caBIG™
cancer Biomedical Informatics Grid ™

NCI Efforts in Healthcare Informatics

Ken Buetow, Ph.D.

Director, Center for Bioinformatics and Information Technology

NCAB Meeting

June 23, 2010

caBIG®: Creating a Worldwide Web of Cancer Research

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 cancer Biomedical Informatics Grid™

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

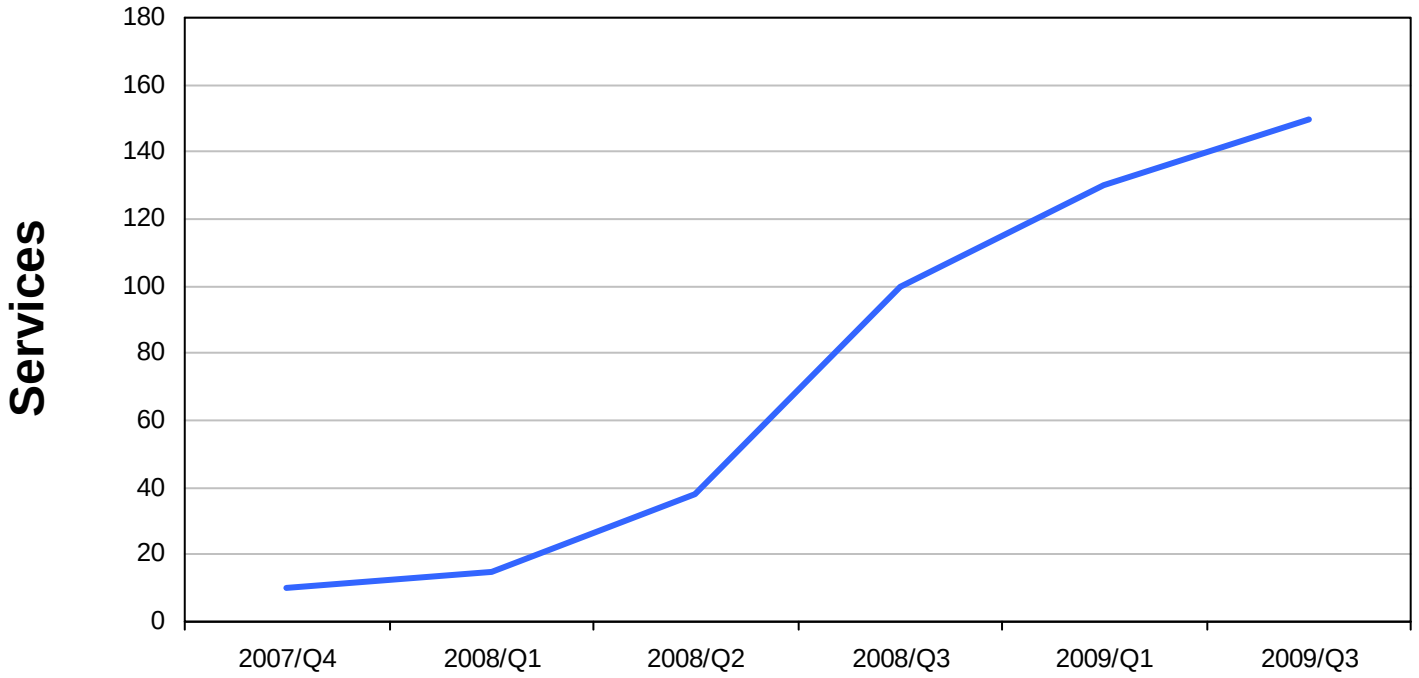
Content

- 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

caBIG cancer Biomedical Informatics Grid™

A Rapidly Expanding Virtual Community

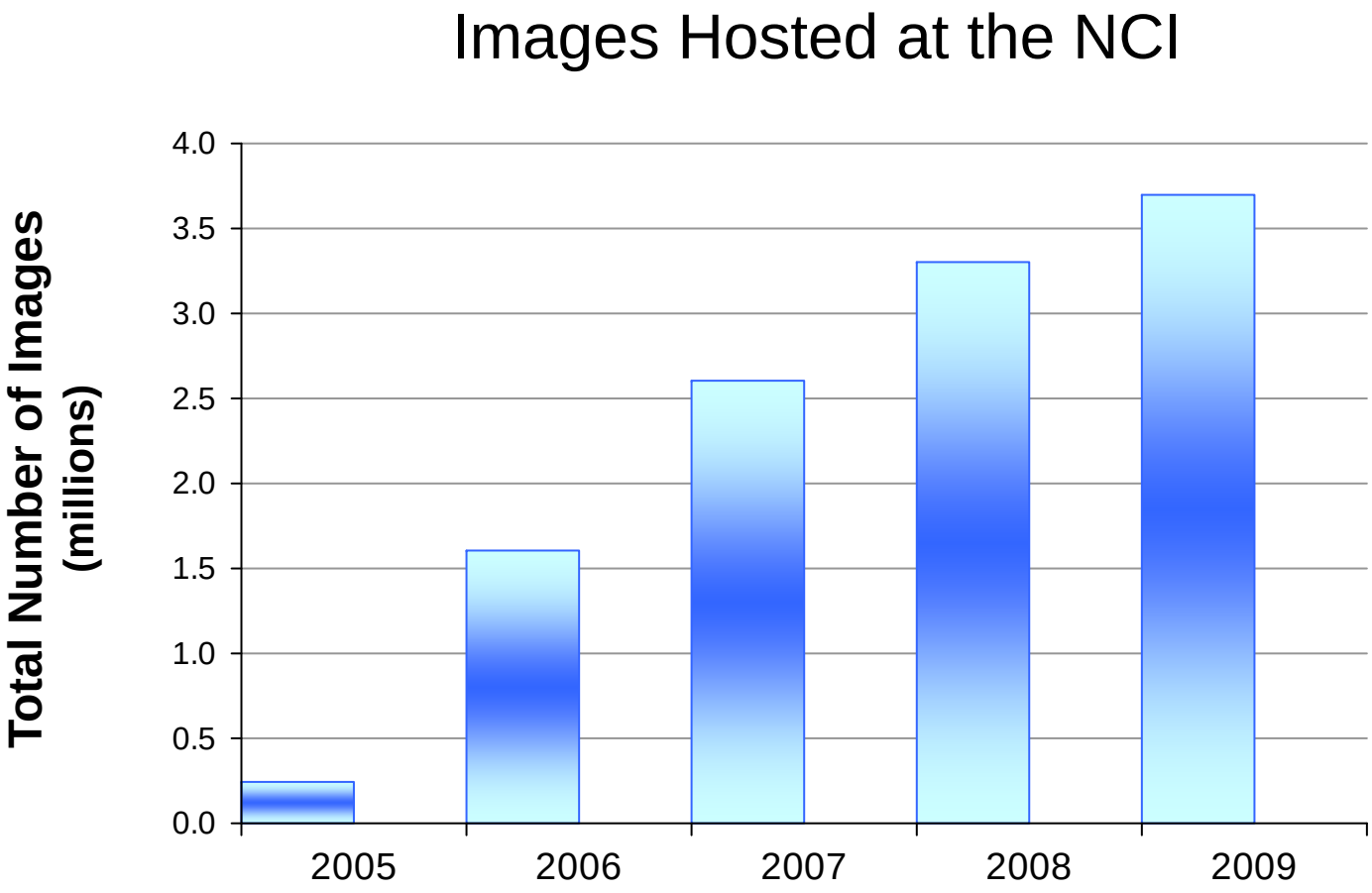
Connecting the Cancer Community via caGrid



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.

caBIG cancer Biomedical Informatics Grid™

Expanding Numbers of Images are Available to Researchers through NBIA



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.

caBIG cancer Biomedical Informatics Grid™

caBIG™
cancer Biomedical Informatics Grid ™

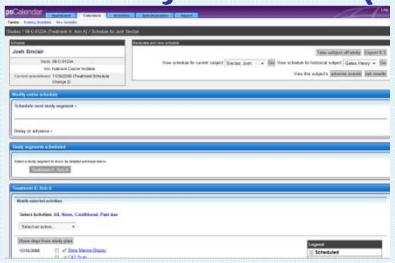
caBIG® Usage Patterns

University of Arkansas: Managing Institutional Clinical Research

Cancer Central Clinical Participant Registry (C3PR)

