Big Data: Implications for Nursing Informatics

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

University of Minnesota
School of Nursing
Driven to Discover℠

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

Geocodes

Acuity
SCHEDULING

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

Data Sources

- CTSA – https://ctsacentral.org/
  - NCATS - https://ncats.nih.gov/
- PCORnet - 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)
- http://www.data.gov/ - Search over 192,872 datasets

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 economy[1]
- Three-legged stool
  - **Data**
  - Technology
  - **Algorithms**

NIH

NIH Big Data to Knowledge (BD2K)

CTSA Clinical & Translational Science Awards®

pcori®

NSF Announces Interagency Progress on Administration's Big Data Initiative

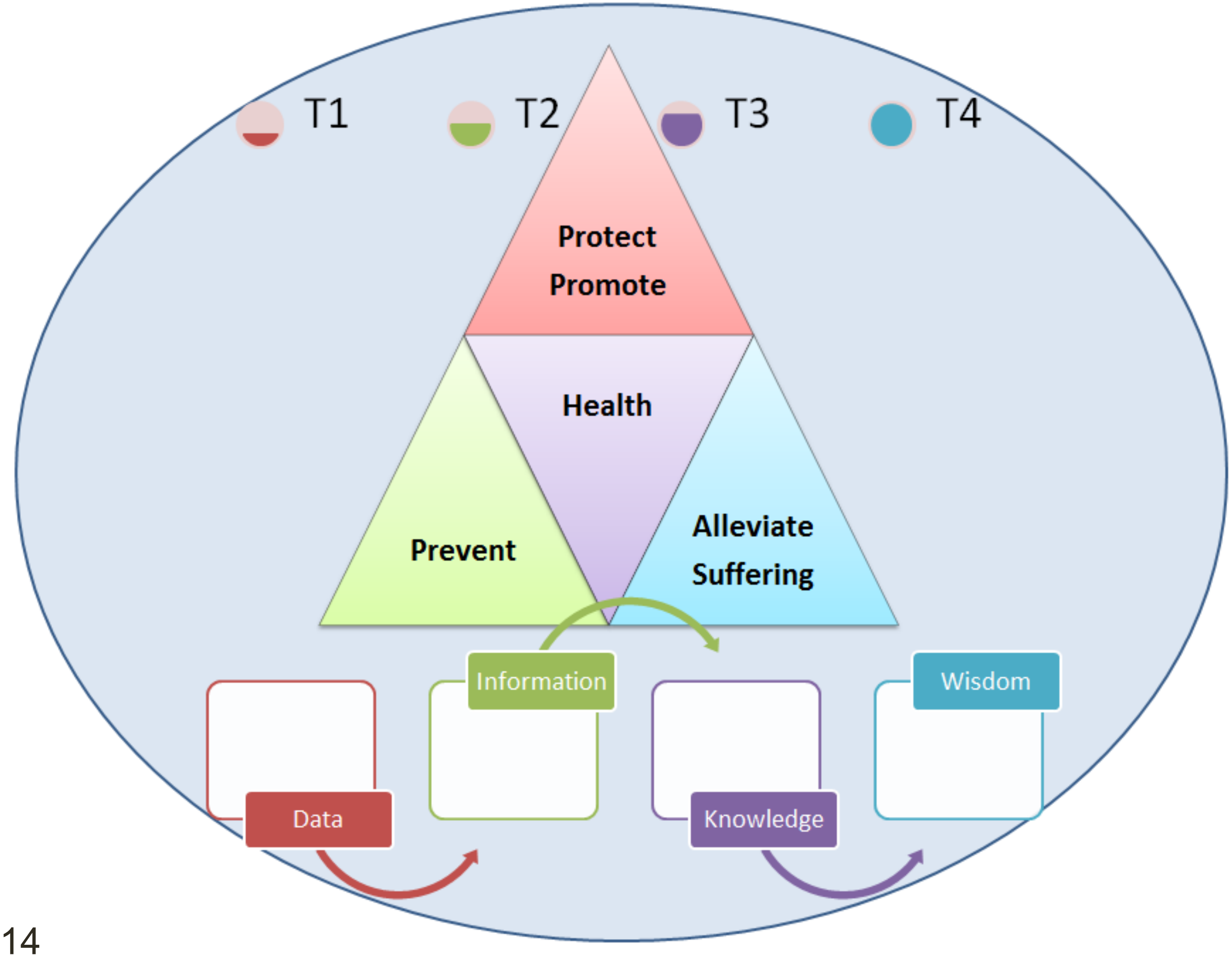
Harnessing the EHR for Research

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

Nursing Research Journal!

http://www.sciencemag.org/site/special/data/ScienceData-hi.pdf

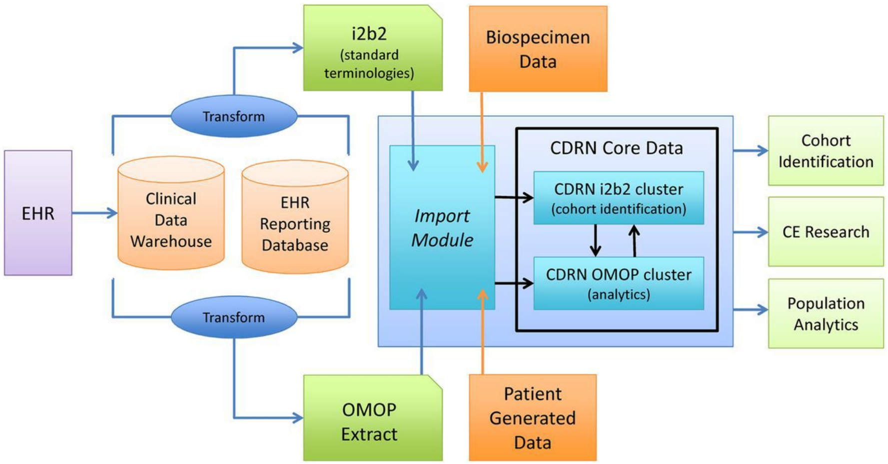
Big Data Analytics for Nursing



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.

PCORNET_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

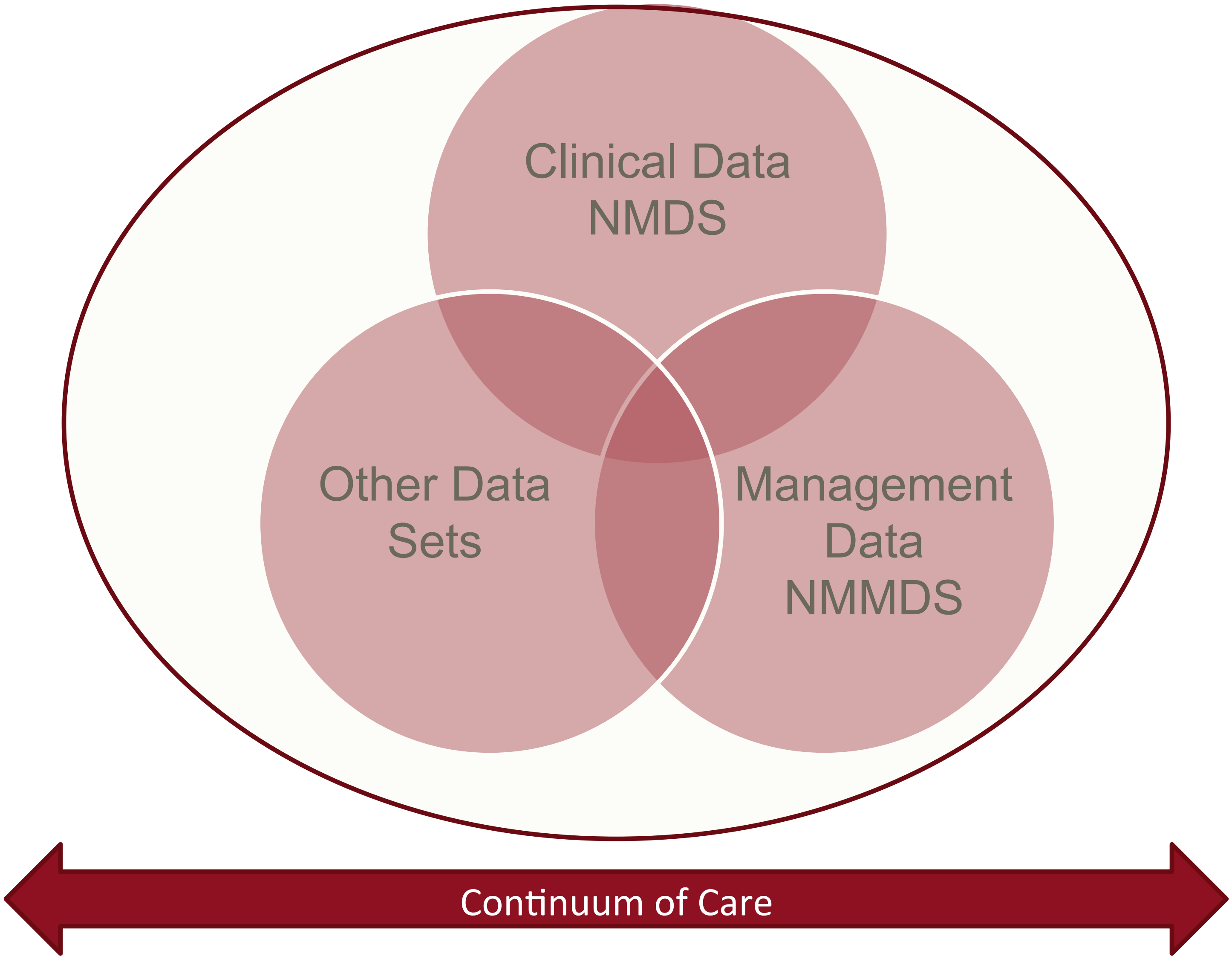
Vital signs (such as height, weight, and blood pressure) directly measure an individual's current state of attributes.

