Georgetown | Lombardi
COMPREHENSIVE CANCER CENTER

PATIENT CARE
RESEARCH
EDUCATION
COMMUNITY

Enabling 21st Century Cancer Research

Integrating Clinical and Molecular Databases to Foster Research and Patient Care

Louis M. Weiner, MD
Director, Lombardi Comprehensive Cancer Center
Chair, Department of Oncology
Georgetown University Medical Center

NCI CCC *A Comprehensive Cancer Center Designated by the National Cancer Institute*

http://lombardi.georgetown.edu
Lombardi CancerLine: 202.444.4000

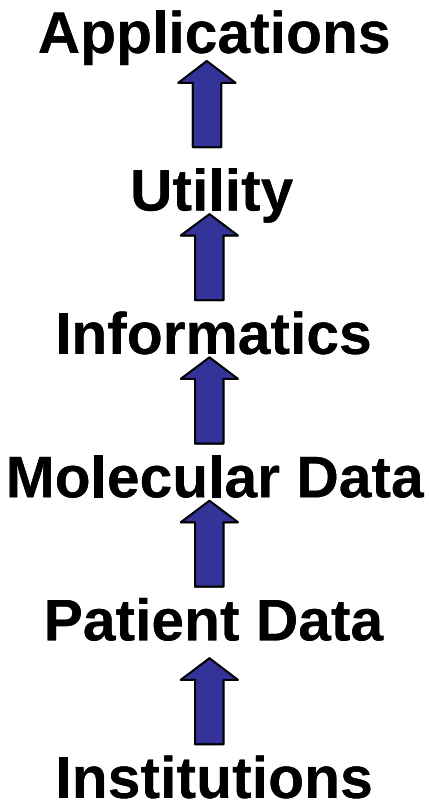
New ways are needed to collect, store, analyze and implement research discoveries

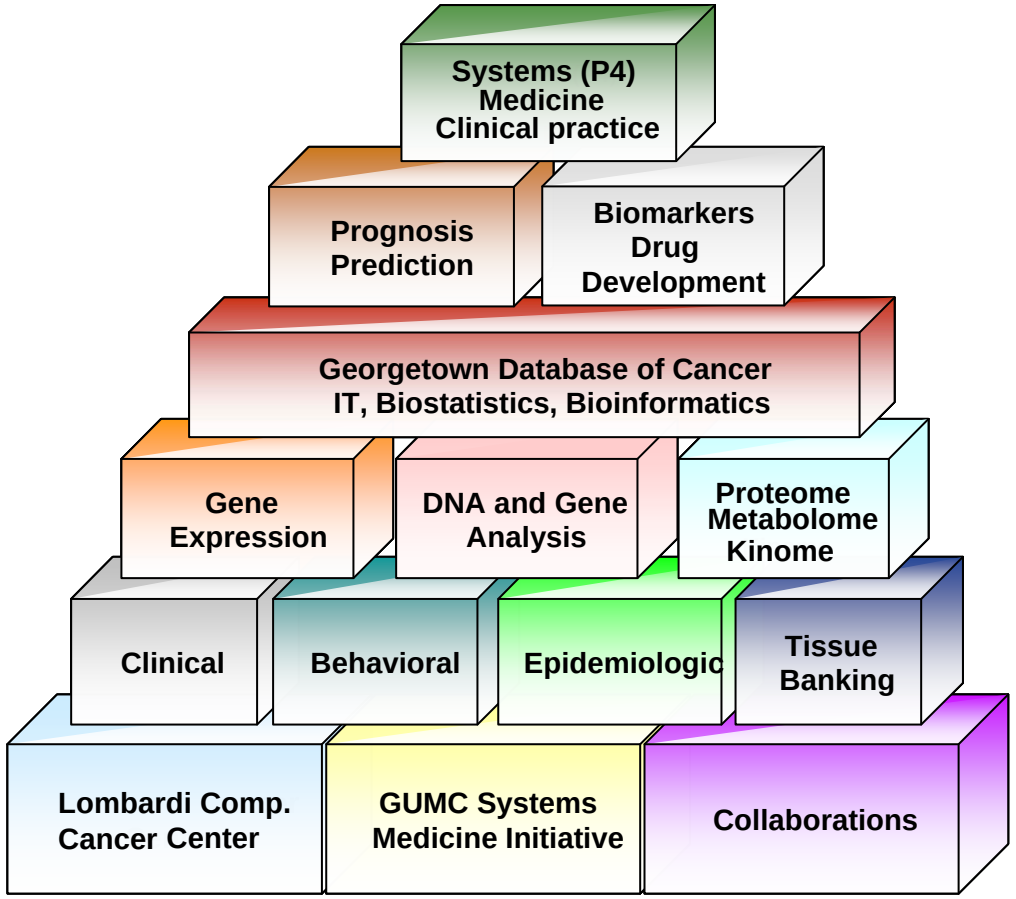
- Drug discovery efforts require massive, complex data repositories
- Shared resources contain huge amounts of unconnected, but potentially vital data
- Only a small fraction of patient-related information is prospectively collected and stored for future use
- Clinical and molecular databases rarely “talk” to each other
- Missed opportunities
  - Personalized medicine: predictive, prognostic
  - Development of cancer biomarkers
  - New drug discovery
  - New drug development and clinical testing

