Big Data: Implications for Nursing Informatics

Bonnie L. Westra, PhD., RN, FAAN, FACMI
Associate Professor and Director,
Center for Nursing Informatics
April 29, 2016
Objectives

• Define big data as it relates to nursing
• Identify challenges in use of nursing data for big data science
• Explore examples of big data nursing research
• Identify strategies for nurse informaticians to share knowledge across settings to create nursing big data
Big Data Sources - Nursing

- Volume
- Velocity
- Variety
- Veracity
- Value

https://infocus.emc.com/william_schmarzo/thoughts-on-the-strata-rx-healthcare-conference/
Data Sources

• CTSA – [https://ctsacentral.org/](https://ctsacentral.org/)
  • NCATS - [https://ncats.nih.gov/](https://ncats.nih.gov/)
• PCORnet - [http://www.pcornet.org/](http://www.pcornet.org/)
  • 13 clinical data research networks (CDRNs)
  • 22 patient powered research networks (PPRNs)
• Optum Labs – 140 million lives from claims data + 40 million from EHRs ([delaney@umn.edu](mailto:delaney@umn.edu))
Big Data & Big Data Science

• Application of math to large data sets to infer probabilities for associations/prediction
• Purpose is to accelerate discovery, improve critical decision-making processes, enable a data-driven economy¹
• Three-legged stool
  • Data
  • Technology
  • Algorithms
NSF Announces Interagency Progress on Administration's Big Data Initiative

Harnessing the EHR for Research

- in areas of eScience such as
  - [data capture],
  - Databases,
  - Workflow management,
  - Visualization
  - Computing technologies.

Nursing Research Journal¹

http://www.sciencemag.org/site/special/data/ScienceData-hi.pdf
Big Data Analytics for Nursing

Protect
Promote

Health

Prevent

Alleviate
Suffering

Information

Data

Knowledge

Wisdom

© Westra, 2014
Nursing Informatics & Translational Science
Requirements for Useful Data

• Common data models
• Standardized coding of data
• Standardize queries
PCORnet CDM Domains, v3.0

CONDITION v2.0
A condition represents a patient’s diagnosed and self-reported health conditions and diseases. The patient’s medical history and current state may both be represented.

DEATH v3.0
Reported mortality information for patients.

DEATH_CAUSE v3.0
The individual causes associated with a reported death.

DEMOGRAPHIC v1.0
Demographics record the direct attributes of individual patients.

DIAGNOSIS v1.0
Diagnosis codes indicate the results of diagnostic processes and medical coding within healthcare delivery.

DISPENSING v2.0
Outpatient pharmacy dispensing, such as prescriptions filled through a neighborhood pharmacy with a claim paid by an insurer. Outpatient dispensing is not commonly captured within healthcare systems.

ENROLLMENT v1.0
Enrollment is a concept that defines a period of time during which all medically-attended events are expected to be observed. This concept is often insurance-based, but other methods of defining enrollment are possible.

ENCOUNTER v1.0
Encounters are interactions between patients and providers within the context of healthcare delivery.

HARVEST v3.0
Attributes associated with the specific PCORnet datamart implementation.

LAB_RESULT_CM v2.0
Laboratory result Common Measures (CM) use specific types of quantitative and qualitative measurements from blood and other body specimens. These standardized measures are defined in the same way across all PCORnet networks.

PCORENET_TRIAL v3.0
Patients who are enrolled in PCORnet clinical trials.

PRESCRIBING v3.0
Provider orders for medication dispensing and/or administration.

PRO_CM v2.0
Patient-Reported Outcome (PRO) Common Measures (CM) are standardized measures that are defined in the same way across all PCORnet networks. Each measure is recorded at the individual item level: an individual question/statement, paired with its standardized response options.

PROCEDURES v1.0
Procedure codes indicate the discreet medical interventions and diagnostic testing, such as surgical procedures, administered within healthcare delivery.

VITAL v1.0
Vital signs (such as height, weight, and blood pressure) directly measure an individual’s current state of attributes.