http://www.pcornet.org/resource-center/pcornet-common-data-model/

Data Standardization

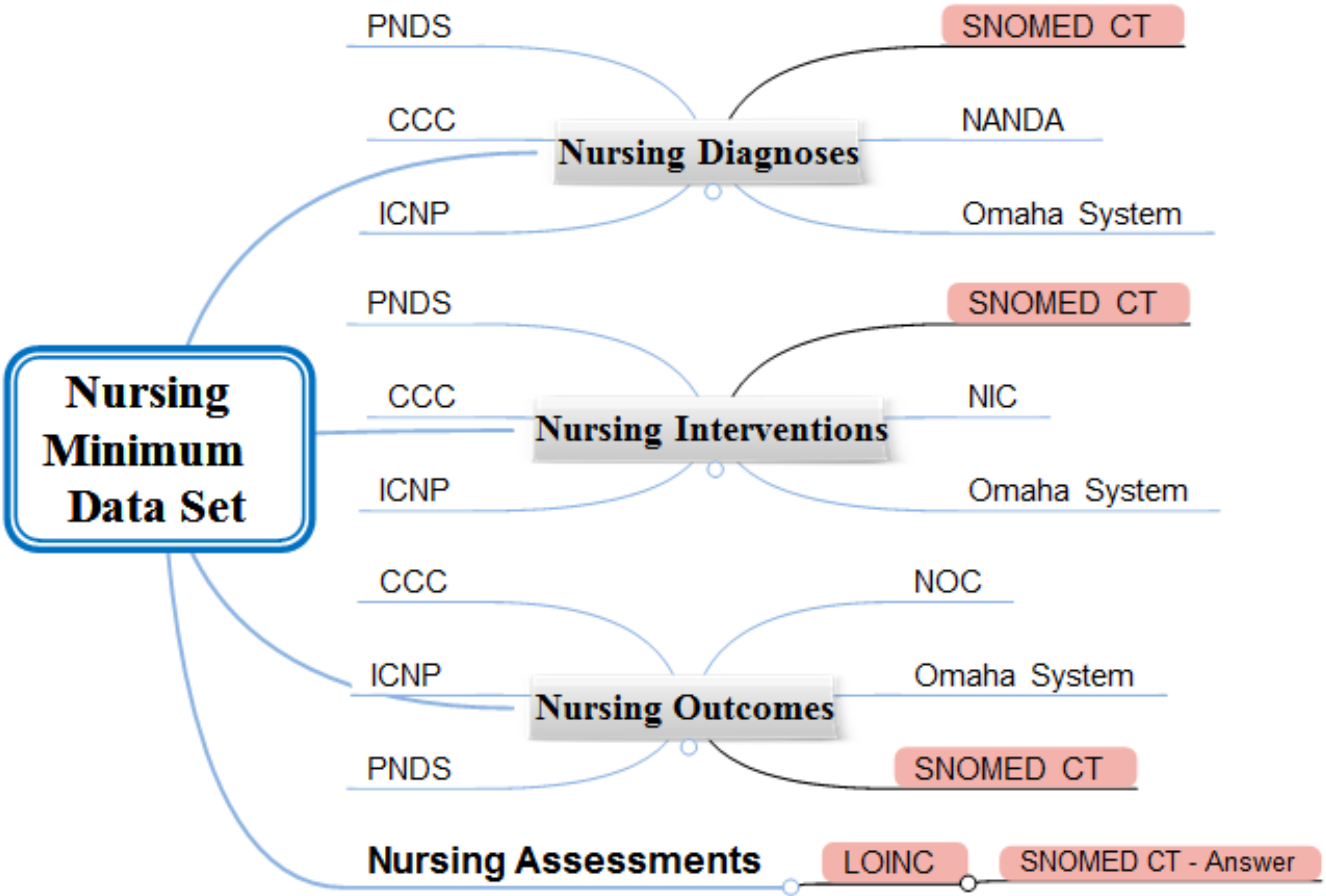
- Demographics – OMB
- Medications - RxNorm
- Laboratory data - LOINC
- Procedures – CPT, HCPCS, ICD, SNOMED CT
- Diagnoses - ICD-9/10-CM, SNOMED CT
- Vital status – CDC
- Vital signs - LOINC

