Metadata Management in Biomedical Research: A Comprehensive Paradigm

ABSTRACT

Biomedical data, of variable quality, is collected and held in many different databases. These databases have been developed for different audience using different terminologies, formats and content. The variability in databases creates a dilemma for potential users of data because data needs to be combined together from multiple sources and analyzed for it to deliver additional value. To address these concerns a multi layer metadata approach is being implemented. The metadata layers help in standardizing the process of data collection, documentation, discovery and use of data. It improves the consistency, and integration of data which in turn increases data transparency and decreases ambiguity.

Metadata is simply data about data. There are distinct classes of metadata that need to be captured within a biomedical environment to allow the user to truly understand the data. This presentation discusses the creation of a comprehensive metadata layer consisting of both contextual and physical metadata models. Develop of metadata standards is critical for data quality. The presentation also discusses the development of the metadata standards for each data element and recording these parameters within a data dictionary that help guide the user as to the limitation on using the data.

BIOGRAPHY

Neera Bhansali
Department of Biomedical Informatics
Moffitt Cancer Center & Research Institute

Neera Bhansali received her doctoral and masters degrees in business from RMIT University, Australia, and BA from Calcutta University, India. Dr Bhansali is an expert in the areas of strategic planning, strategic alignment, data warehousing, data governance and data quality. Over the past twenty years Dr. Bhansali has facilitated transformations and provided strategic direction to organizations in manufacturing, airlines, consulting, media, finance and healthcare sector around the globe. She is the author of the book “Strategic Data Warehousing – Achieving Alignment with Business”. Dr. Bhansali is currently with the Department of Biomedical Informatics, H. Lee Moffitt Cancer Center and Research Institute, Tampa.

Metadata Management in Biomedical Research – a comprehensive paradigm

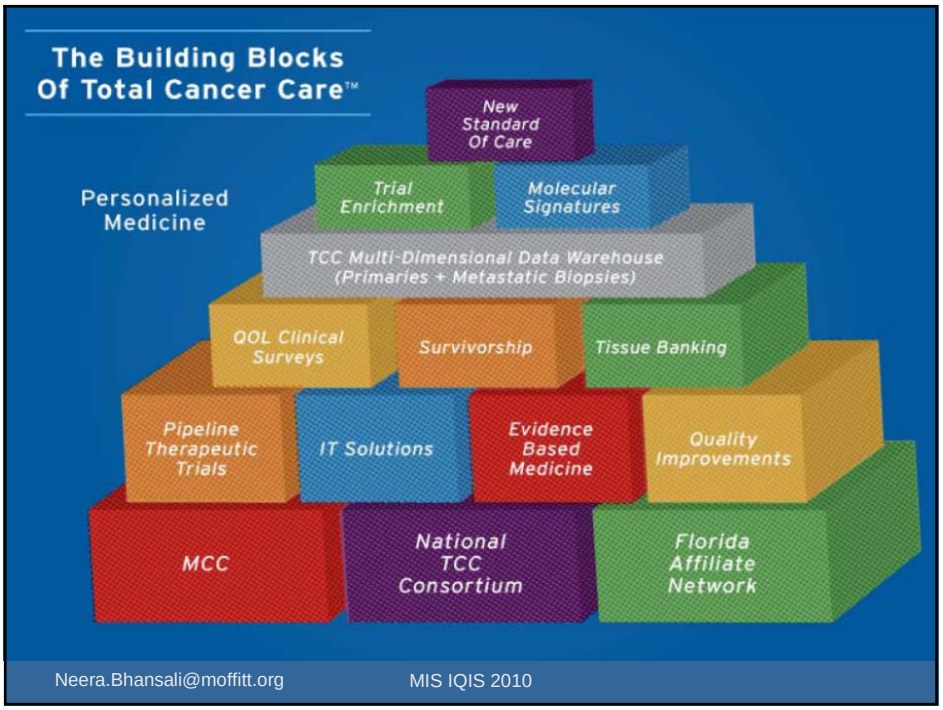
Neera Bhansali, PhD
Dept of Biomedical Informatics
Moffitt Cancer Center & Research Institute
12902 Magnolia Drive
Tampa 33612

Neera.Bhansali@moffitt.org MIS IQIS 2010

Neera.Bhansali@moffitt.org MIS IQIS 2010


A Cancer Patient's Life Journey
Behavioral Oncology
Cancer Prevention
Lifestyle
Education
Diagnosis
Genomics/Proteomics
Imaging Modalities
Nanotechnology
Populations At Risk
Risk Factors
Genetics
Early Detection
Health Disparities
Prognosis
Molecular Oncology
Genomics/Proteomics
Total Cancer Care
Survivorship
Behavioral Research
Psychosocial & Palliative Care
Family Needs
Health Outcomes
Relapsed Disease
Recurrence Therapy
Drug Discovery
Clinical Trials
Treatment
Primary Therapy
• Multimodality
• Target Based
Post Therapy
• Surveillance
Neera.Bhansali@moffitt.org
MIS IQIS 2010

