NIH BECON/BISTIC Symposium (BB2004): Symposium Report:
"Biomedical Informatics for Clinical Decision Support: A Vision for the 21st Century", June 21st and 22nd (2004).

NIH Extramural Co-Chairs: Larry Clarke PhD (NIH/NCI) and Peter Lyster PhD (NIH/NIGMS)

External Co-Chairs: Ruzena Bajcsy PhD (UC-Berkeley) and Michael Vannier MD (University of Chicago).

Organizational Co-Chairs: Mariaileen Sourwine MS (NIH/NIBIB), Anna Retzke BS (NIH/NIBIB), and Theresa Smith BS (NIH/NIBIB).

Final Report: 3/2/05.

1.0 Symposium Goals.

The symposium on "Biomedical Informatics for Clinical Decision Support: Vision for the 21st Century" was conducted on June 21-22, 2004 at the Natcher Conference Center on the NIH Main Campus in Bethesda, Maryland. This document represents the summary and recommendations from panels of extramural scientists. The Symposium was jointly conducted for the first time by the trans-NIH's Bioengineering Consortium (BECON) and the Biomedical Information Science and Technology Consortium (BISTIC); referred to as BB2004. The purpose of the symposium was to identify opportunities, needs, and directions for applying computer science and informatics principles and methods to clinical decision support. Specific areas covered during the meeting included heterogeneous data collection methods, data management (databases and digital libraries), enabling technologies (modeling, software tools, and techniques), and translational informatics as required to support clinical data interpretation and quantitative biomarkers for clinical trials. Approximately 400 people attended the symposium that included academic, government, and industry researchers involved in clinical investigations, imaging and other biosensors including molecular methods, bio molecular sciences, computer science, application specific software tools, and researchers in translational informatics. The primary emphasis focused on the development and assessment of informatics software that is application specific as opposed to the development of medical informatics infrastructure. In section 1.1 an executive summary and meeting recommendations are presented—this is the main report from the meeting. The background reflecting the motivation for the organization of this symposium is described in section 2, the text of the meeting notes from three breakout sessions is provided in section 3. Appendix A is the summary from the satellite meeting on Using Standards to Integrate Biomedical Imaging into Clinical Decision Making. Appendix B is a summary of the satellite meeting on the potential role for public private partnerships to support informatics resources for the scientific and industry community. Finally the meeting website, with updated information and downloadable presentations is: http://www.becon.nih.gov/symposium2004.htm

1.1 Executive Summary and Meeting Recommendations: The symposium recognized that there are ongoing activities at NIH that are engaged in the acquisition and maintenance of clinical and translational research data.   This presents an opportunity to promote the development, optimization, validation, and dissemination of software tools for heterogeneous data integration and clinical decision support, as required for example, for personalized medicine. The following presents the main meeting recommendations that are closely interrelated and should be collectively considered by different NIH ICs, by the NIH roadmap, and other federal agencies.

1. Clinical Data Collection Strategies: Support targeted data collections from NIH-translational research and clinical trials, as a web accessible public resource, specifically for

the development, optimization, and validation of informatics tools for clinical decision support. The developers of the data collections and informatics tools need to be involved from the earliest stage in the design of clinical trials that link multiple data modalities. The strategies for data collection should include provision for access to the data during the course of clinical trials without compromising the goals and integrity of the trials and tools to permit the objective benchmarking of the clinical performance of informatics tools. The development of the targeted data collections and validation methods could be supported, for example, as trans-NIH initiatives through Roadmap activities such as the National Electronics Clinical Trials and Research (NECTAR) network, other NIH informatics initiatives, and through public private partnerships.

2. Harmonization of Data Acquisition Across Biosensors: Support the development of methods to harmonize clinical data collection across multiple bio sensor platforms, such as gene expression, proteomics, and imaging. The goal is to reduce the uncertainty in heterogeneous data collection due to variations in platform and clinical protocol. This may involve data normalization strategies to account for the different performance characteristics of the biosensors employed. In general this effort will require interoperability of software and effective exchange of data. This may require targeted federal support and collaboration with the medical device and pharmaceutical industries.

