Advisory Committee to the Director
Data and Informatics Working Group

Update to the

Advisory Council to the Director

December 8, 2011

Drs. David DeMets and

Lawrence Tabak, Co-Chairs

Membership

Name	Institution	Title
David DeMets, Co-Chair	University Wisconsin - Madison	Professor, Dept. Biostatistics & Medical Informatics
Lawrence Tabak, Co-Chair	NIH	Principal Deputy Director
Russ Altman	Stanford University	Chair, Department of Bioengineering
David Botstein	Princeton University	Director, Lewis-Sigler Institute
Andrea Califano	Columbia University	Chief of Biomedical Informatics
David Ginsburg, ACD Member; Chair, NCBI Needs-Assessment Panel	University of Michigan	Professor, Internal Medicine; HHMI
Patricia Hurn	University of Texas	Associate Vice Chancellor Health Science Research The University of Texas System
Dan Masys	University of Washington	Affiliate Professor, Department of Medical Education and Biomedical Informatics
Jill Mesirov, Ad Hoc Member, NCBI Needs-Assessment Panel	Broad Institute	Associate Director and Chief Information Officer
Shawn Murphy	Harvard University	Associate Director of Laboratory of Computer Science and Assistant Professor , Neurology
Lucila Ohno-Machado	University of San Diego	Chief Division of Biomedical Informatics

Initial Charge to the Working Group

- The Advisory Committee to the Director (ACD) Data and Informatics Working Group (DIWG) will provide the ACD and the NIH Director with expert advice on the management, integration, and analysis of large biomedical datasets. The DIWG will address the following areas:

  - Research data spanning basic science through clinical and population research

  - Administrative data related to grant applications, reviews, and management

  - Management of IT at the NIH

Working Group Progress

✓ August – Kick-off meeting

✓ October - Teleconference

✓ November - Face-to-face meeting at NIH

✓ December - **Interim Report to ACD**

❑ March/April - Face-to-face meeting at NIH

❑ June - Final Report with actionable recommendations to ACD

❑ Ongoing – Monthly teleconference

Research Data Spanning Basic Science through Clinical and Population Research

- Committee Concerns About Initial Charge: time, expertise, and sheer magnitude of the issues

- Refocused Charge

  - The connection and integration of large volumes of "omics" data with other large data sets including clinical and phenotypic data

  - The management and curation of these large data sets, including the use of new and emerging technologies (e.g. clouds)

  - The analysis of these integrated data sets to facilitate the development of more sophisticated predictive models of disease susceptibility and pathobiology

Research Data Spanning Basic Science through Clinical and Population Research (cont.)

- Approach:
  - Subgroup to focus on the three most pressing data types:
    - Imaging Data
    - Molecular Profiling Data
    - Phenotypic Data
  - Supplemented with additional expertise
  - Eye towards common themes

Format for Analysis and Recommendations

		Recommendations	
		Policy	Implementation
Data Type	**Imaging**		
	Molecular Profiling		
	Phenotypic		

Possible Questions to be Answered

- Define the problem space (in terms of scope and the research information lifecycle)

- What issues/challenges is the extramural community faced with when working with this type of data and/or at this point in its life cycle?

- Are the issues/challenges solvable with today's political climate and technology?