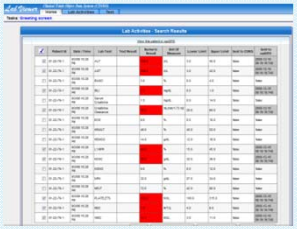
Eligibility is verified and patient is registered to a study

Patient Study Calendar (PSC)



Tracks the patient schedule throughout the study

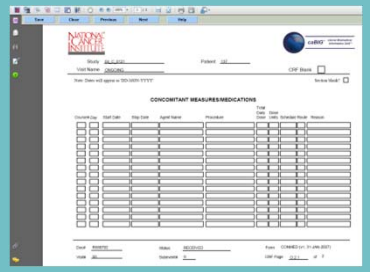
Lab Viewer



Identifies labs, loads them into the CDMS and AE system

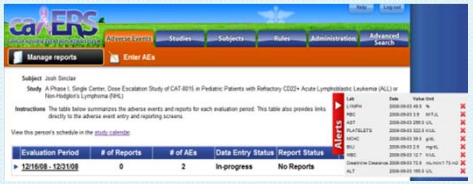
caXchange (Hub)

Vendor Clinical Data Management System



Clinical data is captured

Cancer Adverse Event Reporting System (caAERS)



Identifies and tracks adverse events and any associated schedule changes

(CTODS)

Clinical Trial Object Database System

vendor

caBIG®

caBIG cancer Biomedical Informatics Grid™

Providing Integrated Access to Multidimensional Data

Connecting multiple sources, experiments, and data types

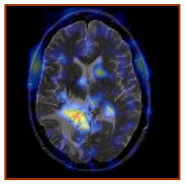
Multiple forms of cancer

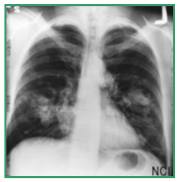
glioblastoma multiforme (brain)

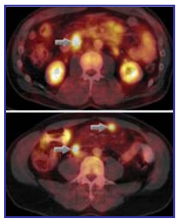
squamous carcinoma (lung)

serous cystadenocarcinoma (ovarian)

(20 total types)





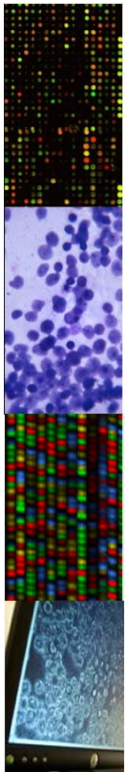


Multiple Organizations

- Biospecimen Core Resource
- Cancer Genomic Characterization Centers
- Genome Sequencing Centers
- Data Coordinating Center

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



caBIG cancer Biomedical Informatics Grid™

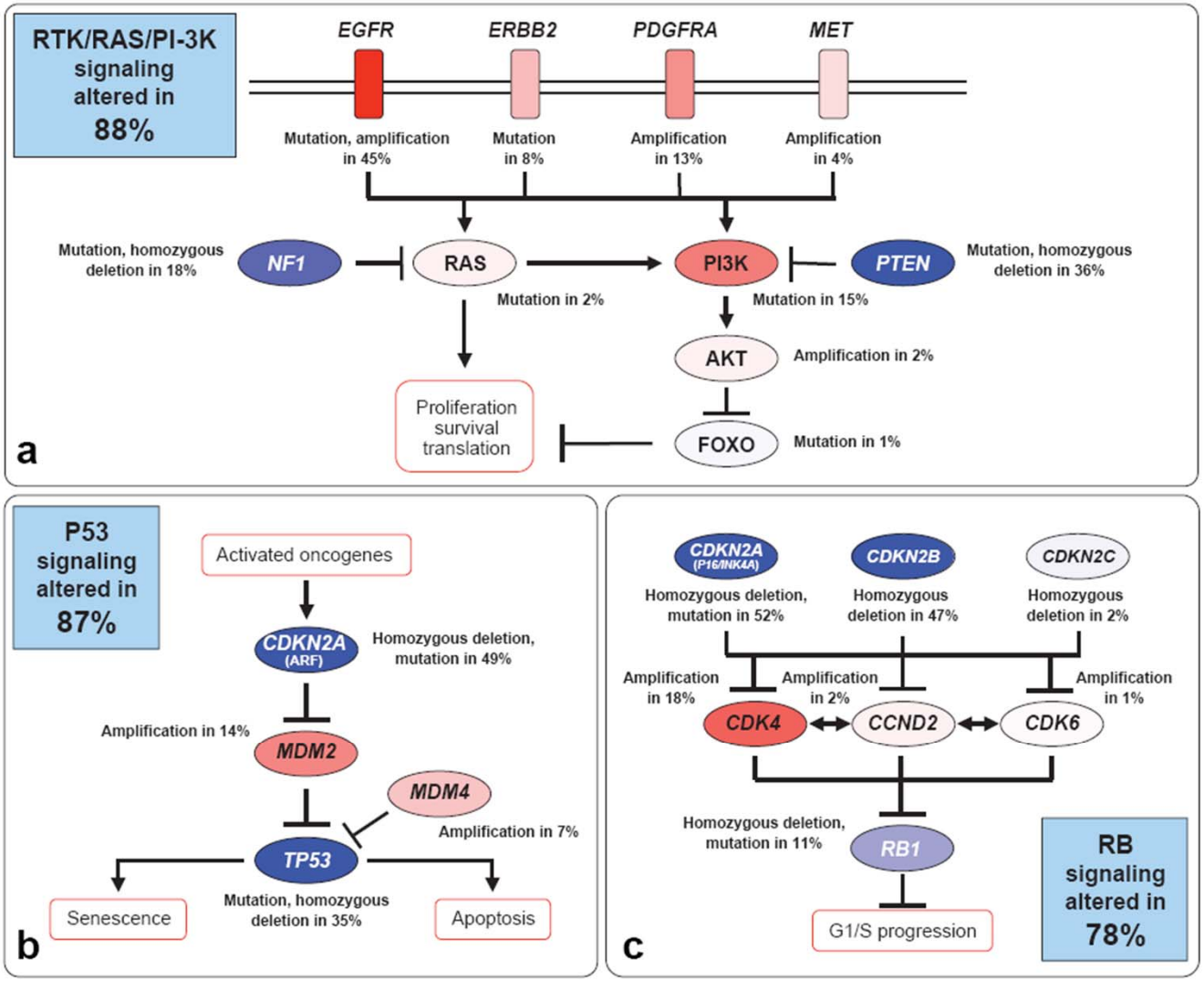
The Cancer Molecular Analysis Portal

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

Calculated Sample Size And Study Duration	Hypothetical HER2+ Prevalence	Required "Screened" Population
1250 ➡ 52 mos	100%	1250
	50%	2500
	25%	5000