Vision – Inclusion of Nursing Data



Challenges

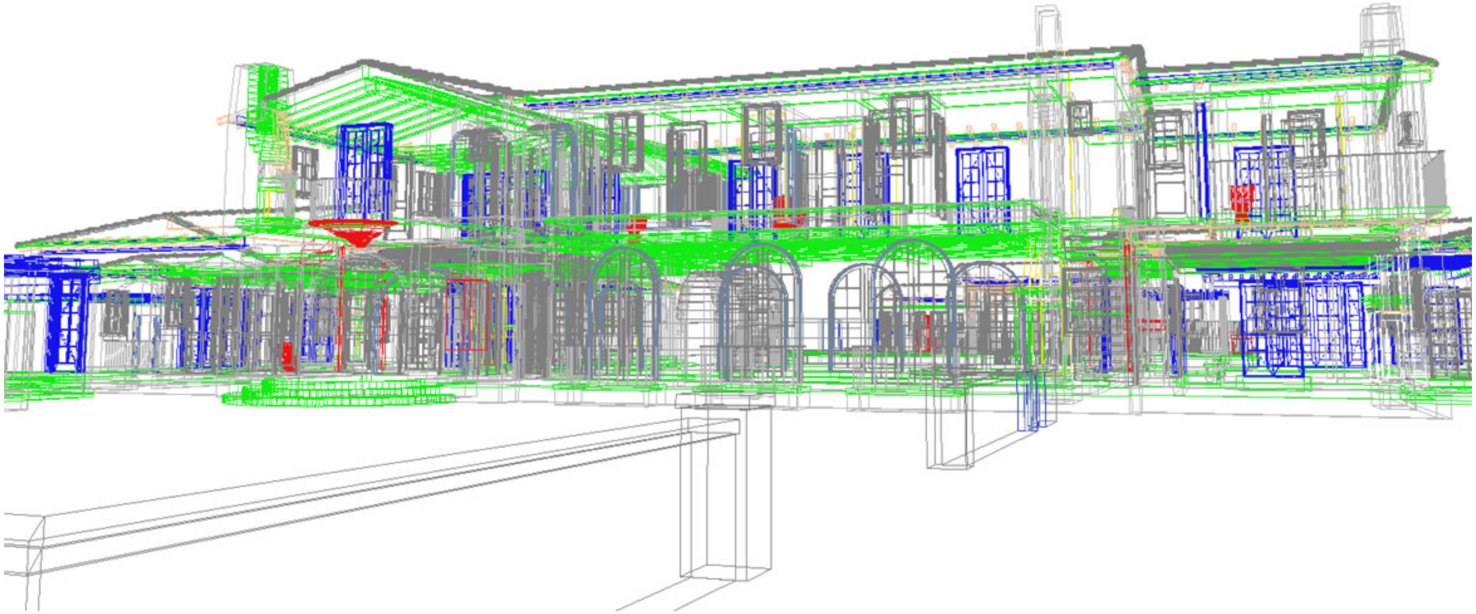
Challenges - Standards



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



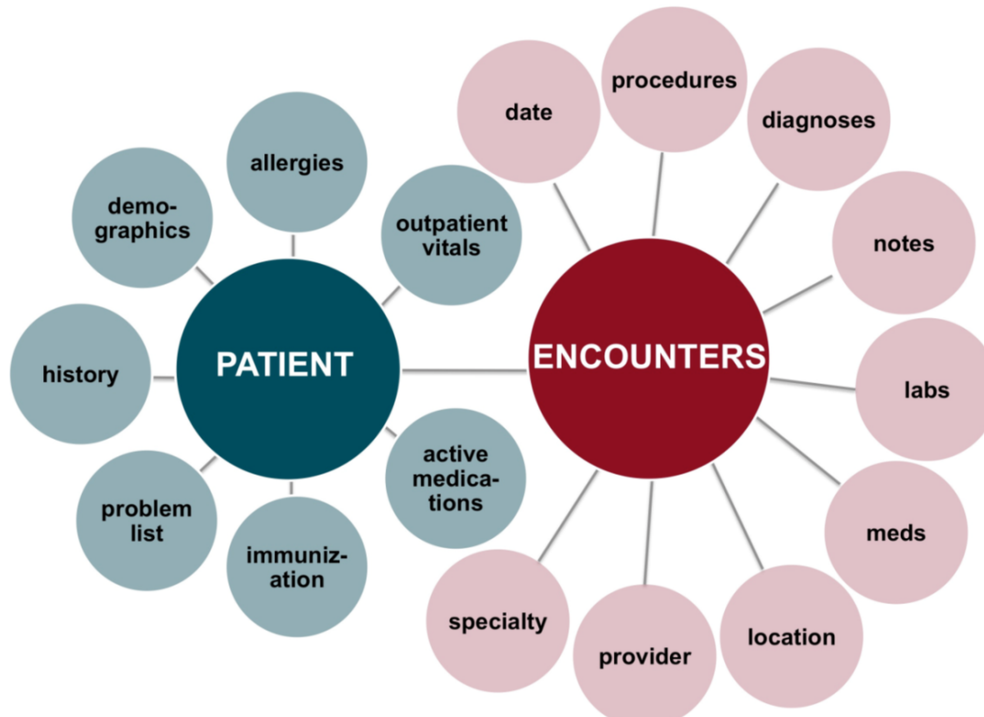
Challenges - Reinventing

UMN Clinical Data Repository

Cohort discovery /recruitment

Observational studies

Predictive Analytics

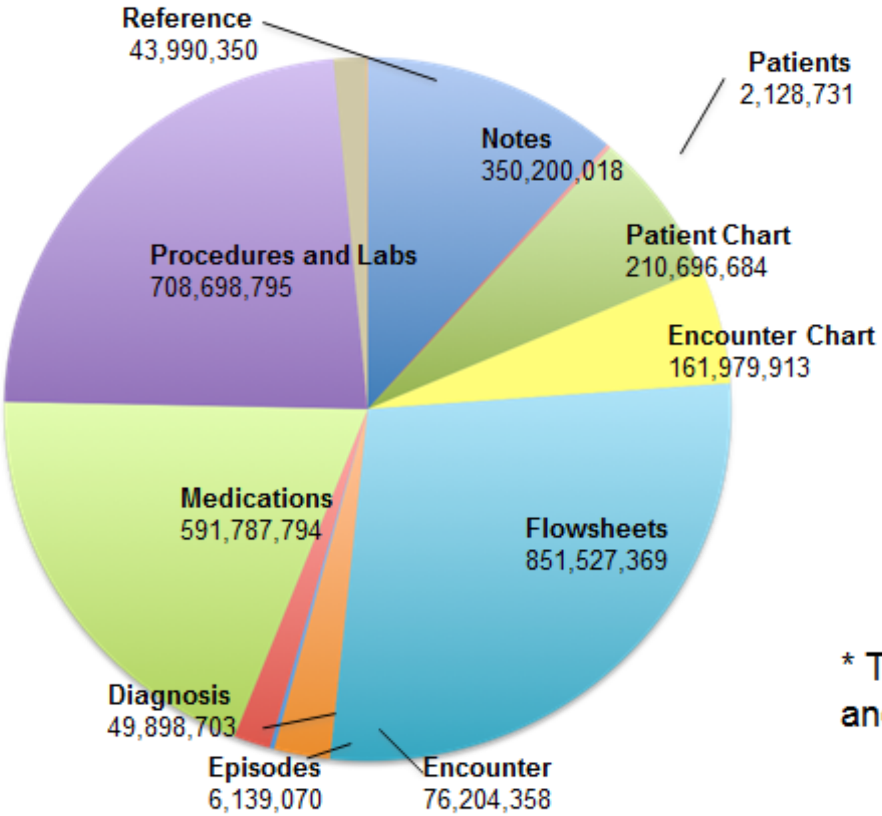


8 Fairview hospitals (data from 2011 and on)
40+ Fairview (from 2005) and UMP clinics (from 2011)

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 or patients and records changes daily

Example Flowsheet

Adult Assessment			*Vital Signs*	*I & O*	*IV Assessr*
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			
Musculoskeletal					
Skin					
Cardiac					
Neuro					
Functional Status					

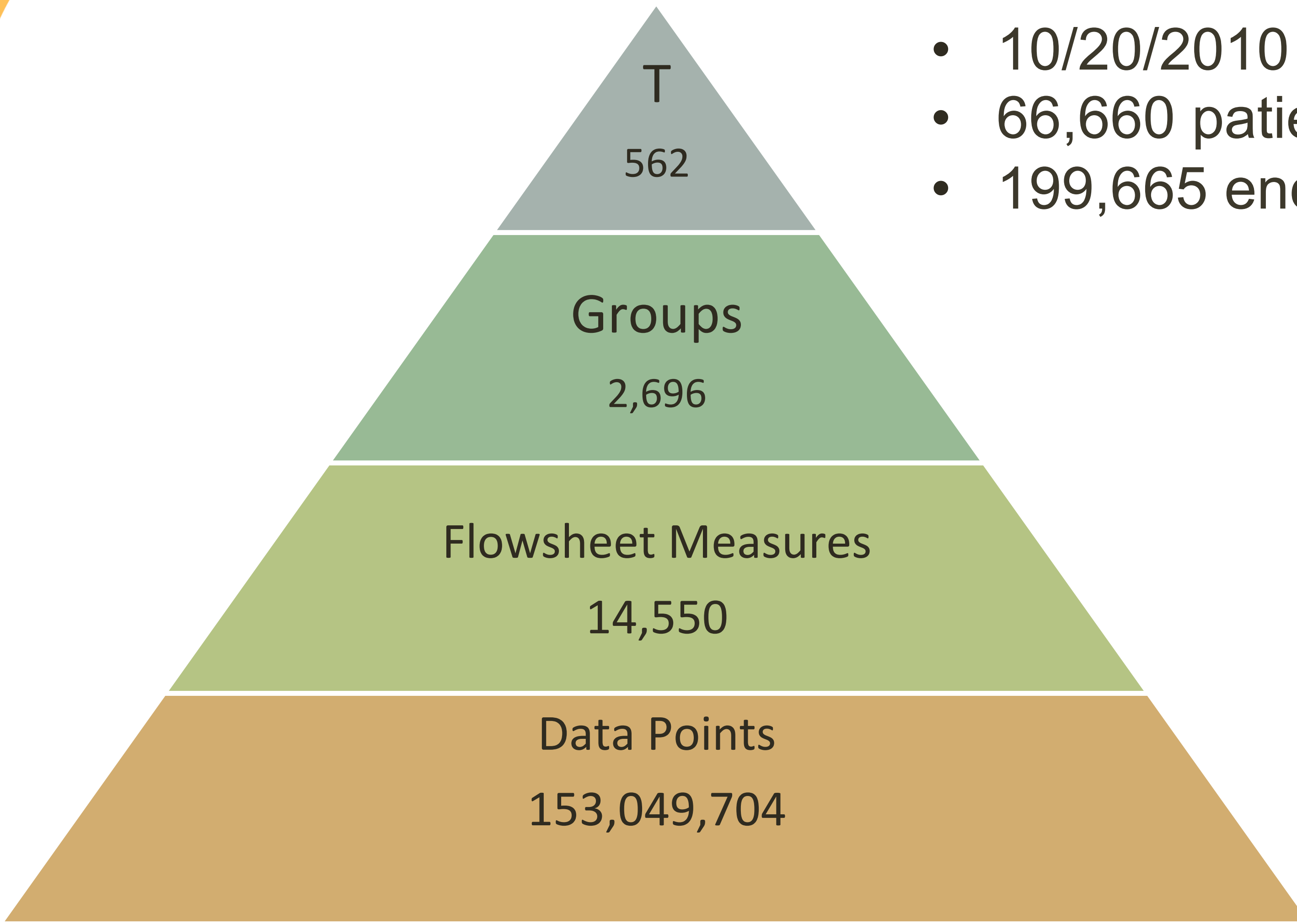
Screens/ Templates

Groups (of questions)

Value Sets/ Answers

Questions (Flowsheet Measures)