3. Development and Evaluation of Translational Informatics Tools: In conjunction with the development of targeted data collections and data collection strategies, increase the support for the development of advanced informatics tools and modeling methods for application-specific clinical decision support. For example, this includes methods for knowledge extraction and data integration from imaging and other bio sensors. The development and optimization of software tools need to accommodate the changing scale of the spatial and temporal resolution of the data collections, their high dimensionality, and the unprecedented data complexity such as that anticipated to realize the ultimate goal of personalized medicine. Access to the database resources is critically required so different communities may engage in this research area without the need to be concerned about the methods for data collection--a well recognized barrier. This will require inter-agency support to engage and motivate the basic domain and computer science communities within academia, national laboratories, and industry to develop the required informatics and modeling tools.

4. Software Engineering Approaches: Require the use of modern software engineering approaches for coding, benchmarking, and version control of software development supported by investigator initiated grants. Support is required, for example, for open source software development, grid computing, web services for sharing software and its deployment, and other software engineering methods to promote software interoperability and extensive reuse.

5. Provide Mechanisms for Dissemination and Regulatory Approval of Software Tools: Reduce the barriers to the commercial development and dissemination of advanced software tools. For example provide mechanisms for dissemination of research tools using SBIR/STTR mechanisms or industry-academic partnerships, and engage NIST and FDA to encourage a means to either harmonize or standardize methods for benchmarking software performance that should accelerate regulatory approval and reimbursement by CMS.

6. Public Private Partnerships: Encourage the development of public private partnerships to leverage NIH and private sector support for informatics resources that will accelerate the delivery of informatics tools for clinical decision support and clinical trails. The resources of the foundation of NIH, for example, can be utilized for this purpose.

7. Demonstration Projects: Support demonstration projects that address the above recommendations for targeted clinical applications. This is intended to increase confidence in

the broader scientific and industry community as to the efficacy of the informatics tools, thus increasing the likelihood of extended research and translation to clinical use.

2.0 Background: Symposium Abstract used to Advertise the Meeting.

Creating, distributing, and analyzing heterogeneous biomedical information poses significant challenges for future medical decision-making. With the rapid development of technologies for high-throughput data production, high-speed communication networks, and the potential for enormous distributed databases of medical and biological reference material, it is becoming increasingly clear that sophisticated new software will be required to harness the potential of these vast new sources of information. Images, genomic data, gene and protein expression, and patient medical records are often managed and analyzed independently of each other. Although this explosive growth in information enriches biology and medicine, it also raises concerns about current and future practices for software optimization and validation, heterogeneous knowledge integration and the development, interpretation, and regulation of new information technologies, and privacy infringement and the potential misuse of information.

This symposium focused primarily on the software tools and approaches that will ultimately deliver the benefits of biomedical information technologies to patients at the time and place where decisions are made regarding risk, diagnosis, treatment, and follow-up. Specifically, this meeting sought to provide a scientific vision of the future where healthcare information technologies may be more fully deployed in the clinical workflow to improve efficiency and outcomes. As healthcare accommodates the individual variation in the population, mass customization using lifelong information records will be needed.

To address this complex array of issues, the NIH Bioengineering Consortium (BECON) and the NIH Biomedical Information Science and Technology Initiative Consortium (BISTIC) jointly organized the symposium, which brought together software researchers and clinical users who are developing programs to support translational research and clinical decision-making. This symposium was the seventh in a series of annual conferences coordinated by the BECON and the first joint symposium with the BISTI Consortium.

The symposium format included four plenary tracks followed by three cross-cutting breakout sessions. The plenary speakers described recent progress, but substantial shortcomings, of current systems. They introduced a vision to address clinical needs that emphasized the potential benefits of biomedical informatics for healthcare providers, patients, researchers, and society.

The breakout sessions identified the expertise and infrastructure needed to satisfy the technological demands for translational research and clinical care specifying what must be done to accomplish these goals, as well as outlining how the work will build on existing technologies to develop new ones to promote success.

