Informatics Needs and Challenges in Cancer Research: Cooperative Groups Perspective

Mitchell Schnall, MD, PhD and Robert L. Comis, MD

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

National Cancer Policy Forum

Institute of Medicine

February 27-28, 2012

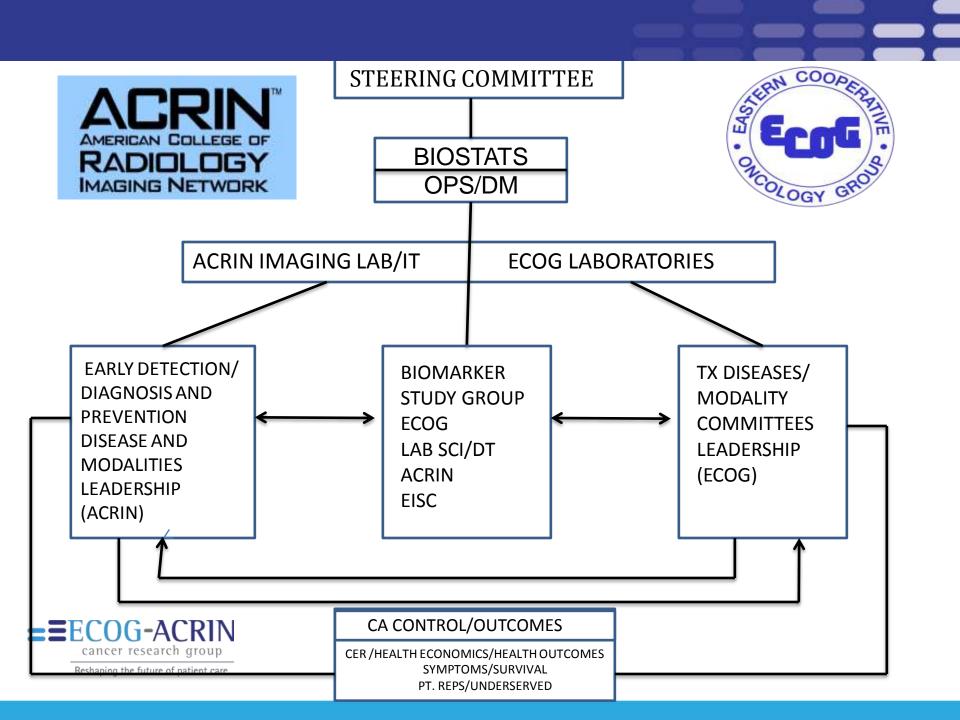# Reconfigured Cooperative Group System

ALLIANCE – CALGB, NCCTG, ACOSOG

COG – Children's Oncology Group

ECOG – ACRIN Cancer Research Group

NRG – NSABP, RTOG, GOG

SWOG

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Cooperative Group Environment

Cross Section of Health Care System

- Academic Centers
- Large Community Practices
- Smaller Community Practices
- Biomedical Scientists
- Geographically Dispersed
- Include international sites/scientist

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Historical Perspective

- Group trials focused on clinical endpoints
- Data almost exclusively from CRF's
- Individual Group database
- Local systems to manage operations
- Ad hoc interfaces for reporting
- Local IT systems for tissue banks

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Historical Perspective

Toward Development of Integrated IT platforms

- Adoption of common data elements
- Establishment of CTSU
- Initiate process of common electronic data capture system

Cooperative Group RDE Project:
Clinical Data Management System Procurement Timeline

- Oct 2005: Proposal for a common system presented by Group Statisticians to Coalition Board
- Feb 2006: Approved in principle and for funding by Coalition Board. Formal technical and business evaluation initiated
- Jan – April 2007: Three vendors analyzed in depth
- Aug 2007: Specifications presented to CTMS Steering Committee
- Sep 2007: Final requirements specifications forwarded to NCI
- Aug 2008: Award announced to Medidata, Inc.