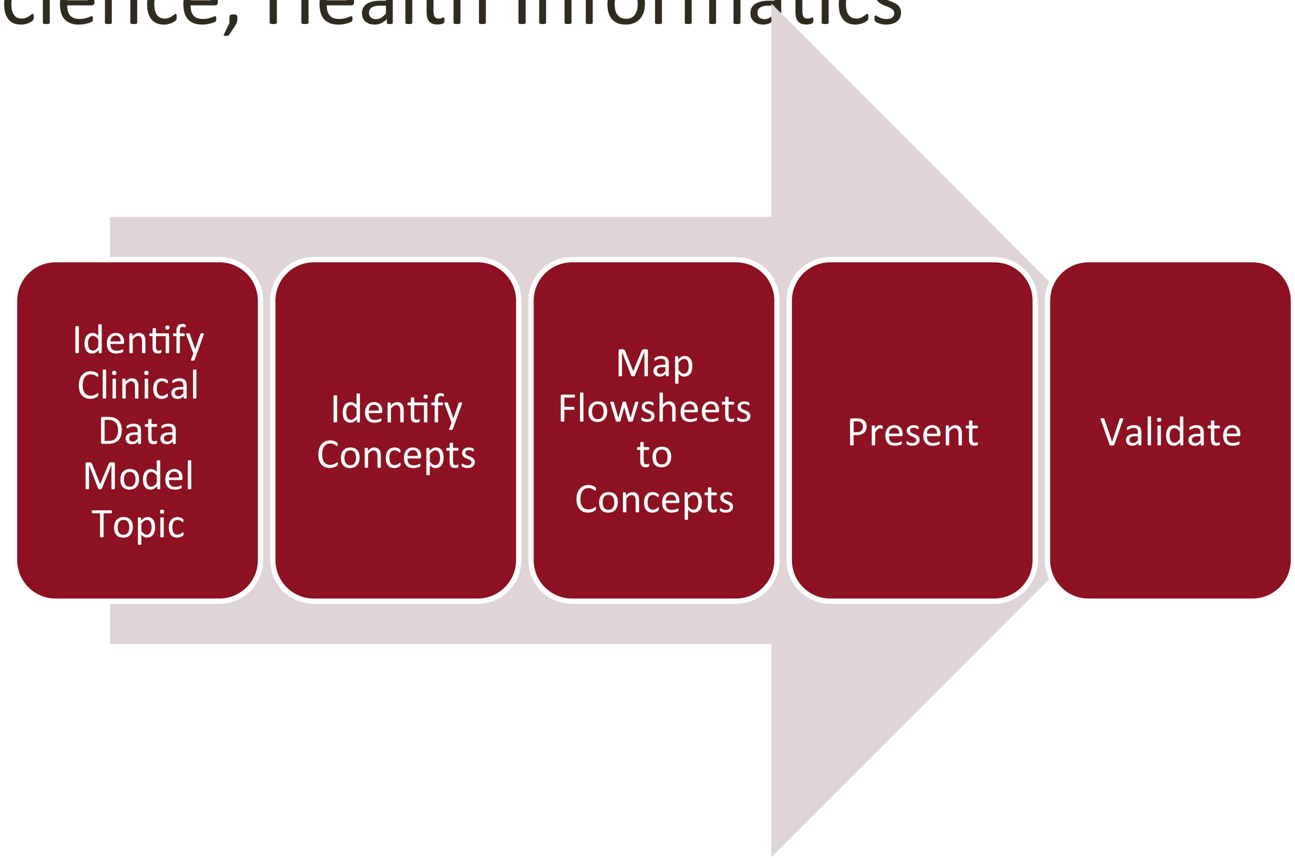
Data Source
Clinical Data Models - Flowsheets



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

UMN CTSI - Extend CDM

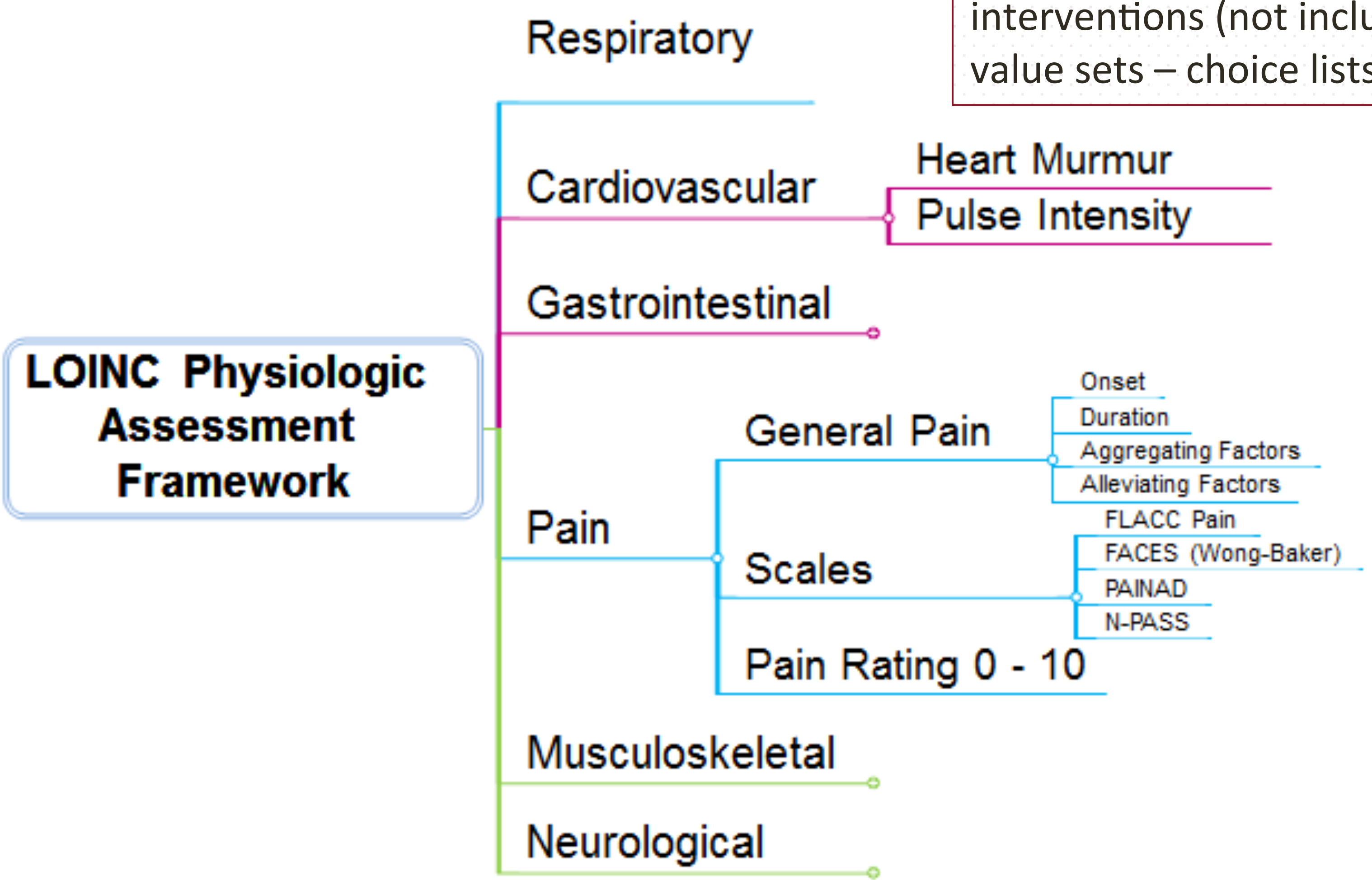
Team: Nursing (DNP/ PhD), Computer Science, Health Informatics



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

Cardiovascular System	Pain
Falls/ Safety	Peripheral Neurovascular (VTE)
Gastrointestinal System	Pressure Ulcers
Genitourinary System/ CAUTI	Respiratory system
Neuromusculoskeletal System	Vital Signs, Height & Weight

Information Model Name	Number Flowsheet IDs Mapped to Observables	Number Information Model Classes/ Observables	
		Classes	Concepts
Cardiovascular System	241	8	84
Falls	59*	4	57
Gastrointestinal System	60	3	28
GI/ CAUTI	79	3	38
Musculoskeletal System	276	9	72
Pain	309	12	80
Pressure Ulcers	104	6	56
Respiratory System	272	12	61
VTE	67	8	16

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[8, 9]