In addition to encouraging increased communication among scientists who develop software, the device, drug and informatics industries, researchers, and clinicians, the symposium aimed to identify major challenges and opportunities that should be addressed by NIH policies and funding programs, including partnerships with the private sector.

3.0 Meeting Notes from Breakout Sessions: The following discussion points developed by the symposium and breakout sessions committees prior to the meeting are outlined below together with a list of recommendations. There is some overlap among the three panels, and we have attempted to remove duplication in the Executive Summary and Recommendations (above):

3.1 Breakout Group 1: Clinical Challenges and Related Software/Informatics Requirements.

Initial Charge to Panel: Maturing data-networks and innovative instrumentation can improve clinical decision-making through quantitative integration and informatics. Such synergism will be

possible only if data structures and standards permit the integration of conventional clinical data and the information obtained from imaging, other bio sensors, and the clinical application of basic research advances in genomics and gene expression technologies. First steps require an analysis of technical problems in hardware and software, quality assurance matters, privacy and other regulatory concerns, and other uncertainties that impact data integration. How to harmonize data collection so that software tools can have general utility remains a major problem.

A critical review of the weaknesses in existing clinical trial infrastructure could provide essential information about current difficulties, and serve to point the way to developing a pathway for advancing quantitative data analysis, data-integration and related informatics methods. A clinical trial demonstration project might provide exemplary data for testing software optimization procedures and clinical validation. Ultimately, clinical acceptance of such decision-making tools will require demonstration of value, cost/benefit trade-offs, ease of use, and cultural acceptability.

3.1.1 *Discussion point:* *Identify and prioritize current informatics challenges posed for quantitative data acquisition and extraction, analysis and data integration using a variety of sensors (imaging and other bio sensors, including molecular profiling/genomic methods), and the methods needed to support the ultimate goal of assisting with clinical decision making.*

Recommendations:

- Develop standards and shareable vocabularies for clinical, imaging, and genomic data – building on such initiatives as NLM's SNOMED licensing
- Establish and fund a "lead institute" at NIH to foster those standards and formalize a process for data sharing
- Foster a handful of demonstrations and projects that effectively combine computer sharable data from clinical imaging and genomic data sources
- Collaborate with, and encourage, current clinical electronic health record (EHR) initiatives through interagency policy initiatives (e.g., reimbursement, research, routine health reporting requirements)
- Efforts to proceed with EHR must have a plan to include office and outpatient practices, harmonizing those data with hospital-based EHR while consistent with HIPAA compliance.
- Continue to encourage computer systems interoperability and favor open-source computer software solutions with policy implementation coordinated by multiple NIH ICs

3.1.2 *Discussion point*: *Review the impact of clinical data-acquisition protocols and methods for the different front-end platform technologies on the implementation of the informatics software tools and identify new methods or technologies required to address barriers to progress.*

Recommendations:

- Continued encouragement of standards development for existing and new technology methodologies (X-ray, CT, NUC/PET, US, MRI, MRS) such as DICOM acquisition standards to further the goal of evolving anatomic, functional and molecular data as integratible information
- An integrated approach should be taken to device procedures, technology platforms, quantification, visualization, and interpretation
- The challenge is not just technologic but also educational, with currently inadequate training for data collection
- Need standards for clinical data, clinical biomarkers and imaging that is funded and coordinated by the several governmental agencies. This effort should include CMS and FDA. Device manufacturers can be encouraged to participate in this process by expectation of validation requirements as new technology emerges.