Systems Medicine Platforms are Needed

Georgetown Database of Cancer (G-DOC)

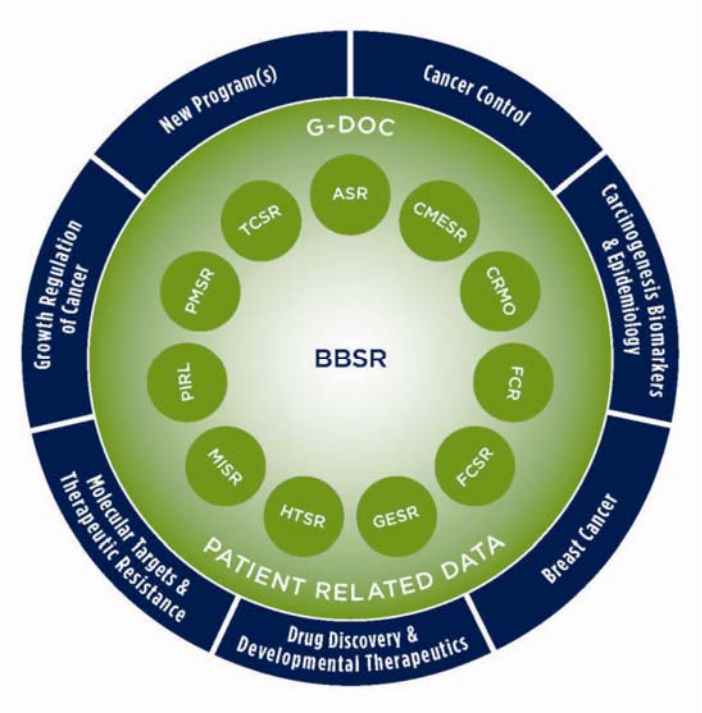
Towards an Integrated Clinical and Molecular Database of Cancer





G-DOC will Integrate Patient and Research Data

Current Configuration of Lombardi Programs and Shared Resources



Evolving Configuration of Lombardi Programs and Shared Resources

G-DOC will connect clinical and laboratory research, and link shared resources to these efforts

Value of G-DOC

- A rich resource for systems-based biology
- Connects multiple data types to yield better understanding of cancer biology better and faster
- Creates an enabling infrastructure to foster collaborations
- Integrates numerous “omics” platforms

Value of G-DOC (ctd)

- Create a Systems Medicine Environment
  - Correlate experimental laboratory data with clinical data (treatment, history, pathology, outcome, etc.)
  - Apply biomarker discoveries to sub-group patients for experimental treatments
  - Utilize clinical data to continuously evaluate outcomes
  - Develop and refine evidence-based patient management strategies at an individualized level

