Big Data at the Interdisciplinary Level to Improve Care

Bonnie L. Westra, PhD, RN, FAAN, FACMI

Associate Professor and Director, Center for Nursing Informatics
Summer Institute for Nursing Informatics
July 20, 2016

http:/z.umn.edu/bigdata

University of Minnesota
School of Nursing
Driven to Discover

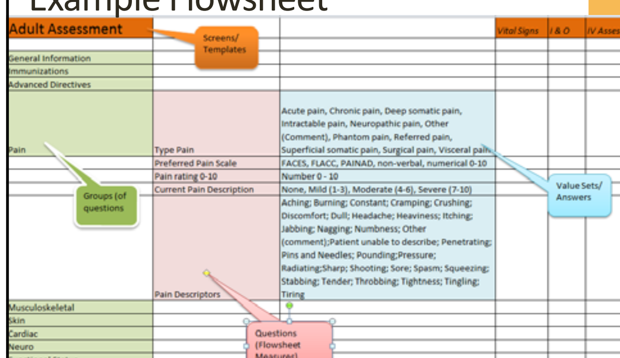
Purpose

- Identify information documented by nurses and interprofessional providers in flowsheet data for big data and data science
- Discuss the challenges of using flowsheet data for research
- Demonstrate use of information models with standardized data to create usable flowsheet data
- Examine an exemplar of interprofessional research that incorporates flowsheet data.

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
- Three-legged stool
  - Data
  - Technology
  - Algorithms

Example Flowsheet



Flowsheets – Interdisciplinary Care

Flowsheet Data Challenges

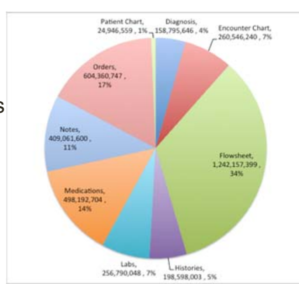
- Lack of standardized codes
- Volume of data – what is essential?
- Duplication for the same assessments, interventions, & goals
  - Different people building screens
  - Software upgrades
  - Discipline/ practice specific needs
  - Typing (spaces, capitalization, plural, etc)
  - Pre-coordination
- No information models exist across EHR
  - Data driven information modeling required

UMN – Academic Health Center

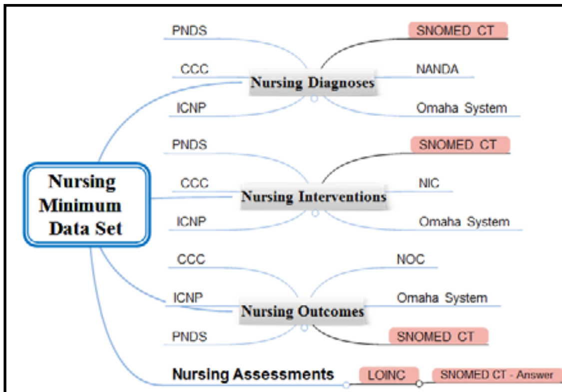
Flowsheets constitute 34% of all data

- 14,564 measure types
- 2,972 groups
- 562 templates
- 1.2 billion observations

- 2,000 measures cover 95% of observations



2,401, 377 patients as of 6/10/16



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

ANA AMERICAN NURSES ASSOCIATION

Position Statement

Inclusion of Recognized Terminologies Supporting Nursing Practice within Electronic Health Records and Other Health Information Technology Solutions

Purpose:

- Support for the use of recognized terminologies supporting nursing practice
- Promote integration of terminologies into information technology solutions
- Facilitate interoperability between different concepts, nomenclatures, and information systems

http://www.nursingworld.org/MainMenuCategories/Policy-Advocacy/Positions-and-Resolutions/ANAPositionStatements/Position-Statements-Alphabetically/Inclusion-of-Recognized-Terminologies-within-EHRs.html

ANA AMERICAN NURSES ASSOCIATION

Position Statement

When exchanging a Consolidated Continuity of Care Document (C-CDA) with another setting for problems and care plans,