4

3.1.3 *Discussion point:* *Identify sources of uncertainty in clinical data collection that may influence the ability to perform reliable quantitative measurements and data integration, and thereby propose methods or solutions that may require research investments for addressing such problems. Sources of uncertainty may be attributable to either human or technologic factors, in particular Human factors arise from: (i) data which is inaccurate at its inception; incomplete data, or invented data. As clinical knowledge evolves on an experiential basis there is often a lack of strong evidence base in many clinical areas which may permit untoward influence of judgment, opinion, and experience; (ii) Technology factors may arise from differing age or versions of technology in use which fail to keep pace with clinical change or merely a lack of willingness to purchase and utilize appropriately accurate technology. Additionally, programming errors, and lack of interoperability may extend the boundaries of uncertainty.*

Recommendations:

- Bring current electronic health record (EHR) users into the research fold to utilize available data and show its utility and increase published research showing differences in outcomes and evidence of efficiencies in time and cost. Move from system information in silos to integration
- Work with government and insurance payers to utilize data already gathered in authorization and payment processes, and detailing and "report card" data already in use by utilizing quality/performance improvement data already available via partnerships with JCAHO and NCQA
- Provide strong leadership and decision making in support of the President's efforts to improve use of effective technology in healthcare.
- Operationalize artificial intelligence focused on systems and their effectiveness in the context of actual healthcare delivery.

3.1.4 *Discussion point:* *Identify existing or planned clinical trials that may be targeted or enhanced as short-term demonstration projects to test the performance of emerging informatics tools in support of clinical decision-making (See NIH Road map: Re Engineering the Clinical Research Enterprise/NECTAR).*

Recommendations:

- NIH clinical trials, sponsored by NCI, NHLBI and other ICs as well as EHR implementations occurring at NIH Clinical Center and in leading healthcare facilities, offer available test beds for demonstration projects and development of 'best practice' models of healthcare management. Investigator commitment and recognition of the experimental opportunities in these informationally enriched environments could accelerate technological refinement. NECTAR cancer networks and caBIG efforts could be leveraged to extend informationally advanced routine care to disease research applications. The structural uniformity required in clinical trials, though different from the more permissive free text found in medical records may offer an edge for promoting data element and vocabulary standards
- Continued emphasis on promotion and evolution of standard data definitions (e.g. HL7, etc) for sponsored clinical trials could hasten adoption and propagation of structural reform. Incentives for data sharing will enhance the need for common data elements.

3.1.5: ***Discussion point:*** *Identify and prioritize future informatics challenges posed when undertaking patient-specific molecular screening, diagnosis and treatment.*

Recommendations:

- Personalized medicine intended to reliably characterize disease risk and choose the particular therapeutic regimens that would be effective for that unique individual, requires gathering disparate information including components of an individual's genetic profile. Some

accommodation must be made culturally and legally to protect patient privacy while striking a balance that permits medical science progress intended to achieve improved human health.

3.1.6 *Discussion point:* *Explore the impact of humanization of research (including drug discovery), expanding clinical information (as outlined in the 2003 AAMC report), and phenotyping of humans (Human Genome Project), thereby leading to proposals for long-term research emphases.*

Recommendation:

- The problem and objectives must be framed so as to be addressable not by one IC but as an overall NIH effort and needs to be open and linked to other efforts like WHO to have the widest impact. Movement toward these goals requires an NIH-wide initiative to identify the necessary financial funding and harmonize the grants processes such that individual investigator and hypothesis-driven research share more cross discipline objectives. Solutions must acknowledge the potential influence of CMS, FDA and industry in the conduct of clinical trials. At present there is no infrastructure in any one place nor demonstration projects that involve effective data standards and data-sharing between the various participants in the clinical trials process.
- Considerably greater emphasis has to be placed on standard methods of recording and exchanging data like HL7 and SNOMED in order to accelerate the timeline.

3.1.7 *Discussion point:* *Develop recommendations for developing a broad consensus for prioritization of identified clinical and research challenges so that informatics tools requirements can be recognized and developed in a timely way [See NIH roadmap: Research teams of the Future].*

Recommendations:

- Electronic health. All communities need to work together to standardize with the same strategy. Provision must be made for incentives for collaborations and disincentives for isolated solutions. This implies empowering a national informatics infrastructure for prioritization.
- An opening wedge to this progress can occur with the increasing dissemination of EHR systems but individual vendor solutions should be capable of data sharing and valid data transfers. Governmental (e.g. NIH and DOD, VA, etc) and health-plan impending deployment of these EHR systems would be beneficial only if they create and use data transfer capable instrumentation.
- Medical schools, as part of their educational mission, have not yet focused on formal clinical e-medical training and informatics and such commitment is needed to change culture.
- Building infrastructure must consider extensibility since it is likely we currently have only 1% of the data for genomics that we will have in the next five years. A 'change management plan' must be inherent in publicly available data banks
- The lessons learned from the Human Genome project is that we need a grand vision that is well articulated and compelling with a major leadership champion. Recent Presidential announced emphasis on this subject should be followed up and sustained and the public convinced of its value.