G-DOC Pilots at Lombardi

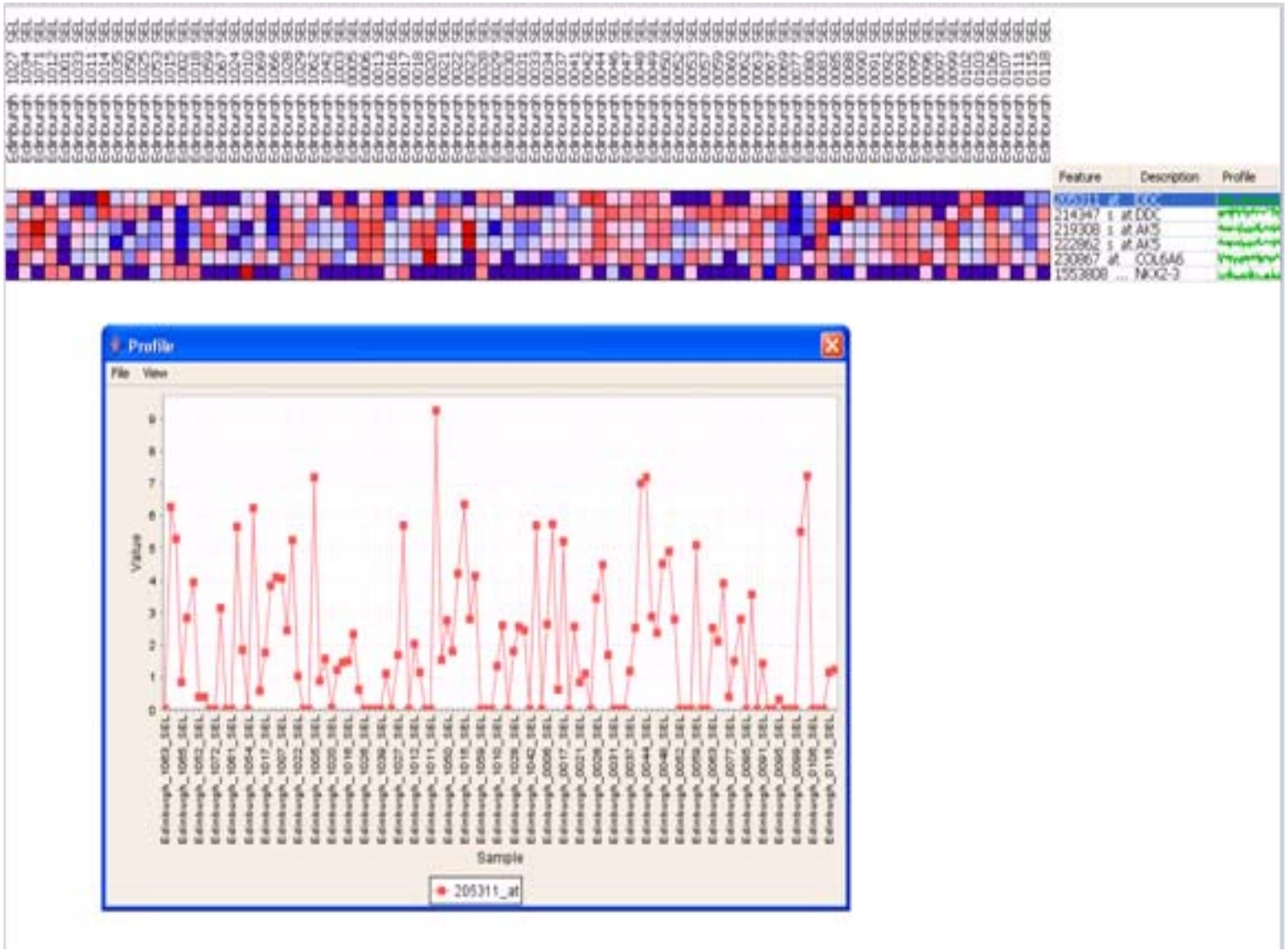
1. Familial Cancer Registry Shared Resource (FCR)
2. Prospective clinical trial in breast cancer
3. **Retrospective analysis of Stage II colon cancer samples and outcomes**
4. Clinical and Molecular Epidemiology Shared Resource (CMESR)

Retrospective analysis of predictors of outcome in Stage II colon cancer

- Tumor sample collection facilitates hypothesis generating studies
- Stage II colon cancer retrospective dataset
  - No chemotherapy
  - Clinical outcomes: 5-yr relapse and survival
- Standard genomics and transcriptomic analysis
  - miRNA, metabolomics
- Matched control samples are available
- Value: integrating multiple platforms