- SNOMED CT® and LOINC® should be used for exchange
- LOINC® should be used for coding nursing assessments and outcomes
- SNOMED CT® for problems, interventions, and observation findings

http://www.fhims.org/press_ulcer.html

Pressure Ulcer Model - Standardized

Type	Display Name	Definition	LOINC Code	LOINC Name	SNOMED CT FSN	SNOMED CT CID
Question	Presence of Skin Alteration				integrity of skin (observable entity)	364535009
Value	Present				impaired skin integrity (finding)	7919002
Value	Absent				absent (qualifier value)	2667000
Panel	Skin Inspection Class	Skin Inspection Class	72284-3	Skin assessment panel		
Question	Skin Color		39107-6	Color of Skin	color of skin (observable entity)	364533002
Value	Normal for Ethnic Group	normal for baseline; variation in ethnic groups; pigmentation (color and pigmentation = appropriate for race)			skin normal color (finding)	297952003
Value	Cyanotic	Cyanosis is a dusky blue appearance of the fingers, lips or mucous membranes, and is abnormal in both light-and dark-skinned individuals. Cyanosis occurs when there is greater than 5 g/dL of deoxygenated hemoglobin in the blood. The earlobes, lower eyelids, lips, oral mucosa, nailbeds, and palmar and plantar surfaces may be especially cyanotic. Dark-skinned individuals have an ashen-gray to pale tint, and the lips and tongue are good indicators of cyanosis.			cyanosis of skin (finding)	119419001
Value	Flushed	Sustained bright red or pink coloration in light-skinned individuals is abnormal. Dark-skinned individuals may have no underlying change in coloration. This may occur because of dilated superficial blood vessels, increased blood flow, febrile states, local inflammatory conditions, or excessive alcohol intake.			flushed complexion (finding)	248213001
Value		Jaundice is a yellow-green to orange cast or coloration of the skin, sclera, mucous membranes, fingernails, and palmar or plantar surfaces in the light-skinned and may appear as yellow staining in the sclera, hard palate, and			jaundice (finding)	

NDNQI - Skin/ Pressure Ulcer

Unique Flowsheet IDs

Actual Data Collected

Skin Inspection
Skin WDL
Inspection Type
Body site Inspected
Skin Color
Skin Temperature
Skin Moisture
Skin Turgor
Skin Integrity
Skin Characteristics
Skin Signs Symptoms / Conditions

Adult Braden Assessment
Braden Sensory Perception
Braden Moisture
Braden Physical Activity
Braden Mobility
Braden Nutritional Intake
Braden Friction and Shear
Braden Score

Pressure Ulcer Assessment
Pressure Ulcer Present
Pressure Ulcer Problems Assessed
Pressure Ulcer Location
Drainage Amount
Drainage Color or Characteristics
Site Assmnt Wnd Base
Pressure Ulcer Description
Anatomical Site . anatomical location
Pressure Ulcer Stage
Wound Size
Wound Skin Edges
Pressure Ulcer State of Healing
Wound Tunneling Length (cm)
Wound Undermining Length (cm)
Pressure Ulcer Dressing
Pressure Ulcer Dressing Status
Peri-wound Assessment

Actual Data Collected

Braden Q Scale
Mobility (Braden Q)
Activity (Braden Q)
Sensory Perception (Braden Q)
Skin Moisture (Braden Q)
Friction and Shear (Braden Q)
Nutrition (Braden Q)
Tissue Perfusion and Oxygenation (Braden Q)
Braden Q Score

Neonatal Skin Condition Assessment
Neonatal Skin Condition Dryness
Neonatal Skin Condition Erythema
Neonatal Skin Condition Breakdown
Neonatal Skin Condition Total Score

Skin Intervention
Pressure Ulcer Treatment
Pressure Redistribution Techniques
Pressure Reduction Devices
Skin Mucous Membrane Protection
Incontinence Protocol in Use
Incontinence Skin Appearance
Incontinence Product
Incontinence Product Application Frequency
Skin Hygiene Care
Skin Treatment / Device / Implant

