NCI Efforts in Healthcare Informatics

Ken Buetow, Ph.D.
Director, Center for Bioinformatics and Information Technology

NCAB Meeting
June 23, 2010
caBIG® is a virtual network of interconnected data, individuals, and organizations that redefines how research is conducted, care is provided, and patients/participants interact with the biomedical research enterprise.
caBIG® Sample Statistics

Community
- 2,300+ participants from more than 700 institutions
  - 56 NCI-designated Centers
  - 16 Community Centers
- 1,100+ attendees at the 2009 caBIG® Annual Meeting
- 16 licensed Support Service Providers to sustain the biomedical community as they deploy caBIG® tools and technology
- 15 countries using caBIG® tools and technology to facilitate

Connectivity
- 40+ applications supporting full continuum of biomedical research
- 120+ “nodes” connected to National Grid via caGrid

Content
- 1.19 million biospecimens available through caGrid
- 3.71+ million medical images stored in the National Biomedical Imaging Archive (NBIA)
- 25,000+ microarray experiments available for research use on caGrid
- Collaborative biomedical research
- 30 peer-reviewed scientific publications featuring or enabled by caBIG® tools and technology in 2009
Since the NCI first deployed caGrid in 2007, a growing number of data and analytical services, hosted by the NCI and a diverse collection of organizations, have been made available to researchers across the globe.
More than 1,100 individuals representing over 800 organizations have registered for access to the National Biomedical Imaging Archive (NBIA) hosted at the NCI, which currently contains more than 3.7 million medical images.
University of Arkansas: Managing Institutional Clinical Research

Cancer Central Clinical Participant Registry (C3PR)
- Eligibility is verified and patient is registered to a study

Vendor Clinical Data Management System
- Clinical data is captured

Vendor
caBIG®

Patient Study Calendar (PSC)
- Tracks the patient schedule throughout the study

Patient Study Calendar

Lab Viewer
- Identifies labs, loads them into the CDMS and AE system

caXchange (Hub)

Cancer Adverse Event Reporting System (caAERS)
- Identifies and tracks adverse events and any associated schedule changes

CTODS
Clinical Trial Object Database System

caBIG® vendor
Multiple forms of cancer

- Glioblastoma multiforme (brain)
- Squamous carcinoma (lung)
- Serous cystadenocarcinoma (ovarian)

Total: 20 types

Multiple data types

- Clinical diagnosis
- Treatment history
- Histologic diagnosis
- Pathologic status
- Tissue anatomic site
- Surgical history
- Gene expression
- Chromosomal copy number
- Loss of heterozygosity
- Methylation patterns
- miRNA expression
- DNA sequence

Connecting multiple sources, experiments, and data types
The Cancer Molecular Analysis Portal
http://cma.nci.nih.gov

- Enables users to access, search, visualize, and integrate genomic data with corresponding clinical information
- Helps find novel correlations between data and observations that would be difficult or impossible to find using conventional analytical tools and methods
- Provides access to, and facilitate analysis of, data from other research studies such as REMBRANDT*, TARGET* GSK Expression, COSMIC mutations, and JHU mutations

* The Cancer Genome Atlas (TCGA)
* Repository of Molecular Brain Neoplasia Data (REMBRANDT)
* Therapeutically Applicable Research to Generate Effective Treatments (TARGET)
GBM Results: Pathways

TCGA: Nature 2008
Patient Selection for HER2 Tx Required Tissue Screen and Allowed Only 1 of 4 Women to Participate

<table>
<thead>
<tr>
<th>Calculated Sample Size And Study Duration</th>
<th>Hypothetical HER2+ Prevalence</th>
<th>Required “Screened” Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1250 → 52 mos</td>
<td>100%</td>
<td>1250</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>2500</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>5000</td>
</tr>
</tbody>
</table>

* Need a obtain a suitable specimen, wait for test results. (Results were obtained in days to weeks)
* Need to screen many patients.

Courtesy H. Kim Lyerly, M.D., Director
Size of Population with Pathway to Inhibit*

<table>
<thead>
<tr>
<th>Population fraction containing signature</th>
<th>100%</th>
<th>50%</th>
<th>25%</th>
<th>12.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Pathway Defect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype A2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype B1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype B2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Size of Population Needed To Screen

1,250  2,500  5,000  10,000

Courtesy H. Kim Lyerly, M.D., Director
Leveraging the Nation’s HIT Investment for Research
The Nation’s Health Information Technology (HIT) Investment

- “Our recovery plan will invest in electronic health records and new technology that will reduce errors, bring down costs, ensure privacy, and save lives.”
  