Initial Data Set

808 agencies, 1,560,508 OASIS records, 888,243 patients

List of patients with and without WOC Nurse

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

Final Data Set

785 agencies, 447,309 patients,

449,243 episodes of care, 0.6% re-admissions

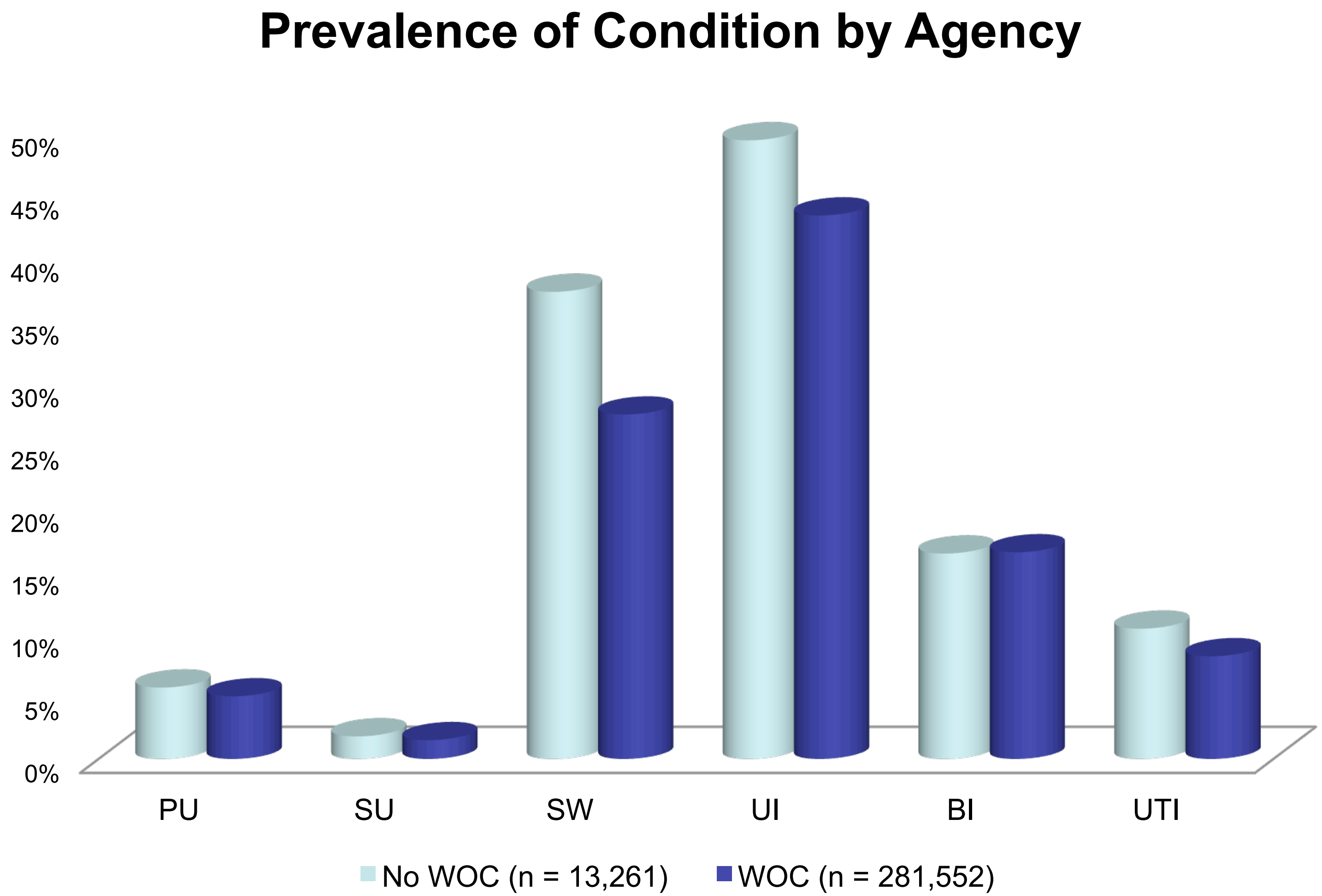
Certified WOC Nurses Influence on Incontinence & Wounds

Outcome Variables	Description
Pressure Ulcers	Total number of pressure ulcers (M0450 a-e)
Stasis Ulcers	Total number stasis ulcers (M0470/ M0474)
Surgical Wounds	Total number of surgical wound (M0484/ M0486)
Urinary Incontinence	Presence/management of urinary incontinence or need for a catheter (M0520)
Urinary Tract Infection	Treated for UTI in past 14 days (M0510)
Bowel Incontinence	Frequency of bowel incontinence (M0540)

Improved/ Not Worse (Stabilize) Outcomes

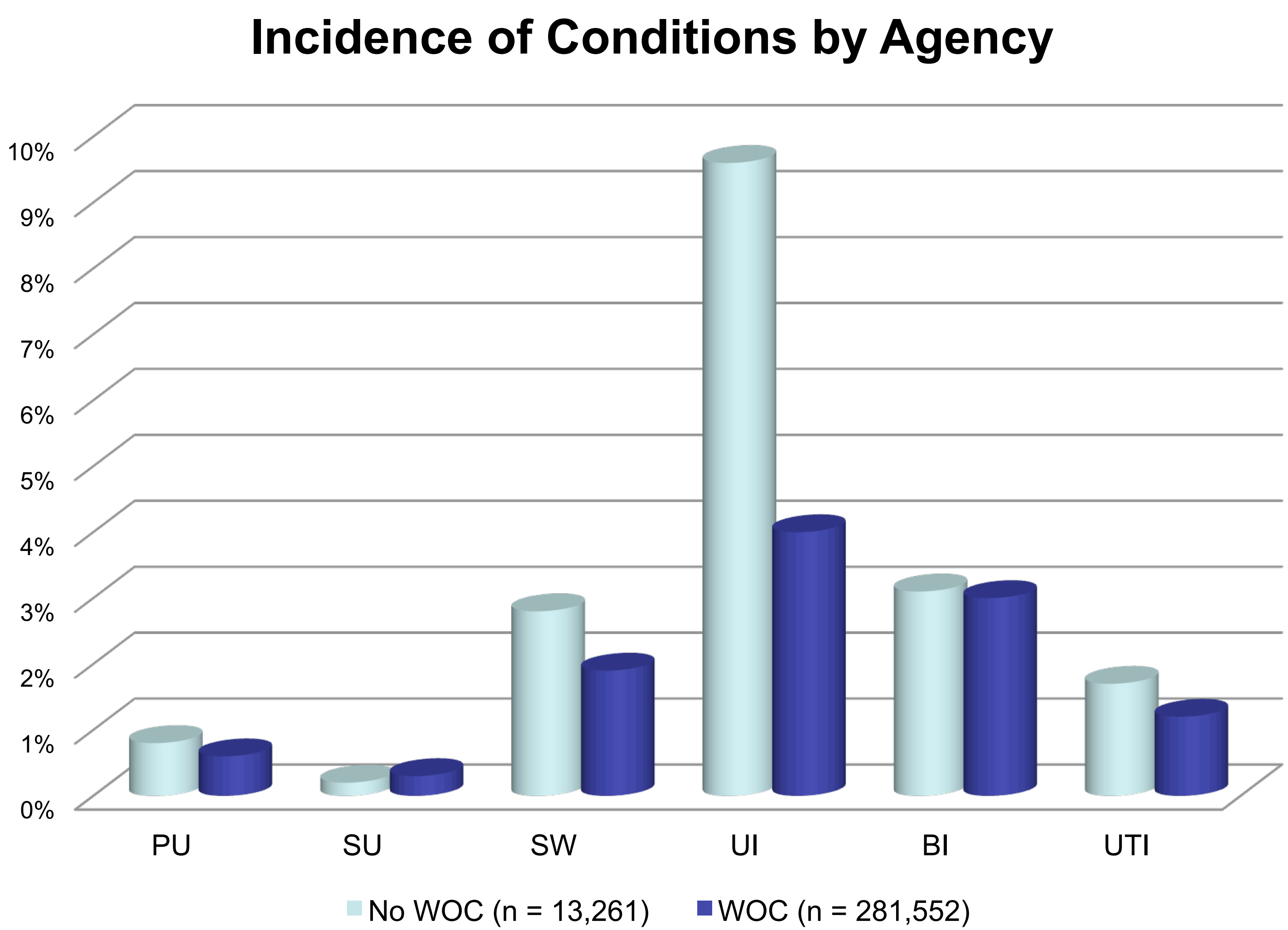
Score	Bowel Incontinence Frequency	Improved	Not Worse (Stabilize)
0	Very rarely /never has BI or has ostomy for bowel elimination		
1	Less than once weekly		
2	One to three times weekly		
3	Four to six times weekly		
4	On a daily basis		
5	More often than once daily		

Aim 1: Prevalence



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

Aim 2: Incidence



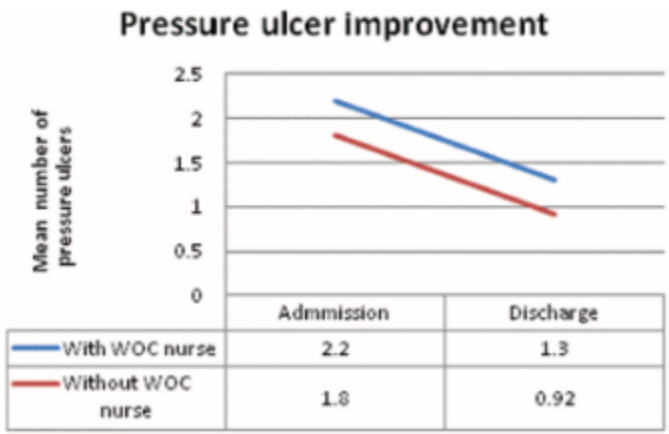
Effect of WOC Nurses on Agency Outcomes

Outcomes Comparing Agencies With and Without a WOC Nurse[a]

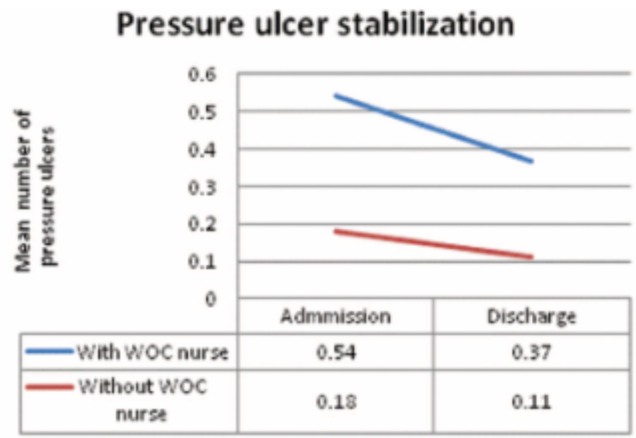
	Improvement		Stabilization	
Outcome Concept	**OR**	**95% CI**	**OR**	**95% CI**
Pressure ulcers	1.9	1.8-2.0	1.29	1.21-1.37
Urinary incontinence	1.4	1.38-1.43	2.3	2.26-2.4
Urinary tract infections	1.4	1.38-1.43	1.2	1.16-1.27
Surgical wounds	1.39	1.36-1.42	1.5	1.46-1.57
Stasis ulcers	1.2	1.1-1.3	Unable to model[b]	
Bowel incontinence	1.14	1.11-1.2	1.16	1.23-1.9

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.

