Research Access to Clinical Data

BRIITE Conference
Fred Hutchinson Cancer Research Center
September 11, 2003
Bill Lober, MD
University of Washington
Not an Oncologist

- Electrical/Computer Engineering
- Boarded in Emergency Medicine
- Faculty in Biomedical and Health Informatics
  - Public Health Informatics
    - Bioterrorism Surveillance
    - Integration of PH Reporting Systems
  - Clinical Informatics
    - Clinical Data Interfaces for Research Systems
    - “Specialty” clinical and research infrastructure
  - Enterprise CIS @ UW Medicine
Bob’s goal

- Look at “IT in an R01 setting”
- How clinical data access is affected by:
  - Funding/Organizational goals
  - Security/Auth/Authz
  - Distribution and dissemination
  - Data representation/management
- Develop a framework to facilitate working group discussions
Bob’s questions

• What issues control/limit/promote research use of clinical data?
• Why is it so darn hard to figure out?

• We spend so much $$ on these clinical systems, why aren’t research apps just the icing on the cake…?
My talk

- Issues in using clinical data
- Organize issues into framework
- Succumb(ed) to pessimism
- Raise difficult issues
  - and then

Duck...
Issues

• What data?
• What uses for it?
• What is hard?
• Research is different, right?
What data are we talking about?

• Demographics
• Labs
• Structured
• Unstructured…
  – Clinic Notes, Path Reports, Images…
Patient Demographics

Age: 53 years
Born: 01-Apr-1950
SSN: 000-64-0027
Occupation: RETIRED

Language: unavailable
Race: Black
Home phone: (206) 111-1111
Work phone: (206) 598-7775

1015 BLOWHOLE COURT
ELECTRIC CITY, WA 99123 9218

Medicaid Pic Number: CL042683PEEK82
Medicare Number: 123456789A

Mother's Maiden Name: HAROLDSON
Father's Name: unavailable

internal PCP - PETER TARCZYHORNOCH
internal PCP - DOUGLAS FAAUW

Next of kin:
JOHN DOE

Patient ID at External Hospitals:
A0016999999 at Northwest Hospital
ATEST1 at Children's Hospital
## Blood Gas Panel 1, ART

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimen Description</td>
<td>Arterial</td>
<td></td>
</tr>
<tr>
<td>Fi (O2) %</td>
<td>100 %</td>
<td></td>
</tr>
<tr>
<td>PEEP (Pos End Expired Air)</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>pH, ART</td>
<td>7.35</td>
<td></td>
</tr>
<tr>
<td>PacO2, ART</td>
<td>60 mmHg</td>
<td>Cerner Testing, Called to Virtual Floor</td>
</tr>
<tr>
<td>PacO2, ART</td>
<td>120 mmHg</td>
<td></td>
</tr>
<tr>
<td>Base Excess, ART</td>
<td>5.2 mEq/L</td>
<td></td>
</tr>
<tr>
<td>Base Deficit, ART</td>
<td>1.1 mEq/L</td>
<td></td>
</tr>
<tr>
<td>Sodium, WBLD</td>
<td>135 mEq/L</td>
<td></td>
</tr>
<tr>
<td>Potassium, WBLD</td>
<td>3.6 mEq/L</td>
<td></td>
</tr>
<tr>
<td>Calcium (Ionized), WBLD</td>
<td>1.23 mmol/L</td>
<td></td>
</tr>
<tr>
<td>Glucose, WBLD</td>
<td>150 mg/dL</td>
<td></td>
</tr>
<tr>
<td>Glucose Specimen Type</td>
<td>Jackson Pratt</td>
<td></td>
</tr>
<tr>
<td>Protein (Total), FLD</td>
<td>40 mg/dL</td>
<td></td>
</tr>
<tr>
<td>Cell Count, Body Fluid</td>
<td>100 cells/mL</td>
<td></td>
</tr>
<tr>
<td>Erythrocyte Sedimentation</td>
<td>60%</td>
<td></td>
</tr>
</tbody>
</table>