  President Obama  
  Address to Joint Session of Congress  
  February 24th, 2009

- **$46 Billion investment:** incentivize “meaningful use” of EHRs

- “With the passage earlier this year of the Health Information Technology for Economic and Clinical Health (HITECH) Act, we have the tools to begin a major transformation in American health care made possible through the creation of a secure, interoperable nationwide health information network.”
  
  Office of the National Coordinator for Health IT  
  August 19, 2009
What’s Needed to Leverage the HIT Investment for Research?

• Oncology-Specific Electronic Health Records as a source of clinical information, annotated biospecimens, images, and molecular data

• The data-sharing infrastructure to capture, aggregate, analyze and appropriately share massive amounts of information from millions of patient-physician encounters

• Ability to prospectively identify sub-groups of patients and collaborate across organizations to test research hypotheses
Oncology-Extended EHR:
A collaborative national effort

• American Society of Clinical Oncologists (ASCO)
  • Began evaluating issue, involving end users
  • Engaged the vendor community through its EHR lab, utilizing unique case scenarios
  • High level requirements document/white paper outlining the issue

• cancer Biomedical Informatics Grid (caBIG®)
  • Vendor technology evaluation
  • Problem assessment
  • Technical Specification
  • Reference Implementation

• NCI Community Cancer Center Program (NCCCP)
  • Oncology EHR Laboratory

• Other domain experts
Collaborations: Vendors

- **Proprietary**
  - Altos Solutions
  - DoX Systems
  - Elekta Impac Software
  - GeniusDoc, Inc.
  - MDLand International
  - MedSym, Inc.
  - Rabbit Healthcare Systems
  - Rational Health Systems
  - Smart ID Works, LLC
  - Varian
  - EPIC
  - Eclipsys (formerly MediNotes)
  - Cerner
  - Sunrise

- **Open Source**
  - ClearHealth/MirrorMed
  - Open MRS
  - Tolven EHR
  - World VistA
  - Medsphere Open VistA
  - Ultimate EMR
  - Torch
  - Open EHR
  - Indivo Health
  - Free Med
  - GNU Med (Germany)
  - Open EMR (2 versions – community and managed)
  - OSCAR (Canada)
  - PrimaCare (Malaysia)
Information collection, decision support, and reporting needs of the oncologist providing patient-focused care in a clinical setting:

- Generate and transmit a treatment plan
- Generate and transmit a treatment summary
- Support oncology-specific documentation
- Support oncology-specific EHR functionality
Leveraging and Extending caBIG®
“Periodic Table of Services”

- **BUSINESS**
  - R: Registration
  - Pt: Protocol
  - Oc: Study Outcomes
  - Po: Patient Outcomes
  - E: Eligibility
  - Ae: Adverse Event
  - Hx: Hx and Physical
  - Dx: Discharge Note
  - Ds: Decision Support
  - Ra: Referral and Authorization

- **CAPABILITY”BUSINESS”**
  - Cr: Credentialing
  - S: Specimen
  - Tp: Treatment Plan
  - I: Image
  - L: Lab
  - Rx: Pharmacy
  - Sc: Scheduling

- **“CORE”**
  - Sd: SDTM
  - Qr: Data Query
  - C: Correlation
  - O: Organization
  - P: Person
  - Pa: Protocol Abstraction
  - D: Disease
  - A: Agent
  - Mp: Master Problem List
  - Ay: Allergy

- **“Infra / UTILITY”**
  - Km: Knowledge Management
  - Cm: Contract Management
  - Ev: Enterprise Vocabulary
  - Va: Validation
  - Tx: Translation
  - Au: Audit
  - Id: Id Management
  - Tr: Trust Management
  - Aa: Authorization Authentication
  - Py: Policy
“Ultra-light” Oncology EHR: An On-ramp to Electronic Health for Community Practices
NCCCP Outcomes Database Project

Objectives: Model IT Infrastructure for Rapid Learning Healthcare System

• Design and implement a system that enables NCCCP sites to aggregate and analyze standard clinical encounter data that can be used to support decision-making for physicians and administrators

• Increase organizational “data liquidity”, integrating and aggregating data across sites and returning it back in a consumable, format to improve policies, practices, and research participation

• Generate a novel research resource of individuals with deep clinical annotation and readily available biospecimens
The Rapid Learning Healthcare System links research and care in a “virtuous circle”:

- Patient outcomes data is combined and aggregated to inform discovery and care in real-time.
- Latest scientific findings are made available to physicians to encourage rapid clinical adoption.
- Discovery is transformed into a natural outcome of patient care.