Data Standardization

- Demographics – OMB
- Medications - RxNorm
- Laboratory data - LOINC
- Procedures – CPT, HCPCS, ICD, SNOMED CT
- Diagnoses - ICD-9/10-CM, SNOMED CT
- Vital status – CDC
- Vital signs - LOINC
Vision – Inclusion of Nursing Data

Clinical Data
NMDS

Management Data
NMMDS

Other Data Sets

Continuum of Care
Challenges
Challenges - Standards
ANA Position Statement – Inclusion of Recognized Terminologies Supporting Nursing Practice within Electronic Health Records and Other Health Information Technology Solutions

http://z.umn.edu/bigdata
Challenges - Architecture
Challenges - Reinventing
UMN Clinical Data Repository

Cohort discovery /recruitment
Observational studies
Predictive Analytics

Data available to UMN researchers via the Academic Health Center Information Exchange (AHC-IE)
2+ million patients
MHealth / Fairview Health Services

AHC-IE - acute & ambulatory clinical data
2+ million patients
4+ billion total rows of unique data

* The number of patients and records changes daily
# Example Flowsheet

## Adult Assessment

### General Information
- Immunizations
- Advanced Directives

### Pain
- **Type Pain**
  - Acute pain, Chronic pain, Deep somatic pain, Intractable pain, Neuropathic pain, Other (Comment), Phantom pain, Referred pain, Superficial somatic pain, Surgical pain, Visceral pain

- **Preferred Pain Scale**
  - FACES, FLACC, PAINAD, non-verbal, numerical 0-10

- **Pain rating 0-10**
  - Number 0 - 10

- **Current Pain Description**
  - None, Mild (1-3), Moderate (4-6), Severe (7-10)

- **Pain Descriptors**
  - Aching; Burning; Constant; Cramping; Crushing; Discomfort; Dull; Headache; Heaviness; Itching; Jabbing; Nagging; Numbness; Other (comment); Patient unable to describe; Penetrating; Pins and Needles; Pounding; Pressure; Radiating; Sharp; Shooting; Sore; Spasm; Squeezing; Stabbing; Tender; Throbbing; Tightness; Tingling; Tiring

### Muscle/Skeletal
- Skin
- Cardiac
- Neuro
- Functional Status

### Screens/Templates

### Value Sets/Answers

### Groups of questions
Data Source
Clinical Data Models - Flowsheets

- 10/20/2010 - 12/27/2013
- 66,660 patients
- 199,665 encounters

Data Points
153,049,704

Flowsheet Measures
14,550

Groups
2,696

T
562
UMN CTSI - Extend CDM

Team: Nursing (DNP/PhD), Computer Science, Health Informatics
Making Data Useful

200K Patient Encounter

Pain Information Model
309 observations
80 Unique concepts – assessments, goals, interventions (not including value sets – choice lists)
## Flowsheet Information Models

<table>
<thead>
<tr>
<th>System</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular System</td>
<td>Pain</td>
</tr>
<tr>
<td><strong>Falls/ Safety</strong></td>
<td>Peripheral Neurovascular (VTE)</td>
</tr>
<tr>
<td>Gastrointestinal System</td>
<td>Pressure Ulcers</td>
</tr>
<tr>
<td>Genitourinary System/ CAUTI</td>
<td>Respiratory system</td>
</tr>
<tr>
<td>Neuromusculoskeletal System</td>
<td>Vital Signs, Height &amp; Weight</td>
</tr>
<tr>
<td>Information Model Name</td>
<td>Number Flowsheet IDs Mapped to Observables</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular System</td>
<td>241</td>
</tr>
<tr>
<td>Falls</td>
<td>59*</td>
</tr>
<tr>
<td>Gastrointestinal System</td>
<td>60</td>
</tr>
<tr>
<td>GI/ CAUTI</td>
<td>79</td>
</tr>
<tr>
<td>Musculoskeletal System</td>
<td>276</td>
</tr>
<tr>
<td>Pain</td>
<td>309</td>
</tr>
<tr>
<td>Pressure Ulcers</td>
<td>104</td>
</tr>
<tr>
<td>Respiratory System</td>
<td>272</td>
</tr>
<tr>
<td>VTE</td>
<td>67</td>
</tr>
</tbody>
</table>
Nursing Big Data Research
Nursing Research