*Note: If you need further information or assistance, please contact Community Services at (206) 398-6066.*
<table>
<thead>
<tr>
<th>ref.</th>
<th>Last ActvDt</th>
<th>Start Dt</th>
<th>CurrentD/CExtMed</th>
<th>Medication Name, Strength, Form</th>
<th>Sig</th>
<th>RefsCat</th>
<th>Left</th>
<th>Prescriber</th>
<th>Location or Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>19-aug-03</td>
<td>19-aug-03</td>
<td>✔️</td>
<td></td>
<td>Trimethoprim/Sulfamethoxazole (Oph) Sol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13-aug-03</td>
<td>13-aug-03</td>
<td></td>
<td></td>
<td>Aspirin 600Mg Sup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13-aug-03</td>
<td>13-aug-03</td>
<td></td>
<td></td>
<td>Docusate 100Mg/Ml</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-jul-03</td>
<td>01-jul-03</td>
<td></td>
<td></td>
<td>Sildenafil 50Mg Tab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01-jul-03</td>
<td>01-jul-03</td>
<td></td>
<td></td>
<td>Sildenafil 25Mg Tab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26-jun-03</td>
<td>25-jun-03</td>
<td></td>
<td></td>
<td>Aldesleukin (18Mu/Ml) 22Mu Sdv inj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-jun-03</td>
<td>25-jun-03</td>
<td></td>
<td></td>
<td>Carboplatin 50Mg inj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-jun-03</td>
<td>25-jun-03</td>
<td></td>
<td></td>
<td>Tacrolimus 0.03% Oint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-may-03</td>
<td>27-may-03</td>
<td></td>
<td></td>
<td>Sulfapyridine 500Mg Tab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-may-03</td>
<td>19-may-03</td>
<td></td>
<td></td>
<td>Beclomethasone (Ovar) 40Mg c/inh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06-may-03</td>
<td>06-jun-03</td>
<td></td>
<td></td>
<td>Amoxicillin 500Mg Cap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27-feb-03</td>
<td>27-feb-03</td>
<td></td>
<td></td>
<td>Botulinum Tox A-Hcp 1Unit/Ea Inj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-feb-03</td>
<td>25-feb-03</td>
<td></td>
<td></td>
<td>Dn-101(Calcitriol) 15Mg Cap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-feb-03</td>
<td>19-feb-03</td>
<td></td>
<td></td>
<td>Propofol 10Mg/Ml</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-feb-03</td>
<td>12-feb-03</td>
<td></td>
<td></td>
<td>Selenium 40Mg/Ml</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-feb-03</td>
<td>12-feb-03</td>
<td></td>
<td></td>
<td>Valacyclovir 500Mg Or Ploch Tab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-jan-03</td>
<td>30-jan-03</td>
<td></td>
<td></td>
<td>Dmethylphenidate/Pilocarpine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28-jan-03</td>
<td>28-jan-03</td>
<td></td>
<td></td>
<td>Bupropion <em>SR</em> Study Tab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17-jan-03</td>
<td>13-nov-00</td>
<td></td>
<td></td>
<td>Investigational Drug Study Mis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>08-jan-03</td>
<td>08-jan-03</td>
<td></td>
<td></td>
<td>Erlotinib Study Tab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>08-jan-03</td>
<td>06-may-03</td>
<td></td>
<td></td>
<td>Amoxicillin 250Mg Cap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-dec-02</td>
<td>12-dec-02</td>
<td></td>
<td></td>
<td>Dolasetron 20Mg/Ml Inj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-nov-02</td>
<td>22-nov-02</td>
<td></td>
<td></td>
<td>Coal Tar/Menthol W/Cond Lotn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-nov-02</td>
<td>24-may-02</td>
<td></td>
<td></td>
<td>idarubicin (1Mg/Ml) 5Mg/Ea Inj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-nov-02</td>
<td>20-nov-02</td>
<td></td>
<td></td>
<td>Docusate 100Mg Cap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-oct-02</td>
<td>07-oct-02</td>
<td></td>
<td></td>
<td>Doxycycline 100Mg Cap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09-oct-02</td>
<td>07-oct-02</td>
<td></td>
<td></td>
<td>Amoxicillin 500Mg cap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ref lib</td>
<td>Major Problem Description</td>
<td>Onset Date</td>
<td>Last Dx Date</td>
<td>Episode Number</td>
<td>Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------</td>
<td>------------</td>
<td>--------------</td>
<td>----------------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>SCHIZOPHRENIA UNSPECIF I see dead people...</td>
<td>~01-Jun-2002</td>
<td>22-Nov-2002</td>
<td>2</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>PRINZMETAL ANGINA teste</td>
<td>~01-Nov-2002</td>
<td>06-Nov-2002</td>
<td>1</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SYST RHEUM ARTHRITIS NE</td>
<td>24-Sep-2002</td>
<td>24-Sep-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>HTN, ESSEN, BENIGN</td>
<td>~01-Jul-2002</td>
<td>23-Jul-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>LBP</td>
<td>~01-Jul-2002</td>
<td>23-Jul-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CEREBRAL ARTERITIS</td>
<td>25-Feb-2002</td>
<td>25-Feb-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>FX LUMBAR VERTEBRA-CLOS</td>
<td>25-Feb-2002</td>
<td>25-Feb-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>TOXIC ENCEPHALOPATHY</td>
<td>~01-Jun-2000</td>
<td>25-Feb-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>OTHER BREAST DISORDERS</td>
<td>~01-Feb-2002</td>
<td>05-Feb-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>RENAL ARTERY ANEURYSM</td>
<td>13-Jan-2002</td>
<td>13-Jan-2002</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ARRHYTHMIA Onset of SVT 1998</td>
<td>~01-Sep-2001</td>
<td>24-Sep-2001</td>
<td>1</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>COLON CA</td>
<td>~01-Sep-2001</td>
<td>24-Sep-2001</td>
<td>unreported</td>
<td>Active</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# ALERT! History of Harm/Violence