Skin Color – Choice Lists

ashen	no abnormal color
blanchable	non-blanchable
blue	other (see comments)
color consistent w/ ethnicity	pale
dusk	red
dusky	ruddy
ecchymotic	shiny
flushed	waxen
gray	white
jaundice	yellow
mottled	

Skin Inspection Choices

abrasion	excoriation	pressure ulcer
blister	fragile	pressure ulcer(s)
body piercing	inci	ra
burn(s)	incision	rash
cracked	incision(s)	rash(s)
cut(s)	intact	scab
cyst	itchy	scar
drain/device	mass	skin tear
ecchymosis	other (see comments)	subcutaneous emphysema (specify)
erosion	petechiae	tattoo
		wound

Requirements for Useful Data

- Common data (information) models
- Standardized coding of data
- Standardized 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.

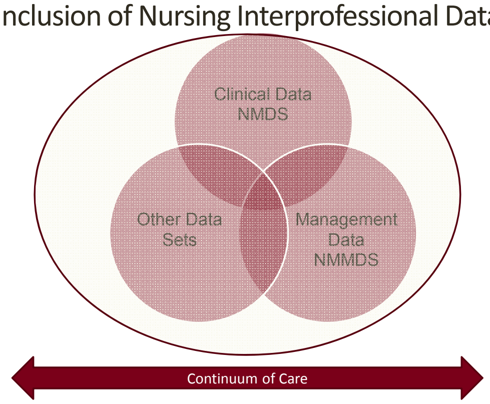
PRO_CM v2.0
Patient-Reported Outcome (PRO) Common Measures (CM) are standardized measures that are defined in the same way across all PCORnet networks. Each measure is recorded at the individual item level: an individual question/statement, paired with its standardized response options.

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/

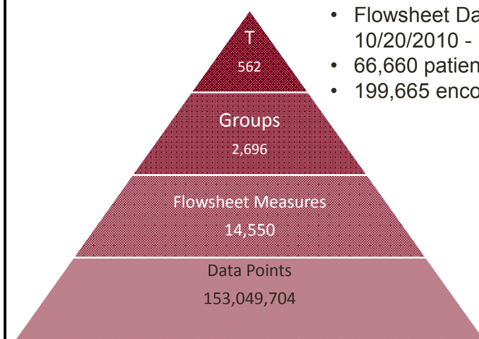
Vision
Inclusion of Nursing Interprofessional Data



20

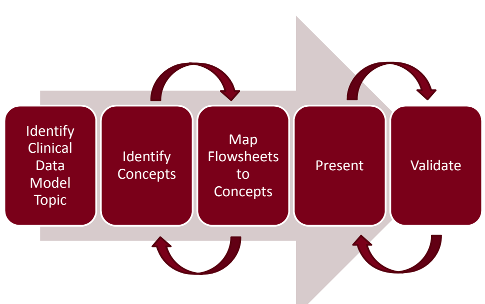
Creating Information Modes Flowsheet Data

Sample Data Source - Clinical Data Models



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

Information Model Development Process



Flowsheet Information Models

Pain	Cardiovascular System
Falls/ Safety	Gastrointestinal System
Peripheral Neurovascular (VTE)	Neuromusculoskeletal System
Genitourinary System/ CAUTI	Respiratory system
Pressure Ulcers	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
Vital Signs/ Anthopometrics	85	10	48

