PHIS+

(Pediatric Health Information Systems)

Ram Gouripeddi MD, MS
University of Utah

Bridging the Terminology Gap in Pediatrics
September 21, 2012
Acknowledgements

• FURTHeR development was supported by the NCRR and the NCATS, NIH, through Grant UL1RR025764 and supplement 3UL1RR025764-02S2. This project was funded under grant number R01 HS019862-01 from the AHRQ, U.S. Department of Health and Human Services (HHS). The opinions expressed [in this document] are those of the authors and do not reflect the official position of AHRQ or the HHS.

• PHIS+ Team members across multiple institutions
• FURTHeR Team members
• Apelon
• Augment Children’s Hospital Association’s (CHA) existing electronic database of administrative data - Pediatric Health Information System (PHIS) with clinical data to conduct Comparative Effectiveness Research studies.

• UU Biomedical Informatics - Informatics Partners

• Agency for Healthcare Research and Quality (AHRQ) PROSPECT funded project.
PHIS+ Overview

3 Data Streams
- Laboratory
- Microbiology
- Radiology

4 CER Studies
- Pneumonia
- Appendicitis
- Osteomyelitis
- Gastroesophageal Reflux Disease

5 Years Data
- 2007 – 2011
- 2009 – Development
- 2012....
The PHIS+ Process

Pediatric Research in Inpatient Setting (PRIS) Sites

1. Cincinnati Children’s Hospital Medical Center (CCHMC)
2. Children’s Hospital Boston (CHB)
3. Children’s Hospital of Philadelphia (CHOP)
4. Children’s Hospital of Pittsburgh (CHP)
5. Primary Children’s Medical Center, Intermountain Healthcare (PCMC)
6. Seattle Children’s Hospital (SCH)
FURTHeR = Federated Utah Research & Translational Health e-Repository
FURTHeR as a CER Infrastructure

On-the-fly query capability replaced with a data file adapter.

- PHIS+ Sites providing data as HL7 like text files.

Ability to store results to a physical database added.
# Electronic Data Sources for PHIS+ Hospitals

<table>
<thead>
<tr>
<th>Site</th>
<th>Lab Information System</th>
<th>Electronic Medical Record</th>
<th>PHIS+ Lab Datasource</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCHMC</td>
<td>Cerner Millenium</td>
<td>Epic</td>
<td>Epic Clarity</td>
</tr>
<tr>
<td>CHB</td>
<td>Cerner Pathnet</td>
<td>Cerner</td>
<td>In-house data warehouse</td>
</tr>
<tr>
<td>CHOP</td>
<td>Meditech</td>
<td>Epic</td>
<td>In-house data warehouse</td>
</tr>
<tr>
<td>CHP</td>
<td>Sunquest</td>
<td>Cerner</td>
<td>Cerner PowerInsight*</td>
</tr>
<tr>
<td>PCMC</td>
<td>Sunquest</td>
<td>In-house system</td>
<td>In-house data warehouse</td>
</tr>
<tr>
<td>SCH</td>
<td>Cerner Pathnet</td>
<td>Cerner</td>
<td>Cerner PowerInsight</td>
</tr>
</tbody>
</table>

Narus et. al, Federating Clinical Data from Six Pediatric Hospitals: Process and Initial Results from the PHIS+ Consortium. AMIA 2011
Developmental Process Overview

Narus et. al, Federating Clinical Data from Six Pediatric Hospitals: Process and Initial Results from the PHIS+ Consortium. AMIA 2011
Modeling & Terminology Analysis

• How do we want to model the data?
• Is it useful and generic enough to support all CER studies
• What is available with the sites?
  – Coded information
  – Free text entries (Micro)
  – Consistency of codes, multiple codes
  – Other associated metadata
  – Granularity
• What tests to include?

<table>
<thead>
<tr>
<th>Metadata Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Battery/Panel Name/Code</td>
</tr>
<tr>
<td>Battery/Panel Description</td>
</tr>
<tr>
<td>Local Test Name</td>
</tr>
<tr>
<td>Local Test Code</td>
</tr>
<tr>
<td>Test Description</td>
</tr>
<tr>
<td>LOINC Code</td>
</tr>
<tr>
<td>Test Value Type</td>
</tr>
<tr>
<td>Test Value Sample Data</td>
</tr>
<tr>
<td>Test Start Date Format</td>
</tr>
<tr>
<td>Test End Date Format</td>
</tr>
<tr>
<td>Specimen</td>
</tr>
<tr>
<td>Units of Measure</td>
</tr>
<tr>
<td>Reference Range</td>
</tr>
<tr>
<td>Interpretation Codes</td>
</tr>
<tr>
<td>Test Status Codes</td>
</tr>
<tr>
<td>Comments</td>
</tr>
</tbody>
</table>
# Lab Data Model

## Data Submission Model
- Sequence Number
- Hospital Number
- Campus ID
- Patient ID
- Billing Number
- Encounter Type
- Event ID
- Battery Code
- Test Code
- Test Name
- Test Description
- Test Specimen
- Value Type
- Value
- Units
- Start Date/Time
- End Date/Time
- Reference Range
- Reference Range Low
- Reference Range High
- Interpretation
- Comments