- What research benefits are not being realized, because of each of the issues/challenges?
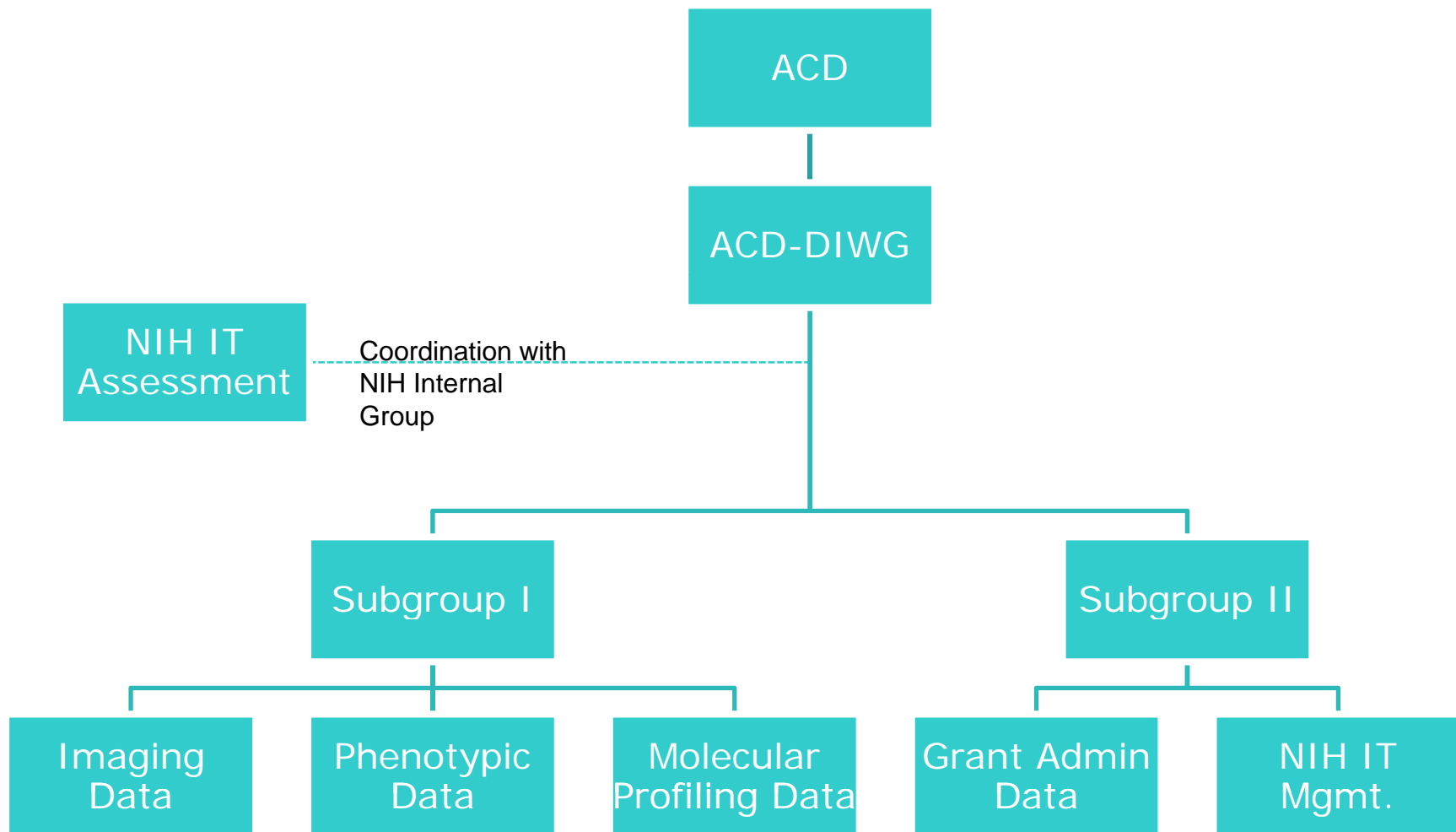
Possible Questions to be Answered (cont.)

- What concrete recommendations can the committee make to NIH for NIH action? (policy, investment in standards, investment in science, investment in infrastructure, new collaborations, updates to current policies, standards, investments, and collaborations, or other)

- What factors will stand in the way of these recommendations? What kind of resistance will NIH encounter?

- What further analysis and expertise does the committee require to complete its work?

- Others?

Grant Administrative Data and NIH IT Management (NIH "On-Campus" Issues)

- Proposed Approach: form second (sub)group supplemented with expertise

- Charge

  - Administrative data related to grant applications, reviews, and management

  - Management of information technology (IT) at the NIH

- Committee Concerns: time and expertise

Committee Structure

ACD

ACD-DIWG

NIH IT Assessment

Coordination with NIH Internal Group

Subgroup I

Subgroup II

Imaging Data

Phenotypic Data

Molecular Profiling Data

Grant Admin Data

NIH IT Mgmt.

Proposed Next Steps

- Subgroups develop policy recommendations

  - Subgroup 1: initial call to establish commonalities, divide into data type teams

  - Subgroup 2: initial call to present landscape, decide how to proceed

  - Hold workshops on proposed policy recommendations, one per policy area

- Publish a Request for Information on data challenges to aid policy recommendation development

Questions?