Example Mapping Pain

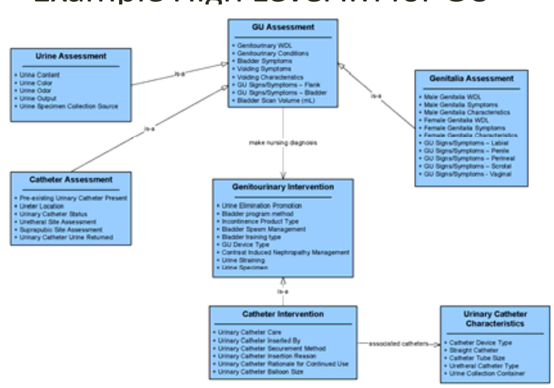
Type	Information Model Concept Name	Value Type	Flow Measure ID
Question	Current Pain	8	1770009571, 1510000901, 677485, 1450101723, 1071000059, 1050000363, 1770011583, 1770011588, 1770011593, 1770011598, 1770011603, 3040110359
Value	Complains of pain/discomfort		
Value	Denies pain/discomfort		
Value	Sleeping; pt not able to self report		
Value	Unable to assess		
Value	Yes		
Value	No		
Value	Other (comment)		
Question	Pain Type	8	3040100009, 3040100010, 605339, 600455, 671125, 606588
Value	Acute pain		
Value	Chronic pain		
Value	Deep somatic pain		
Value	Intractable pain		
Value	Neuropathic pain		
Value	Other (Comment)		
Value	Visceral pain		
Question	Pain Rating 0 - 10 Scale	8	3040110433, 13, 18, 301080, 3040100517, 673797, 301130, 301180, 7060860, 7060910, 7061000, 6183, 7061120, 671197, 675152, 602137, 604909, 601149, 604425, 3040110434, 3040102228, 3040102230, 3040102231, 3040102234, 3040102237, 3040102240, 3040102243, 671197, 675152, 607036, 1770004196,
Value	0		
Value	1		
Value	2		
Value	3		
Value	4		
Value	5		
Value	6		
Value	7		
Value	8		
Value	9		

Value type: 1 = Numeric, 2 = Free Text, 8 = Choice list

Example of Mapping GU

RowNum	Question	Type	Information Model Concept Name	Value Type	ID
1		Panel	Genitourinary Assessment		
2		Question	Genitourinary WDL	8	606179
3	1	Value	WDL		
4	1	Value	WDL Except		
5		Question	Genitourinary Conditions	8	674599, 675571, 605415
6	5	Value	acute renal failure		
7	5	Value	benign prostatic hyperplasia		
8	5	Value	burning		
9	5	Value	chronic kidney disease		
10	5	Value	decreased output		
21	5	Value	prostatitis		
22	5	Value	pyelonephritis		
23	5	Value	recent urinary tract infection		
24	5	Value	recurrent urinary tract infection		
25	5	Value	renal calculi		
26	5	Value	retention		
27	5	Value	urgency		
28		Question	Bladder Symptoms	8	676482
29	28	Value	bladder distended		
30	28	Value	bladder fullness		
31	28	Value	bladder pressure		
32	28	Value	bladder spasm		
33	28	Value	flank pain		
34	28	Value	other (see comments)		
35	28	Value	urinary retention		
36		Question	Voiding Symptoms	8	674075
37	36	Value	burning		
38	36	Value	decreased output		

Example High Level IM for GU



Next Steps

- Validating information models with other organizations
  - Allina, Cedars Sinai, Duke, Kaiser Permanente, MediComp, NY VNS
  - Beginning with Pain, then other quality measures
  - You are welcome to join also!
- Extracting their EHR flowsheet data into specific Excel spread sheets
- Searching for concepts in the Pain IM
- Finalize a generalizable model – comparing across organizations

- Resources
  - Pain Information Model for Validation. xls
  - High Level Clinical Information Models from Flowsheet Data

A Data Mining Approach to Determine Sepsis Guideline Impact on Inpatient Mortality and Complications

Bonnie L. Westra, PhD, RN, FAAN, FACMI
Lisiane Pruinelli, MSN, RN, PhD-C; Pranjul Yadav, PhD-C;
Alexander Huff; Jakob Johnson; Maribet McCarty, PhD, RN;
Vipin Kumar, PhD; Connie W. Delaney, PhD, RN, FAAN, FACMI;
Michael Steinbach, PhD; György J. Simon, PhD

University of Minnesota
School of Nursing
Driven to Discover
Computer Science & Engineering