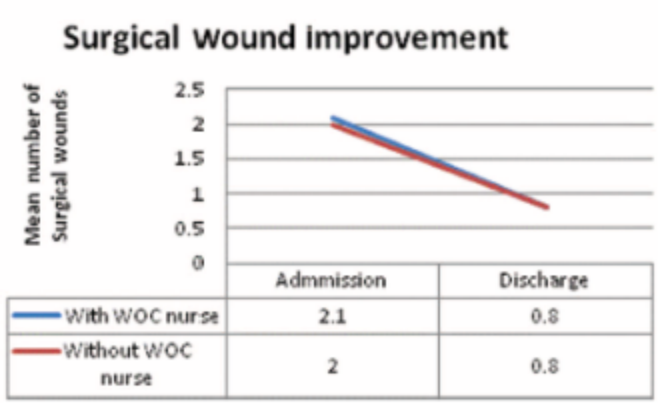
Individual Patient Outcomes



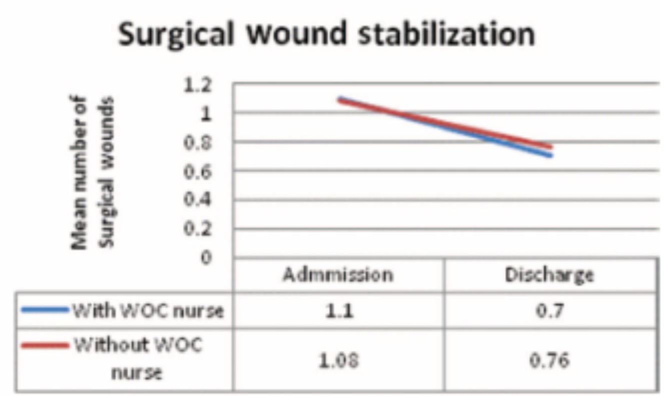
	Admmission	Discharge
With WOC nurse	2.2	1.3
Without WOC nurse	1.8	0.92



	Admmission	Discharge
With WOC nurse	0.54	0.37
Without WOC nurse	0.18	0.11



	Admmission	Discharge
With WOC nurse	2.1	0.8
Without WOC nurse	2	0.8



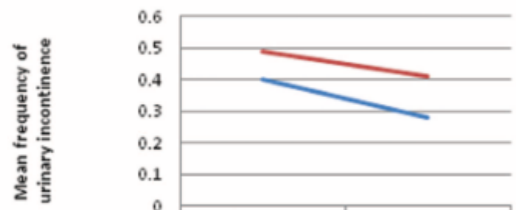
	Admmission	Discharge
With WOC nurse	1.1	0.7
Without WOC nurse	1.08	0.76

Individual Patient Outcomes

Urinary incontinence improvement

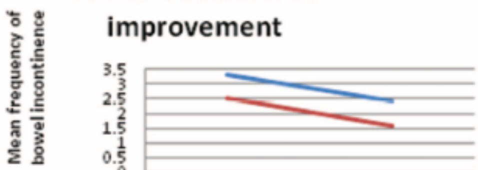
	Admission	Discharge
With WOC nurse	1.1	0.7
Without WOC nurse	1.08	0.76

Urinary incontinence stabilization



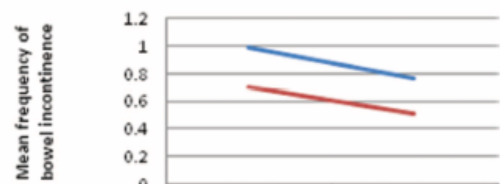
	Admission	Discharge
With WOC nurse	0.4	0.28
Without WOC nurse	0.49	0.41

Bowel incontinence improvement



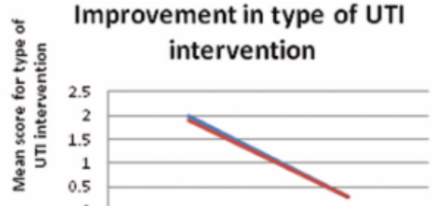
	Admission	Discharge
With WOC nurse	3.3	2.4
Without WOC nurse	2.5	1.6

Bowel incontinence stabilization



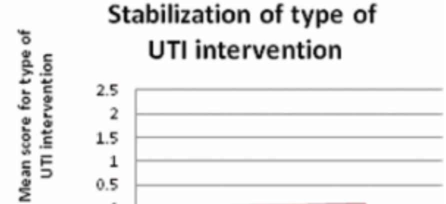
	Admission	Discharge
With WOC nurse	0.99	0.77
Without WOC nurse	0.70	0.51

Improvement in type of UTI intervention



	Admission	Discharge
With WOC nurse	2	0.3
Without WOC nurse	1.9	0.28

Stabilization of type of UTI intervention



	Admission	Discharge
With WOC nurse	0.01	0.05
Without WOC nurse	0.01	0.06

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

Score	Label	Description
0	INDP	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)
1	DEVICE	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
2	SUPERV	Able to walk only with the supervision or assistance of another person at all times
3	CHAIR_I	Chairfast, unable to ambulate but is able to wheel self independently
4	CHAIR_NI	Chairfast, unable to ambulate and [not independent] to wheel self
5	BED	Bedfast, unable to ambulate or be up in a chair

Note. Scores are based on Outcome and Assessment Information Set question M0700 Ambulation/Locomotion.

Comparison of Outcomes by Group

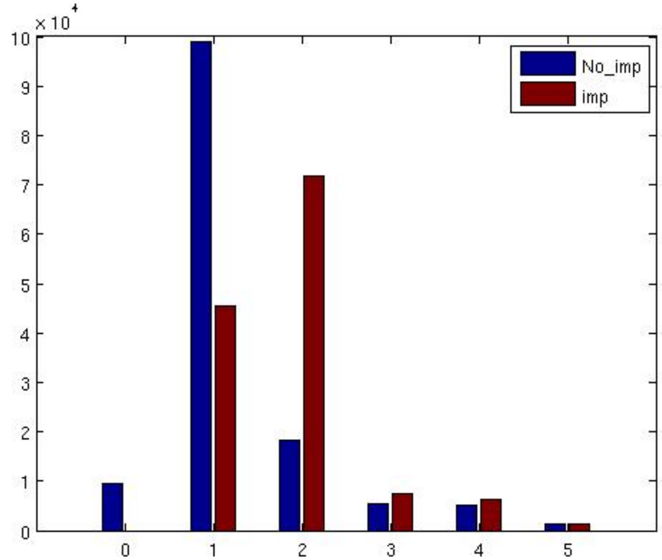
TABLE 2. Mobility Scores at Admission and by Change in Mobility at Discharge From Home Healthcare

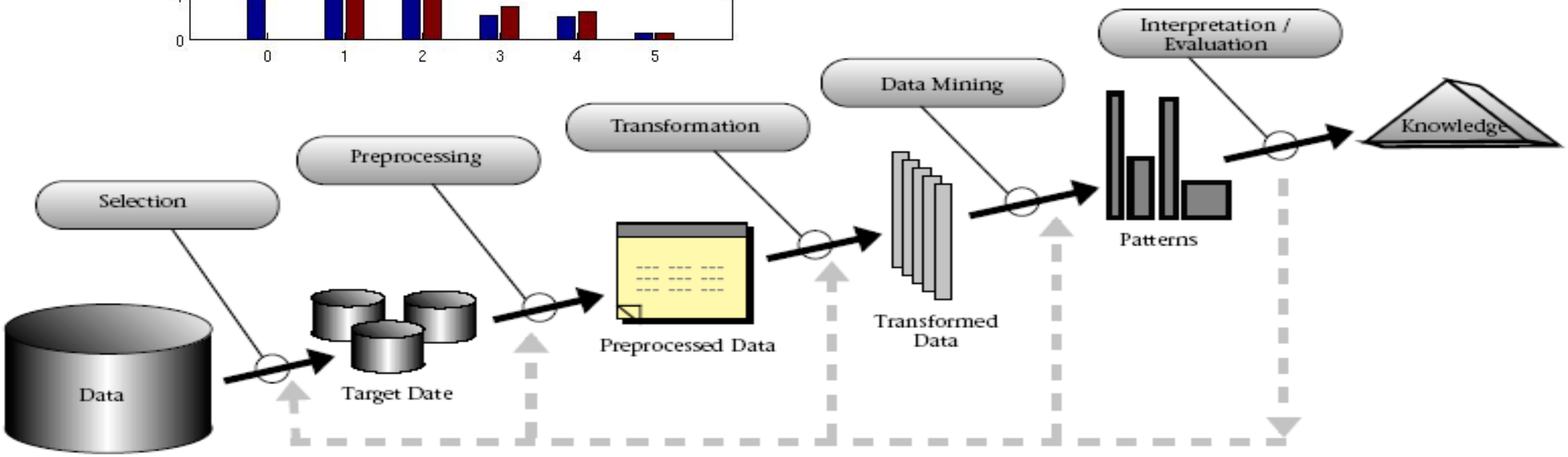
		Total (N = 261,035)		No improvement[b] (n = 128,920)		Improvement[c] (n = 132,115)	
Score[a]	Label	n	(%)	n	(%)	n	(%)
1	INDP	144,615	(55.4)	99,119	(68.5)	45,496	(31.5)
2	DEVICE	89,860	(34.4)	18,129	(20.2)	71,731	(79.8)
3	SUPERV	12,669	(4.9)	5,322	(42.0)	7,347	(58.0)
4	CHAIR_I	11,339	(4.3)	5,163	(45.5)	6,176	(54.5)
5	CHAIR_NI	2,552	(1.0)	1,187	(46.5)	1,365	(53.5)
All		261,035	(100.0)	128,920	(49.4)	132,115	(50.6)

[a]Scores are based on Outcome and Assessment Information Set question M0700 Ambulation/Locomotion. [b]Mobility outcome = 0. [c]Mobility outcome = 1.