## PATIENT HAS REMINDERS BUT ... **ACT NOW**

### Transcripts

<table>
<thead>
<tr>
<th>Service Date</th>
<th>Dictate Date</th>
<th>Type (click to view)</th>
<th>Clinic</th>
<th>Status (click for dtl)</th>
<th>Author; Authenticator</th>
</tr>
</thead>
<tbody>
<tr>
<td>31-dec-2003</td>
<td>08-sep-2003</td>
<td>Admit Note</td>
<td></td>
<td>Authenticated</td>
<td>Tan, Effie</td>
</tr>
<tr>
<td>10-sep-2003</td>
<td>10-sep-2003</td>
<td>Nursing Record/Note</td>
<td></td>
<td>Unauthenticated</td>
<td>Walden, Blake C</td>
</tr>
<tr>
<td>10-sep-2003</td>
<td></td>
<td>Pathology Report</td>
<td></td>
<td>Authenticated</td>
<td>Chou, MD, David</td>
</tr>
<tr>
<td>10-sep-2003</td>
<td>10-sep-2003</td>
<td>Clinic Note</td>
<td>3-Endocrine</td>
<td>Transcribed</td>
<td>Test, Position2</td>
</tr>
<tr>
<td>09-sep-2003</td>
<td></td>
<td>Lit Hi Cath Retro</td>
<td></td>
<td>Authenticated</td>
<td>Chou, MD, David</td>
</tr>
<tr>
<td>07-sep-2003</td>
<td>07-sep-2003</td>
<td>In Error</td>
<td>3-Gen Int Med</td>
<td>In Error</td>
<td>Position Test, PDS Imaging, Position Test, PDS Imaging</td>
</tr>
<tr>
<td>07-sep-2003</td>
<td>07-sep-2003</td>
<td>In Error</td>
<td></td>
<td>In Error</td>
<td>Tan, Effie, Tan, Effie</td>
</tr>
<tr>
<td>06-sep-2003</td>
<td>06-sep-2003</td>
<td>In Error</td>
<td></td>
<td>In Error</td>
<td>Reilly, Dominic Francis</td>
</tr>
</tbody>
</table>
ALERT! History of Harm/Violence

PATIENT HAS REMINDERS DUE 17-SEP-04

CLINIC NOTE

PAIN CENTER

IDENTIFICATION: Mrs. Doe is a 46-year-old female, previously seen by myself on 7/17/98.

SUBJECTIVE: The patient has returned complaining of symptoms of menopause and to follow up on her lab examinations. She has been complaining of hot flashes since this summer, as well as increasing irritability and difficulty with sleep.

