Opportunities and Obstacles in Using Routinely Collected Electronic Data for Health Services Research

The Example of Identifying Infection Risk Factors and Outcomes

CIRAR Seminar Series
Thursday, October 2, 2008

Sherry Glied, Professor and Chair
Department of Health Policy and Management, Mailman School of Public Health
David Caplan, Director of Data Analytics,
Information Services – Business Solutions Group, New York Presbyterian Hospital

The Electronic Medical Record: Projections and Concerns for Policy

- Potential for long-term cost savings
- Possible perverse consequences of decision support: flagging patient safety issues has not necessarily reduced incidence
- Limitations of ICD-9 coding for reimbursement

Little discussion on the challenges of using electronic data for research purposes…

Exciting Possibilities for Research Using Electronic Data

- Post-marketing pharmaceutical research
- Clinical decision support systems
- Policy and payment changes
- Study of hospital phenomena (e.g. antimicrobial resistance)

Obstacles: Through the Lens of Antimicrobial Resistant Infection

1. Cooperation from Institutions and Individual Departments
2. Timing Issues
3. Integrating Disparate Datasets
4. Limits of Automated Case Finding
5. Definition Variation
6. Construct Validity
7. Data Reliability
8. Free Text
9. Time Stamps

"Distribution of the Costs of Antimicrobial Resistance" (Grant R01 NR010822): Aims

1. Estimate hospital costs of resistant infections
2. Assess extent to which patients other than the initially infected patient experience higher costs due to cases of antibiotic resistance
3. Examine the distribution of the costs of resistant infection between hospitals, third-party payers, and patients
4. Model the impact of alternative payment policies on incentives to invest in reducing resistant infections
5. Evaluate whether observed patterns of investment in the control of resistant infections are related to the distribution of the costs of infection

Performance Sites

New York Presbyterian Hospital

Columbia campus

- Milstein/Presbyterian Hospital – 700 bed tertiary
- Morgan Stanley Children's Hospital – 257 bed children's hospital

