Data Science and the Nursing Doctorate*S*

Patricia Flatley Brennan, RN, PhD, FAAN
University of Wisconsin-Madison

Support for this presentation came from the AHRQ and the Office of the Vice Chancellor for Research and Graduate Education, University of Wisconsin-Madison

The Brennan Speech Effect?

Seeing Patterns in Big Data

October 21, 2012
AACN mid-year Washington DC
October 22, 2012
Hurricane Sandy

January 31, 2015
AACN Doctoral Conference San Diego
January 31 – February 2, 2015
Chicago blizzard
19.3 inches of snow

January 21, 2016
AACN Doctoral Conference Naples
January 22, 2016
Winter Storm Jonas
East Coast expects record snows

Core Content

1. Evaluate data science perspectives and methods for their relevance to nursing
2. Select core learning content necessary for future nurse scientists and leaders in advanced nursing practice
3. Recognize the key roles that academic leaders, faculty, and students themselves play here

Audience participation opportunities abound!

It's been almost a year… what have you done?

If you heard my talk last year…

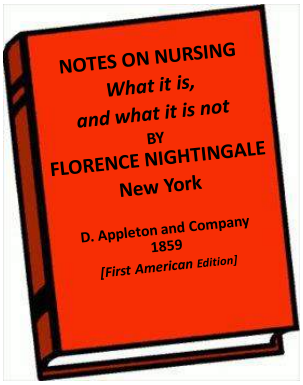
1. Have you engaged in any personal learning?
2. What course development or curriculum integration about data science has occurred?
3. What research partnerships have you initiated or do you plan to initiate around big data?

If you missed my talk last year …

1. What have you heard about big data and data science?
2. What is your school or University doing about data science?
3. Are your students asking for more learning in this area?

What did we do at Wisconsin?

1. Published last year's talk (Brennan & Bakken, JNS September, 2015)
2. Determined that for the School of Nursing, our big data strategy will focus on using data science approaches to characterize possible future states of a person at the present time and create a suite of nursing interventions to achieve desirable future states or avoid bad ones
3. Made contact with our Center for Predictive Computing, one of the 11 BD2K Computing Centers



Consider the variety of sources & the volume of data needed to convey
the essence of Nightingale's nursing

1. Ventilation and warming
2. Health of houses
3. Petty management: whatever a patient *can* do for himself…
4. Noise
5. Variety
6. Taking food
7. Beds and bedding
8. Light
9. Cleanliness of rooms and walls
10. Personal cleanliness
11. Chattering hopes and advice
12. Observations of the sick

And she had no computers,
no EHR,
and
only primitive statistics
to help her !!!

We've got big data
and
data science to help us!

Big data, BD2K,
data sci
a
Prec Medicine
itiative

So much progress, so fast, so far to go

What is biomedical Big data ?

- ***Biomedical Big Data is*** **more than just very large data or a large number of data sources.** Big Data refers to the **complexity, challenges, and new opportunities** presented by ***the combined analysis of data***. In biomedical research, these data sources include the **diverse, complex, disorganized, massive, and multimodal data being generated by researchers, hospitals, and mobile devices around the world**.
- ***Biomedical Big Data is*** diverse and complex. It includes **imaging, phenotypic, molecular, exposure, health, behavioral, and many other types of data.** These data could be used to discover new drugs or to determine the genetic and environmental causes of human disease.
- ***Biomedical Big Data*** faces many challenges. The **unwieldy** amount of information, **lack of organization and access to data and tools,** and insufficient training in data science methods make it difficult for Big Data's full power to be harnessed.
- ***Biomedical Big Data*** provides **spectacular opportunities**. Big Data methods allow researchers to maximize the potential of existing data and enable new directions for research. Biomedical Big Data can increase accuracy and supports the development of precision methods for healthcare.

NIH https://datascience.nih.gov/bd2k/about/what

Data Science at NIH

*Centers of Excellence for Big Data Computing ⓘ
*The Precision Medicine Initiative
*Training Programs
*The Commons

The Precision Medicine Initiative

- Precision Medicine Initiative®
  - Create a million participants for genomic assay
  - Leveraging the strengths of healthcare provider organizations (HPOs) with existing relationships with potential participants and the other opening enrollment directly to volunteers who are not part of a participating HPO.