OBJECTIVE: Blood pressure 104/80. Pulse 72. Physical examination was deferred, except for pelvic, which shows normal external genitalia, normal vaginal mucosa without lesions. No cervical discharge. A repeat Pap was done.

LABORATORY DATA: Laboratories done 7/98, Pap was negative, but showed no endocervical cells. Her labs dated 12/14/98, showed FSH of 127, post menopausal is greater than 30 million IU/ml.

ASSESSMENT/PLAN:
1. I discussed hormone replacement with Mrs. Doe. She has no history of breast cancer and is willing to undergo the slight increase in breast cancer risk. She has no family history or personal history of thromboembolic disease. She is interested in the benefits, including decreasing her symptoms, her cardiovascular benefits, as well as osteoporosis. She understands that she might have some spotting and may require further intervention in the future. We have discussed various modalities and she prefers to be on a continuous Premarin/Provera combination, as she has not had a hysterectomy in the past. We will consider cyclic if this is not sufficient in the future.
U6999999  TS-00-00010  Rodney Alan Schmidt MD

PATIENT FUNNY TEST  Collected: 24-apr-2000
DOB: 01-apr-1950  Received: 24-apr-2000
Sex: Female

** DEMOGRAPHICS DRAWN FROM PATHOLOGY REPORT **

PATIENT:  TESTA, PATIENT FUNNY
MRN:  U6999999  (UWMC)
DOB:  Dec 1 1966
SEX:  F

**** THIS IS A REVISED OR CORRECTED REPORT ****

MATERIALS RECEIVED:
Label  Consult Accession No  Blocks/Slides  Description
A  UF-00-33  OB,OS  Brain, whole

NEUROPATHOLOGY ACCESSION NUMBER:  NP-44444

CLINICAL DATA:

CROSS DESCRIPTION:

NEUROPATHOLOGICAL DIAGNOSIS:
Dx

This report was amended for demographic reasons. No change was made to the
diagnosis on 1/1/2001.

Demographic Correction
Amended Report
Electronically signed 01/11/2001

In compliance with HIPAA regulations, the pathologist’s signature on this report
indicates that the case has been personally reviewed, and the diagnosis made or
confirmed by, the Attending Pathologist.

NEUROPATHOLOGY ACCESSION NUMBER:
NP-44444
Zztest, Pt 1.  
ID: 006999999  
22-OCT-1996 09:11:00  
UW MEDICAL CENTER

- Vent. rate: 67 BPM
- PR interval: 168 ms
- QRS duration: 100 ms
- QTc: 408/431 ms
- P-R interval: 32, 58, 65

Referred by:  
Confirmed by: JOHN R. BLACKMON, M.D.
What uses does it have?

- Identify Patients
- Link Patients
- Outcomes Data
- Treatment Data
- Enrollment
- Patient Care?
What is difficult?

• Data Access
• Privacy
• Data Model
• Meaning
• Timeliness
• Linking research results
Research is different, right?

- Timeliness of data
- Patient identified data
- ...

Timeliness

- No
  - Retrospective analysis
- Yes
  - Patient Enrollment
  - Experiment/Lab Management Systems
Patient Identification

• No
  – De-identified specimens

• Yes
  – Short term linkages to integrate data
  – Link across studies, across lifetimes
  – Familial studies
  – Re-identify tissue from banks?
  – … just how good an identifier is DNA…?
Research Systems are different, right?

- Purpose built
- Small user community
- ...
Built for the right purpose?

• Clinical Systems
  – Origins in financial accounting/billing
  – Episodes based on billable encounters
  – Track patients across episodes/sites (hard)

• Research Systems
  – Organize data from experiments/studies
  – Episodes based on funded studies
  – Track patients across studies (hard)
Purpose: Research Only?