Cornell campus

- New York Hospital – 797 bed tertiary

Allen pavilion

- 212 bed community

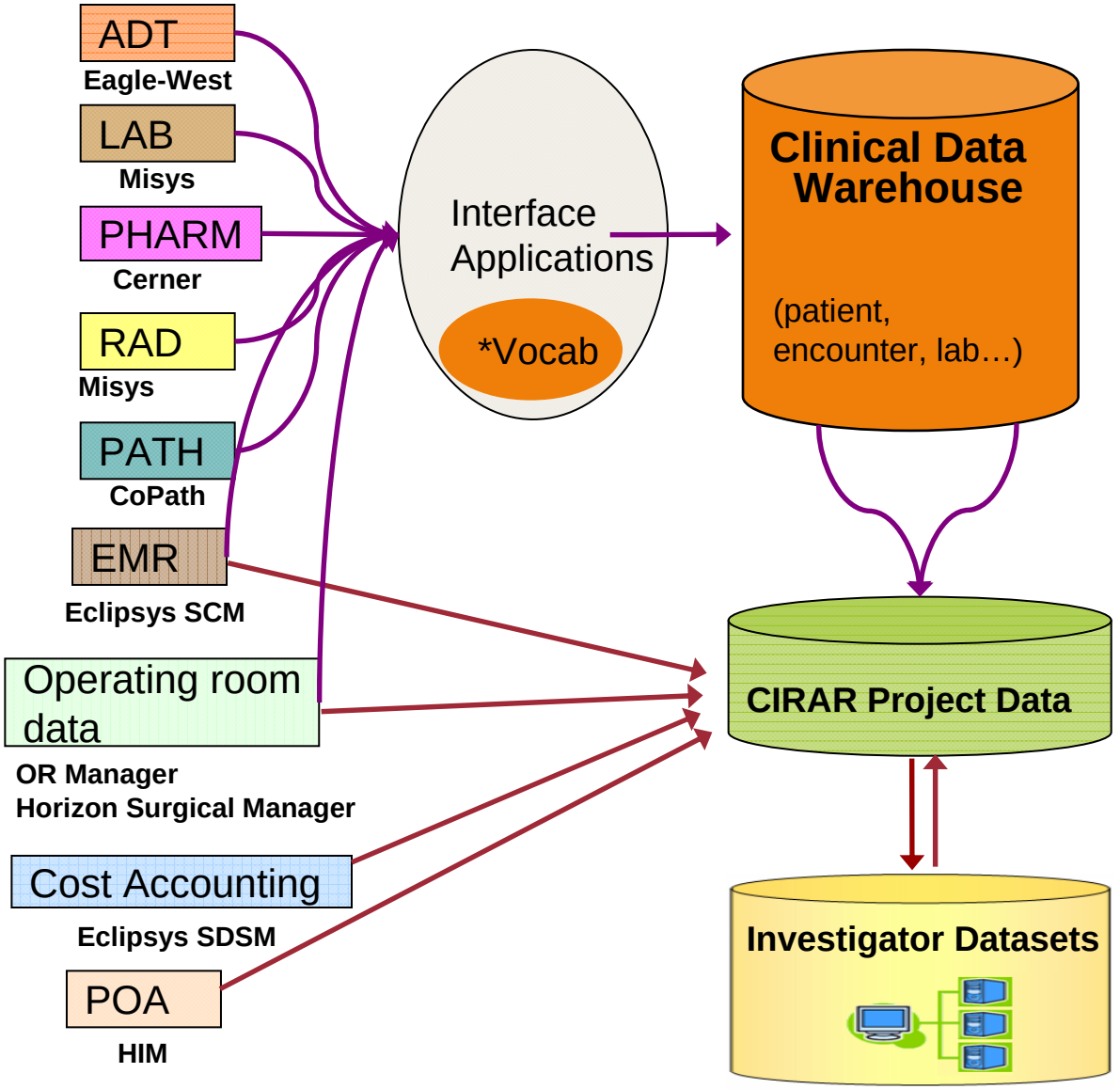
Methods

- Data linkage
- Identify cases resistant infection 2005-2009
- Match cases to susceptible and non-infected controls
- **Aim 1** – compute length of stay, costs, breakdown of costs
- **Aim 3** – determine primary payer expenditures; incidence of costs

Data Needs

- **Organism** – Enterococci, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Streptococcus pneumoniae
- **Sensitivity/resistance pattern**
- **Infection Site** – surgical, bloodstream, urinary tract, pneumonia
- **Infection Source** – hospital vs. community acquired
- **Risk factors for infection** – age, trauma, burns, diabetes, renal failure, malignancies, endotracheal intubation, central venous catheters, urinary catheters, mechanical ventilation, invasive procedures, immunosuppressive agents, length of hospital stay, stay in ICU
- **Risk factors for Resistant Infection** - prior hospitalization, exposure to long-term/broad-spectrum antibiotics transplantation, gastrointestinal surgery

High Level Data Architecture: Columbia (MHB, MSCH, TAP)