• Severe Sepsis Compliance guidelines and impact on patient complications and mortality
• Unanticipated ICU admissions for elective surgery patients
• Patient and nurse staffing factors associate with CAUTI
• Factors associated with urinary and bowel Incontinence improvement
• Predicting hospitalization for frail elders
• Demonstrate value of Wound, Ostomy, Continence Nursing for improving wounds and incontinence
• Improvement in managing oral medications
Home Care EHR De-Identified Data

Initial Data Set
808 agencies, 1,560,508 OASIS records, 888,243 patients
List of patients with and without WOC Nurse

<table>
<thead>
<tr>
<th>Reason for Removing Records</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incomplete episode records</td>
<td>464,485</td>
</tr>
<tr>
<td>Assessment outside study dates</td>
<td>125,886</td>
</tr>
<tr>
<td>Incorrect type of assessment</td>
<td>51,779</td>
</tr>
<tr>
<td>Masked or missing data</td>
<td>16,302</td>
</tr>
<tr>
<td>Duplicate records</td>
<td>2,748</td>
</tr>
<tr>
<td>Age &lt; 18 or primary dx related to pregnancy/ complications</td>
<td>822</td>
</tr>
</tbody>
</table>

Final Data Set
785 agencies, 447,309 patients, 449,243 episodes of care, 0.6% re-admissions
## Certified WOC Nurses Influence on Incontinence & Wounds

<table>
<thead>
<tr>
<th>Outcome Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure Ulcers</td>
<td>Total number of pressure ulcers (M0450 a-e)</td>
</tr>
<tr>
<td>Stasis Ulcers</td>
<td>Total number stasis ulcers (M0470/M0474)</td>
</tr>
<tr>
<td>Surgical Wounds</td>
<td>Total number of surgical wound (M0484/M0486)</td>
</tr>
<tr>
<td>Urinary Incontinence</td>
<td>Presence/management of urinary incontinence or need for a catheter (M0520)</td>
</tr>
<tr>
<td>Urinary Tract Infection</td>
<td>Treated for UTI in past 14 days (M0510)</td>
</tr>
<tr>
<td>Bowel Incontinence</td>
<td>Frequency of bowel incontinence (M0540)</td>
</tr>
</tbody>
</table>
# Improved/ Not Worse (Stabilize) Outcomes

<table>
<thead>
<tr>
<th>Score</th>
<th>Bowel Incontinence Frequency</th>
<th>Improved</th>
<th>Not Worse (Stabilize)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Very rarely /never has BI or has ostomy for bowel elimination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Less than once weekly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>One to three times weekly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Four to six times weekly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>On a daily basis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>More often than once daily</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Aim 1: Prevalence

Prevalence of Condition by Agency

Pressure Ulcer (PU), Stasis Ulcer (SU), Surgical Wound (SW), Urinary Incontinence (UI), Bowel Incontinence (BI), Urinary Tract Infection (UTI)
Aim 2: Incidence

Incidence of Conditions by Agency

No WOC (n = 13,261)  WOC (n = 281,552)
Effect of WOC Nurses on Agency Outcomes

<table>
<thead>
<tr>
<th>Outcome Concept</th>
<th>Improvement OR</th>
<th>95% CI</th>
<th>Stabilization OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure ulcers</td>
<td>1.9</td>
<td>1.8-2.0</td>
<td>1.29</td>
<td>1.21-1.37</td>
</tr>
<tr>
<td>Urinary incontinence</td>
<td>1.4</td>
<td>1.38-1.43</td>
<td>2.3</td>
<td>2.26-2.4</td>
</tr>
<tr>
<td>Urinary tract infections</td>
<td>1.4</td>
<td>1.38-1.43</td>
<td>1.2</td>
<td>1.16-1.27</td>
</tr>
<tr>
<td>Surgical wounds</td>
<td>1.39</td>
<td>1.36-1.42</td>
<td>1.5</td>
<td>1.46-1.57</td>
</tr>
<tr>
<td>Stasis ulcers</td>
<td>1.2</td>
<td>1.1-1.3</td>
<td>Unable to model</td>
<td></td>
</tr>
<tr>
<td>Bowel incontinence</td>
<td>1.14</td>
<td>1.11-1.2</td>
<td>1.16</td>
<td>1.23-1.9</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; OR, odds ratio.

