 2014), separating the responsibilities of the entity hosting the data warehouse and the governing entity (board of directors).

Governance structure. In the charter, the Consortium governs issues such as ownership of intellectual property, financial implications, data confidentiality, and use of data. Figure 1 is a graphic description of the governance structure. The Consortium executive board appoints and oversees an operational board of directors who define the structure, duties, and processes for the research data warehouse. Each entity providing data has representation on the board. The operational board of directors may appoint committees as needed.

Consortium model. A consortium model is two or more parties working together to achieve goals that would be unreachable by any one party. Knowledge, skills, and resources are shared between consortium entities. A memorandum of understanding (MOU) is a document delineating items agreed upon by participating parties. For this project a MOU, drafted by University of Colorado legal staff with input from NVRDW governance task force members, outlined key agreements between collaborating universities including support for nursing research and oversight of the NVRDW Board of Directors. Deans of the collaborating universities reviewed, edited, and signed the MOU.

Three universities (University of Colorado, University of Minnesota, and University of Kansas) provide executive oversight of the NVRDW. One representative from each university compose an executive board with oversight of the operational board decisions. The data repository, as well as personnel to support data collection, analysis, and reporting, is located on the University of Colorado Anschutz campus. Key processes include data security, Health Insurance Portability and
Accountability Act (HIPPA) compliance, data transfer from participating institutions, setup and data cleaning of extracted data, and initial analysis using local HIPPA-compliant servers as well as high-performing computers. Multiple health systems provide data for nursing research. The first hospitals providing data are located in the West and Southwest United States. Several additional systems from across the country may provide data in the future. An industry partner located in Kansas City assists with data extraction.

Board member structure. The board of directors is responsible for ensuring the mission, strategy for development, maintenance, and use of the data stored in the NVRDW. After reviewing duties of a board member (care, loyalty, and obedience), the workgroup determined diversity was key. Board members should have a working knowledge of nursing and big data, but each member should bring a different skill set and perspective.

Core stakeholders appointed during year 1 include nurse leaders from hospital, community, academia, and industry; a data scientist; and an attorney (non-voting). The board will grow during year 2 with the addition of a health economist, nursing terminologist, data warehouse technology expert, and journal editor. Members serve 3-year, rotating terms.

Decisions deferred to operating board. The NVRDW governance task force worked through key decisions and came to consensus that many decisions should be deferred to the operating board of the warehouse. Input from practice sites contributing data is imperative for several operating decisions. Decisions such as processes for evaluating requests for access to data, merits of research proposals, addition/deletion of data plus review of data use agreements, cost for data use, and dissemination will be constructed by the board of directors or appointed operating committees. Because of the nature of the data, a combination of both clinical and operational data as well as the direct linkage between individual nurses caring for each patient, this allows a range of new studies and analyses that have only recently been possible due to improvements in EHR technology and high-performance computers. This ability requires oversight for those researchers as well as clinical partners on how to provide access to the warehouse in the future.

The operational board of directors will meet at least quarterly to:
- Serve as stewards of the data.
- Monitor alignment to the vision of the NVRDW.
- Set research priorities.
- Review the financial status, develop process for budgeting and monitoring of expenditures.
- Appoint operational committees or task forces as needed.
- Develop the processes for access to data.
- Review and approve recommendations of committees.
- Address conflicts of interest.
- Coordinate progress with the Nursing Knowledge: Big Data Science Conference planning committee.
- Plan for board of directors’ succession

Charter Development

The IBM model (Ballard et al., 2014) for research database governance provided a guiding framework for charter development. Four key concepts are included in the IBM model: supporting disciplines, core disciplines, enablers, and outcomes. Supporting disciplines include data architecture, classification/metadata, and audit information logging/reporting, which are foundational in the model to support core disciplines. Core disciplines include data quality management, information life cycle management, and information security and privacy. Enablers enhance the core disciplines and include organizational structures and awareness, policy, and stewardship. Outcomes in the model are data risk management/compliance and value creation.

The NVRDW governance task force developed the charter through consensus decision-making. Task force members wrote various parts of the charter in a document located on Dropbox™. Members reviewed document drafts, refined definitions, clarified concepts, designated roles, and discussed decision-making/accountability during bi-weekly meetings.

Lessons Learned

The diverse workgroup offered multiple perspectives that added rich content to the charter. Use of complexity leadership allowed exchange of valuable viewpoints and consideration of carefully considered alternative solutions. The workgroup was volunteer and accomplished MOU and charter development in a little over 1 year. If resources such as budget for legal counsel, project manager, and workgroup members were available, the timeline might be condensed. Competing priorities of each workgroup member’s “real job” contributed to the extended timeline.

Access to examples of other data warehouse governance documents was limited but interviews with individuals participating in data warehouses provided insight into the charter requirements and potential pitfalls. The governance charter workgroup had to address potential downsides, especially the issue of permission to use data once imported into the data warehouse. The charter allows all data to be accessible once imported without requiring permission from each contributing organization. This provides for a seamless process when providing the data to an approved researcher.

The NVRDW governance task force was a public/private partner-
The research data warehouse hosts sensitive patient and nurse characteristic data. Strict data procedures are necessary due to the sensitive nature of data. Practice sites de-identify data prior to transfer to the warehouse. Secure file transfer protocol is used to transfer data from practice sites to the warehouse. Researchers use HIPPA-compliant high-speed computers to store and analyze data.

**Conclusion**

Use of big data in nursing is emerging. Publishing and sharing knowledge and lessons learned from the NVRDW workgroup can assist future nursing research teams as they develop collaborative models and research data warehouses. Business and managerial acumen is necessary to take on governance and board development work. The collective open minds of several diverse partners including academia, practice, and industry allowed for robust discussion, freethinking, and led to development of an innovative governance structure for a nursing value research data warehouse. $\$\$

**Future Implications for Nursing**

Access to big data can be expensive or elusive for nursing students and researchers. Use of big data for research is complicated and requires high-speed computers. The nursing data warehouse infrastructure can provide researchers access to extracted and anonymized data plus high-speed computational capability for use in nursing value studies. The operational board of the NVRDW develops and enforces procedures for accessing data for research. Researchers will access data from the warehouse to advance the science of nursing value by increasing the speed of knowledge generation since data follows a standard model and is accessible without institutional review board approval.

**REFERENCES**


continued on page 251
Research Data Warehouse Governance

continued from page 212