## PHIS+ Lab Model
- Generic Specimen Code
- Specific Specimen Code
- Specimen Type
- Body Site
- Units of Measure
- Interpretation Code

## Standard Terminology
- LOINC
- SNOMED
- HL7
Example Mappings

Laboratory Test

- C Reactive Protein (801679)
- C- Reactive Protein (801582)
- CRP Test (700111)

- C Reactive Protein (CRPT)
- C Reactive Protein (8726)
- C-Reactive Protein

Unit of Measure

- Nanogram/Decilitre (258805003)
- NG/DL
- ng/dL
- ng per dL

Laboratory Test

Unit of Measure
Microbiology Data

Patient → Specimen * → Culture 1

Patient → Specimen 1 → Culture * → Micro-organism 1

Patient → Specimen 1 → Culture * → Anti-microbial Susceptibility 0..* → Micro-organism 1

Patient → Specimen 1 → Culture * → Anti-microbial Susceptibility 0..* → Micro-organism 1
Modeling and Terminology

• Team: Pediatric Infectious Diseases and Hospital Medicine Specialists, Information Technology Personnel, Microbiologists, Informaticists.

• Specimen: Specimen Type & Body Site

• Micro-organisms
  – Species vs. genus levels (e.g., *Staphylococcus aureus* vs. *Staphylococcus* species).
  – Microbiological classification terms (e.g., coagulase negative *Staphylococcus*, Gram-negative rods)
  – Negation & Qualifiers – “Not abnormal culture result”
    • “no organism or growth seen”, “organism X not seen” or “normal flora”, “no growth seen after two days”.

• Preliminary and Final Reports with appropriate date-times
  – Gram-positive cocci → *Staphylococcus aureus*.
PHIS+ Microbiology Model

Submission Data Model

PHIS+ Micro Data Model

<table>
<thead>
<tr>
<th>PDE</th>
<th>Patient Data Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpDE</td>
<td>Specimen Data Element</td>
</tr>
<tr>
<td>CDE</td>
<td>Culture Data Element</td>
</tr>
<tr>
<td>ODE</td>
<td>Organism Data Element</td>
</tr>
<tr>
<td>SDE</td>
<td>Susceptibility Data Element</td>
</tr>
</tbody>
</table>
# PHIS+ Microbiology Model

<table>
<thead>
<tr>
<th>Submission Model</th>
<th>PHIS+ Micro Model</th>
<th>Standard Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Element</strong></td>
<td><strong>Field Name</strong></td>
<td><strong>Field Name</strong></td>
</tr>
<tr>
<td>Culture (CDE)</td>
<td>Specimen</td>
<td>Specimen</td>
</tr>
<tr>
<td>Body Site</td>
<td>Body Site</td>
<td>Body Site</td>
</tr>
<tr>
<td>Local Culture Code</td>
<td>Culture</td>
<td>Culture</td>
</tr>
<tr>
<td></td>
<td>Stain</td>
<td>Stain</td>
</tr>
<tr>
<td>Culture Result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture Normalcy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism (ODE) Susceptibility (SDE)</td>
<td>Organism Code</td>
<td>Organism Code</td>
</tr>
<tr>
<td>Local Antibiotic Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Susceptibility Test Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Susceptibility Test Interpretation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gouripeddi R et al. Federating Clinical Data from Six Pediatric Hospitals: Process and Initial Results for Microbiology from the PHIS+ Consortium
Example Mappings

Culture Study

Blood Culture

BCX

Blood Culture

CULTURE, BLOOD

5862

4349382

747233

Organism

Staphylococcus aureus

Staphylococcus aureus

SA

SAUR

4

750570

SA
# Radiology Data Model

<table>
<thead>
<tr>
<th>Harmonized Model Field</th>
<th>PHIS+ Radiology Model</th>
<th>Standard Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence Number</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital Number</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campus ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Billing Number</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession Number</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encounter Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Code</td>
<td></td>
<td>Radiology Test Code</td>
</tr>
<tr>
<td>Test Name</td>
<td></td>
<td>CPT</td>
</tr>
<tr>
<td>Test Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start Date/Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>End Date/Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local CPT Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Report Text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Free Text Fields*
Integrating NLP Extracts

Results
Examination: Chest, one view
Clinical Indication: Extracorporeal life-support, pneumonia
Comparison: **DATE
Findings/impression: Endotracheal tube is in unchanged position; it terminates at the level of T4-T5
Other lines and tubes in unchanged positions.