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Cooperative Group RDE Project:
Procurement Timeline (cont'd)

- Sept 2008: Protests to US GAO from all three unsuccessful finalists; automatic 100-day stay of all work under contract

- Dec 2008: GAO determined solicitation not sufficiently clear about government's license requirements; recommended solicitation be amended. Final proposal revisions solicited from offerors in the competitive range in support of a new award determination

- Jul 2010: Litigation continues unabated

- April 2011: Program officially initiated

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Background for the Initiative

- The National Cancer Institute (NCI) Cancer Therapy and Evaluation Program (CTEP) purchased licensing rights from Medidata Solutions (Medidata) to use and distribute Medidata Rave®
  - Rave is a Clinical Data Management System (CDMS) software application, to facilitate the conduct of clinical research throughout the NCI-supported clinical research enterprise
  - Rave is a web-based system for capturing, managing and reporting clinical research data that enables the user to record patient information using forms customized per study (visit, lab and adverse event data)
  - Program officially initiated **April 1, 2011**

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Collaboration Across the Stakeholders

- Program is managed by CTEP with the assistance of the CTSU and Medidata
- Eight priority one working groups were established in November 2010 to provide a collaborative forum for the review and discussion of project wide implementation decision points
  - ~260 participants from the stakeholders in the working groups
- Stakeholder groups have dedicated hundreds of hours over the past year both to the working groups and training!

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Training- what does it take to deploy Rave?

- Over 250 group representatives participated in a 2 day classroom introduction to Rave
- ~ 300 group representatives are currently engaged in training on Rave being offered via classroom, webinar and eLearning
  - Metrics include:
    - ~120 reps trained in Study Design and Study Build
    - 64 instructor led webinars to date
    - 36 classroom sessions held to date lasting 2-5 days

Training for the groups will continue through August 2012!

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

The GBC Reporting Tool- A Web-Based Catalog of NCI Cooperative Group Trial Biospecimens

Dave Billiter, MBA, PMP, Mark Watson, MD, PhD, and the NCI Group Banking Committee Informatics Subcommittee
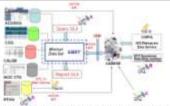
BACKGROUND

The NCI Cooperative Group Banks provide an unparalleled resource of biospecimens associated with multi-institutional, Phase II / III therapeutic trials. A past limitation to more generalized use of this unique and highly valuable biospecimen resource has been the lack of a comprehensive inventory across all nine NCI-funded cooperative groups and their corresponding biorepositories.

One charge of the Informatics Subcommittee of the NCI Group Banking Committee (GBC) is to achieve Cooperative Group Bank harmonization by establishing common vocabularies for biospecimen data exchange. Accordingly, the GBC has developed the Group Bank Reporting Tool (GBC-RT), a web-accessible data warehouse that integrates the inventories of over ten, diverse cooperative group banks into a single data source. When fully operational, the GBC-RT will be connected with the NCI Cooperative Group Bank public web site, and will greatly facilitate and accelerate the search and utilization of Cooperative Group trial biospecimens for novel translational cancer research.

APPROACH

Although some Group Banks are migrating to caGrid-enabled systems, each bank currently uses a unique system and data model for biospecimen tracking. Therefore, a set of ETL tools and web services were developed to allow each bank to map a minimum set of common data elements for periodic migration to a central data warehouse (GBRT). The data store has a web-based user interface for query and data reporting (A). Plans to caGrid-enable the data warehouse will allow for future points of integration with the NCI's Biospecimen Resource Locator II and clinical trial annotation services (B), and create a caGrid-accessible data service (C). Once potentially suitable specimens are identified across the Group Banks, investigators use a common Request for Specimens form to make requests from specific banks (D).

DATA MAPPING

To ensure caBIG® compatibility, the GBC-RT uses the Common Biospecimen Model- CBM V0.9 (A)- for mapping individual Group Bank data elements to a common, minimal data set (B). A web-based interface (C) allows bank administrators to specify data mappings from their data to the CBM, to allow regular, semi-automated data upload.

SYSTEM FEATURES

- Microsoft .NET 3.5 (ASP.NET) Framework
- SQL Server 2008 Database
- Business Objects 2008 Reporting Engine
- Secure Web Services for Data Upload
- Role-Based User Sign-On and Access / Functionality
- CBM-based data model / caDSR-based permissible values
- Transaction Log and Full Auditing Capability
- Drill-down and Filtering Report Functionality

QUERY AND REPORTING

Role-based access provides users with a list of pre-defined query options (left) which can be refined using a set of controlled vocabularies for specific query parameters (center). The results of each query (right) can be exported and gives reference to the specific group bank(s) that may be contacted directly about specific biospecimen availability.