NIH Precision Medicine Initiative

Precision Medicine Initiative® Cohort Program Biobank (U24)	RFA-PM-16-004
Precision Medicine Initiative® Cohort Program Coordinating Center (U2C)	RFA-PM-16-001
Precision Medicine Initiative® Cohort Program Healthcare Provider Organization Enrollment Centers (UG3/UH3)	RFA-PM-16-002
Precision Medicine Initiative® Cohort Program Participant Technologies Center (U24)	RFA-PM-16-003

The NIH Commons

not to be confused with the eRA Commons

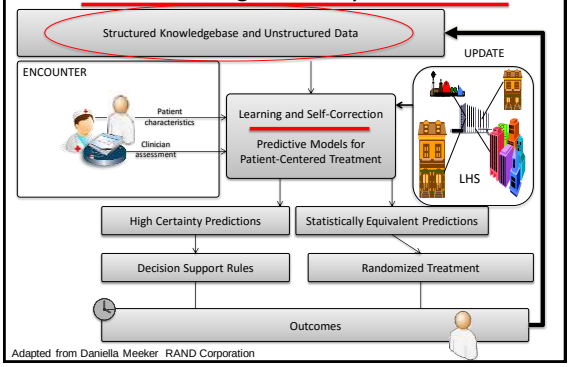
A **computing environment**, which support access, utilization and storage of digital objects.

Public **data** sets that adhere to *Commons* Digital Object Compliance principles.

Software services and tools that enable provisioning, indexing, sharing, and connectivity

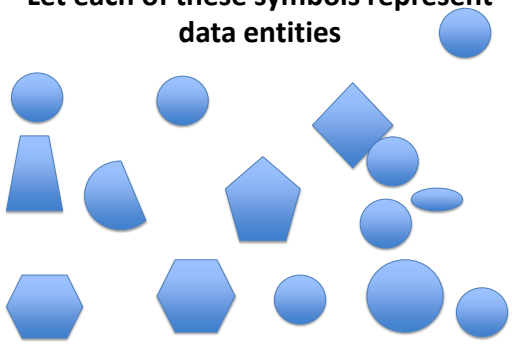
A set of **Digital Object Compliance** principles that describes the properties of digital objects that enables them to be findable, accessible, interoperable and reproducible (FAIR).

Big Data & Data Science enables a Learning Health System

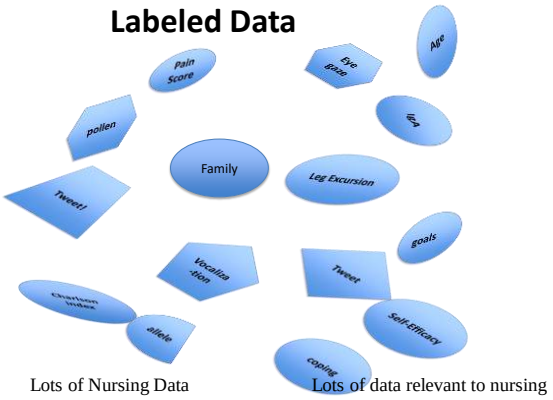


Adapted from Daniella Meeker RAND Corporation

Let each of these symbols represent data entities



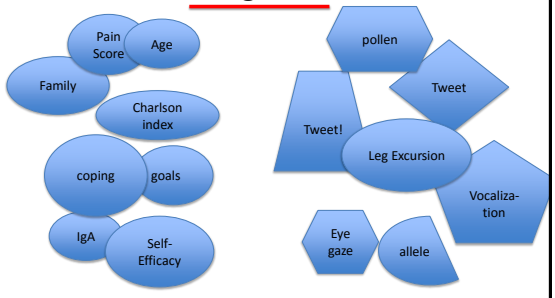
Labeled Data



Lots of Nursing Data

Lots of data relevant to nursing

Growing Nursing Knowledge out of Big Data



Lots of Nursing Data

Lots of data relevant to nursing

Classical Statistics Data Science	Approaches
• Concepts and variable defined **prior** to collecting data	• Concepts and variable defined prior to **using** data
• Data = *variables*	• Data *Streams*, not single variables
• Investigator controls the data collection	• Metadata & ingestion controls data availability
• Integrity built in *collection process*	• Integrity built in during *cleaning process*
• Data analysis strategies **predicated** on key assumptions (e.g. data is i.i.d)	• Analysis strategies rely on many models and are **robust** to most assumptions
• **Goal: Inferences**	• **Goal: Classification, insights, guidance**

Nursing, Big Data and the Patient Experience

What questions would *you* answer if you had immediate, complete access to all of the data that you wanted?