*a ORs weighted by the propensity score for having a WOC nurse.

*b Unable to model due to more than 99% stabilization across all subjects.
Individual Patient Outcomes

**Pressure Ulcer Improvement**
- With WOC nurse: Admission 2.2, Discharge 1.3
- Without WOC nurse: Admission 1.8, Discharge 0.92

**Pressure Ulcer Stabilization**
- With WOC nurse: Admission 0.54, Discharge 0.37
- Without WOC nurse: Admission 0.18, Discharge 0.11

**Surgical Wound Improvement**
- With WOC nurse: Admission 2.1, Discharge 0.8
- Without WOC nurse: Admission 2, Discharge 0.8

**Surgical Wound Stabilization**
- With WOC nurse: Admission 1.1, Discharge 0.7
- Without WOC nurse: Admission 1.08, Discharge 0.76
Individual Patient Outcomes

- **Urinary Incontinence Improvement**
  - Admission: 1.1
  - Discharge: 0.7
- **Urinary Incontinence Stabilization**
  - Admission: 0.4
  - Discharge: 0.20

- **Bowel Incontinence Improvement**
  - Admission: 3.3
  - Discharge: 2.4
- **Bowel Incontinence Stabilization**
  - Admission: 0.99
  - Discharge: 0.77

- **Improvement in Type of UTI Intervention**
  - Admission: 2
  - Discharge: 0.3
- **Stabilization of UTI Intervention**
  - Admission: 0.01
  - Discharge: 0.05
Lessons Learned

- Obtaining data
- Tracking WOC nurse patient visits
- Data quality
  - Matching patients start and discharge
  - Duplicate patient records
  - Encrypted data
  - Missing data
- Selecting variables - theory and domain expertise
- Type of analysis - Research question, structure of the data
Mobility Outcomes

• Discover patients and support system characteristics associated with the mobility outcomes

• Find new factors associated with mobility besides current ambulation status during admission (OR = 5.96)

• In each subgroup of patients defined by current ambulation status during admission (1-5)

• To compare the predictors across each patient subgroup to find the consistent biomarkers in all subgroups and specific factors in each subgroup
# Mobility Outcome

## TABLE 1. Mobility Scores

<table>
<thead>
<tr>
<th>Score</th>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>INDP</td>
<td>Able to independently walk on even and uneven surfaces and climb stairs with or without railings (i.e., needs no human assistance or assistive device)</td>
</tr>
<tr>
<td>1</td>
<td>DEVICE</td>
<td>Requires use of a device (e.g., cane, walker) to walk alone or requires human supervision or assistance to negotiate stairs or steps or uneven surfaces</td>
</tr>
<tr>
<td>2</td>
<td>SUPERV</td>
<td>Able to walk only with the supervision or assistance of another person at all times</td>
</tr>
<tr>
<td>3</td>
<td>CHAIR_I</td>
<td>Chairfast, unable to ambulate but is able to wheel self independently</td>
</tr>
<tr>
<td>4</td>
<td>CHAIR_NI</td>
<td>Chairfast, unable to ambulate and [not independent] to wheel self</td>
</tr>
<tr>
<td>5</td>
<td>BED</td>
<td>Bedfast, unable to ambulate or be up in a chair</td>
</tr>
</tbody>
</table>

*Note. Scores are based on Outcome and Assessment Information Set question MO700 Ambulation/Locomotion.*
Comparison of Outcomes by Group