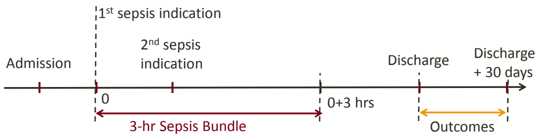
Purpose

- Demonstrate a methodology for using EHRs data to estimate the compliance with the Surviving Sepsis Campaign (SSC) guideline recommendations
- Estimate the effect of the SSC individual and combined recommendations on the prevention of in-hospital mortality and sepsis-related complications

Data and Cohort Selection

- De-identified EHR data from a Midwest health system was transferred to a CDR/UMN through a CTSA
- Timeframe: 1/1/09 – 12/31/11
- Inclusion criteria:
  - ICD-9 diagnosis = 995.92 and 785.5*
- Exclusion criteria:
  - Patients with cardiogenic shock
  - Patients with no antibiotic therapy

Study Design



- Baseline
  - In patients with at least 2 indications of sepsis, the time of the first sepsis indication
- Pre-existing complications assessed at admission
- Patient's labs and vitals at baseline
- Outcomes
  - Complications assessed at discharge
  - Updated for 30 days after discharge for complications

Baseline

- Sociodemographics
  - Age
  - Gender
  - Race/ ethnicity
  - Payer (Medicaid for low income)
  - Charlson Index of Comorbidity
- Vital signs
  - Heart rate (HR)
  - Respiratory Rate (RR)
  - Temperature (Temp)
  - Mean arterial pressure (MAP)
- Laboratory results
  - Lactate
  - White Blood Cell Count (WBC)
- Level of severity for each outcome complication

Next study - Urinary output, Dialysis, Level of Consciousness, Weight, IV, CVP

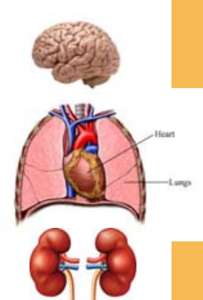
Sepsis TimeZero

At least 2 of the following criteria:

- MAP < 65
- HR > 100
- RR > 20
- Temp < 95 or > 100.94
- WBC < 4 or > 12
- Lactate > 2.0

Baseline/ Outcomes

- Complication (in hospital and up to 30 days after discharge)
  - Cerebrovascular
  - Respiratory
  - Cardiovascular
  - Kidney
- In-hospital mortality



3 Hour Bundle

- Lactate checked between admission and 0+3 hrs
- Blood culture drawn between admission and 0+3 hrs
- Broad-spectrum ABx administered between admission and 0+3 hrs
- Crystalloid fluid administered if MAP < 65 or lactate $\geq$ 4 between admission and 0+3 hrs

Data Preparation/Transformation

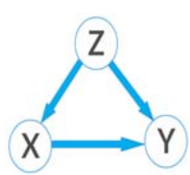
- Mapping the SSC guidelines to EHR data elements
- Data quality assessment based on literature and domain knowledge
  - Baseline missing values
  - For each data element, we evaluated range and created rules for suitable range to be included
    - Compared with other values i.e. MAP and SBP/ DBP and health conditions
  - Determine the use of one or more flowsheet measures for vital signs
- Created ICD grouping / severity for complications - respiratory, cardiovascular, cerebrovascular, renal

Method: Design

X : SSC guidelines
- Interventions

Y : Outcomes
- Mortality
- Complications

Z : Confounders
- Demographics such as age, gender
- Comorbidities
- Laboratories test results



Causal Model

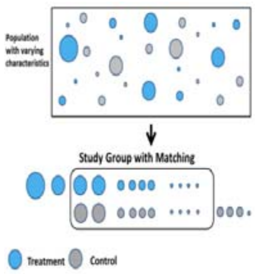
Our aim is to estimate the causal effect of X on Y, while simultaneously adjusting for the effects of Z (i.e. Confounders).