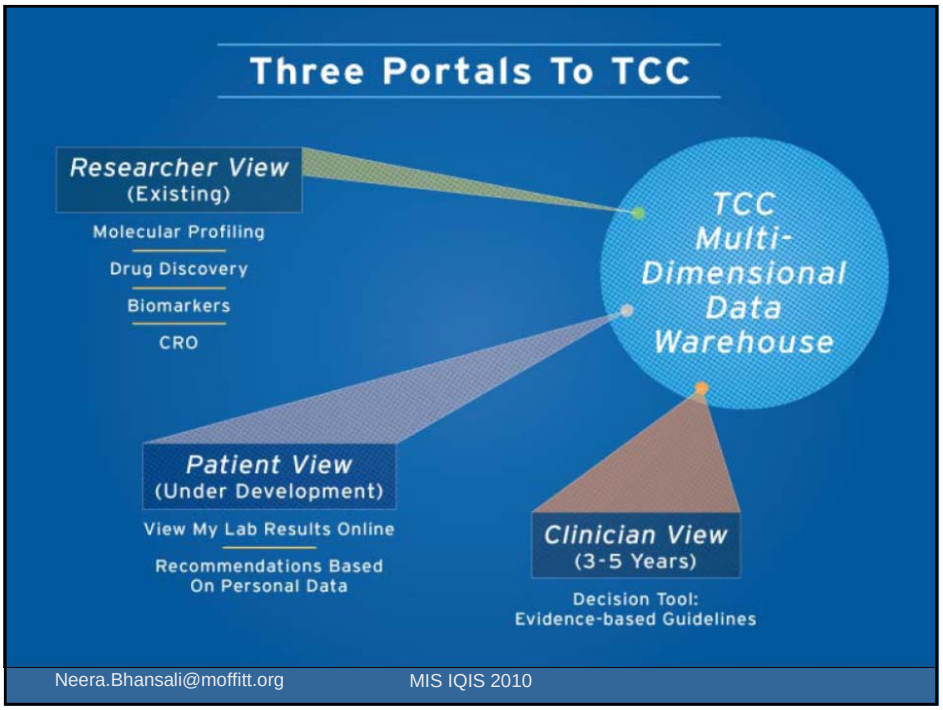

Total Cancer Care Goals
• Identify the needs of the individual patient
• Identify markers to predict needs and risks
• Develop methods of early detection
• Match the right treatment for the right patient
• Improve the performance of clinical trials through molecular profiling and matching
• Create evidence-based guidelines
• Raise the standard of care
Neera.Bhansali@moffitt.org
MIS IQIS 2010

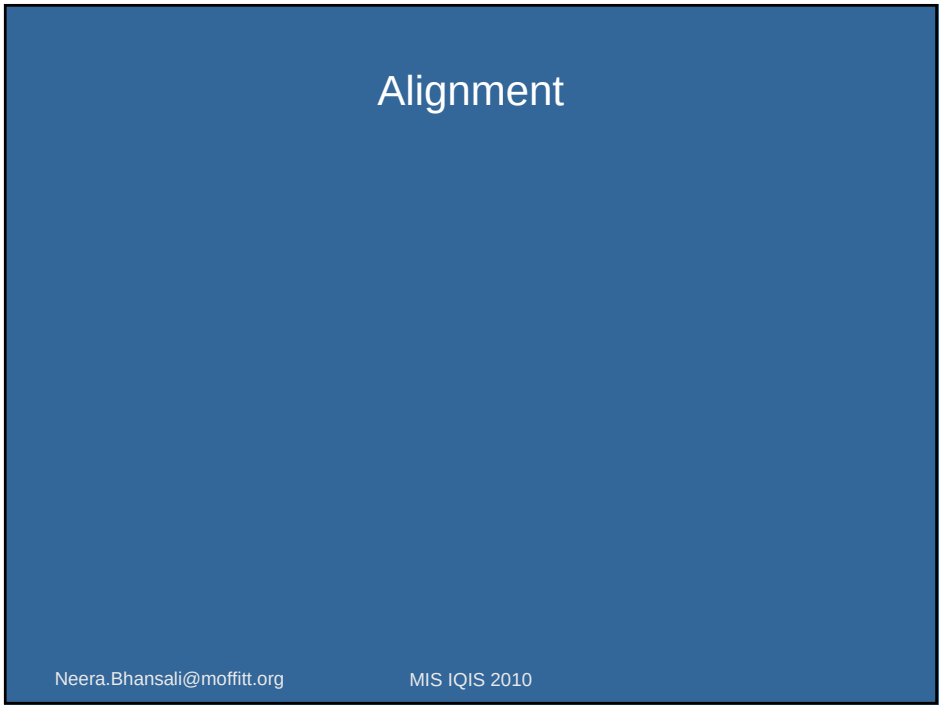


Approach: Total Cancer Care Personalized Medicine Project (PCC)

- A large prospective translational research project with patient consent
- Collect, relate, and interpret clinical data and molecular data from thousands of patients
  - Tumor, blood, urine samples
  - Clinical data (risk factors, therapies, outcomes)
- Identify molecular signatures for prognosis
- Personalize therapy and follow-up
  - Right care, right patient
- Information->evidence->knowledge->wisdom

Neera.Bhansali@moffitt.org MIS IQIS 2010


Three Portals To TCC
Researcher View
(Existing)
Molecular Profiling
Drug Discovery
Biomarkers
CRO
TCC
Multi-
Dimensional
Data
Warehouse
Patient View
(Under Development)
View My Lab Results Online
Recommendations Based
On Personal Data
Clinician View
(3-5 Years)
Decision Tool:
Evidence-based Guidelines
Neera.Bhansali@moffitt.org
MIS IQIS 2010


Alignment
Neera.Bhansali@moffitt.org
MIS IQIS 2010

Architecture for Moffitt Data Warehouse

Source Systems

Multiple Dimensions of Oncology Data

Multiple Dimensions of Oncology Data in Data Dictionary

Metadata

Data Dictionary

Standardized Vocabularies

Standardized Data Elements

Contextual Data Dictionary

Physical Layer Data Dictionary

Establish RACI for Data Elements

Challenge

- Deploying standard rules and common language to easily share information
- Building processes and tools to collect, analyze, integrate and disseminate information in a consistent manner