<table>
<thead>
<tr>
<th>Score</th>
<th>Label</th>
<th>Total (N = 261,035)</th>
<th>No improvementb (n = 128,920)</th>
<th>Improvementc (n = 132,115)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INDP</td>
<td>144,615 (55.4)</td>
<td>99,119 (68.5)</td>
<td>45,496 (31.5)</td>
</tr>
<tr>
<td>2</td>
<td>DEVICE</td>
<td>89,860 (34.4)</td>
<td>18,129 (20.2)</td>
<td>71,731 (79.8)</td>
</tr>
<tr>
<td>3</td>
<td>SUPERV</td>
<td>12,669 (4.9)</td>
<td>5,322 (42.0)</td>
<td>7,347 (58.0)</td>
</tr>
<tr>
<td>4</td>
<td>CHAIR_I</td>
<td>11,339 (4.3)</td>
<td>5,163 (45.5)</td>
<td>6,176 (54.5)</td>
</tr>
<tr>
<td>5</td>
<td>CHAIR_NI</td>
<td>2,552 (1.0)</td>
<td>1,187 (46.5)</td>
<td>1,365 (53.5)</td>
</tr>
<tr>
<td>All</td>
<td></td>
<td>261,035 (100.0)</td>
<td>128,920 (49.4)</td>
<td>132,115 (50.6)</td>
</tr>
</tbody>
</table>

---

*a* Scores are based on Outcome and Assessment Information Set question M0700 Ambulation/Locomotion. *b* Mobility outcome = 0. *c* Mobility outcome = 1.
OASIS data extracted from EHRs from 270,634 patients served by 581 Medicare-certified home health agencies. Standardizing data, identifying patients, imputing missing values, binarizing data into 98 variables.

Data Mining Techniques

- Identify risk variables significantly associated with mobility outcomes - varied among the groups
- Group the single predictors based on whether they cover same or different patient group
  - Clustering
    - Based on similarity of patients
    - Not discriminative
    - High frequency variables got merged
  - Pattern mining based approach
    - Discriminative
    - Coherence (similarity of patients)
Subgroup Variability
Clustering Groups

Similarity between each pair of 99 variables for SUPERV group

- Help for finance and healthcare
- Problem with toilet and transferring
- No ADL problems
- Healthy Physiological Elderly
- Functional Deficiency
- No Problem with Dressing upper body and grooming
- Behavioral and Psychological problems
- Paid Help

Variables
Improvement Group 2

Older adults with no problems in daily activities

Healthier physiological and psychosocial elderly

Household Management
No Improvement Group 2

- Incapable to toilet and transfer
- Paid Help
- Frail patients with functional deficiency
- Help with financial and legal matters
- Cognitive deficits and behavioral problems
Lessons Learned

- Transform data into binary variables
- Selection of variables – remove if
  - Too little variation or high inter-correlations of predictors
- Medical diagnoses used to describe patients, not predict
- Analysis by subgroup
- Interpretation of results is critical – requires domain experts
- Different clusters point to the need to tailor interventions for subgroups
- **Lack of standardized interventions** precluded understand how care provided effects outcomes
Sharing Experiences
Nursing Knowledge: 2015 Big Data Science Conference

School of Nursing

Nursing Knowledge: 2015 Big Data Science Conference

Demonstrating the Value of Sharable and Comparable Nurse-generated Data

z.umn.edu/bigdata
Vision

• Better health outcomes from the standardization and integration of the information nurses gather in electronic health records
  • EHR increasingly the source of insights and evidence
  • Used to prevent, diagnose, treat and evaluate health conditions.
• Other IS - Rich contextual data about patients (including environmental, geographical, behavioral, imaging data, and more)
• Lead to breakthroughs for the health of individuals, families, communities and populations.
Create a National Action Plan

- Implement nursing information in EHRs and other information systems using standardized language
  - Streamlined, essential, evidence-based, actionable, and demonstrates value of nursing’s contribution to health
- Standardize nursing informatics education to build an understanding and competences
- Influence policy and standards for documenting and coding nursing information in health care knowledge systems
- Use standardized nursing data with other data sources for business analytics and research
Conferences*