…so what is it that nurses' deal with?

- Signs
- Symptoms
- Observations of Daily Living (ODL)
- Images
- Biomarkers
- Family dynamics
- Patient experience
- Population phenotypes
- Biological specimens
- Every-day living spaces
- ….

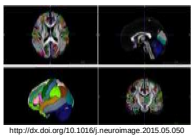
….but we don't have to do it all ourselves!

In fact we should not!

…Lessons from the Big Data Scientists!

Three case studies

- How do we scale an infrastructure? (PSCANNER & PCORnet)



http://dx.doi.org/10.1016/j.neuroimage.2015.05.050

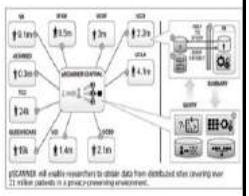
- Can we predict who will get Alzheimer's Disease
- Which newborn(s) status is changing?

http://doi.ieeecomputersociety.org/10.1109/MC.2013.157

PCORI Infrastructure: pSCANNER

AHRQ – R01 HS19913-01 (Ohno-Machado); Patient Centered Outcomes Research Institute - CDRN 1306-04819 (Ohno-Machado); NIA – 1RC4AG039115-01 (Docttor)

- Tighter partnerships with patients
  - Define preferences for research direction
  - Define preferences for data use
- Infrastructure for Patient Generated Data
- Standards-based analytic data management
- Adding privacy-preserving analytic options by leveraging "Big Data" distributed computing platforms

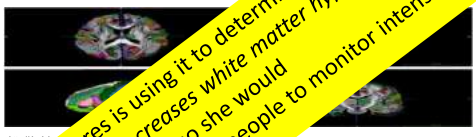


What is PCORnet?

AHRQ – R01 HS19913-01 (Ohno-Machado); Patient Centered Outcomes Research Institute - CDRN 1306-04819 (Ohno-Machado); NIA – 1RC4AG039115-01 (Docttor)

- Patient Centered Outcomes Research Institute network-of-networks
  - Objective to create a data infrastructure for patient-centered outcomes research
  - 18 "Patient Powered Research Networks"
  - 11 "Clinical Data Research Networks"
    - Health systems that have partnered to share **research governance practices**
    - Harmonize **data to a common standard**
    - **Software** to manage study and data transfer workflow