- Research systems
  - Clinical uses?
  - Clinician/Researchers
    - Disease Specific Views
    - Diabetes system
  - Research Group Information Systems
    - GI Path
<table>
<thead>
<tr>
<th>Date</th>
<th>Study</th>
<th>patient role</th>
<th>Study Form</th>
<th>Form Status</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/18/03</td>
<td>VM ERCP</td>
<td>anonymous</td>
<td>VM Form</td>
<td>Not Started</td>
<td>De-identify</td>
</tr>
<tr>
<td>8/18/03</td>
<td>Familial Pancreas</td>
<td>familial</td>
<td>Pancreatic Form</td>
<td>Incomplete</td>
<td></td>
</tr>
<tr>
<td>8/18/03</td>
<td>R01 Colon</td>
<td>Consented</td>
<td>R01 Form</td>
<td>Completed</td>
<td></td>
</tr>
</tbody>
</table>

Enrolled Studies

Date: [ ] (select study to add) [ ] (select role) Add

Help Notes: Use this screen to add studies and study forms.
### Pancreatic Disease Data Collection Form

#### Demographic and Lifestyle data

<table>
<thead>
<tr>
<th>Patient Age:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth Place:</td>
</tr>
<tr>
<td>City</td>
</tr>
<tr>
<td>State</td>
</tr>
<tr>
<td>Country</td>
</tr>
<tr>
<td>Marital Status:</td>
</tr>
<tr>
<td>Never Married</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Widow(er)</td>
</tr>
<tr>
<td>Separated</td>
</tr>
<tr>
<td>Divorced</td>
</tr>
<tr>
<td>Race:</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Ashkenazi Jew</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Height:</td>
</tr>
<tr>
<td>feet</td>
</tr>
<tr>
<td>inches (OR cm)</td>
</tr>
<tr>
<td>Current Weight:</td>
</tr>
<tr>
<td>lbs (OR kg)</td>
</tr>
<tr>
<td>Weight Prior to Pancreatic Cancer:</td>
</tr>
<tr>
<td>lbs (OR kg)</td>
</tr>
<tr>
<td>Education:</td>
</tr>
<tr>
<td>Grade School</td>
</tr>
<tr>
<td>Junior School</td>
</tr>
<tr>
<td>High School</td>
</tr>
<tr>
<td>Technical</td>
</tr>
<tr>
<td>College</td>
</tr>
<tr>
<td>Tobacco:</td>
</tr>
<tr>
<td>Never Use</td>
</tr>
<tr>
<td>Current User</td>
</tr>
<tr>
<td>Ex-user</td>
</tr>
<tr>
<td>if Ex-user Cessation Date:</td>
</tr>
<tr>
<td>Tobacco Type and Amount:</td>
</tr>
<tr>
<td>Cigarettes &lt;= 1 pack/day</td>
</tr>
<tr>
<td>Cigarettes &gt; 1 pack/day</td>
</tr>
<tr>
<td>Cigars</td>
</tr>
<tr>
<td>Pipe</td>
</tr>
<tr>
<td>Chewing Tobacco</td>
</tr>
<tr>
<td>Alcohol:</td>
</tr>
<tr>
<td>Never drink</td>
</tr>
<tr>
<td>Drink less than 1 drink/day</td>
</tr>
<tr>
<td>Drink approximately one drink/day</td>
</tr>
<tr>
<td>Drink 2 or more drinks/day</td>
</tr>
<tr>
<td>Former Drinker</td>
</tr>
<tr>
<td>Number of Drinks</td>
</tr>
<tr>
<td>Coffee:</td>
</tr>
<tr>
<td>Never Drink coffee</td>
</tr>
<tr>
<td>1 cup or less/day</td>
</tr>
<tr>
<td>2-5 cups/day</td>
</tr>
<tr>
<td>6 or more cups/day</td>
</tr>
<tr>
<td>Type of Coffee:</td>
</tr>
<tr>
<td>Regular</td>
</tr>
<tr>
<td>Decaffeinated</td>
</tr>
<tr>
<td>Espresso</td>
</tr>
</tbody>
</table>
Clinical Data in Research Systems

- Avoid double entry
- Link with clinical system for updates
- Links back to clinical system?
- Patient self reported data
  - Common in clinical research
  - Soon in lab research?
Patient self reported data

• Biomedical research
  – Cells and molecules
  – Test tubes and culture dishes

• People:
  – Large aggregates of cells
  – Capable of answering questionnaires
  – Contributing “clinical data” to research
Quality of Life Survey

ENTER Unumber

ORDERED BY LAST NAME  UNNUMBER

CANCEL  SUBMIT
Quality of Life Survey
DEMO VERSION

* U. number 6999999  Visit Number  
First Name Smith  
Last Name John  
Patient's Age 55  GENDER ◆ male ◆ female  

Today's Date: September 11, 2003

BEGIN ➔
Quality of Life Survey

DEMO VERSION (database entries will not be saved)

What is this project all about?