PUBLIC ACCESSIBILITY

To further promote biospecimen accessibility, the GBC-RT will be integrated within the public-facing NCI Cooperative Group Bank website, which is also currently under development.

FUTURE GOALS

- caGrid Connectivity
- Biospecimen Annotation with Clinical Trial (Protocol) Information Using NCI Enterprise Data Services
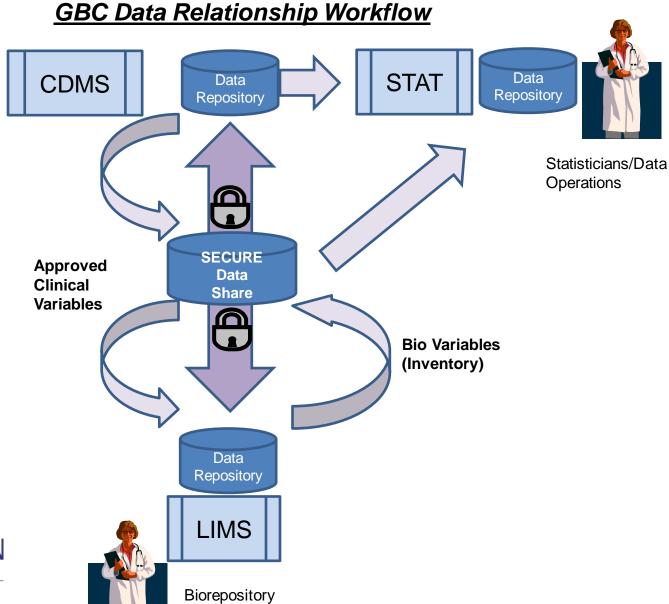- Development of a Common Biospecimen Submission and Tracking Tool

ACKNOWLEDGEMENTS

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

GBC Data Relationship Workflow

CDMS

Data Repository

STAT

Data Repository

Data Manager, CRA, etc.

Statisticians/Data Operations

SECURE Data Share

Approved Clinical Variables

Bio Variables (Inventory)

Data Repository

LIMS

Biorepository Personnel

ECOG-ACRIN
cancer research group

Reshaping the future of patient care
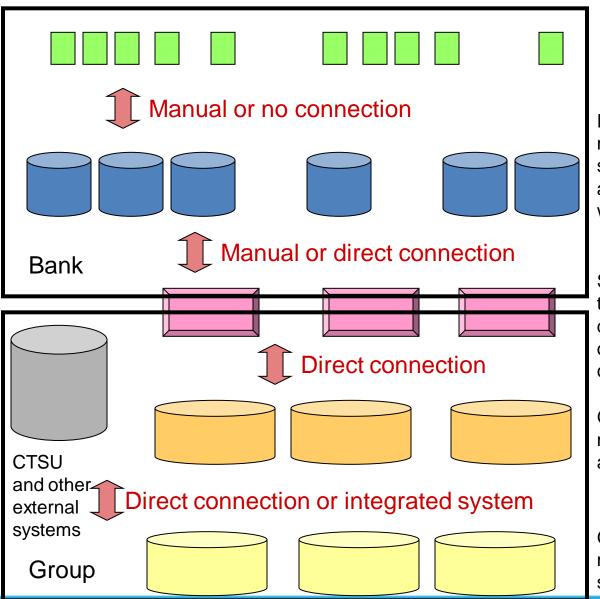
Relevant Data Components