** Medical Entity Dictionary (MED)*

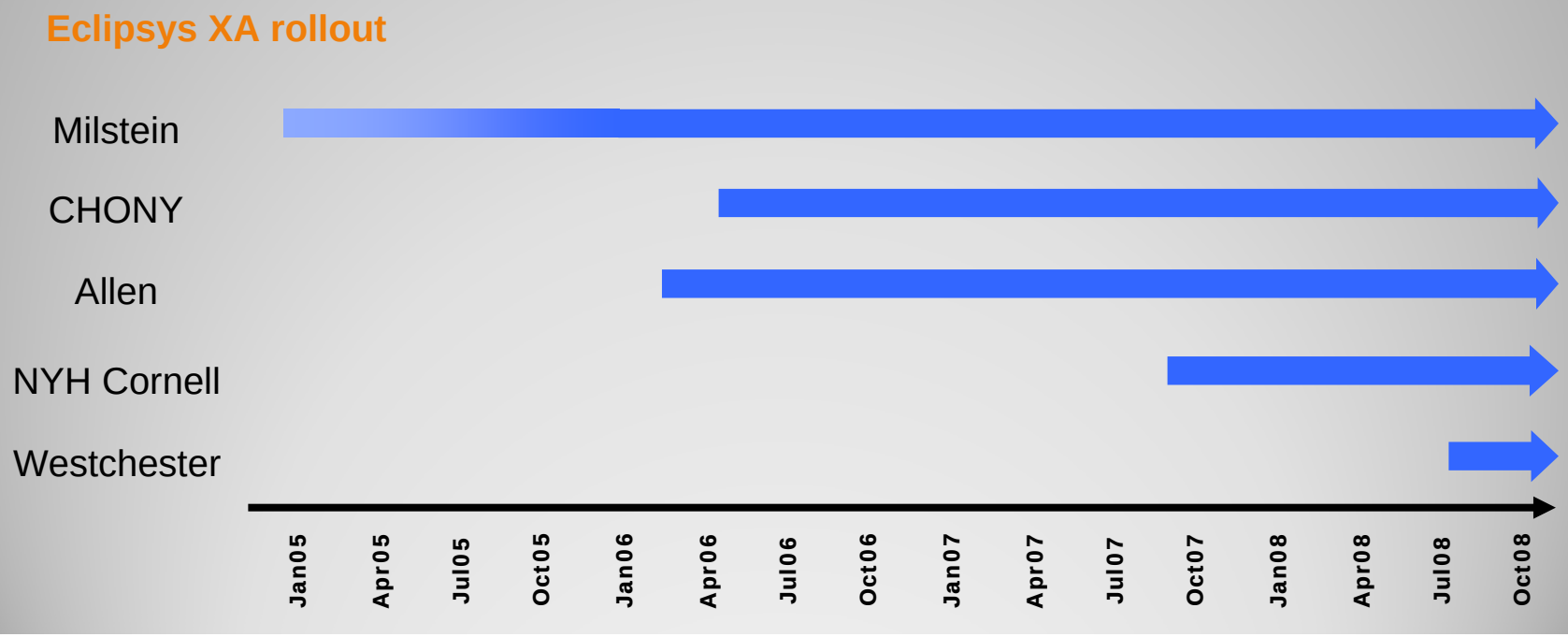
Obstacle #1: Cooperation from Institutions and Individual Departments

- Gaining access to a single department's data requires cooperation and approval from multiple parties
  - **Data Owners**
  - **Data Custodians**