Overall Steps

OASIS
EHRs f
serve
certif



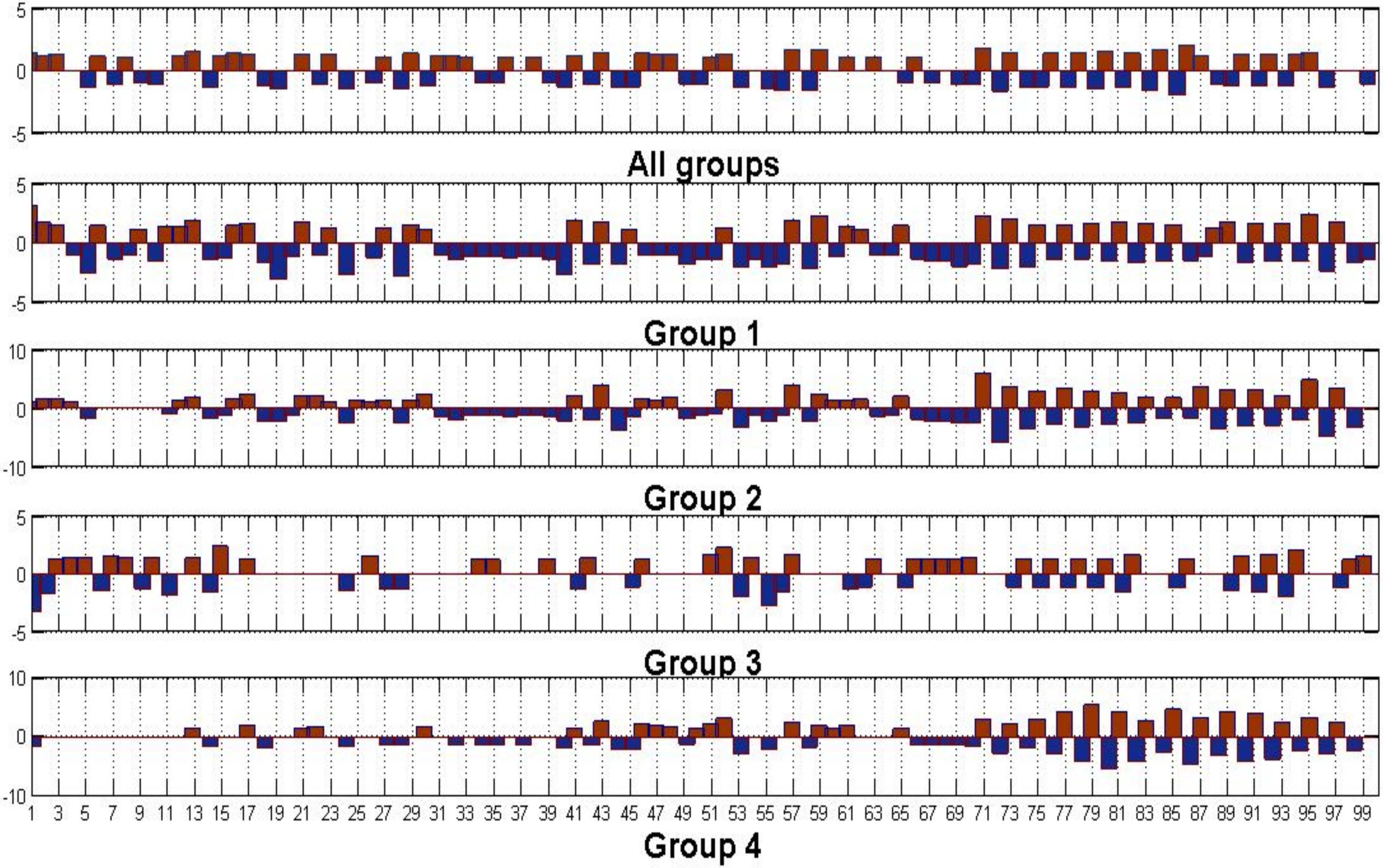


Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, pp. 37 – 54. http://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf. P. 41