We are developing better and more efficient ways to provide care to our patients. We are asking you to answer a quality of life/symptom questionnaire on this computer. We will go over this information with you when you are done and use it to plan ongoing care for you while you are receiving treatment.

We are also asking your permission to put your answers to the questions into a research database. The next screen will give you information you will need to decide if you give your permission for your answers to be used in the research database. If you do not want your answers to appear in the research database, your answers will only be in the clinical care database.

NEXT
Quality of Life Survey

DEMO VERSION (database entries will not be saved)

How the database works?

Your answers will be electronically recorded in a secure clinical database that is used for your ongoing care. The computer will graph your answers. The nurse or doctor will go over this graph with you and then it will be placed in your radiation care notebook.

We would also like to use your answers in a secure research database. If you give permission, we will electronically record your answers in the research database. Your answers will be grouped with those of other patients with cancer. We hope that the results of this project will help us better care for patients with a diagnosis of cancer. Your name will not appear in the research database and all other pieces of information that could identify you will be automatically removed from the displayed data.

Are there any risks for this project?

Some people feel that providing information for research is an invasion of privacy. Some of the questions asked on the screen may make you uncomfortable or upset. You don't have to answer every question, just skip past those questions if you wish.
Please select one:

- I give permission for my answers to be in the research database table and also used by the nurses for my care in the clinic.

- I give permission for my answers to be used by the nurses only for my care in the clinic.

- I do not want to use the computer to answer any questions.
Think about what each statement says, then select the one statement that most closely indicates how you have been feeling during the past week including today.

**NAUSEA**

- I seldom if ever have nausea
- I have nausea once in a while
- I have nausea fairly often
- I have nausea half the time at least
- I have nausea almost continually
Quality of Life Study

- Research data about a person
- Session response clinically useful
- Population responses are research

- What about longitudinal individual responses?
- Will this type of data be integrated with genome data…
Purpose Built?

• Research systems
  – Clinical uses?
  – Clinician/Researchers
  – Research Lab Management Systems

• Clinical systems use for research
  – Outcomes research
  – Population research
Unique to Research

- Tissue/specimen tracking
- Anonymous patient/specimens
- Distance from patients
- Different device/machine interfaces
Central Questions

How are research and clinical systems the same? Different?

Clinical data should flow to researchers, but are their data that should flow back to clinical systems?
Revisit: What is difficult?

• Data Access
• Privacy
• Data Model
• Meaning
• Timeliness
• Linking research results
Data Access

- Institutional boundaries
- Multiple sites in an organization
  - Technical and organizational barriers
- Multiple sides of an organization
Seattle Cancer Care Alliance

• Fred Hutch
• UW Medicine
  – UW Medical Center
  – School of Medicine
  – Harborview Hospital
  – UW Medicine Neighborhood Clinics (UWPN)
• Children’s Hospital
• SELU Clinic Building
Cancer Research Consortium

• Fred Hutch
• UW Medicine
  – UW Medical Center
  – School of Medicine
  – Harborview Hospital
  – UW Medicine Neighborhood Clinics (UWPN)
• Children’s Hospital
• SELU Clinic Building

* Disclaimer: I could be wrong…
Keeping Score

• 2 major cancer organizations
  – large overlap
  – parts of some entities excluded
• 4 Medical Records units (helpful)
• 3 IRBs (reciprocity)
• 3 IT groups (friendly, knowledgable)
• Large herd of Cats
  (with independent funding!)
  (and advanced degrees!)
Data Access

• Organizational Agreements
  – Intended use
  – Privacy safeguards
  – Security

• But, with whom?
Privacy

• HIPAA Privacy regulations
• Patient Data
  – Protected Health Information
  – De-identified
  – Limited Data Set
• Accounting of Disclosures
Protected Health Information