Mutual Benefit

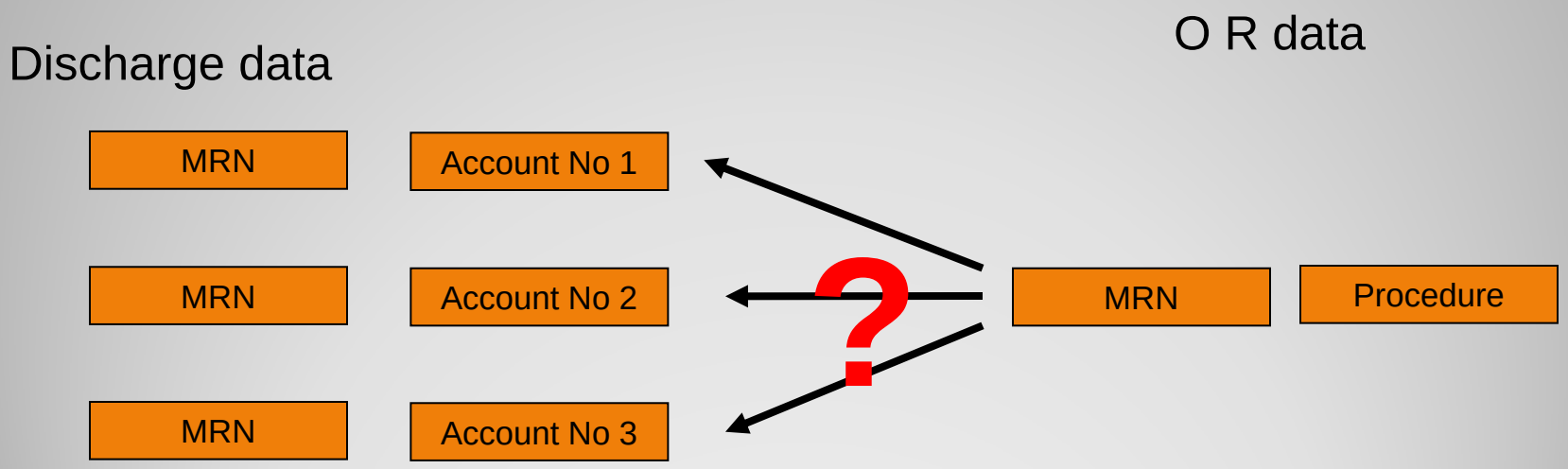
Obstacle #2: Timing Issues

- Dates for which data are available vary between different data systems



Obstacle #3: Integrating Disparate Datasets

- Departments use different methods to identify patients and encounters



Obstacle #4: Limits of Electronic Logic

- How do the results of an automated case finding program using electronic data compare with the "gold standard" of a clinician's or epidemiologist's judgment?
- How should the validity of electronic data be assessed?

Obstacle #5: Definition Variation

- Variable definitions can vary between data sources
- Available variable definitions may not be appropriate for your project

EXAMPLE:

(1) Defining Infections

(2) Operating Room Procedure Time

Definition Variation
Defining MRSA Infection

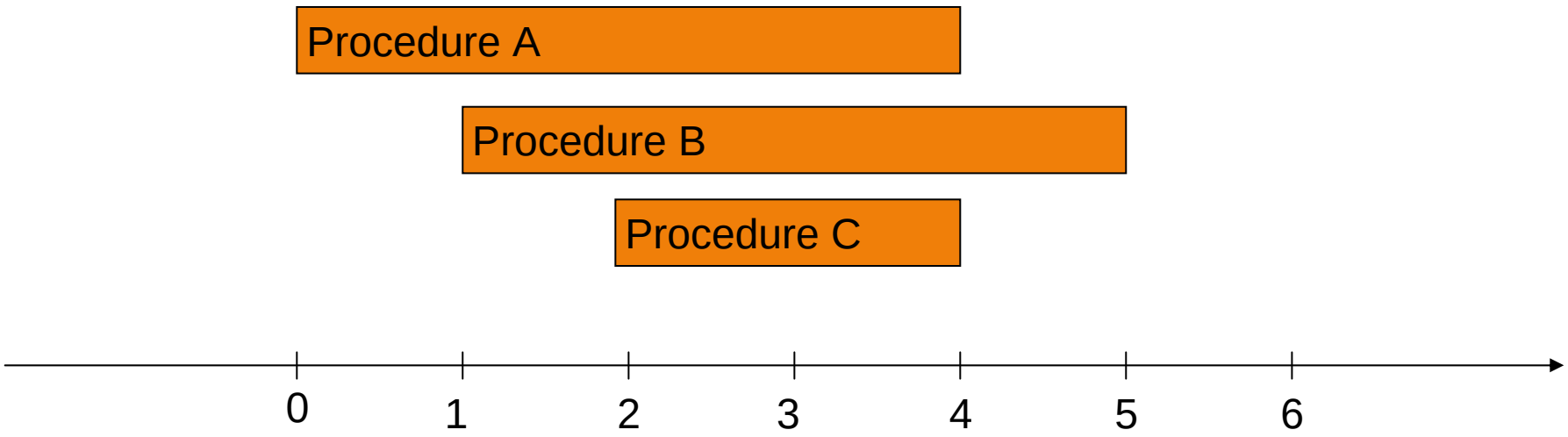
Epi-portal

- Culture positive for MRSA
- Culture positive for Staph. aureus and resistant to Oxacillin or Methicillin

Project

- Positive blood culture for organism excluding common skin contaminants (i.e., diphtheroids {*Corynebacterium* spp}, *Bacillus* {not *B. anthracis*} spp., *Propionibacterium* spp., coagulase-negative staphylococci, viridans group streptococci, *Aerococcus* spp., *Micrococcus* spp.)
- No positive culture of same organism from different infection site in past 30 days
- Two or more positive blood cultures collected within 2 days of each other drawn on separate occasions
- Fever >38°C

Definition Variation
Operating Room Procedure Time



Total procedure time

A + B + C =10?