Easy to use interface

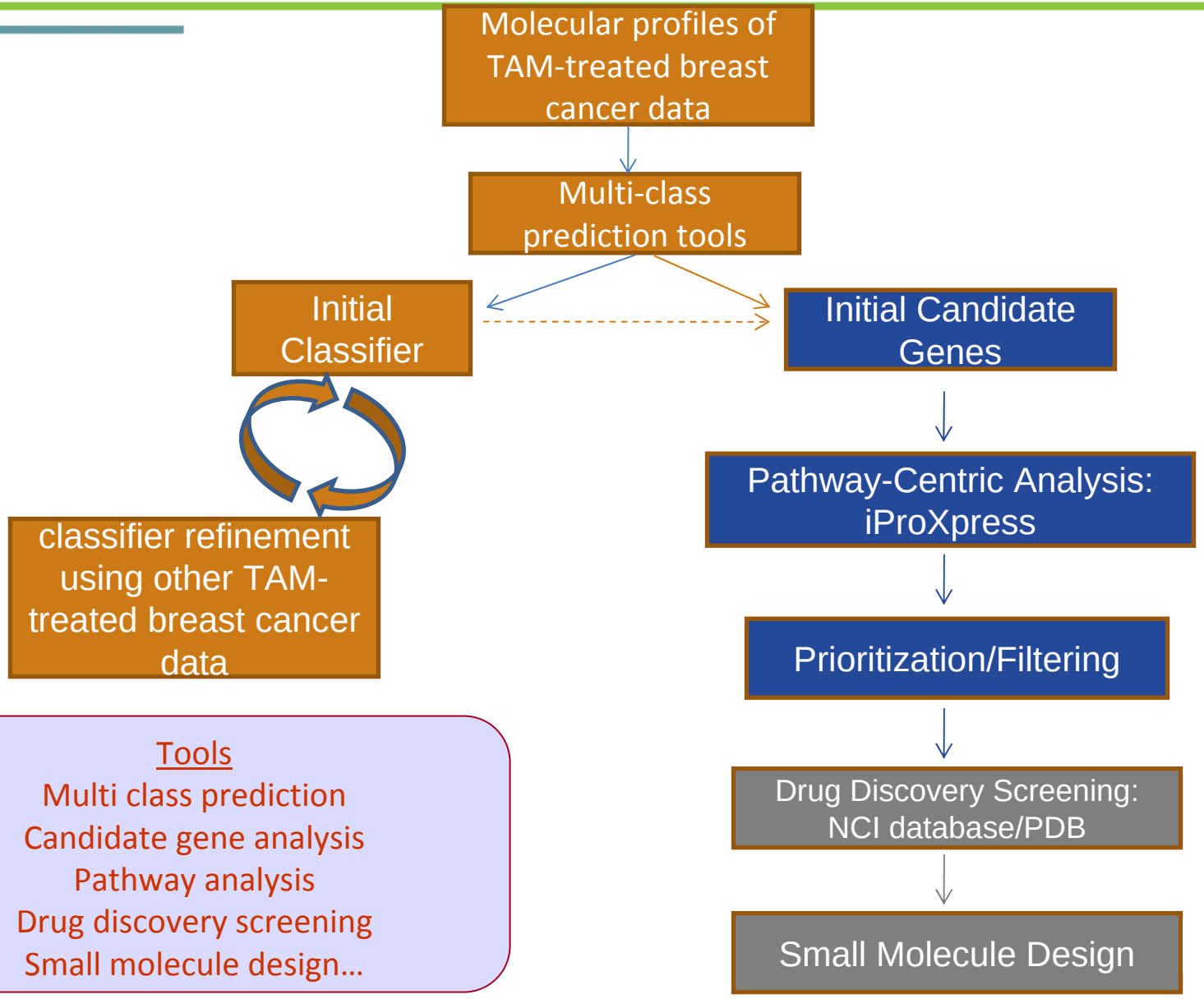
(Built for Physician Researchers; Integration of caBIG tools)



Lombardi - caBIG®

- Lombardi-NCICB collaboration predates the release of caBIG (SpiderWeb)
- Lombardi hosted the 2007 Architecture/Vocabularies and Data Elements Face-to-Face meeting
- Active participant across the entire caBIG program
  - "Funded developer" and "funded adopter" in the Integrative Cancer Research Workspace
  - "Funded developer" in the Architecture Workspace
  - "Funded adopter" in the Clinical Trials Management Workspace
  - "Funded participant" in all of these workspaces
- Developed one of the first reference implementations for caGrid: (gridPIR)

Improving Outcomes in ER+ Breast Cancer



Check out the G-DOC poster!