- Relates to the past, present, or future physical or mental health or condition of an individual; or
- Relates to the provision of health care to an individual; or
- Relates to the past, present, or future payment for the provision of health care to an individual; and
- Identifies the individual; or with respect to which there is a reasonable basis to believe the information can be used to identify the individual.
Limited Data Set requires

- Excluding direct identifiers for patient, relatives, employers, or household members
- Data sharing agreement
  - Establish permitted uses of data
  - Establish recipients of data
  - Establish no use or further disclosure beyond data user agreement
  - Appropriate safeguards, transitive property, etc....
Limited Data Set excludes

- 1) Names;
- 2) Postal address information, other than town or city, State, and zip code;
- 3) Telephone numbers;
- 4) Fax numbers;
- 5) Electronic mail addresses;
- 6) Social security numbers;
- 7) Medical record numbers;
- 8) Health plan beneficiary numbers;
- 9) Account numbers;
- 10) Certificate/license numbers;
- 11) Vehicle identifiers and serial numbers, including license plate numbers;
- 12) Device identifiers and serial numbers;
- 13) Web Universal Resource Locators (URLs);
- 14) Internet Protocol (IP) address numbers;
- 15) Biometric identifiers, including finger and voice prints; and
- 16) Full face photographic images and any comparable images.
De-identified

• 1) all LDS data elements, plus…
• 2) geography smaller than state, or 3 digit zip (if > 20K people)
• 3) all elements of dates except year (birth, admit, d/c, death, and age > 89 yrs.)
• 4) any other unique identifying number, characteristic, or code (sequence data?!)
De-identified

OR – a demonstrated “very small” risk of re-identification by expert using generally accepted methods.

May assign re-identification number

Ref: https://know1.mcis.washington.edu/manuals/amc_privacyPolicy/
Accounting of Disclosures

• Patient has right to receive an accounting of when PHI disclosed outside of
  – Treatment, Payment, or Operations
  – LDS disclosures (data use agreement)
Revisit #2: What is difficult?

- Data Access
- Privacy
- Data Model
- Meaning
- Timeliness
- Linking research results
Data Model

• Common clinical model
  – Visit oriented
  – Patient-centric

• Common research model
  – Study oriented
  – Specimen-centric
  – Subset view of patient
Data Model

• Interesting research data
  – Pedigrees
  – Specimens, incl. anonymous
  – Tumor/Dz registry links
  – Sequences/expression data, etc.
  – Next year’s research tests
Missing from Clinical and Research

- Lifetime view
  - + family, population
- Common data model
- Comprehensive view
Different Representations?

• Operational, stable
  – Fixed model
  – Relational tables
  – Event driven
• Adaptable, dynamic
  – Extensible
  – Entity-Attribute-Value model?
  – Graph/Network model?
Master Patient Index

• Clinical
  – Multi-system deterministic matching
  – Multi-institutional, maybe

• Research
  – Multi-institutional
  – Probabilistic matching
Data Integration

• Vendor C – “buy from us”
• HL7 messaging
• HL7 RIM
  – eHealth initiative, HL7 EMR ballot
• Distributed Queries/Multiple Schemas
  – Mediated schema
  – PTP discoverable schemas
Meaning

• Syntax
  – Message structure
  – Granularity
  – HL7 Reference Information Model?

• Semantics
  – Vocabularies
  – SNOMED?
Interfacing

- Event Driven
  - HL7 messaging
- Query Driven
  - Reporting
  - “Research”
Timeliness

• Real Time
  – “clinical”
• Near Real Time
  – BT surveillance
• Research Retrospective Time
  – Study data analysis
Research Time?

- When do researchers care about real time?
  - Patient enrollment
  - Clinical decision making
  - Experiment/lab management systems
Revisit #3: What is difficult?

• Data Access
• Privacy
• Data Model
• Meaning
• Timeliness
• Linking research results
Why? - Structural Challenges

- Centralized support, decentralized activity
- Different needs for overlapping data
- Different meanings of same data
- Different performance issues with aggregate vs. individual queries
- Uniform security and privacy concerns
Approach – Central Services?

• Authentication/Authorization/Auditing
• Master Person Index, De-identification
• Individual/aggregate query support
• Query/Result Annotation
• Extensible, secure storage
• Protocol/Guideline support
• …
Questions

???