3.1.8 *Discussion point:* *Develop recommendations for more rapid physician acceptance of informatics software tools for data interpretation and clinical decision-making.*

Recommendations:

- Develop tools in conjunction with early electronic health record (EHR) adopters to move the decision support to the physician-patient interface. These tools need to be managed by researchers to accomplish updates as the data from research is analyzed

- Continue to increase use and training in technology in medical school, residency, and post residency training through partnerships with educators and professional specialty organizations across all levels of healthcare providers
- Survey offices/facilities which are not adopting technology to identify reasons and what factors would improve rate of adoption. Human factors monitoring, some of which may be accomplished through remote network processes may be employed to identify bottlenecks, both user-originated and attributable to quality of service of software/network.
- Provide software research to assist physicians and facilities in identifying what is available and what the utility and efficiency factors are for different products
- Increase literature and seminar topics related to use of software with focus on patient outcomes
- Consider partnerships with insurance and/or drug detailers to provide computer basics to isolated practices.
- Increase data feedback of interest to physicians – personal, regional, specialty, etc, whatever data is available for analysis and incorporate information frameworks that address the totality of the health-providing environment, incorporating new and currently un-incorporated data components such as patient preference and consent. These items may create a demand pull if they also incorporate elements which satisfy the patient's ability and demand for information
- Encourage development of recursive adaptive information systems that evolve and accommodate user-specific knowledge structures
- Explore and make recommendations for how test-beds, model systems and the clinical trials infrastructure (e.g. the NIH roadmap, NECTAR, caBIG) can accommodate and incorporate the more recent advances in informatics software tools, using them to develop 'best practices' benchmarks

3.2 Breakout Group 2: Databanks for assessment of application-specific software and data-integration and other informatics tools.

Initial Charge to Panel: The ability to produce effective software tools for data analysis and clinical decision support is limited by the availability of the appropriate test datasets in reliable public databases. Effective datasets will require databases designed with shared data models and transparent interoperable data management. The goal of this session is to identify and assess the current status of available databases, to examine their underlying data models, and to define short and long term recommendations for the development of well-designed and interoperable data or database management systems

3.2.1 *Discussion Point:* *Identify the tools needed to facilitate the establishment of databases including tools for maintenance, for validation of data, and for effective and efficient usage (for both depositing data and accessing data). Develop recommendations for how the methods for image database collection can be better coordinated and integrated with clinical trials across NIH IC's, e.g., NIH Roadmap Re-Engineering the Clinical Research Enterprise/NECTAR, NIH Roadmap New Pathways to Discovery, NIH Roadmap National Centers for Biomedical Computing (NCBC), and efforts from the NCI Center for Bioinformatics.*

Specific recommendations for the national database resources include:

- Trans-NIH processes need to be developed to permit the collection of targeted data sets from on going clinical trails, and potentially linked to the current NIH NECTAR roadmap activity.
- The targeted data sets would include images, electrical signals, genomic and all other relevant patient data, together with truth information necessary to benchmark the performance of software tools.
- Security and quality control needs to be addressed as this database may serve as a standard for benchmarking software performance.
- The validation of input data and consensus process for the determination of truth files is critically important for the acceptance of performance standards.

- NIH support to maintain and update this resource is critically required.
- Leveraging of NIH support through public private partnerships is encouraged.