Lombardi COMPREHENSIVE CANCER CENTER AT GEORGETOWN UNIVERSITY

gdoc

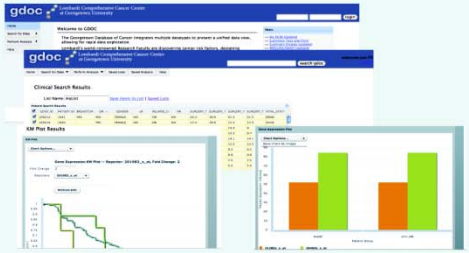
Georgetown Database Of Cancer

Abstract

G-DOC is an informatics framework that allows researchers, using unified data portals, to access and analyze clinical and research data across multiple trials and studies. The framework can be used to import data from multiple studies to access biomedical research data to perform analysis, generate ad hoc queries and customized reports. Moreover it provides a mechanism for integrating and aggregating biomedical research data and provides access to a variety of data types (e.g. tissue annotations, gene expression, SNPs, clinical trials data, health service related population data, etc.) in an integrated fashion. This framework, as envisioned by the Lombardi Comprehensive Cancer Center (LCCC) leadership, will facilitate systems medicine by providing easy identification of trends and patterns in these integrated datasets that ultimately results in better targeted/personalized therapies for cancer and in the future, other diseases.

The G-DOC application framework is being developed in phases based on requirements collected and specific study-trials that are contributing to the current phase requirements are the FCR (Family Cancer Registry), translational breast cancer pre-operative trial and the CMESR (Clinical and Molecular Epidemiology Shared Resource) studies. The alignment of G-DOC's goals with caBIG® and G-DOC development team member's experience in caBIG have naturally resulted in adoption of appropriate caBIG tools and standards. This system embeds a version of caIntegrator that is customized to fit the needs of LCCC researchers and plans to utilize a local LexEVS instance for standard vocabularies used by LCCC studies.

In addition, G-DOC team is actively exploring adoption of caTissue Suite and caArray as part of G-DOC framework to handle tissue and array information at LCCC.

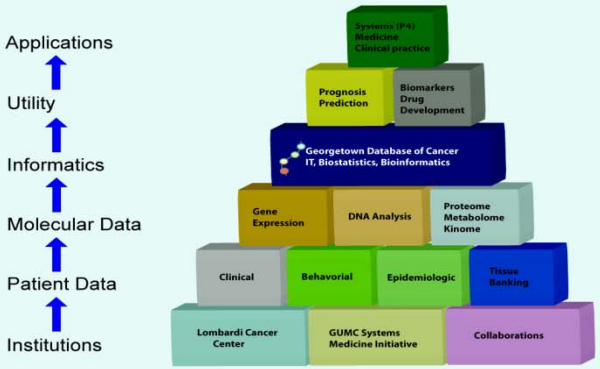
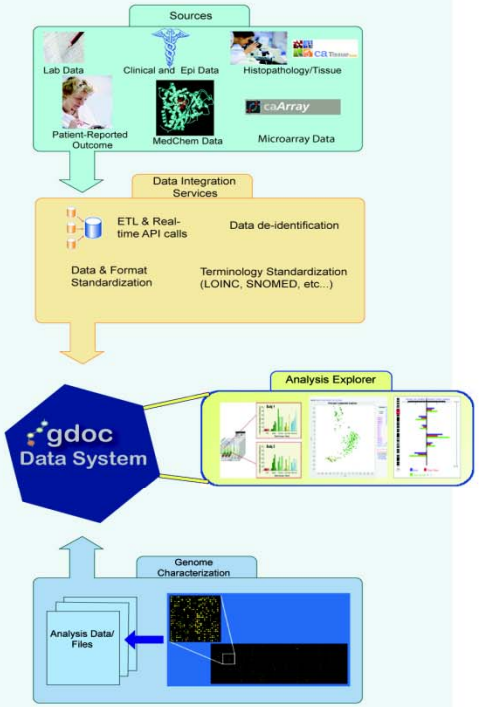


Motivation

* A rich resource for systems-based biology
* Connects multiple data types to yield better understanding of cancer biology better and faster
  -- Researchers are able to correlate experimental laboratory data with clinical data (treatment, history, pathology, outcome, etc.)
* Customize treatment for each individual patient – a platform for systems medicine
* Creates an enabling infrastructure to foster collaborations
* Integrates numerous LCCC "omics" platforms

Research Programs and Core

Building Blocks



GDOC Informatics