Start of A to end of B = 5?

Obstacle #6: Construct Validity

- Do the data actually represent what you believe them to represent?

EXAMPLE: Has the patient received an antibiotic in the past six months?

Physician Order Written?	**YES**
Patient Charged for Drug?	**YES**
Drug Administered to Patient?	**NO**

Obstacle #7: Data Reliability

- A variable of interest may not be well coded
- Are missing/misclassified data systematic or random?

EXAMPLE: Urinary Catheterization

		Clinical data (Eclipsys XA)		
		Present	Absent	
Discharge data (ICD 9)	Present	11	1	12
	Absent	7	5	12

Obstacle #8: Free Text

- Do free text queries reliably capture the variable of interest?

EXAMPLE: **Nursing Home Residency**

Structured Field: **Patient admitted from Nursing Home?**
Yes / No

Free Text: **Patient's Home Address**
1060 Amsterdam Ave. Apt 8C
1060 Amsterdam Avenue #8C
1060 Amsteerdam Apt. 8C

1060 Amsterdam Avenue is a Nursing Home

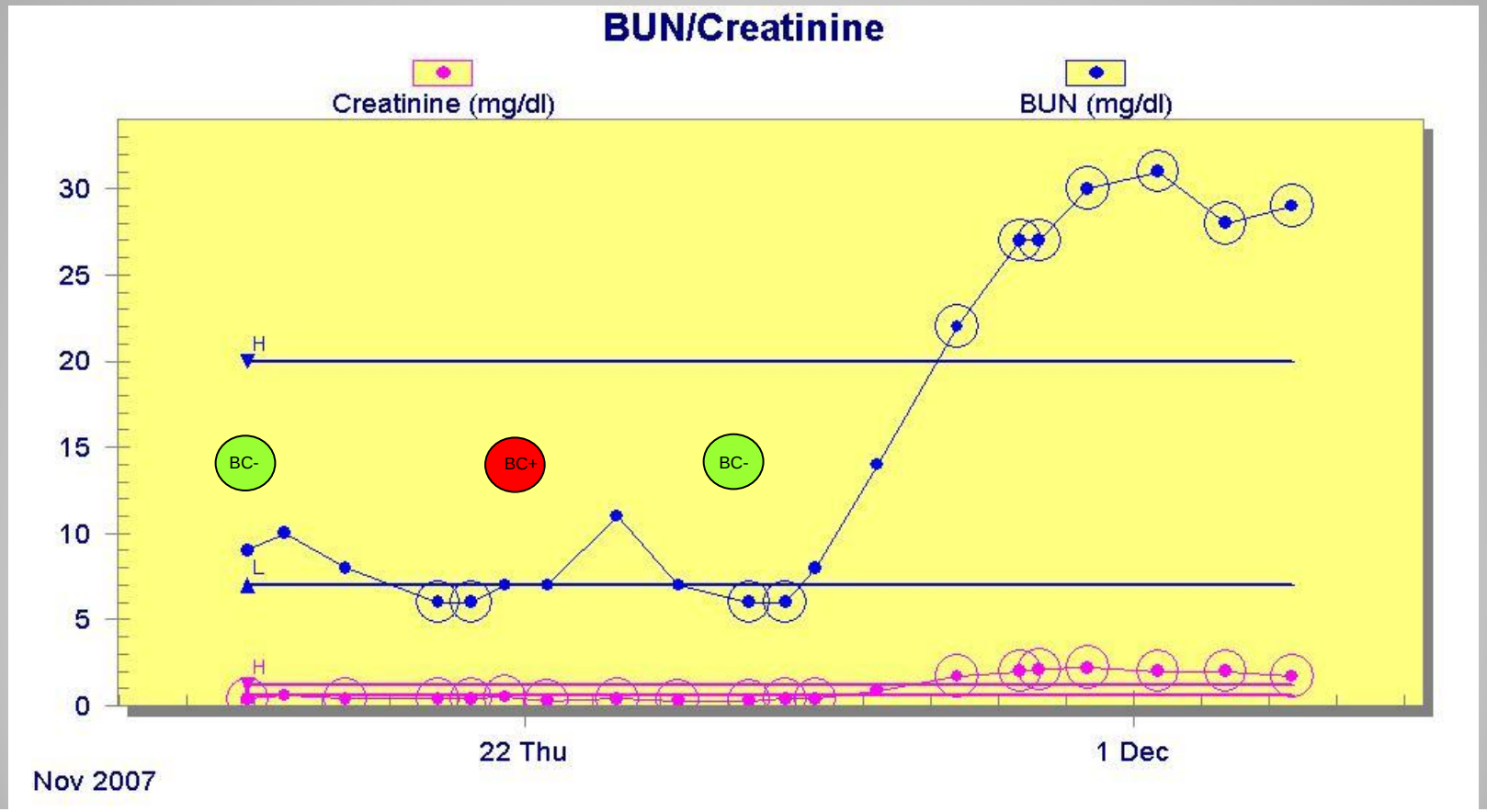
Obstacle #9: Time Stamps

- Is it possible to determine the sequence of events that occurred during a patient's stay?
- Can your data establish that a supposed risk factor occurred before disease onset?
- Was a condition present on admission, or did it develop during the patient's hospital stay?

EXAMPLE: Renal Failure

Time Stamps
Renal Failure

Acute renal failure Adm: 17 Nov, Disch: 4 Dec POA flag = No



BC: Blood culture

National Issues: Comparisons Across Hospitals