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

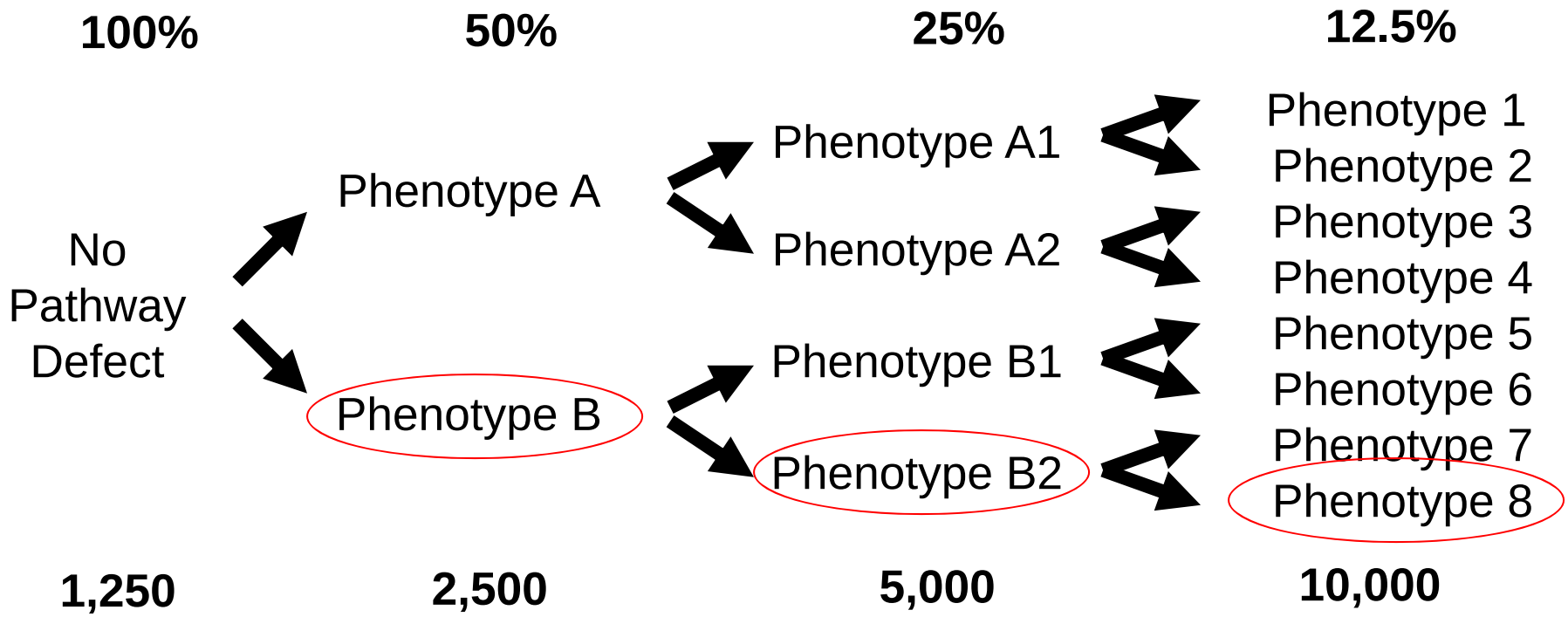
** Need to screen many patients.*

Courtesy H. Kim Lyerly, M.D., Director Duke Comprehensive Cancer Center

caBIG cancer Biomedical Informatics Grid™

Size of Population with Pathway to Inhibit*

Population fraction containing signature



Size of Population Needed To Screen

Courtesy H. Kim Lyerly, M.D., Director

Duke Comprehensive Cancer Center

caBIG

caBIG™
cancer Biomedical Informatics Grid ™

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

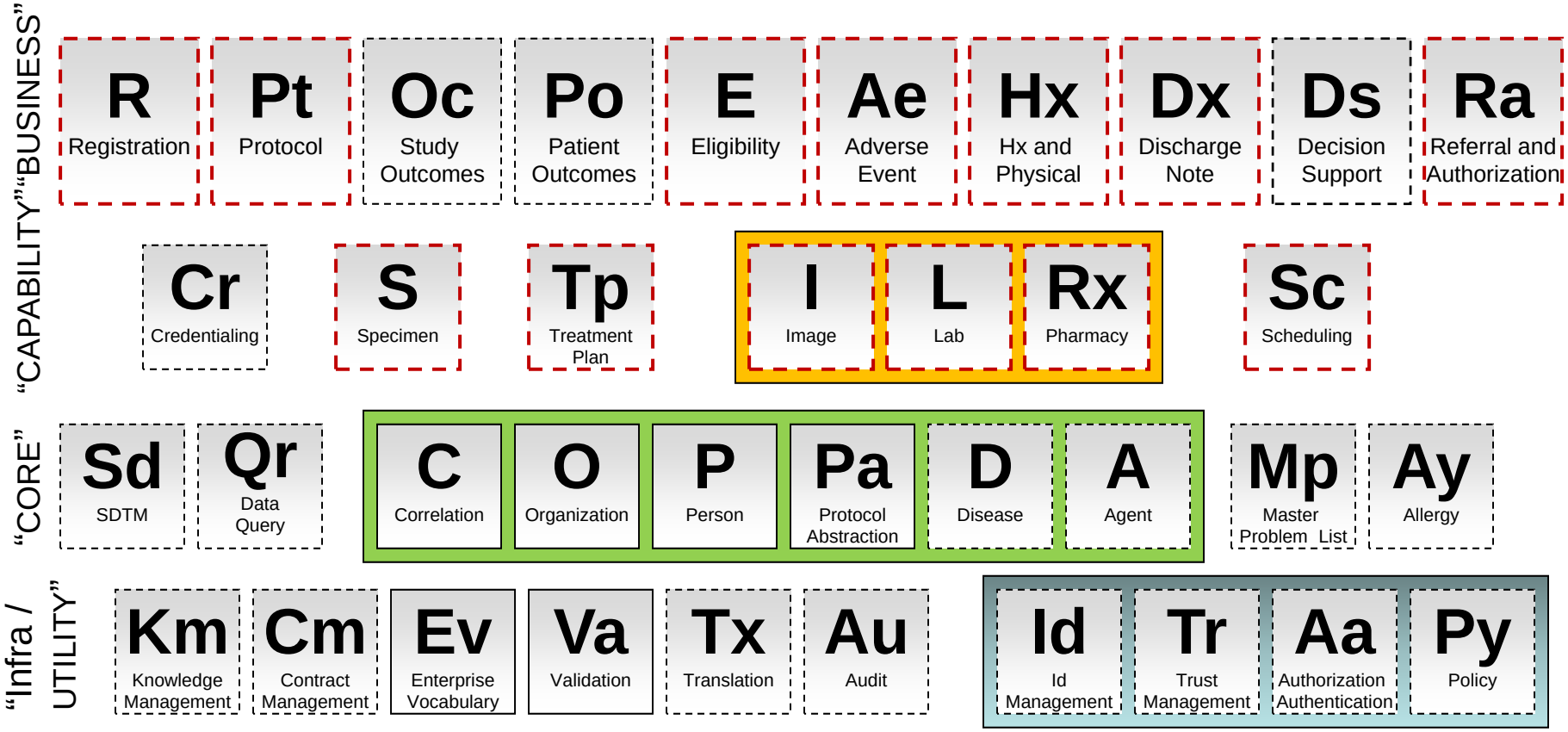
- **Proprietary**
  - Altos Solutions
  - DoX Systems
  - Elekta Impac Software
  - GeniusDoc, Inc.
  - MDLand International
  - MedSym, Inc.
  - Rabbit Healthcare Systems
  - Rational Health Systems
  - Smart ID Works, LLC
  - Varian
  - EPIC
  - Eclipsys (formerly MediNotes)
  - Cerner
  - Sunrise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

Clinical Oncology Requirements for the EHR: (CORE):
Functional Requirements

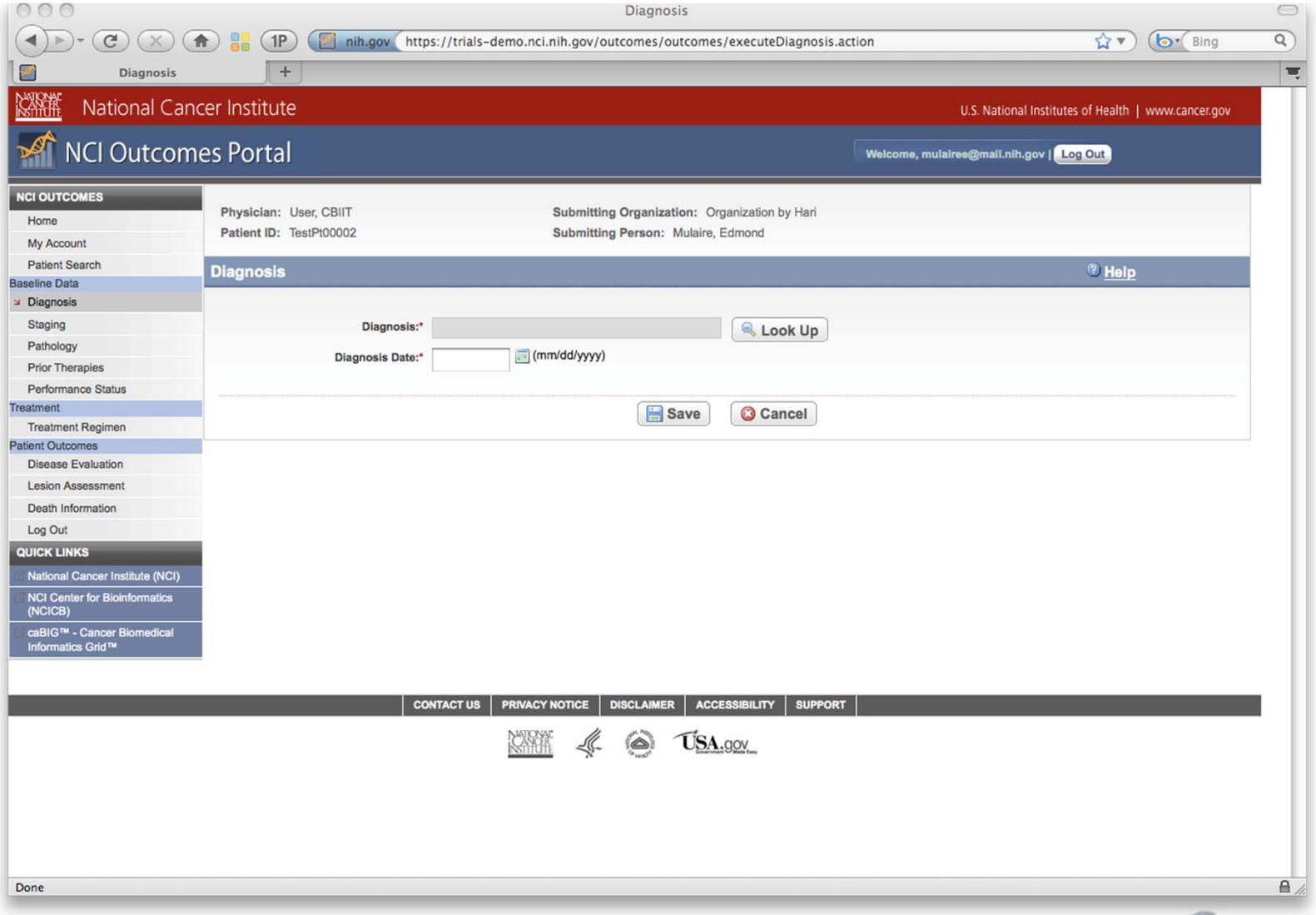
Information collection, decision support, and reporting needs of the oncologist providing patient-focused care in a clinical setting:

- 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"



“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

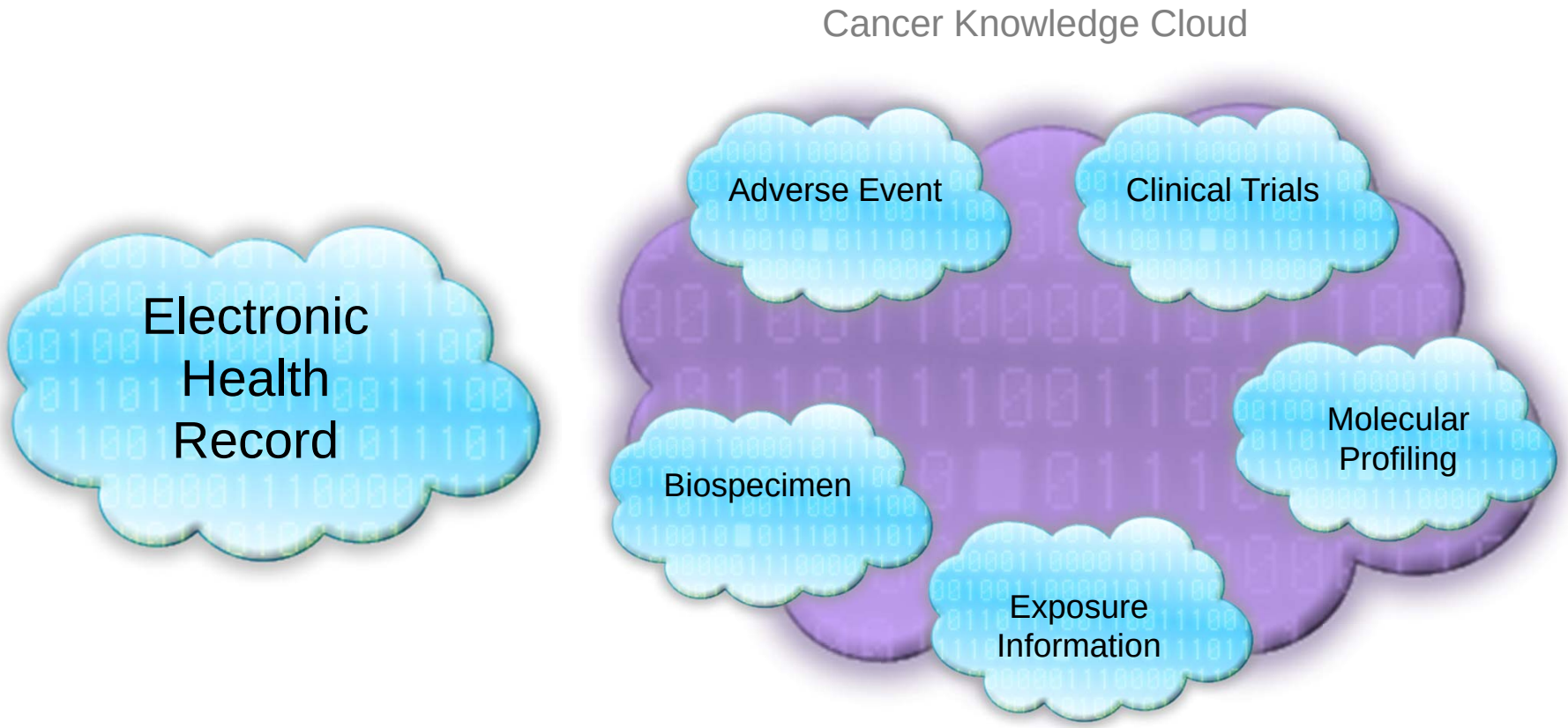
Rapid Learning Healthcare System

The Rapid Learning Healthcare System links research and care in a “virtuous circle”:

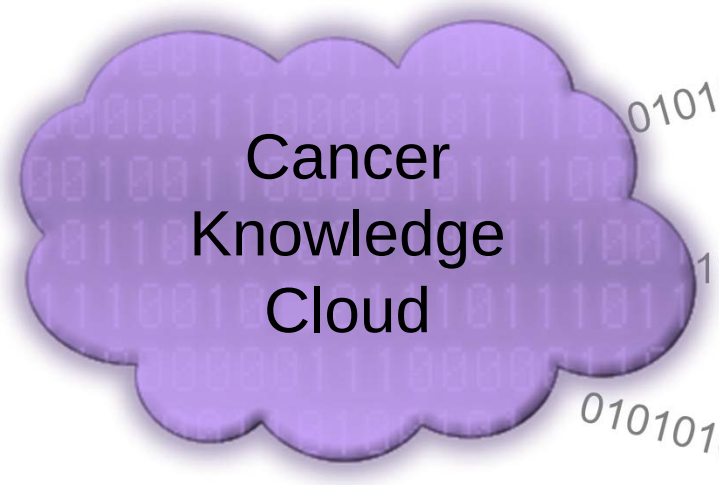
- Patient outcomes data is combined and aggregated to inform discovery and care in real-time
- Latest scientific findings are made available to physicians to encourage rapid clinical adoption
- Discovery is transformed into a natural outcome of patient care

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



Researchers Can Query the Data in the Cancer Knowledge Cloud



01010101010101010101010101

10101010101010101010101

01010101010101010101010101

Epidemiologists

- Query data to seek correlations among genes, environment, outcome
- Develop standing online cohorts of volunteers

Basic Researchers

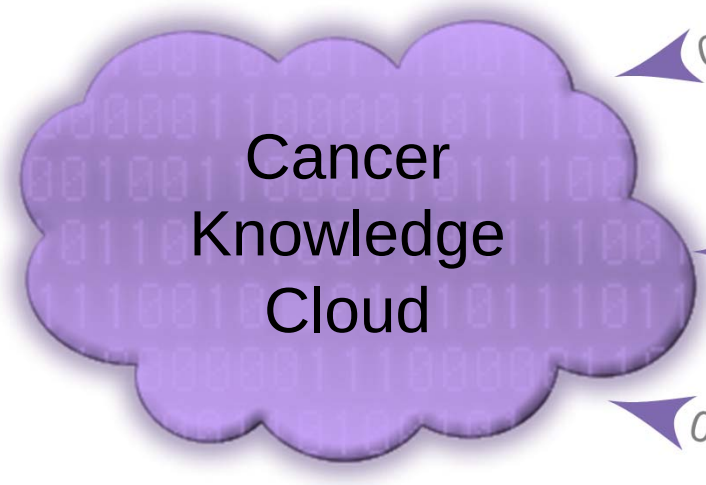
- Generate new hypotheses
- Identify biomarker-outcome correlations
- Validate biomarkers *in silico*