Some banks have legacy data sets of biospecimen collections

Manual or no connection
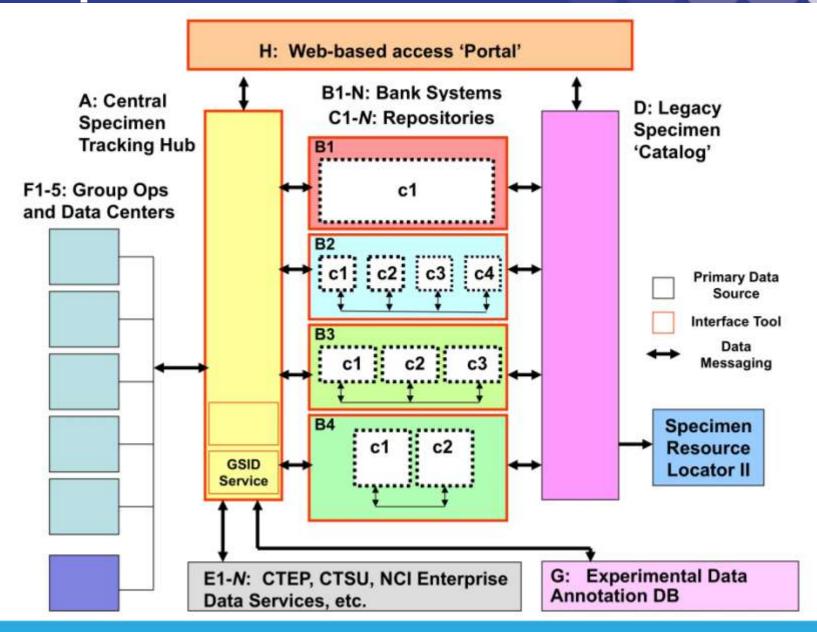
Each bank has its own biospecimen management system. Some systems integrate multiple banks and some systems are integrated with institutional systems

Bank

Manual or direct connection

Some groups use 'specimen tracking systems' which connect data entry at clinical sites with trial operations systems and in some cases, bank inventory systems

Direct connection

Group CTMS (Trial operation management systems)- enrollment and tracking

CTSU and other external systems

Direct connection or integrated system

Group

Group CDMS (Clinical data management)- data annotation and statistics

Proposed GBC Informatics Network

Current Challenges/Opportunities

Increasing role of correlative science in group trials

- Multiple data streams contribute to study database

- Complex analysis/modeling

- Interface to promote and support scientific collaboration

- Cooperative group realignment is driving reconfiguration of Group IT infrastructure

ECOG-ACRIN
cancer research group

Reshaping the future of patient care

Unique Opportunity in Cooperative Groups

Biosamples / Images associated with structured clinical information
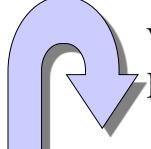
- Diverse patient population from a network committed to team science
- Clinical information collected in the context of trials is structured by the CRF
- Adoption of CDE's
- Track record of long term follow up

ECOG-ACRIN
cancer research group
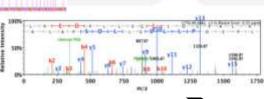Reshaping the future of patient care

10^1

Vetted by Domain Knowledge Expertise

Integrating Information in Disease x Outcome Categories
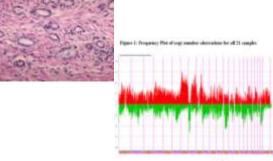
Convergence of large, diverse, scaled, fused non-linear data streams

Facilitated by Machine Learning

10^{-10}
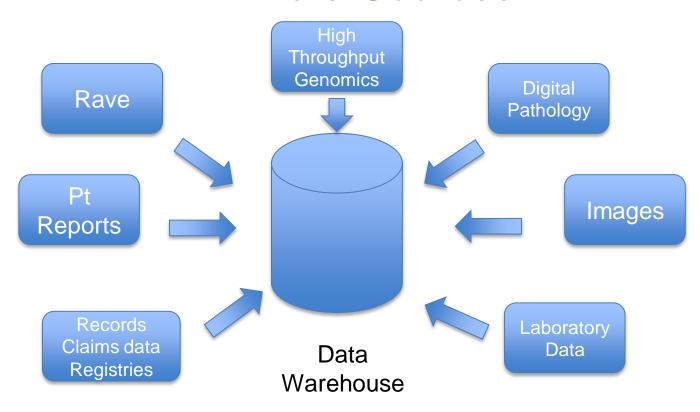
ECOG-ACRIN Vision

Integrated Data Warehouse

- CRF data
- Imaging data
- Tissue/Specimen repository inventory
- Digital pathology
- "Omics"