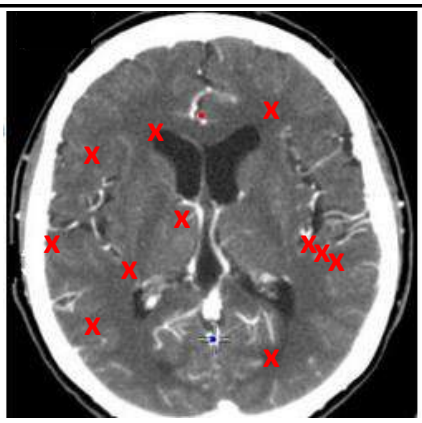
Can we predict who w… Alzheimer's Di…

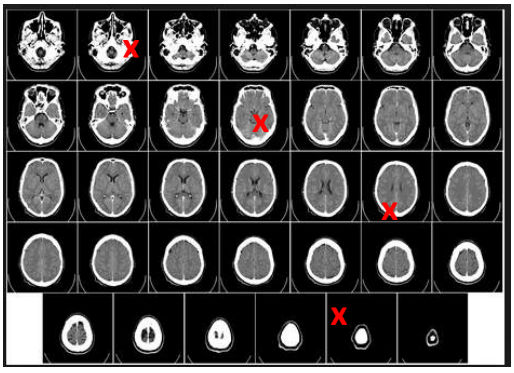


http://dx.doi.…

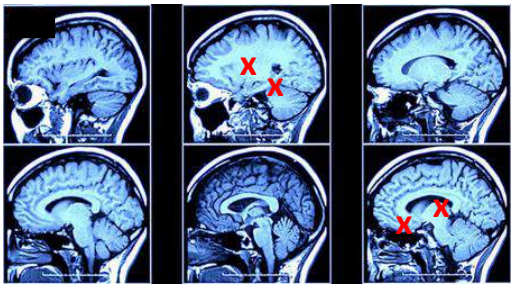
Elisa Torres is using it to determine whether *physical activity increases white matter hyperintensities-* so she would use it to target which people to monitor intensely

- …: Create a score based …y tangles inside the …oid plaques outside the
- …uld we do with this information?





Now where???



www.livescience.comCredit: Dreamstime

Which baby is going bad?

- ECG provides 1000 readings a second to construct a single waveform. That's 86.4 M reading a day to derive heart rate
- Wave form impedance, needed to assess breathing, adds another 5.4 M data points a day
- Data Science helps create the models for clinical use

http://doi.ieeecomputersociety.org/10.1109/MC.2011.157

Nursing's needs

- Information at the point of care
  - Visualization utilities
  - Analytics
  - Capture tools for efficient data acquisition
- Information at the point of knowledge generation
  - Data management and analytics
- Workforce development
  - Big data investigators
  - Researchers with the skill and insight to use big data
  - Clinicians with the skill and the ability to use the products of big data
- A seat at the table!

Beyond volume- Big Data & the Nursing Research Agenda

- Vision recognition
- Complex simulation
- Internet of things
- Public data sets
- Quantum computing – massive parallel processing of unstructured data
- i2B2 – judy warren's group

What can big data bring us?

Smart Organizations

Smart Practice

Smart Patients

Nursing Science participation in Data Science

- Concept identification and verification
- Population Engagement
- Phenotype explication
- Basic Analytics
- Complementary Analytics
- Interpretation
- Translation
- Evaluation

Three key issues

- Data science is a philosophy, a suite of methods and an approach
- Nursing doctoral programs include methods courses that are necessary but not sufficient to leverage data science for nursing
- There are three fundamental differences between a data science approach to knowledge generation and those we traditionally use in nursing:

Difference between Data Science & traditional Nursing Research

1. The nature of the questions
2. The manner & POINT IN THE PROCESS when variables are defined
3. How one knows when the work is done

Core learning content necessary for future nurse scientists and leaders in advanced nursing practice

Core Content for all Doctoral Programs

1. Data science approaches complement and augment survey, experimental, interpretive, epidemiological, and health services research strategies familiar to nurses
2. Data science is not just about ***more***… it is about a new paradigm of discovery when the question to be answered arose after the data appeared
3. It's more important to "work with", rather than "do the same as..."
4. Nurses help patients integrate data science into their lives

Integrating DATA SCIENCE into the DNP Curriculum

DNP and Data Science

- Practice-facing responsibilities
  - Data acquisition
  - Integrating structured (data = variables) and unstructured (data streams) at the point of care
  - Institutional transformation – leveraging data science for evidence-based practice
- Institutional leadership
  - Scanning for new approaches to practice
  - Change management
  - Clinical information system design & evaluation

Guiding Principles for Big Data in Nursing: Using Big Data to Improve the Quality of Care and Outcomes

Source: HIMSS CNO-CNIO Vendor Roundtable

Guiding Principles for Big Data in Nursing
http://www.himss.org/big10

Promote Standards and Interoperability

1. Nurses should promote the use of **standardized and accepted terminologies** that address the documentation needs of the entire care team regardless of care setting
2. Nurses should recommend consistent use of **research-based assessment scales and instrument**s that are standardized through an international consensus body.
3. **The ANA-recognized nursing terminologies** should be consistently updated and made available to international standards organizations for translation and complete, comprehensive mapping.
4. **Minimize use of free text** documentation.

Advance Quality eMeasures

5. eMeasures must ensure **the data to be collected and measured are aligned with the clinician's workflow**
6. To **advance nursing sensitive quality** eMeasures, paper measure sets must be evaluated
7. New quality must have **testing and piloting** that consider all stakeholders
8. Clinical quality eMeasures must **support evidence-based, cost effective care** that follows clinical practice guidelines and minimizes the negative impact on clinicians' workflow.