- Overcoming obstacles 1-9 within a single hospital does not ensure that the data will be comparable to that of another hospital

EXAMPLE: Differences in ICD-9 coding practices across hospital

National Issues: ICD-9 CM Documentation and Coding Variation
"CMS Won't pay for Medical Errors" Patient Example

Admit Source:

Admit Type: EMERGENCY/URGENT

MS DRG: 834

Acute leukemia w/o major O.R. procedure w MCC

Revised = 835, "w CC"

NYDRG: 576
ACUTE LEUKEMIA W MAJOR CC

APRDRG: 690
ACUTE LEUKEMIA

Mort. Risk: 3 **Sev. Level:** 3

Discharge Disp: ROUTINE

Financial Group: MCR FEE FOR SERVICE

Coded Date: 4/1/2008

Clin. Case Type HEMATOLOGIC MALIGNANCIES

Service Line: ONCOLOGY

CMI: 4.59

Rev 2.58

LOS: 36

Principal Diagnosis: 205.00
ACUTE MYELOID LEUK W/O REMISSN

Secondary Diagnosis:

		POA:	MS CC:
284.1	Pancytopenia	POA	CC
285.9	Anemia, unspecified	POA	---
287.4	Secondary thrombocytopenia	POA	---
288.03	Drug induced neutropenia	POA	---
311	Depressive disorder, not elsewher	POA	---
401.9	Essential hypertension; unspecifie	POA	---
414.01	Coronary atherosclerosis; of nativ	POA	---
707.03	Decubitus ulcer; lower back	HAC	MCC
V10.46	Personal history of malignant neo	POA	---
V45.82	Percutaneous transluminal corona	POA	---

# of Secondary Diagnoses: 10 **# HACs:** 1

Principal Procedure: 99.25
INJECT CA CHEMOTHER NEC

Secondary Procedures:

38.93	VENOUS CATHETER NEC
41.31	BONE MARROW BIOPSY
87.41	C.A.T. SCAN OF THORAX
99.04	PACKED CELL TRANSFUSION
99.05	PLATELET TRANSFUSION

# of Secondary Procedures: 5

National Issues: A Challenging Research Patient!