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Data Sources

Guiding Principals

- Embrace relevant standards
- Deep annotation
- Query interface
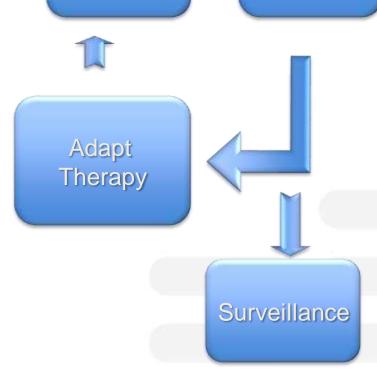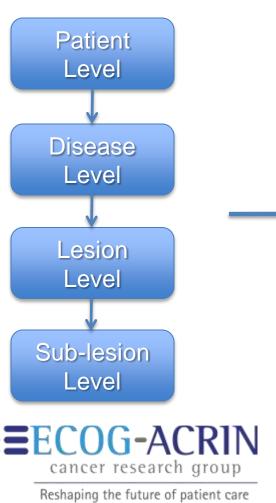- Data derived from ECOG-ACRIN specimens/images added to warehouse

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Diagnostics and the clinical care path

Prevention → Detection → Characterization → Optimized Treatment → Response Assessment

Adapt Therapy → Optimized Treatment

Response Assessment → Adapt Therapy

Response Assessment → Surveillance

Data Sources
- History
- Physical Exam
- Laboratory /Pathology
- Genetics/Expression Profile
- Imaging

ECOG-ACRIN
cancer research group
Reshaping the future of patient care

Informatics Structure

Triad
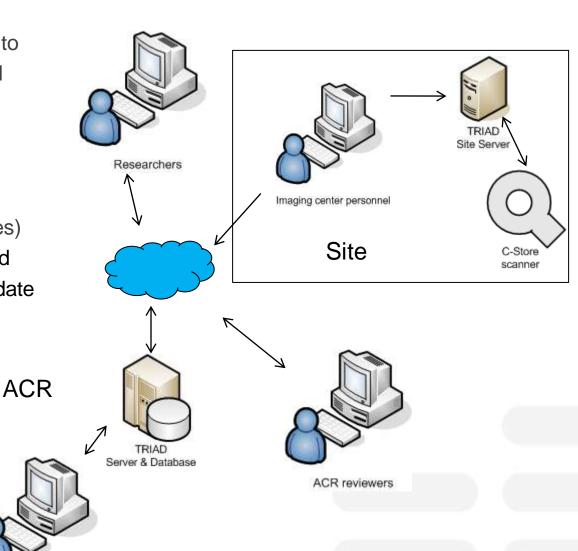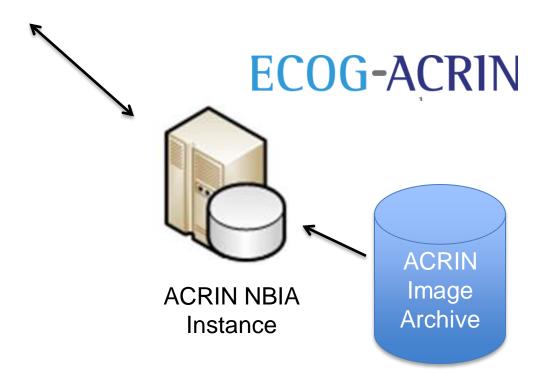
- Standards-based tool developed to facilitate the exchange of medical images and support clinical trials workflow.
- GCP compliant
- In production at over 200 sites
- Over 100,000 cases (23 terabytes)
- AIRP – improved non-DICOM and large-file support (e.g. accommodate Pathology studies)
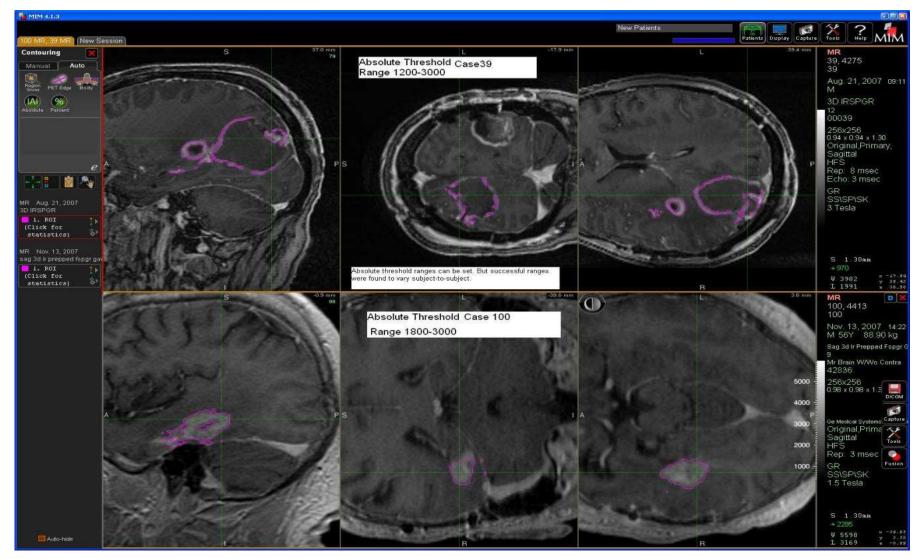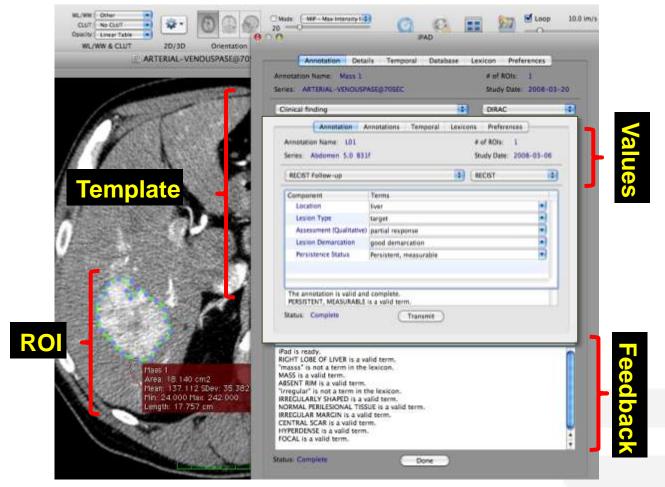- Integration with Medidata RAVE