Bilateral diffuse pulmonary opacification has increased, with near complete opacification of both lungs. Subcutaneous edema persists. The heart size cannot be evaluated.
Resident: **NAME[CCC, DDD], MD, MS
I have personally reviewed this study and agree with the report above.
SEP TXL Radiologist: **NAME[EEE, FFF], MD

NLP: Annotation & Information Extraction

Structured Clinical Data

Data for CER Study

Corpora Export based on Clinical Research Criteria

FURTHeR

Data
No Match to a Standard Terminology

• < ~ 0.1%
• Pseudo-extensions of SNOMED, LOINC - PHISPLUS
• Actively managed.
• Submission to IHTSDO & Regenstrief Institute.
• E.g. Strains of Enterococcus faecalis, Lyme Index Value, Immune Status Ratio
## PHIS+ CER Database - 2009

### Laboratory

<table>
<thead>
<tr>
<th>Site</th>
<th>Records</th>
<th>LOINC Lab Test Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6040786</td>
<td>427</td>
</tr>
<tr>
<td>B</td>
<td>1680867</td>
<td>390</td>
</tr>
<tr>
<td>C</td>
<td>3872170</td>
<td>394</td>
</tr>
<tr>
<td>D</td>
<td>834836</td>
<td>323</td>
</tr>
<tr>
<td>E</td>
<td>2757546</td>
<td>751</td>
</tr>
<tr>
<td>F</td>
<td>3036648</td>
<td>473</td>
</tr>
<tr>
<td>Total</td>
<td>18222853</td>
<td>2758 (1202)</td>
</tr>
</tbody>
</table>

### Radiology

<table>
<thead>
<tr>
<th>Site</th>
<th>Records</th>
<th>CPT Radiology Procedure Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>164490</td>
<td>281</td>
</tr>
<tr>
<td>B</td>
<td>225490</td>
<td>327</td>
</tr>
<tr>
<td>C</td>
<td>188886</td>
<td>451</td>
</tr>
<tr>
<td>D</td>
<td>90020</td>
<td>453</td>
</tr>
<tr>
<td>E</td>
<td>44121</td>
<td>358</td>
</tr>
<tr>
<td>F</td>
<td>106453</td>
<td>270</td>
</tr>
<tr>
<td>Total</td>
<td>819460</td>
<td>2140 (524)</td>
</tr>
</tbody>
</table>

### Microbiology

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>101379</td>
<td>114</td>
<td>70</td>
<td>113</td>
<td>57</td>
<td>97</td>
</tr>
<tr>
<td>B</td>
<td>149075</td>
<td>58</td>
<td>42</td>
<td>56</td>
<td>58</td>
<td>85</td>
</tr>
<tr>
<td>C</td>
<td>84414</td>
<td>179</td>
<td>46</td>
<td>162</td>
<td>59</td>
<td>99</td>
</tr>
<tr>
<td>D</td>
<td>51703</td>
<td>110</td>
<td>34</td>
<td>145</td>
<td>57</td>
<td>75</td>
</tr>
<tr>
<td>E</td>
<td>38299</td>
<td>130</td>
<td>56</td>
<td>160</td>
<td>59</td>
<td>76</td>
</tr>
<tr>
<td>F</td>
<td>60306</td>
<td>264</td>
<td>71</td>
<td>121</td>
<td>51</td>
<td>89</td>
</tr>
<tr>
<td>Total</td>
<td>485176</td>
<td>855 (451)</td>
<td>319 (95)</td>
<td>757 (203)</td>
<td>341 (74)</td>
<td>521 (136)</td>
</tr>
</tbody>
</table>

*Distinct standard codes across all sites in ().
File Processing & Terminology Mapping

- Ask sites to provide a metadata file for each load.
- Analyze the metadata file with what is currently available in the Terminology server.
- Add new/edit metadata.
- Process data load with FURTHeR.
- Is the log error free?
  - Yes: Processing Complete.
  - No: Ask additional metadata from sites if necessary & fix data errors.
- Analyze log file for unmapped terms & data errors.
- Check mappings.
Conclusion

CER Infrastructure to Federate & Integrate Data from Disparate Sources.

- Validation Phase

Standard Terminologies in the Federating Models.

- HITSP Recommendations

Most Local Terms to Standards Mappings possible.

Terminology choice depends on

- Metadata availability
- Modeling Choices

Future Steps

- Provide for Versioning of local and standard terminologies.
- Semi-automated Mapping esp. For Free Text Entries.
References

• Gouripeddi R, Mitchel JA, Narus SP et al. Federating Clinical Data from Six Pediatric Hospitals: Process and Initial Results for Microbiology from the PHIS+ Consortium AMIA Annu Symp Proc. 2012 (Accepted).


THANK YOU
Software Architecture Overview

Parse → Build Entities → Resolve Dependencies → Translate → Persist → Terminology Server

Micro

Courtesy: Phillip Warner & Peter Mo
Microbiology Dependency Resolution

Abstract Factory
PDE, CDE, ODE, SDE

Build Entity
- PDE
- CDE
- ODE
- SDE

Has Dependencies

Generate List

Dependency Resolution
- Hashset
- EHCache
  - During Build
  - Before Translation

Courtesy: Phillip Warner & Peter Mo
Data Quality Measures

- Parse errors
  - Newlines
  - Date formats
- Improperly formatted lines
- Orphans lines
- Duplicate lines
- Missing data for required fields