Clinical Researchers

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

caBIG cancer Biomedical Informatics Grid™

New Knowledge from Research Is Fed into the Cancer Knowledge Cloud



01010101010101010101010101

01010101010101010101010101

01010101010101010101010101

Epidemiologists

New links to behaviors and exposures that increase / decrease risk of disease or disease reoccurrence

Basic Researchers

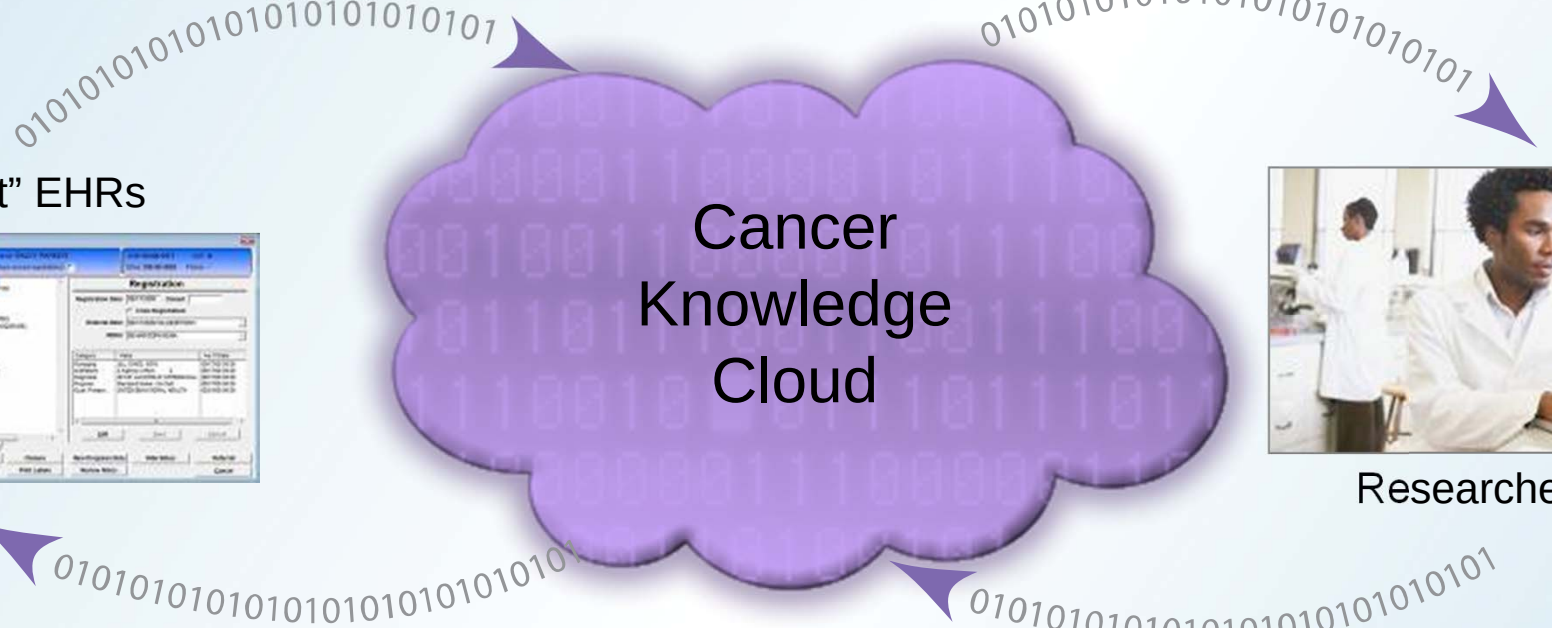
New drug targets

Clinical Researchers

Targeted drugs for molecularly-defined sub-groups

caBIG cancer Biomedical Informatics Grid™

Virtuous Circle From Smart EHRs Through Research and Back to Clinical Care (Rapid Learning Healthcare System)



"Smart" EHRs

Researchers

caBIG cancer Biomedical Informatics Grid™

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

Analyze

Patient Info

Data Subset

	Attribute	Value
☑	Gender	Male
☑	Age Group	65-69
☑	Race/Ethnic Grou	Black
☑	Family History Fl	Y
☑	Prostate Size Fla	Y
☑	PSA Flag	Y
☑	PSA Change Flag	Y
☑	Alcohol	Moderate
☑	Smoker	No
☑	Clinical Stage	2
☑	Biopsy Stage	NA
☑	Incontinent	N

Analyze | Reset me

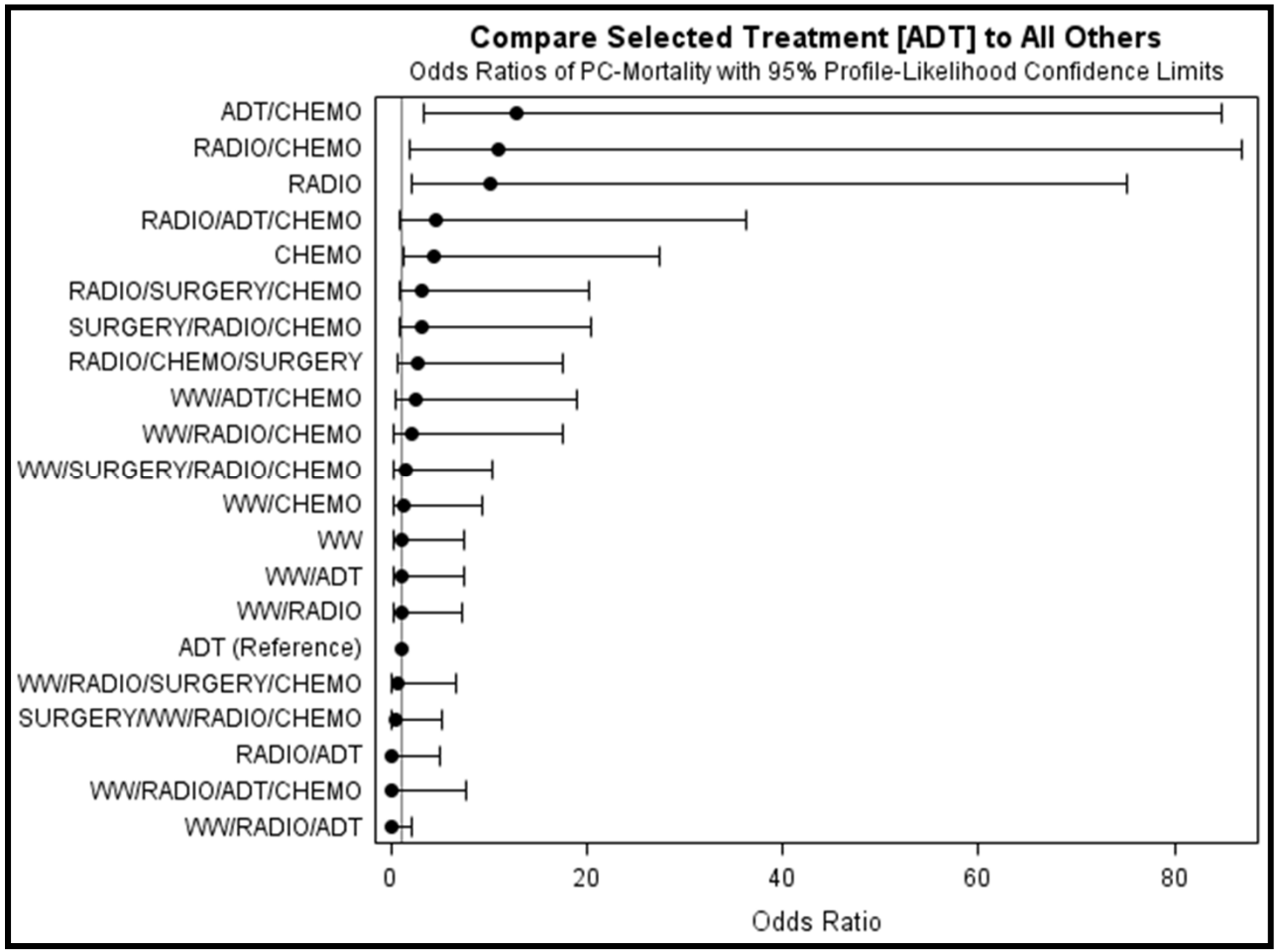
Chart | Report | Patient Detail

Analysis Results for Selected Patient Group

Treatment Sequence	N	Erectile Dysfunction	ED n	Mean PSA Change	Mean Change ECOG	Mean Change Karnofsky	Mean Survival (YRS)
RPS	121	55%	(119)	-3 (117)	-3(89)	-2.5 (50)	14.2 (44)
IMRT	60	20%	(40)	+3.2 (30)	-.6 (25)	-1.2(15)	8 (35)
IMRT/LHRH Analog	20	19%	(20)	+9.2 (18)	-2 (11)	-1.3 (5)	11(9)
RPS/LHRH Antagonist	70	46%	(60)	-4.2 (60)	-3.2 (58)	-2.2 (16)	17 (40)
EBRT/Brachy	58	36%	(58)	-5.2 (50)	-3.8 (48)	-2.9 (21)	10 (30)
WatchfulWaiting	70	46%	(60)	-4.2 (60)	-3.2 (58)	-2.2 (16)	17 (40)
WatchfulWaiting/Novantrone	20	46%	(18)	-4.2 (18)	-3.2 (18)	-2.2 (16)	17 (11)
RoboticLRP/Firmagon	70	46%	(60)	-4.2 (60)	-3.2 (58)	-2.2 (16)	17 (40)
PP/IMRT	70	46%	(60)	-4.2 (60)	-3.2 (58)	-2.2 (16)	17 (40)

caBIG cancer Biomedical Informatics Grid™

Comparing Mortality Due to Prostate Cancer *For Androgen Deprivation Therapy Compared to All Other Treatments*