3.2.2 Discussion Point: *Identify database design requirements including the need for open-source tools to facilitate extraction of data from the database. Databases are not just for data storage – information retrieval vs. information extraction. Identify means for integration across databases.*

Recommendations:

- The development of repositories of representative imaging and clinical data is critical for evaluation and testing of new algorithms, as well as providing capabilities for integrating computerized analysis of existing and emerging imaging modalities and other data sources into clinical trials. There are two stages in the development:

- First Stage: Specific informatics tools required for data collection.
  - ➢ Support development of the informatics tools and infrastructure
  - ➢ Develop standards for user-friendly deposition of data by contributor (of primary data, metadata, and truth) and downloading data by others
  - ➢ Implement rapidly using open source software development practices and grid technologies
  - ➢ Should be applicable to multiple types of disease, images and other related data.
  - ➢ Define inclusion/exclusion criteria for the data and redefine as necessary
  - ➢ Sources of data can be from clinical trials or pre-existing individual lab databases.
  - ➢ Require "certified" collection processes at each contributing institution to help ensure integrity of data, metadata, and truth.
  - ➢ Should be started in parallel with the Second Stage.

- Second Stage: Actual design and establishment of databases -- Characteristics
  - ➢ Each database needs an expert oversight committee and core support to check truth and to check unexpected results from users
  - ➢ Database should be a repository, analogous to a tumor bank (i.e. continuously updated)
  - ➢ Should be based on disease rather than imaging modality (include multi-modality and other data sources from different bio sensors)
  - ➢ Should contain data and metadata including well-defined, multiple truths for the tasks that are related to the given disease, e.g., multimodality images of the breast for detection, diagnosis, treatment planning, response to therapy, etc.
  - ➢ Determine database-specific minimal information standards for data submission similar to the MIAME standards for gene expression data.
  - ➢ Include normal cases to serve as controls (e.g., these are the cases that may cause false positives)
  - ➢ Include acquisition data (e.g., physical quality indices such as for image data)
  - ➢ Include longitudinal data where appropriate (interval change)
  - ➢ Include raw data as well as processed and/or reconstructed data for different biosensors
  - ➢ Contributors and users need to commit to sharing
  - ➢ Make incrementally available-- change database with changing technology
  - ➢ Include ability to put derived data and results back into the database in an open source fashion
  - ➢ Need new systems architecture workflow management and software tools to integrate and leverage existing clinical imaging infrastructure of hospitals and imaging centers.
  - ➢ Make deposit into a repository as a requirement for grant and/or publication

> ➢ Use database for other purposes, such as retrieval of similar cases: such retrieval will help in patient management and education

- Other database issues to be addressed
  - ➢ Lifespan of each database
  - ➢ Decision on database priority
  - ➢ Links to other databases
  - ➢ Access to the databases
  - ➢ Standardize protocols
  - ➢ Data ownership
  - ➢ Continued funding mechanisms
  - ➢ Open source toolkit for evaluation of application-specific software tools and could be included in NLM ITK. Informatics tools may be task specific e.g., detection vs. diagnosis. Image guided treatment of disease or assessment of drug therapy.
  - ➢ Clinical validation: Start with ROC analysis since already is a tested, and is available as freeware
  - ➢ Educate users on proper use of the evaluation tools and also on how to do evaluation studies for both application-specific and data-integration software (i.e., type of observers, training of observers).

3.2.3 *Discussion Point:* *Explore the need a culture change with grass roots motivators/initiators required to encourage common methods for reporting results of NIH funded grants. Explore relationships between academic investigators and industry in terms of sharing databases and role of NIH in support of these databases.*

Recommendations:

- Investigators and industry need to share databases
- NIH needs to support non-hypothesis driven database development
- The data-producing community needs to embrace new publication standards similar to the genomics community policy on depositing sequence data into public repositories
  - ➢ Have associate editors suggest in letter to editor of relevant journals
  - ➢ Have some investigators "set the example" for data deposition to public databases
  - ➢ Requirements in publications: work with journals and reviewers to require that descriptions of the database and evaluation methods used (including scoring) are given. Devise a checklist.
  - ➢ Allow "database publications" that do not include hypothesis-driven research.
  - ➢ In the future, upload images and code that are linked to the publication.
- Institutions need to recognize that database development publications should be used in promotion
- Recognize that the ultimate national biomedical databank would feature the deposition and storage of all patient biomedical data

