Automating the Contextualization of Population-Based CDSS in Tethered EHR/PHRs

Hadi Kharrazi, MHI, MD, PhD
Johns Hopkins University
hkharrazi@jhsph.edu
Disclosures

- Hadi Kharrazi reports having a financial relationship with Indiana University from Aug 2008 to Aug 2012.

- Hadi Kharrazi reports having a financial relationship with Johns Hopkins University since Aug 2012.

- Hadi Kharrazi reports that he has no financial relationship with any commercial interest.

Speaker: Hadi Kharrazi, MHI, MD, PhD

Health Sciences Informatics Grand Rounds

Title: Automating the Contextualization of Population-Based CDSS in Tethered EHR/PHRs

Lecture Objectives

- Describe the types of HIT solutions (e.g., EHR, PHR) and how they complement Population Health IT (PopHI)
- Explain the relationship among CPGs, CDS languages, and CDS systems
- Demonstrate how CDSS can be automated and applied in the context of PopHI
- Explain how population-based CDSS can bridge the gap between clinical informatics and PubHI
- Discuss the opportunities and challenges of population-based CDSS

Speaker: Hadi Kharrazi, MHI, MD, PhD

The presentation in a nutshell

- Introduction

- Background
  - Clinical Informatics (CI)
  - Public Health Informatics (PubHI)
  - Health Information Exchange (HIE) role in PubHI
  - Population Health Informatics (PopHI) - bridging the gap between CI and PubHI
  - Clinical Decision Support (CDS) continuum and PopHI

- Regenstrief Institute / INPC / Indiana HIE

- Previous Work
  - Developing renewable CDS systems
  - Application and results of CDS automation in INPC RMRS
  - Contextualizing CDSS in CHI Platforms (PHR)
  - Population-based CDSS

- Center for Population Health IT (CPHIT) / Future Work

- Q & A
Introduction
Introduction

- **Hadi Kharrazi**, MHI, MD, PhD

“You are a doctor who is lost in cyberspace!” *(my wife)*

Current affiliation:  
JHSPH-HPM (Aug 2012)  
Associate Director CPHIT  
*Director: Dr. J. Weiner*

Previous affiliation:  
Indiana University  
Dalhousie University

Research interests:  
Population Health IT  
- *CDSS/QM in HIE environments*  
- *Personal Health Records (PHR)*
Introduction

Biomedical informatics as a basic science

Basic Research

Applied Research

Molecular Research

Health Research

Bioinformatics

Imaging Informatics

Clinical Informatics

Consumer Health Informatics

Public Health Informatics

PopHI

Biomedical informatics methods, techniques, and theories
Background
Background > Clinical Informatics (CI)

Clinical Information Systems
- Administrative
- Accounting
- Clinical Inpatient
- Clinical Outpatient
- Laboratory
- Imaging

CDSS
- EHR
- EMR
- JHH EPR
- QM, KD...

If patient has x, y then do z

Non-clinical
Background > Public Health Informatics (PubHI)

Periodic public reports and alerts

EHR

1

2

Registries

Commun. Diseases

Reports

Claims

Non-clinical

Syndromic Surveillance

Quality Measures

Public Health

© Hadi Kharrazi @ JHSPH-HPM
Background > Health Information Exchange (HIE)
Background > Population Health Informatics (PopHI)

Population Health Info.

Population Health Analytics

Population Health Data-warehouse

Personal Health Records

Insurance Data

Rx Data

Admin data repositories

PubHI to CI

CI to PubHI

PubHI to CI

CI to PubHI

Clinical Informatics

EHRs

1...n

HIE

Public Health Informatics

A

B

C

Public Health
If patient has x, y, has gene w, lives in Baltimore, is high risk, is Hispanic, has insurance q, and more population health rules..., and you don’t have an MRI then do z’

If patient has x, y and then do z

If patient has x, y and you don’t have an MRI then do z’

If patient has x, y and is in high risk population then do z’

If patient has x, y and is Hispanic then do z’

If patient has x, y and lives in Baltimore then do z’

If patient has x, y, and gene w then do z’

If patient has x, y, has gene w, lives in Baltimore, is high risk, is Hispanic, has insurance q, and more population health rules...