Outcomes for Selected Treatment Plan

Analyze

Patient Info

Select a Treatment

Select	Historical Treatment Plans for Prostate Cancer	Count
○	Active Surveillance	5554
○	Radiotherapy – Proton Beam ☐ Chemo yyyy	332
◉	Robotic Prostatectomy ☐ Chemo xxxx	1024
○	Robotic Prostatectomy ☐ Radiotheraphy – Proton	289
○	Conventional Prostatectomy ☐ Chemo	3078
○	Radiotherapy-Conformal ☐ Active Surveillance	663
○	Radiotherapy-IMR ☐ Chemo yyyy	50
○	Chemo xxxx ☐ Chemo zzzz	43

Breakdown Report

Analyze

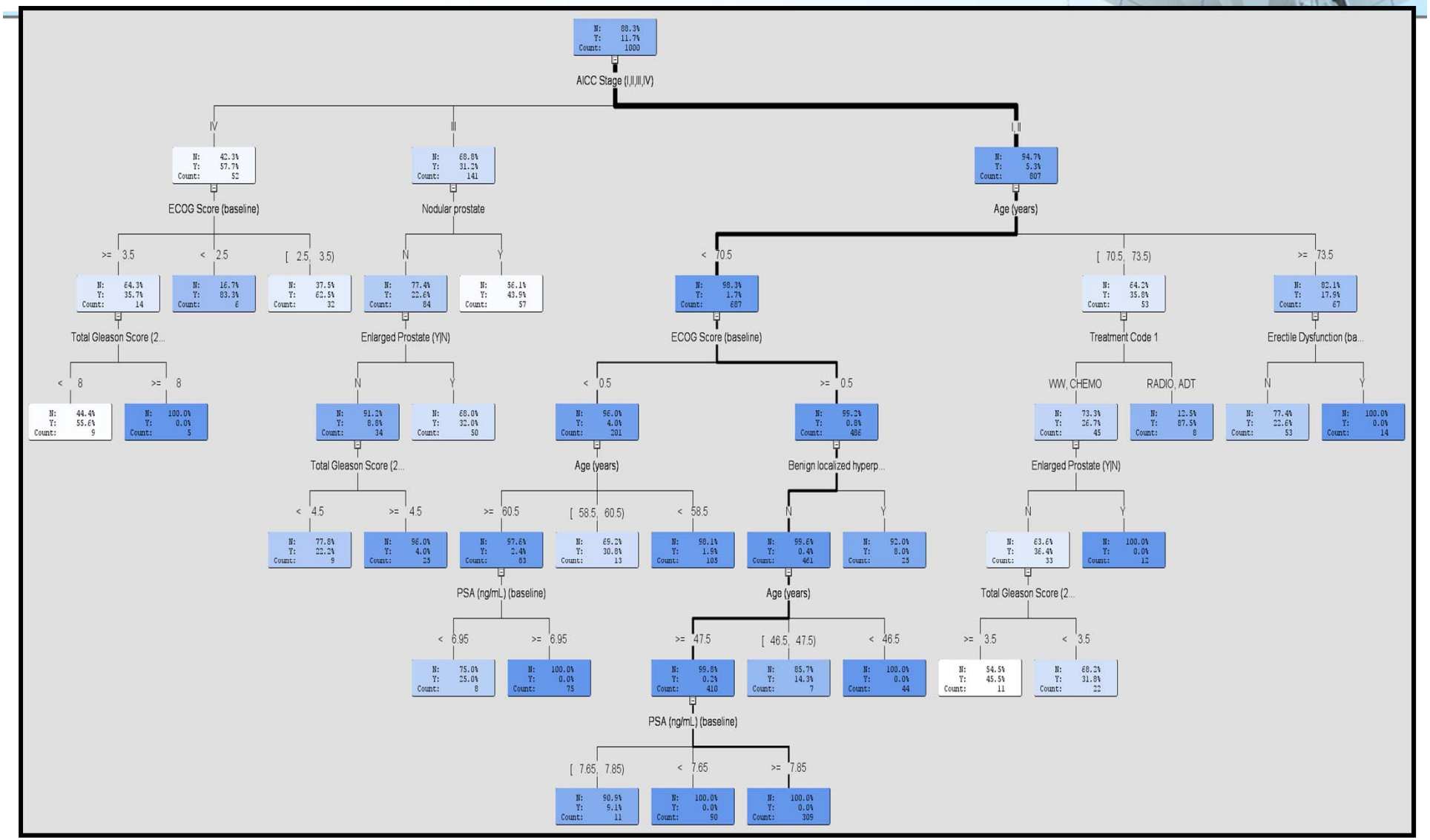
Chart | Summary Report

For Prostate Cancer Treatment Plan: Robotic Prostatectomy -> Chemo xxxx

Factor	Value	N	Mortality Rate	Mean PSA Change	Mean Prostate Siz	Mean EGOG	Incontinenence
Sex	Male	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Unknown	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Other	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
Race	Non-Hispanic Wh	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Non-Hispanic Bla	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Hispanic	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Asian	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Other	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	Unknown	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
Enlarged Prostat	Yes	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	No	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
Family History	Yes	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	No	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
PSA	<3	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	>=3	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
Age Group	50-59	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	60-69	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	70-79	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	80+	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
Clinical Stage	1	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	2	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	3	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)
	4	nnn	xxx%	xx(nnn)	xx(nnn)	x(nnn)	xx(nnn)