*From the Institute of Medicine (IOM) report, “A Foundation for Evidence-Driven Practice: A Rapid Learning System for Cancer Care”*
The Electronic Health Information Is One of Many information Sources within the Biomedical Community
Epidemiologists
- Query data to seek correlations among genes, environment, outcome
- Develop standing online cohorts of volunteers

Basic Researchers
- Generate new hypotheses
- Identify biomarker-outcome correlations
- Validate biomarkers in silico

Clinical Researchers
- Seek clinical trial participants
- Enrich clinical studies with appropriate sub-groups
- Identify new indications

Researchers Can Query the Data in the Cancer Knowledge Cloud
New Knowledge from Research Is Fed into the Cancer Knowledge Cloud

- **Epidemiologists**: New links to behaviors and exposures that increase / decrease risk of disease or disease reoccurrence
- **Basic Researchers**: New drug targets
- **Clinical Researchers**: Targeted drugs for molecularly-defined sub-groups
Virtuous Circle From Smart EHRs Through Research and Back to Clinical Care (Rapid Learning Healthcare System)
Outcomes Analysis:
Subset Data and Analyze Treatments

- Automatically find cohort of patient that is similar to selected patient based on prognostic factors
  - Example: Prostate cancer
    Age group, sex, race, ethnicity, family history, personal history, PSA, prostate hypertrophy, urinary or rectal incontinence, Gleason score, TNM, and stage
  - Obtain values from selected patients and find cohort with the same values
  - Look for all treatment combinations found in this subset of data
  - Analyze outcomes per treatment combination
Treatment Outcomes for Prostate Cancer
Comparing Mortality Due to Prostate Cancer for Androgen Deprivation Therapy Compared to All Other Treatments
Decision Tree: Predicted Change in Mortality

[Diagram of a decision tree showing various decision points based on factors such as ACC stage, ECOG score, total Gleason score, age, PSA levels, and treatment code.]
Result: Researchers have unprecedented access to huge depth and breadth of resources
Integrating Research and Care: The ISPY-2 Trial
The I-SPY TRIAL  (Investigation of Serial studies to Predict Your Therapeutic Response with Imaging And molecular analysis):

A national study to leverage biomarkers in predicting response to combinatorial therapy for women with Stage 3 breast cancer.

(PI Laura Esserman, UCSF)
Projected I-SPY 2 study sites
Accrual: Anticipate 800 patients over 3–4 years

Enroll ~20 patients per month

Participating Sites: 15–20 across US and Canada
I-SPY Adaptive Trial:
Introduce several new agents for a given profile

On Study

Randomize

HER 2 (+)

Taxol + Trastuzumab*

Taxol + Trastuzumab* + New Agent A

Taxol + Trastuzumab* + New Agent B

Taxol + Trastuzumab* + New Agent C

Surgery

Learn, Adapt from each patient

HER 2 (-)

Taxol

Taxol + New Agent C

Taxol + New Agent D

AC → Surgery

*Or Equivalent
**I-SPY Adaptive Trial:**
Introduce several new agents for a given profile

![Diagram showing treatment options based on HER 2 status and randomization]

- **HER 2 (+)**
  - Taxol + Trastuzumab*
  - Taxol + Trastuzumab* + New Agent A
  - Taxol + Trastuzumab* + New Agent B
  - Taxol + Trastuzumab* + New Agent C
- **HER 2 (-)**
  - Taxol + New Agent F
  - Taxol + New Agent G

*Or Equivalent

Learn, Adapt from each patient

On Study

Randomize

Surgery
• caBIG® is connecting the cancer community nationally and internationally to enable a wide spectrum of discovery and clinical research activities

• caBIG® is leveraging new opportunities in HIT in support of increasingly complex research studies
  • caBIG®-compatible oncology-extended Electronic Health Record
    • ASCO and the NCI are collaborating to create an oncology-enhanced EHR using caBIG® standards for interoperability as a source of data for research use

• caBIG® National Infrastructure can bridge research and care
  • Online capabilities permit us to capture and share information on an unprecedented scale
  • Common IT infrastructure supports Rapid Learning Healthcare System