If patient has x, y, then do z

If patient has x, y, has gene w, lives in Baltimore, is high risk, is Hispanic, has insurance q, and more population health rules...
Regenstrief Institute / INPC / Indiana HIE
• Established in 1969 by philanthropist Sam Regenstrief
• Regenstrief receives $3 million per year in core support from the Regenstrief Foundation
• Annual operating budget of approximately $23 million derived from grants and contracts
• Logical Observation Identifiers Names and Codes (LOINC) system, a standard nomenclature that enables the electronic transmission of clinical data from laboratories
• Regenstrief Medical Records System (RMRS) was developed 35 yrs ago. RMRS has a database of 6 million patients, with 900 million on-line coded results, 20 million full reports including diagnostic studies, procedure results, operative notes and discharge summaries, and 65 million radiology images.
• Indiana Network for Patient Care (INPC) was created in 1996. It is a city-wide clinical informatics network of 11 different hospital facilities and more than 100 geographically distributed clinics and day surgery facilities
The Indiana HIE (IHIE) includes (as of mid-2011):

- Federated Consistent Databases
- 22 hospital systems → ~70 hospitals
- 5 large medical groups and clinics & 5 payors
- Several free-standing labs and imaging centers
- State and local public health agencies
- 10.75 million unique patients
- 20 million registration events
- 3 billion coded results
- 38 million dictated reports
- 9 million radiology reports
- 12 million drug orders
- 577,000 EKG tracings
- 120 million radiology images
Regenstrief Institute > PubHI > ELR Data Completeness

Electronic Laboratory Reporting (INPC)

Passive Reporting (Health Dept.)

Passive Reporting (Hospital)

Overhage JM, Grannis S, McDonald CJ. Am J Public Health. 2008 Feb
Regenstrief Institute > PubHI > Timeliness of Data (CC vs ICD9)

![Histograms showing different lags (A, B, C, D) for days from visit with frequency distribution.](image_url)

Copyright Regenstrief Institute
Regenstrief Institute > PubHI > Syndromic Surveillance Service (SSS)
Previous Work
Previous Work > Developing Renewable CDSS

Clinical Practice Guidelines

Computerized CPG (CARE, Arden syntax)

Manual Programming for each CPG

One patient at the time

Relational DB

Popup alert

Patient has diabetes and has not had an eye exam in two years. Based on guideline xxx you may want to consider asking for an eye consultation.

```php
$results = mysql_fetch();
$patient_id = $results['patient_id'];
mysql_query ('SELECT * FROM lab WHERE patient_id = $patient_id and lab_term = 'HBA1c');
$results = mysql_fetch();
$lab_result = $results['lab_term'];
IF ($lab_result > 9)
  ECHO “Patient has a high HBA1c level”;
```
Previous Work > Developing Renewable CDSS

- **Issues with current CDSS**

  - Often successful CDS RCTs are not replicable (renewable) in other clinical settings.
  
  - CDS knowledge interchange standards have existed for as long as clinical data interchange standards (about 2 decades), yet sharing of computable, actionable knowledge in practice is extremely limited
  
  - Current computerized knowledge languages representing CPGs are not standardized: Arden syntax, CARE, GELLO, GLIF
  
  - The majority of the current CDSS systems are hard coded, as experienced in a simpler format in CHI studies, which make them very limited due to maintenance issues.
Previous Work > Developing Renewable CDSS

Clinical Practice Guidelines

Computerized CPG (CARE, Arden syntax)

Manual Programming for each CPG

One patient at the time

Relational DB 1

CDSS SQL 1

NQF QM eMeasures

Automated for all CPGs

All patient at the time

Populay Health Informatics

CDSS SQL 2

Relational DB 2

CDSS SQL...

Relational DB...

CDSS SQL n

Relational DB n
Previous Work > Developing Renewable CDSS

- **Phases to Develop Renewable CDSS**

  1. **Lexical Integration** (mapping of idiosyncratic codes to standard terms)
  2. **Syntactic Mapping** of Different Concrete Representation Languages
     - Parse original content as XML (e.g., Chaperon)
     - XSLT Transformation (e.g., SAXON)
  3. **Semantic Integration** of Procedural Knowledge into Declarative Rules (SQL commands)
  4. **Pragmatic Integration** (query enhancements, implementation and evaluation)
Previous Work > Developing Renewable CDSS

CPG Consortium

CPG

NQF

CARE Arden

XML

Lexical Integration

Syntactic Mapping

Semantic Integration

XSLT trans.