TRACHEOSTOMY W LONG TERM MECHANICAL VENTIL

Mort. Risk: 4 **Sev. Level:** 4 **LOS:** 192 **VAR:** 141.9

Principal Diagnosis: 710.0
SYST LUPUS ERYTHEMATOSIS

Principal Procedure: 84.11
TOE AMPUTATION

Secondary Diagnosis:		POA:	: MS CC	CMS Tgt:
038.0	Streptococcal septicemia	POA	MCC	0
112.5	Candidiasis; disseminated	HAC	MCC	0
286.9	Other and unspecified coagulation	POA	CC	0
344.09	quadriplegia and quadriparesis; ot	POA	MCC	0
403.91	Hypertensive renal disease; unspe	POA	CC	0
440.24	Atherosclerosis of the extremities	POA	CC	0
486	Pneumonia, organism unspecified	HAC	MCC	0
518.81	Acute respiratory failure	HAC	MCC	0
567.29	Other suppurative peritonitis	HAC	MCC	0
584.9	Acute renal failure, unspecified	HAC	MCC	0
585.6	End stage renal disease	POA	MCC	0
599.0	Urinary tract infection, site not sp	HAC	CC	4.
707.03	Decubitus ulcer; lower back	HAC	MCC	5.
707.14	Ulcer of heel and midfoot; chronic	POA	CC	0
785.52	Shock without mention of trauma;	HAC	MCC	0
995.92	Systemic inflammatory response s	POA	MCC	0
997.1	Cardiac complications	HAC	CC	0
997.3	Respiratory complications	HAC	CC	0
998.11	Hemorrhage complicating a proce	HAC	CC	0
998.12	Hematoma complicating a proced	HAC	CC	0
999.31	Infection due to central venous ca	HAC	CC	6.
999.39	Infection following other infusion,	HAC	CC	0
V46.11	Dependence on respirator, status	HAC	CC	0

# of Secondary Diagnoses: 23 **# HACs:** 15

- Medpar research file from claims (only 8 secondaries)
- Relationships to Clinical Databases (ie CLABSI, UTI, Pressure Ulcer)?
- UTI is a CMS “HAC” only when in combination with catheter-related UTI code (rarely documented)
- Acute Renal Failure: Risk or outcome depends on timing
- Variation across hospitals

National Issues

Acute Renal Failure ICD-9 Coding Variation Across AMCs

(Major Chest Procedures DRG; Q1 2008 Medicare)

BARNES JEWISH
BRIGHAM & WOMENS
CEDARS-SINAI
DUKE
EMORY
JOHNS HOPKINS
MASS GENERAL
MAYO CLINIC
METHODIST HOUSTON
MOUNT SINAI NYC
NORTHWESTERN
NYPH COLUMBIA
NYPH CORNELL
NYU
RUSH
ST LUKES HOUSTON
STANFORD
UCLA
UCSF
UNIV CHICAGO
UNIV MICHIGAN
UNIV PENNSYLVANIA
VANDERBILT
WASHINGTON

← (HIGHER) **Severity – Adjusted Payment Per Case** (LOWER) →

AMC:	1	2	3	4	5	6	7	25	26	27	28	29	30
5845	9%		7%	5%			2%		4%	3%	2%		
5849	14%	15%	7%	10%		3%	4%	3%		3%	2%	8%	
Total	23%	15%	14%	15%	0%	3%	7%	3%	4%	5%	3%	8%	0%

5845 Acute Renal Failure with Tubular Nephrosis

5849 Acute Renal Failure Not Otherwise Specified

Summary

- Although data access obstacles and analytic considerations have slowed the pace of progress, we have overcome!
  (Thanks to cooperation and goodwill from collaborating institutions and programmers!)

- Our team looks forward to the many research opportunities to come from routinely collected electronic data