3.2.4 *Short term Recommendations.*

(i) Create a comprehensive inventory of existing databases with corresponding infrastructure (a database of databases). Summarize their success stories and challenges. This will allow cross fertilization among the various investigators who come from different scientific areas
- Include information on use, contributors, allowed users, lifetime, limitations, and other descriptors
- Include broad categories of imaging data, physiological signal data, genomics, etc.
- Include even restricted/limited access databases (e.g., Mayo Clinic/IBM database)
- Create as a "Pubmed" of databases with appropriate interfaces
- Similar to the clinical trial research inventory performed for NECTAR (which is via contract)

- Example (non complete) list of existing databases and sources of support discussed at the meeting:
  - ➢ Physionet   NCRR
  - ➢ Biomedical Informatics Research Network (BIRN--Brain MRI and Networked Infrastructure)   NCRR
  - ➢ Lung Image Database Consortium (LIDC)   NCI
  - ➢ Medical Informatics Europe   EU
  - ➢ caBIG (cross database integration)   NCI
  - ➢ Genomic databases   NIH
  - ➢ FDA databases   FDA CDRH
  - ➢ Microarray (GEO/ArrayExpress)   NCBI/EBI

(ii) Aim to incorporate and create shared grid-based national databanks for depositing existing and new data.
- Centers for databank development -- similar to the feasibility studies for NECTAR
  - o Should be an NIH Roadmap activity funded by NIH and private contributions distributed by the Foundation for the NIH (FNIH)
- Fund limited number of feasibility/demonstration projects for a limited number of specific-type databases
  - o Example of a specific database is the LIDC database -- focused on a specific task (e.g., ECG, CAD)
  - o Relate to tasks in Breakout Session 1
- Fund limited number of feasibility/ demonstration project for a nonspecific-type database
  - Example of a non-specific database is PubMed
  - Might be all patient cases with annotation from records from two hospitals for two years
  - Might be a general imaging database
  - Test with various "appropriate clinical questions"
- Each would incorporate the development of attributes from all 4 issues described in the list of attributes to be developed for databanks in genera, as described below.
- Requires multidisciplinary team due to broad range of attributes (ultimate users, developer, software, hardware)

3.2.5 *Long term Recommendations.*

(i) Aim to incorporate and create a shared national resource of federated databanks for depositing existing and new data.
- Develop methods for incorporation of "private" databases, on which a publication was based, into the national databank or set of federated databanks
- Create a continuing process by which more data enters the databanks
  - ➢ Specific (e.g., LIDC) versus nonspecific (e.g., PubMed) databases
- Requires multidisciplinary teams especially in determining "truth" characteristics
- Need to incorporate the developed attributes from the short term feasibility projects for this national databank
- Funding of new data/databases from grants
- Funding of main infrastructure via contract with NIH and FNIH.

(ii) Contributor agreements and rewards for contributing. Recommendations:
- Credit databases in research publications
- Credit system for those who contribute data
- Debit system for those who don't contribute but want to use (grant fees, university research)
- Line item in RO1 budgets to help in the continued maintenance of the national resource

(iii) The following is a list of attributes to be considered and developed for databanks in general.
- Databank elements/entry descriptors

- Appropriate standard semantics/annotation/CDEs/ontologies from controlled vocabularies
  1. Biomedical objects
  2. Common data elements (CDEs)
  3. Controlled vocabularies
- Databanks contain data, metadata, and sometimes outcome truth
- Metadata (examples) & clinical reason
  1. Clinical info, structured reports
  2. Associated image data, genomic data
  3. Diagnostic or therapeutic outcome data
- Treat the semantics/annotation of "truth" as another descriptor of the database entry
- Use layered truth, i.e., e.g., actionable region --> lesion --> cancerous lesion
- Characteristics of the data acquisition system (e.g.., physical characteristics of an imaging system)
- Methods to handle changing metadata over time (updates or new entries)