Method : Estimation Metric

- Metric for estimating the causal effect of intervention on outcome was ATT
- ATT is defined as the **Average Treatment** effect on the **Treated**.
- ATT = $E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$
  - E is the Expectation operator
  - Y_1 represents the outcome when intervention was prescribed
  - Y_0 represents the outcome when no intervention was prescribed

Method : Bootstrapped PSM

- **Propensity score matching** (**PSM**) was used for estimating the causal metric
- **PSM** is a statistical matching technique to estimate the effect of a treatment, policy or other intervention by accounting for the covariates that predict receiving the treatment
- Caliper Matching was used to match patients from treatment (intervention was prescribed) to that of control (intervention was not prescribed)
- Bootstrapping was used in conjunction with PSM to estimate the CI's



Baseline Features (n=177)

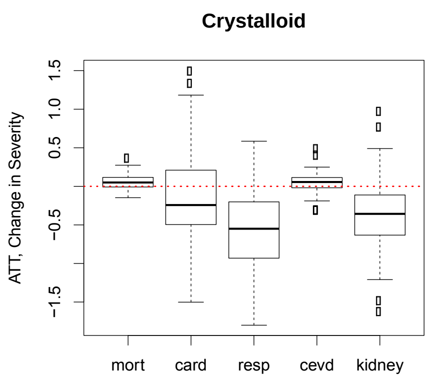
Characteristics	Patient Count	Condition (Severity Score)	Patient Count		
	Count/Mean (IQR)		Count (range of score2)		
Age (years)	61 (51-71)	Cardiovascular (0-12)	109 (0-2)	32 (3-4)	36 (5+)
Gender (Male)	102	Cerebrovascular (0-4)	158 (0)	19 (1)	
Race (Caucasian)	97	Respiratory (0-6)	156 (0-2)	21 (3+)	
Ethnicity (Latino)	11	Kidney (0-12)	121 (0-1)	56 (2+)	
Payer (Medicaid)	102	Charlson Index (0 – 17)	92 (0-1)	58 (2-4)	27 (5+)

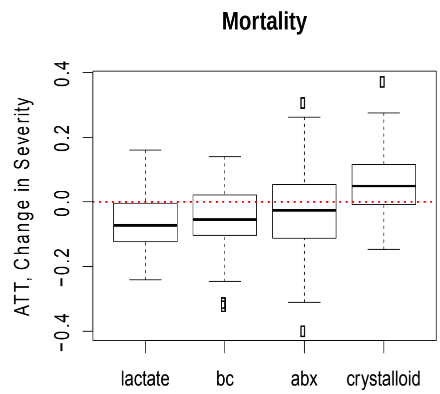
Results of 0+ 3Hrs - 3 Hour Rules

Rules Description	**Patient Count / %**			
	Y	N	% Compl	N/A
1. Lactate checked?	100	77	57%	0
2. Blood culture done?	97	80	55%	0
3. Broad spectrum antibiotic?	151	26	85%	0
4. Crystalloid fluid resuscitation for hypotension (MAP <65) or if Lactate > 4?	97	41	70%	39

Effect of Elements on Outcomes

	Mortality	Cardio-vascular	Respiratory	Cerebro-vascular	Kidney
Lactate checked	**-0.066**	0.617	**-0.004**	0.192	**-0.204**
Blood Culture checked	**-0.051**	0.514	0.534	0.003	**-0.115**
Antibiotic (Broad Spectrum)	**-0.024**	0.167	0.278	0.119	0.159
Crystalloid	0.058	**-0.136**	**-0.561**	0.045	**-0.354**





Limitations

- Data from a single center
- Small sample size
- No comparison group other than patients with sepsis who received or did not receive recommendations
- We may misestimate "TimeZero", as the majority of patients were in the ED when "TimeZero" occurred
- Timeframe of data - under vs over coding for sepsis

Summary

- Flowsheet data extends clinical data available for research
- Standardization is needed to make flowsheet data useful
- JAMIA feedback (revisions needed – but....)
  - "The models are intended to facilitate navigation and extraction of data to support cohort discovery and comparative effectiveness research. This is a very important and relevant objective."
  - "This work is incredibly important to standardize patient data (vitals, assessments, etc) that will be vital to any number of research and quality improvement activities. A set of standardized information models can allow local variation but provide a standard to aggregate and share data."

Questions?

Bonnie L. Westra, PhD, RN, FAAN, FACMI
westr006@umn.edu