Generating Renewable CDSS

CDSS SQL 1

Relational DB 1

CDSS SQL 2

Relational DB 2

CDSS SQL...

Relational DB...

CDSS SQL n

Relational DB n

Pragmatic Integration
Previous Work > Developing Renewable CDSS

Language Specific

Language Free
Previous Work > Developing Renewable CDSS

BEGIN BLOCK CHOLESTEROL

DEFINE "MALE_PT" {VALUE} AS "SEX" IS = 'M'

DEFINE "CHOLESTEROL LAST" AS LAST "CHOL TESTS" [BEFORE "DATE"] EXIST

IF "CHOLESTEROL LAST" WAS AFTER 5 YEARS AGO THEN EXIT CHOLESTEROL ELSE CONTINUE

IF "MALE_PT" EXISTS AND "AGE" WAS >= 35 THEN REMIND Male patients over the age of 35 should have their total cholesterol measured every 5 years. AND ORDER CHOLESTEROL (SCREEN) AND SAVE "UNO" AS "CHOL_SCN" {INTEGER} AND EXIT CHOLESTEROL

END BLOCK CHOLESTEROL

Sample CARE Rule Block
Previous Work > Developing Renewable CDSS

```xml
<?xml version="1.0" encoding="UTF-8"?>
<rule xmlns="http://regenstrief.org/xcare">
  <let name="AGE">
    <and>
      <filter>
        <reference name="BIRTH"/>
        <and>
          <not-equal op="NE" component="VALUE">
            <literal type="integer" value="0"/>
          </not-equal>
          <less-than op="LT" component="VALUE">
            <literal type="time" value="now"/>
          </less-than>
        </and>
      </filter>
      <division>
        <subtraction>
          <literal type="time" value="now"/>
          <reference name="BIRTH"/>
        </subtraction>
        <literal type="real" value="365.25"/>
      </division>
    </and>
  </let>
</rule>
```

Intermediate XML
Previous Work > Developing Renewable CDSS

SELECT id, sys_id_value, numeric_value, time_value, text_value, boolean_value, code_value, data_type, institution_id, patient_id, service_code, primary_time
FROM(
    select t_from.* from (SELECT
        id, sys_id_value, numeric_value, time_value, text_value, boolean_value, code_value, data_type, institution_id, patient_id, service_code, primary_time,
        myrank
    FROM(
        SELECT * FROM (SELECT
            id, sys_id_value, numeric_value, time_value, text_value, boolean_value, code_value, data_type, institution_id, patient_id, service_code, primary_time,
            rank() OVER (PARTITION BY institution_id, patient_id ORDER BY primary_time DESC) myrank
        FROM (select TO_NUMBER(NULL) as id, To_NUMBER(NULL) as sys_id_value, 1 as numeric_value, DATE_OF_BIRTH as time_value, TO_CHAR(NULL) as text_value, To_CHAR(NULL) as boolean_value, TO_CHAR(NULL) as code_value, 'T' as data_type, ...

**SQL Statement (adaptable for new schemas)**
Previous Work > Developing Renewable CDSS

**Limitations**

- The process does not start directly with free text CPG. Knowledge Representation languages such as CARE or Arden are required.

- If originating CPG does not conform to a standard terminology, lexical integration is required again.

- Fuzzy or probabilistic representational languages are currently not incorporated.

- Mismatched data granularity or ignored data sets will create database integration complicated.

- Each CDSS SQL is mutually exclusive of other CDSS thus limiting the merger of SQLs that can result in a combined CDSS.
Previous Work > Developing Renewable CDSS

- **Application Areas**

  - Clinical: CDSS reminder rules (e.g., CPOE integration) → popup CDSS alerts for providers

  - Population-based:
    - Healthcare quality measures (e.g., NQF, ACO) → Quality monitoring dashboard
    - Consumer Health Informatics (e.g., tethered PHRs)
    - Public Health Informatics (PHI) (e.g., notifiable complex conditions)
    - Transition of care (e.g., CCD-based CDSS for inpatient to outpatient)
Previous Work > CDSS Application in EHR

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Variable</td>
<td>42,666,350 records</td>
</tr>
<tr>
<td>Patient</td>
<td>453,518 (unique)</td>
</tr>
<tr>
<td>Institution</td>
<td>118 providers</td>
</tr>
<tr>
<td>Care KRL</td>
<td>74 Rules (SQL)</td>
</tr>
</tbody>
</table>