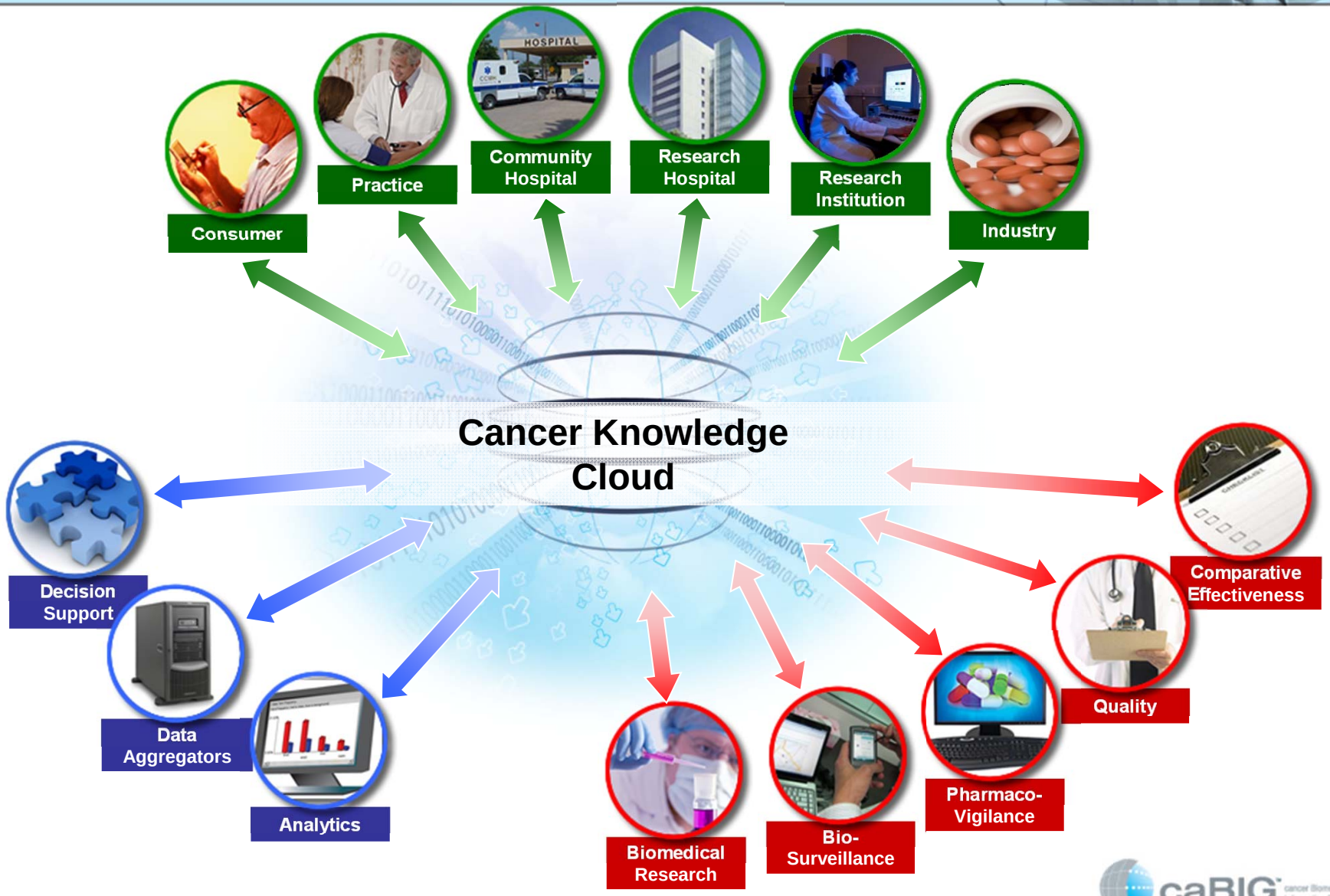
caBIG cancer Biomedical Informatics Grid™

Decision Tree:
Predicted Change in Mortality



caBIG cancer Biomedical Informatics Grid™

Result: Researchers have unprecedented access to huge depth and breadth of resources



caBIG cancer Biomedical Informatics Grid™

caBIG™
cancer Biomedical Informatics Grid ™

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):

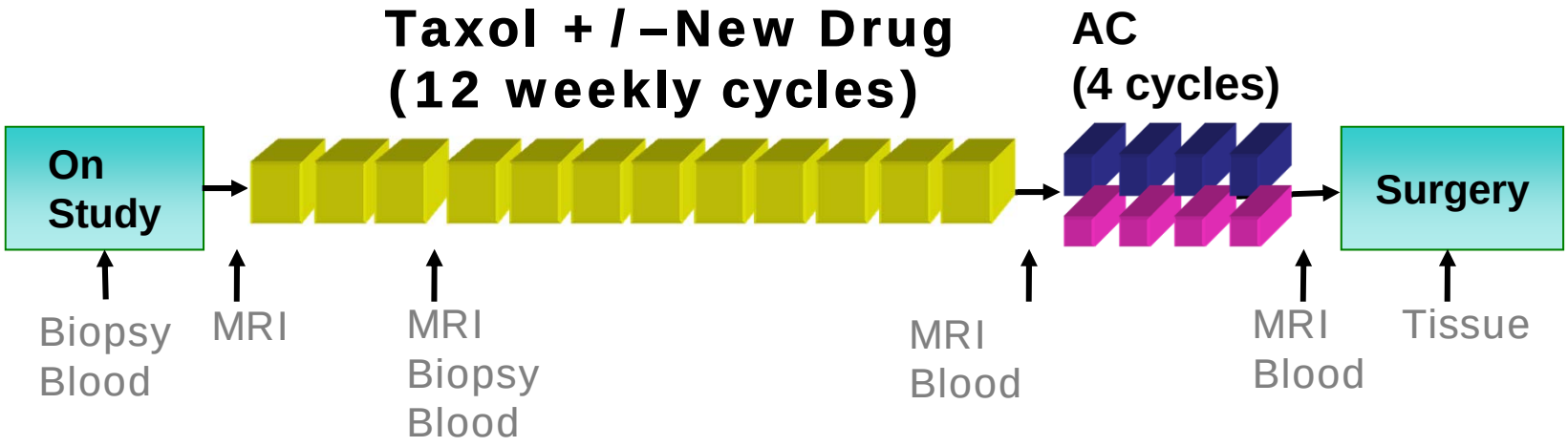
A national study to leverage biomarkers
in predicting response to combinatorial therapy for
women with Stage 3 breast cancer.

(PI Laura Esserman, UCSF)

caBIG cancer Biomedical Informatics Grid™

Projected I-SPY 2 study sites

I-SPY Adaptive Trial Outline



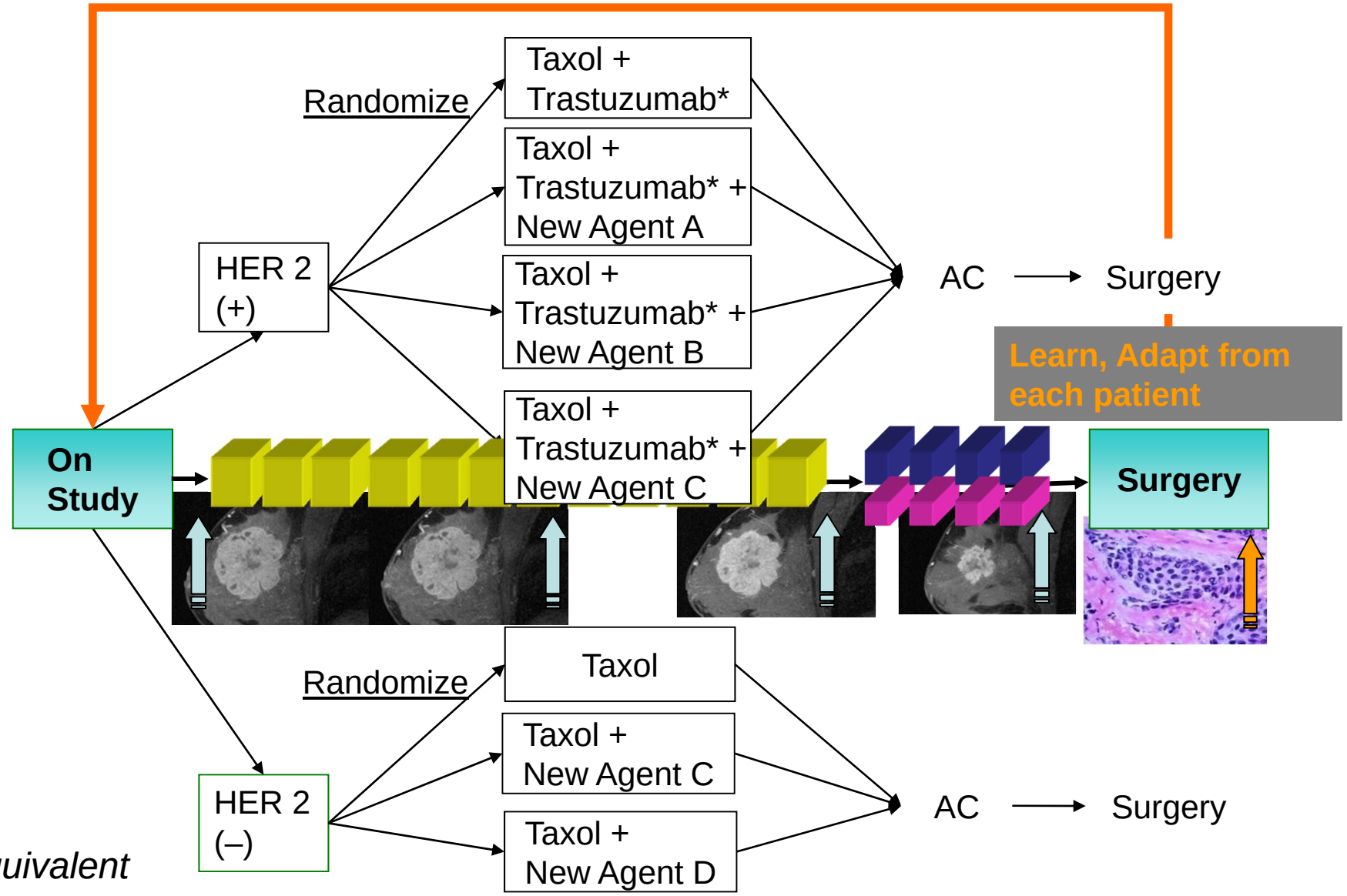
Accrual: Anticipate 800 patients over 3–4 years