Leverage Nursing Informatics Experts

9. Nurse informaticists bring **concept representation, design, implementation and optimization of health IT** to health care
10. To achieve the desired outcomes**, nurse informaticists should have formal informatics training** education and certification.

Integrating DATA SCIENCE into the PhD in Nursing Program

Rules of the Road for Handling Big Data

Hilary Mason, bitly
OSEMN

- Obtain
- Scrub
- Explore
- Model
- iNterpret

…and big data will tell you what to do!

Smart Handling of Big Data

- Scheme on write vs schema on read
  - Traditional sql requires schema on write
  - Hadoop creates *schema on read* – no gatekeeping rules until the data are read
- Analytic workflow is distributed and allows for interim results
  - Converge on a consistent answer – allows for rapid engagement & handles consistently updating data sets

Wither post-docs, PhD's in other disciplines, and masters' prepared analysts?

Sample Exercise: Data Science

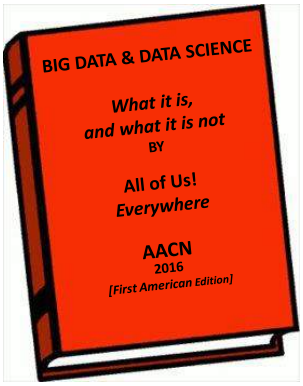
A student has been hired as a research intern at the Wisconsin State Department of Health. The Child & Family Services team wants to expand asthma services, but can only do so in one county. Should they expand services in Milwaukee County or Dane County?
What other data might help with this decision?

Data Science:
Charting an optimal
breast cancer screening protocol

- Consider the case of a 54 year old woman perimenopausal with 20 years use of BCP, 3 positive 2nd degree relations who has a sedentary lifestyle and 13 mammograms
- Three tools:
  - Data mining to find important relationships
  - Simulation for policy, capacity, population
  - Optimization: "smartly" choose among 1,000's of pathways

Nursing-inspired Algorithmic Research

- Algorithms are the sequence of steps used to implement models
- Developing models is largely the purview of mathematicians and statisticians
- Algorithms are like a bridge between the theory and the solution
- Nurses' expertise in understanding patient phenomena can drive new algorithms



DATA SCIENCE
~~NIGHTINGALE~~ PRINCIPLES FOR NURSES

* The sick person is an individual with individual needs.
*** BIG DATA is an art and a science.**
*** Nurses should spend their time caring for patients, not cleaning.**
* Prevention is better than cure.
*** The nurse must work as a member of a team.**
* The nurse must use discretion, but must follow the physician's orders.
* Self-discipline and self-evaluation are important.
* A good nursing program encourages individual development of the nurse.
* The nurse should be healthy in mind and body.
*** Teaching is a part of nursing.**
*** Nursing is a specialty.**
*** A nurse does not "graduate", but continues to learn throughout his or her career.**
*** The nursing curriculum should include both theoretical knowledge and practical experience WITH DATA SCIENCE CONCEPTS.**

DATA SCIENCE
What it is and What it is not

What it is…
- A philosophy
- An approach to knowledge generation
- A "stack" of tools to transform big data into knowledge
- A set of methods, data transport and storage protocols, visualization tools

What it is not….
- Application of known statistical models to well-defined data sets (even very large ones!)
- A solution to everything
- A replacement for all the methods we already know
- Only the purview of genomics researchers
- Irrelevant to nursing

Recognize & prepare for the key roles that academic leaders, faculty, & students themselves play here

Deans

Advocate at the campus level –
investments in data science
must extend to nursing
Set the standard for innovation
Source necessary resources

Faculty

Maintain Knowledge Currency
Establish Partnerships
Foster discovery among students
Integrate content and exercises

PhD students

Critically appraise the manner in which
phenomena are characterized
Maintain currency in analytics & approaches
Try out new lines of inquiry

DNP Students

Foster technology transfer –
help institutions and colleagues adopt
clinical solutions arising from data science
Guide patient engagement & protect patient rights
Devise nursing interventions
informed by data science