RMRS sample database was used in this research

a subset of the INPC database
### Previous Work > CDSS Application in EHR

<table>
<thead>
<tr>
<th>Rules</th>
<th>Population</th>
<th>One patient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4_1.sql</td>
<td>59</td>
<td>0.18</td>
</tr>
<tr>
<td>CHOLESTEROL_1.sql</td>
<td>180</td>
<td>0.11</td>
</tr>
<tr>
<td>CHOLESTEROL_2.sql</td>
<td>1</td>
<td>0.12</td>
</tr>
<tr>
<td>DIABETES_1.sql</td>
<td>541</td>
<td>0.56</td>
</tr>
<tr>
<td>DIABETES_2.sql</td>
<td>1561</td>
<td>0.47</td>
</tr>
<tr>
<td>DIABETES_3.sql</td>
<td>182</td>
<td>0.40</td>
</tr>
<tr>
<td>DIABETES_4.sql</td>
<td>121</td>
<td>0.40</td>
</tr>
<tr>
<td>DIABETES_5.sql</td>
<td>1620</td>
<td>1.51</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TCL_2.sql</td>
<td>67</td>
<td>0.12</td>
</tr>
<tr>
<td>TCL_3.sql</td>
<td>69</td>
<td>0.18</td>
</tr>
<tr>
<td>TCL_4.sql</td>
<td>61</td>
<td>0.17</td>
</tr>
<tr>
<td>TETANUS_SHOT_1.sql</td>
<td>1</td>
<td>0.07</td>
</tr>
<tr>
<td>WHOLE_1.sql</td>
<td>127</td>
<td>0.16</td>
</tr>
<tr>
<td>WHOLE_2.sql</td>
<td>129</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Average (sec)</strong></td>
<td><strong>1249</strong></td>
<td><strong>0.73</strong></td>
</tr>
</tbody>
</table>

**Runtime per second**

* each row shows average of 20 individual runs
Previous Work > CDSS Application in EHR

Institute 1: All CDS Hits (regardless of missed proportion)

Applying all applicable CARE rules for Institute-1

Some rules are never hit for institute 1 (not shown in this diagram)
Previous Work > CDSS Application in EHR

Number of CARE rule hits against the RMRS sample database based on each healthcare provider / institute (zero-hit institutes are removed)

*Academic/Primary hospitals have the highest rates of hits*  
This peak might be simply because of their higher population of patients (needs population adjustment)
Previous Work > CDSS Application in EHR

Number of CARE rule hits against the RMRS sample database based on ever healthcare provider institute (population adjusted percentage)

Percentage of rule hits is suddenly smaller in the left side of the diagram (i.e., primary hospitals)

Percentage of rule hits is larger in the right side of the diagram (i.e., other hospitals)

Percentage > creates higher patient risk at a certain institute
Raw number > forms internal policies / economical road maps
Previous Work > CDSS Application in EHR

Institute versus Rules (Raw Numbers)
*(zero rules / zero institutes are removed)*
Previous Work > CDSS Application in EHR

Institute 1: Percentage of Hit / Miss Rules

Missed applicable CARE rules for institute-1
(CDSS-generated QM fingerprint)

Manual process of deciphering the rules (hit-means-miss OR hit-means-hit-only)
The lower the better (less missed)
Previous Work > CDSS Application in EHR

Institute 2: Inverse Percentage of Hit / Miss Rules

Missed applicable CARE rules for institute-2
(CDSS-generated QM fingerprint)

Manual process of deciphering the rules (hit-means-miss OR hit-means-hit-only)
The lower the better (less missed)
The effect size of the missed rules depends on the population size.
Previous Work > CDSS Application in EHR

Web Interface
(http://aurora.regenstrief.org:8080/ClinicalKnowledgeHub/)
Previous Work > CDSS Application in EHR

ATTENTION: The Indiana Network for Patient Care information system has determined that Josephine Test is a diabetic patient, who has no evidence of a microalbumin urine examination completed within the past 12 months. National recommendations for the care of those with diabetes suggest that all patients receive yearly microalbumins to ensure proper renal health.