Enroll ~20 patients per month

Participating Sites: 15–20 across US and Canada

I-SPY Adaptive Trial:

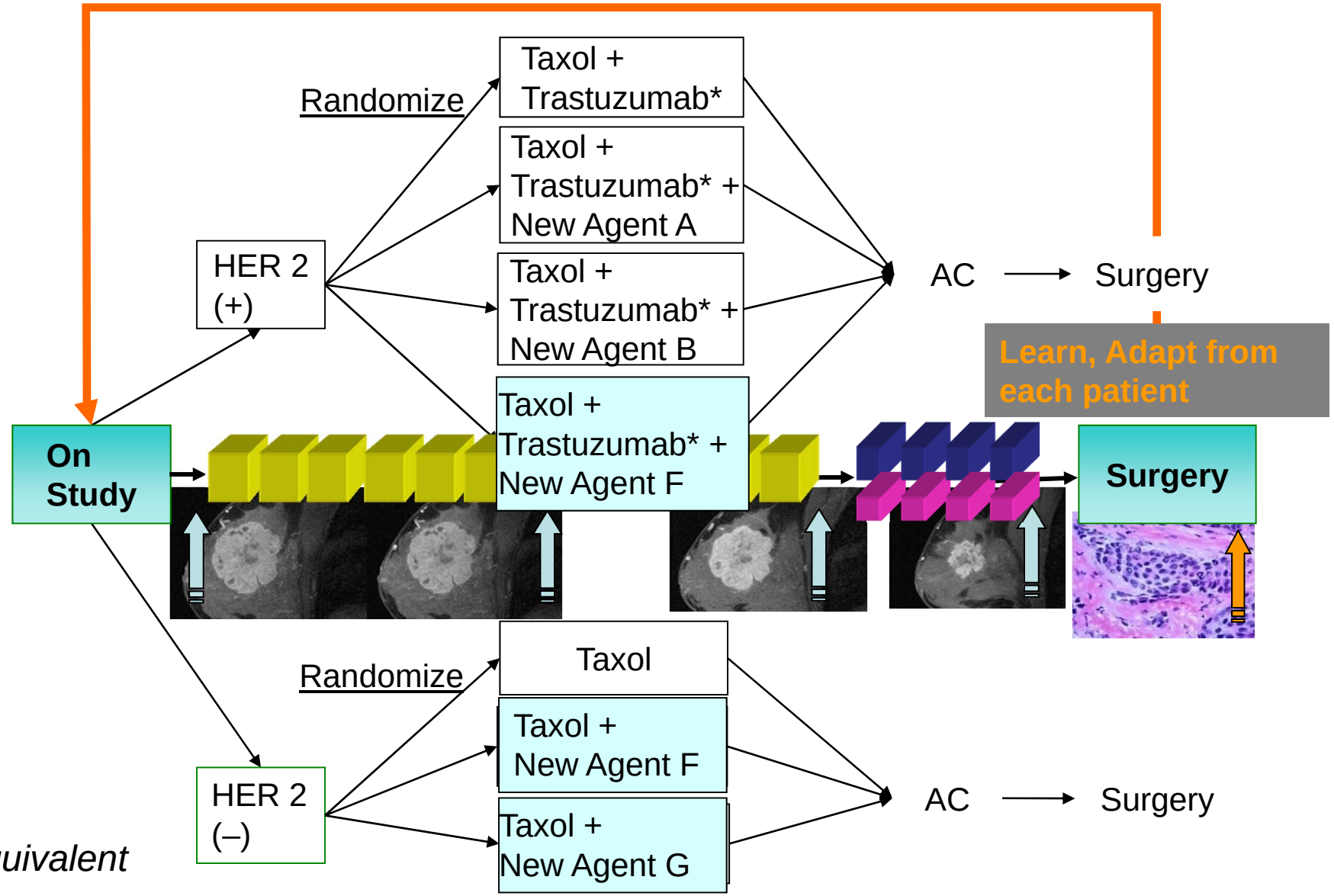
Introduce several new agents for a given profile



**Or Equivalent*

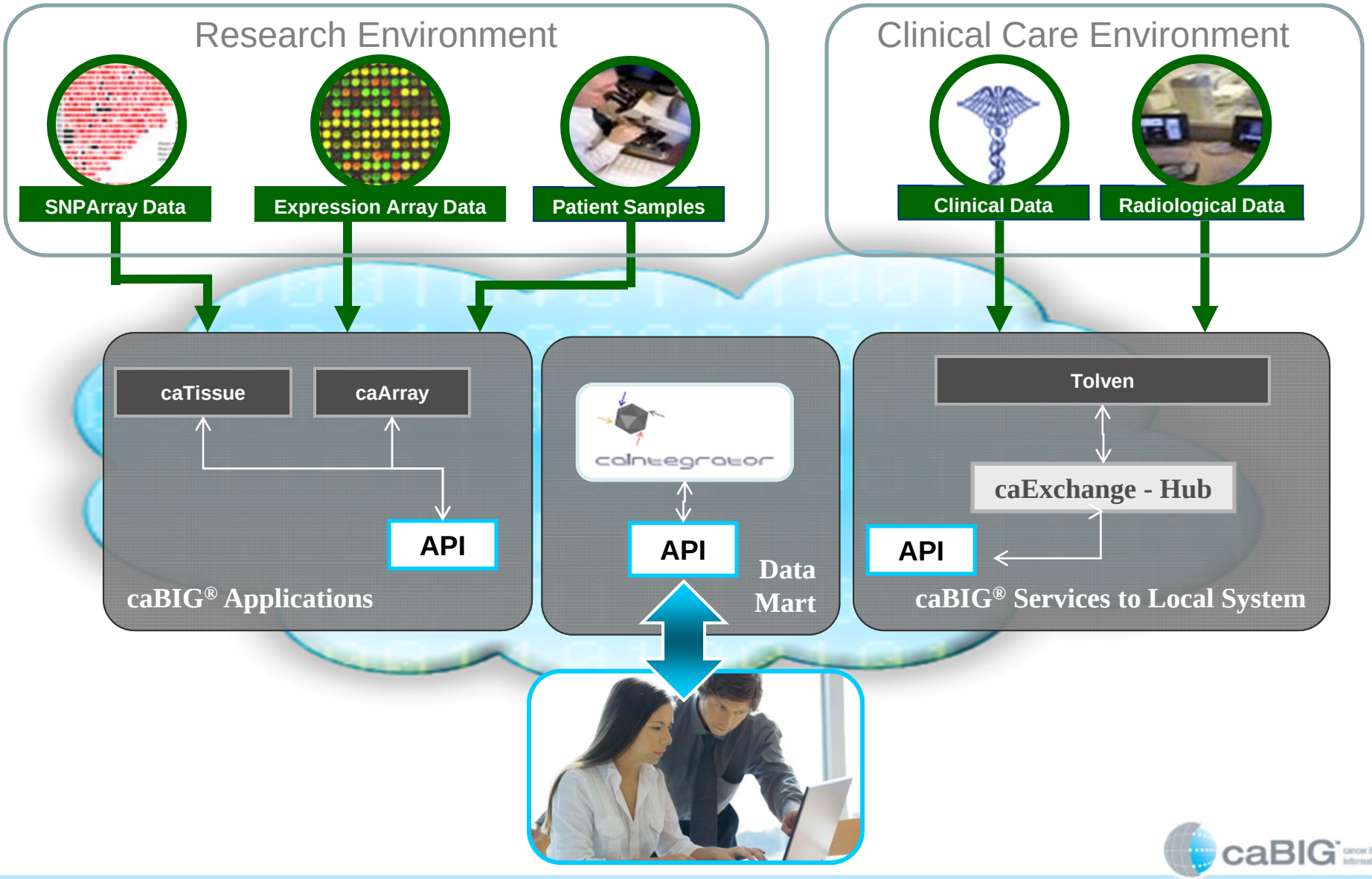
I-SPY Adaptive Trial:

Introduce several new agents for a given profile



**Or Equivalent*

I-SPY TRIAL IT Infrastructure



Summary

- **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