• Working conferences with virtual workgroups taking action between annual meetings
• All focused on the same vision
• Strategic inclusion of stakeholders
  • Practice - leaders
  • Industry, particularly software vendors
  • Professional organizations
    • National – nursing, interprofessional, informatics
  • Academia

*Proceedings: http://z.umn.edu/nbd2k
2014 – 2015 Accomplishments

• **Education**
  - Surveyed accreditation, certification and credentialing programs influencing informatics
  - Faculty resources for teaching informatics available: [http://www.nursing.umn.edu/continuing-professional-development/nnideepdive/](http://www.nursing.umn.edu/continuing-professional-development/nnideepdive/)

• **Science of NBD2K**
  - Completed NMMDS updates/ LOINC coding for public distribution
    - [http://z.umn.edu/nmmds](http://z.umn.edu/nmmds)
  - Started NINR Nursing Informatics SIG
  - Created “Big Data Checklist for Chief Nurse Executives”

• **Quality Measures**
  - Forwarded eMeasure for Pressure Ulcers
2014 – 2015 Accomplishments

• Health IT Policy
  • Guiding Principles for Big Data in Nursing: Using Big Data to Improve the Quality of Care and Outcomes. [http://www.himss.org/big10](http://www.himss.org/big10)
  • ANA Position Statement: Inclusion of Recognized Terminologies Supporting Nursing Practice within Electronic Health Records and Other Health Information Technology Solutions.
  • ANA/ANI/AAN Informatics expert panel collaborated on appointments and comments on policies

• Harmonization/Standardization of Nursing Data/Models
  • Focus on care coordination & standardization
  • Integrate PNDS into data and model standards
2014 – 2015 Accomplishments

• Value of Nursing
  • Created data model to demonstrate value of nursing at individual nurse level

• LOINC/ SNOMED CT Minimum Assessment
  • Developed minimum physiological assessment encoded with LOINC/ SNOMED CT

• Workforce Data
  • Develop dissemination plan for Implementation Guide for NMMDS

• Transform Nursing Documentation
  • Develop a set of recommendations for leveraging EHRs and clinical intelligence tools to promote evidence based, personalized care across the continuum
New Workgroups

• Social/Behavioral Determinants of Health
  • Develop a toolkit to support inclusion of this data into electronic health records, including expected CMS Meaningful Use program requirements

• Nursing & Care Coordination
  • Identify nursing implications related to “big data” associated with “care coordination.”

• Connect Nursing Informatics Leaders
  • Provide a platform for emerging and expert informatics nurses to discuss opportunities to enhance nursing knowledge

• mHealth Data
  • Explore the use of mobile health data by nurses, including nursing- and patient-generated data, and incorporating mHealth data use into workflows
You Are Invited to Get Involved

• Working groups –
  • Contact Lisiane Pruinelli pruin001@umn.edu

• Nursing Knowledge: 2016 Big Data Science Conference
  • June 1-3, 2016
  • Minneapolis, Minnesota
  • Registration open!
    Early bird discount through April 1, 2016
  • http://z.umn.edu/bigdata
Summary

• Big data is increasing
• Existing and newer methods for data analysis
• Big data science useful to address practice questions
• Lessons learned
  • Data quality – originates in practice
  • Standardized data and common data / information models needed for usable data
• “Takes a village” – combined expertise important
Acknowledgments

• Multiple funding for studies:
  • Grant Number 1UL1RR033183 from the National Center for Research Resources (NCRR) of the National Institutes of Health (NIH) to the University of Minnesota Clinical and Translational Science Institute (CTSI).
  • Two NSF grants: NSF IIS-1344135 and NSF IIS-0916439 as well as a University of Minnesota Doctoral Dissertation Fellowship.
  • Wound, Ostomy, Continence Nursing Association
Questions?
Bonnie L. Westra, PhD, RN, FAAN, FACMI
westr006@umn.edu
References