Please note your response below and fax this form back to (317) 555-5545:

☐ Test ordered
☐ Don’t agree with recommendation
☐ Patient refused test
☐ Patient not under my direct care
☐ Test already completed
☐ Reminder is incorrect

Preventive Care Quality Improvement → Reminder Forms

Copyright Regenstrief Institute
Previous Work > CDSS Application in PHR
Previous Work > CDSS Application in PHR

**TELCARE Framework**
- 2-way Cellular Glucose Meter
- Telcare Web Portal and Database

**IU Health Informatics Framework**
- IU Health Informatics CDSS Database
- Knowledge Representation Syntax (derived from CPG)

**Stakeholders**
- Patient
- Caregiver
- Physician

**CHI Framework (Tx Group)**

Automatic generated SQL commands
Dynamically generated CDSS SQL commands
### Previous Work > Population-based CDSS

<table>
<thead>
<tr>
<th>Item</th>
<th>Traditional</th>
<th>CDSS-based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Focus</strong></td>
<td>n-Patient-at-a-time</td>
<td>1-Patient-at-a-time</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>Statistical (probabilistic)</td>
<td>CPG (rule-based)</td>
</tr>
<tr>
<td><strong>Updatability</strong></td>
<td>Delayed (Hard coded)</td>
<td>Immediate (Dynamic)</td>
</tr>
<tr>
<td><strong>HIE Integration</strong></td>
<td>Possible</td>
<td>Already Exists</td>
</tr>
<tr>
<td><strong>Portability</strong></td>
<td>Major Change Req.</td>
<td>Minimal Change Req.</td>
</tr>
<tr>
<td><strong>Data Source</strong></td>
<td>EHR Subsets (limited)</td>
<td>Complete EHR, PHR, ...</td>
</tr>
<tr>
<td><strong>Contextualization</strong></td>
<td>Complex</td>
<td>SQL Driven</td>
</tr>
</tbody>
</table>

**Traditional PubHI versus CDSS-based PopHI**
Previous Work > Population-based CDSS

Simple statistics based on numerical statistics

Consider the potentials with Population-based CDSS integration
CPHIT / Future Work
The Johns Hopkins Center for Population Health Information Technology (CPHIT, or “see-fit”)

- The **mission** of CPHIT is to improve the health and wellbeing of populations by advancing the state-of-the-art of Health Information Technology and e-health tools used by public health agencies and private health care organizations.

- CPHIT’s **focus** will be on the application of EHRs, PHRs and other digitally-supported health improvement interventions targeted at communities, special need populations and groups of consumers cared for by integrated delivery systems.

- **Director:** Dr. J. Weiner

- [www.jhsph.edu/cphit](http://www.jhsph.edu/cphit)
CPHIT / Future Work

CPHIT Organizational Linkages

- JH Health System
- JHU Academic Departments and other R&D centers
- CPHIT
- External PH/IDS Orgs.
- JH Healthcare Solutions, LLC
- Business Partners
- Industry Foundations Government
- Government
- Foundations
- Industry
- Business
- Partners
- CPHIT
- JH Health System
- JHU Academic Departments and other R&D centers
- CPHIT
- External PH/IDS Orgs.
- JH Healthcare Solutions, LLC
- Business Partners
- Industry Foundations Government
- Government
- Foundations
- Industry
- Business
- Partners

Copyright Dr. J. Weiner
CPHIT / Future Work

- **Meaningful Use**

HIT-Enabled Health Reform

- Stage 1: 2009 HITECH Policies
- Stage 1: 2011 Criteria (Capture & Share)
- Stage 2: 2013 Criteria (CDSS)
- Stage 3: 2015 Criteria (Improved Outcome)

Population Health Informatics
Federated Consistent Databases

(1) Data gathered centrally in separate physical files, “mirrors” of remote sites;
(2) Standardized at the time it comes in.

Population-based CDSS across databases
Automating Population-Based CDSS ...
CPHIT / Future Work

- **CPHIT’s initial R&D Priorities:**
  - Natural language processing (NLP) and pattern recognition that improve population based interventions.
  - Effective approaches for integration of consumer-focused m-health/PHRs/e-health with provider-focused EHRs.
  - Development of “e-measures” of population health status, risk, safety/quality using EHR/HIT systems.
  - EHR based tools to identify high risk populations for preventive and/or chronic care outreach.
  - Approaches for integrating current functions of public health agencies into interoperable EHR networks.
  - Approaches for applying HIT to population-based evidence-generation / research.
Thank you

Q & A