Researchers

Imaging center personnel

TRIAD Site Server

C-Store scanner

Site

ACR

TRIAD Server & Database

ACR reviewers

ACRIN HQ Staff

ECOG-ACRIN
cancer research group
Reshaping the future of patient care
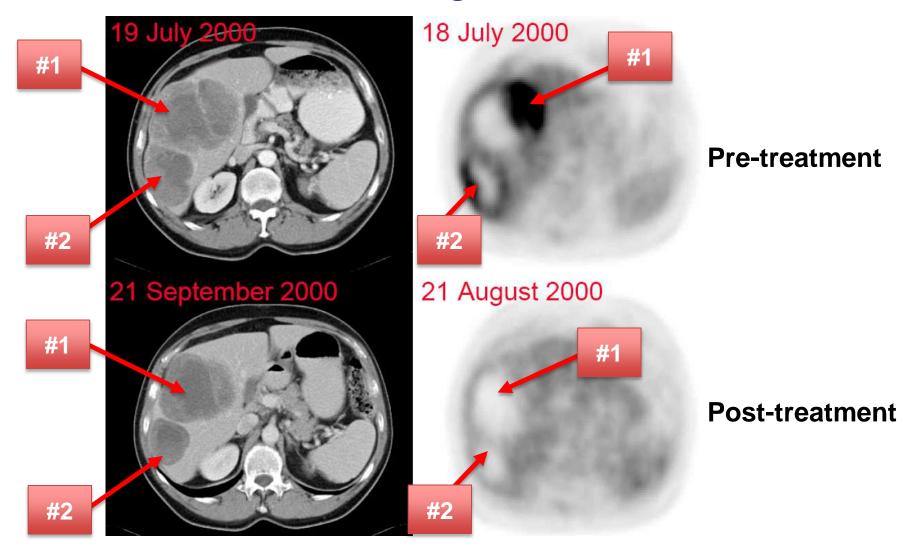
Annotated Functional Imaging

Annotation Image Markup (AIM standard) ePad implementation

Preserve lesion indexing across modalities / time

Pathology / Imaging linkage