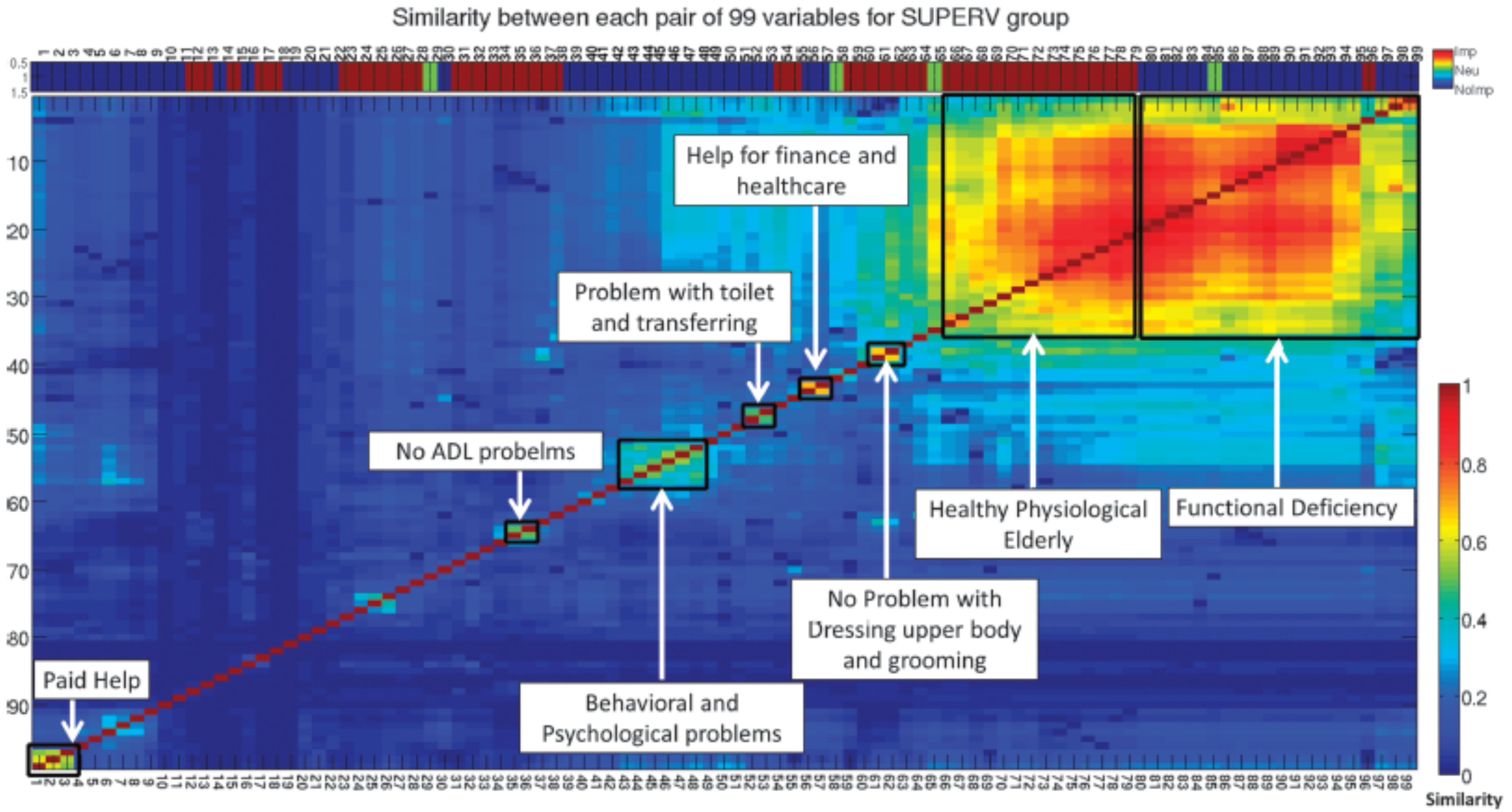
Data Mining Techniques

- 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



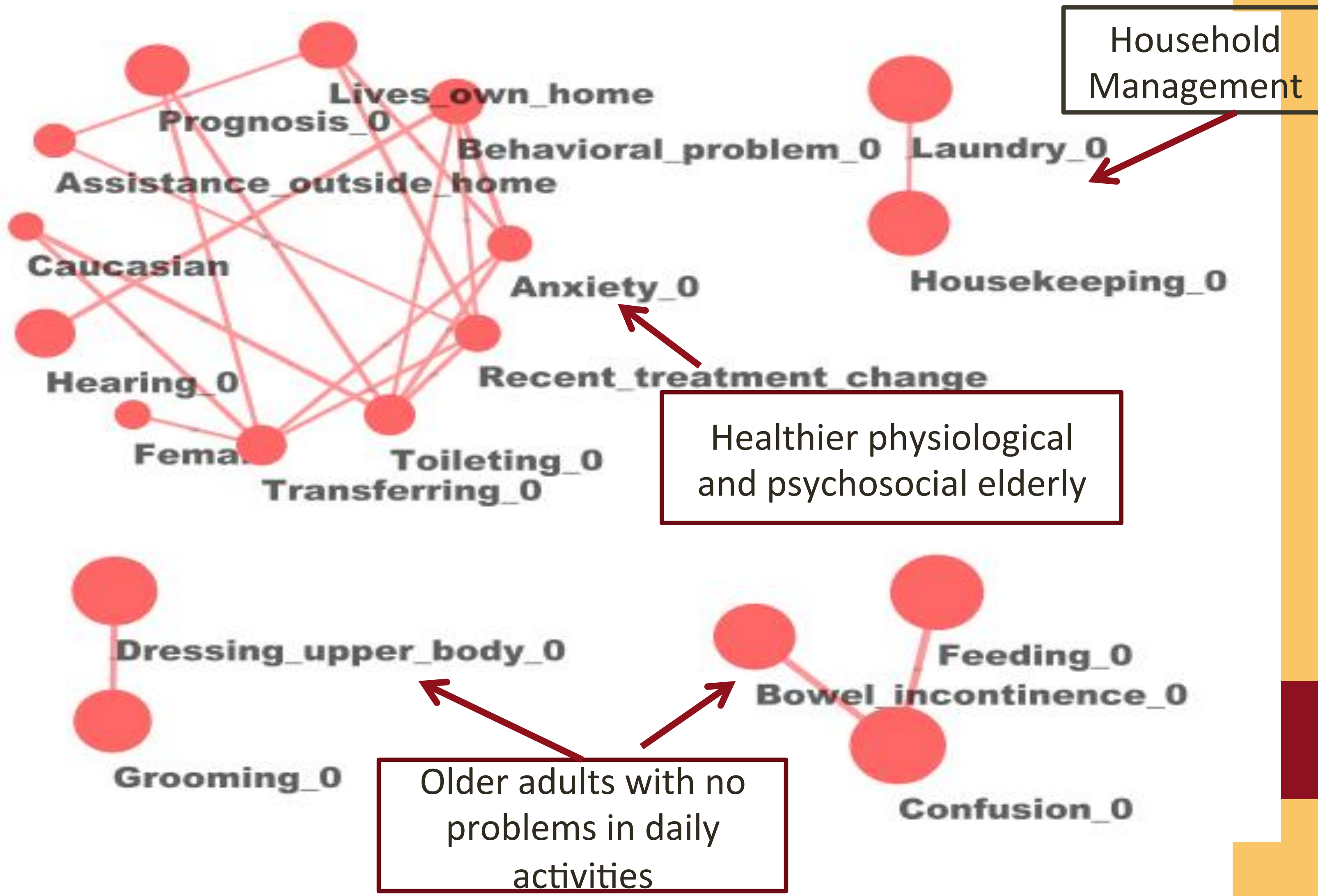
Clustering Groups



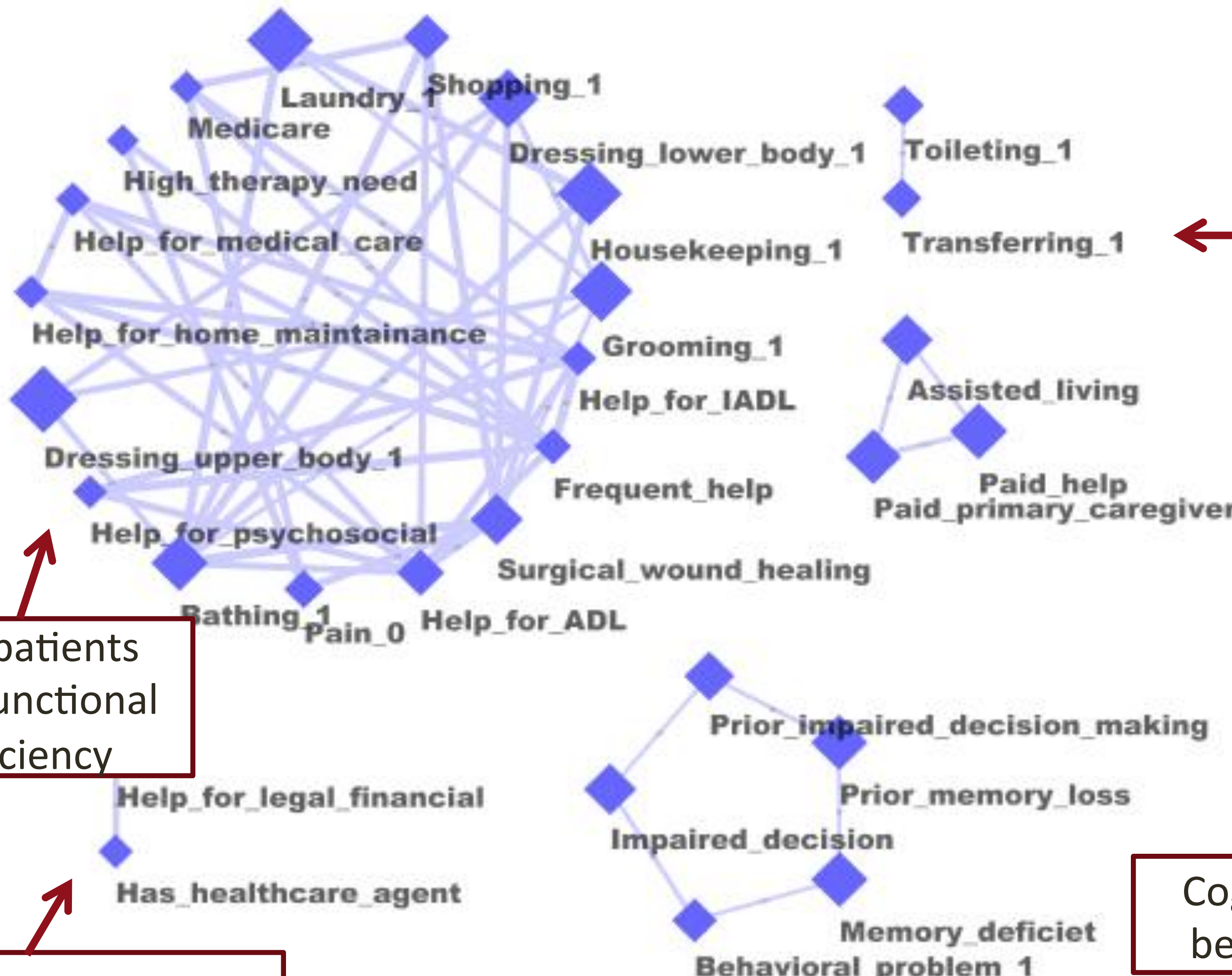
Variables

Variables

Improvement Group 2



No Improvement Group 2



Incapable to toilet and transfer

Paid Help

Frail patients with functional deficiency

Cognitive deficits and behavioral problems

Help with financial and legal matters

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

University of Minnesota
Driven to Discover℠

Nursing Knowledge:
2015 Big Data Science Conference

School of Nursing

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/
- **Science of NBD2K**
  - Completed NMMDS updates/ LOINC coding for public distribution
    - 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
  - 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

- Westra BL, Christie B, Johnson SG, et al. Expanding interprofessional EHR data in I2B2. *AMIA Joint Summits on Translational Science proceedings AMIA Summit on Translational Science*. 2016.
- Dey S, Cooner J, Delaney CW, Fakhoury J, Kumar V, Simon G, Steinbach M, Weed J, Westra BL. Mining Patterns Associated With Mobility Outcomes in Home Healthcare. Nurs Res. 2015 Jul-Aug; 64(4):235-45. PubMed PMID: 26126059.
- Westra BL, Latimer GE, Matney SA, Park JI, Sensmeier J, Simpson RL, Swanson MJ, Warren JJ, Delaney CW. A national action plan for sharable and comparable nursing data to support practice and translational research for transforming health care. J Am Med Inform Assoc. 2015 May;22(3):600-7. PubMed PMID: 25670754.
- Westra BL, Dey S, Fang G, Steinbach M, Kumar V, Savik K, Oancea C, Dierich M. Interpretable Predictive Models for Knowledge Discovery from Homecare Electronic Health Records. Journal of Healthcare Engineering. 2011; 2(1):55-74.