- Databank infrastructure—should not be simply list retrieval; but rather needs intelligent knowledge extraction [semantics/annotation].
  - Input interface
  - Internal organization (note needs to be able to handle image data)
  - Intelligent retrieval based on -- First searchable annotation -- Then intelligent feature extractions
  - Retrieval (web-based, others)
  - Open source
  - Quality control (authenticity of input data/metadata/truth, integrity of maintained data)
  - Integrity of database development to include ethical standards
  - Handling of IRB & HIPAA issues; and associated road blocks
  - Security (privacy issues, varying limited access rights for input, browsing, and retrieval)
  - Links to source, e.g., the clinical trial from which data/images came
  - Flexible/dynamic/expandable/scaleable/robust database
  - Flexible data entry including new modalities; changing truth, expandable
  - Ongoing maintenance (bugs, new metadata, elimination of old, curation)
  - Oversight and advisory committees
  - Ability to reuse data
  - Customer involvement and support
  - Linkage with FDA

- Interoperability (linking among) databases
  - Being considered under caBIG initiative for some cancers and can be translated
  - Common language/structures/ontologies (e.g., UMLS)

3.3 Breakout Session 3: Software Tools for Modeling, Data Analysis, Data Integration, and Work Flow.

Initial Charge to Panel: During this breakout session the state of existing software tools (for feature extraction, data integration, pattern recognition, analysis and interpretation, presentation, etc) for various physical and physiological measurements, including biomedical imaging, and molecular (genomic) profiling was discussed. As important integral parts of software and data sharing for clinical decision support systems, issues of data integrity, software validation, usability, testing, distributed resources, etc was addressed and recommendations for future direction and requirements in achieving the goals set was made. This session discussed bringing

the state of the art in software development to clinical decision applications in ways which is rewarding to both disciplines. The topics of discussions included:

During this breakout session the state of existing software tools (for feature extraction, data integration, pattern recognition, analysis and interpretation, presentation, etc) for various physical and physiological measurements, including biomedical imaging and molecular (genomic) profiling was discussed. As important integral parts of software and data sharing for clinical decision support systems, issues of data integrity, software validation, usability, testing, distributed resources, etc was addressed. The topics of discussions included:

3.3.1 Discussion Point: *Software Sharing: identify the methods to study and establish guidelines and requirements for software validation, usability, testing, and performance measures.*

Recommendations:

- Data at the NIH repository can be organized in a way so that it becomes benchmark tests. Algorithms can be tested with regard to stored images and data and then published. Good organization of the data is important. Open source is important so researchers can talk to each other and advance each other's work. There are problems getting things into the market. Formulating problems, formulating data and then testing is important.
- NIH should clarify and articulate policies related to software sharing. Specifically, NIH should define its position in relation to:

  ➢ open source software

  ➢ software deposited in public clearing-houses

  ➢ maintenance of shared software

  ➢ mechanisms to combine academic and industry expertise

  ➢ social issues of software and data sharing, including integrity, provenance, equal access, and records of access

- NIH should formally recognize the importance of key (but often neglected) software development goals. Such software should perform with acceptable efficiency in clinical environment and meet real-world accuracy and usability constraints. In this regard there are workable business models for community development of tools [e.g., visualization tool kit (VTK)]. For example: NIH should:

  ➢ Host a workshop to identify needs/models for community-based software development.
  ➢ Host an online archive/forum to help foster the community of software developers.
  ➢ Provide incentives for contributing data into the funded domain as identified deliverables (e.g., data used for evaluation).
  ➢ Support related efforts to create appropriate taxonomies, identify best-practices from other communities, etc.
  ➢ Help to carry products "across the chasm" by supporting early deployment stage where value/utility can be demonstrated.
  ➢ Promote development and dissemination of interoperable software and portable data, with provisions for data portability, translation, open transfer, etc.

3.3.2 Discussion Point: *Software Validation, Usability, and Testing: discuss the development and dissemination of interoperable software and portable data to include data portability, translation, open transfer, as well as the necessity of languages and ontologies for managing software as well as data.*

Recommendations:

- NIH should study and establish guidelines for what constitutes acceptable validation, testing, and usability evaluation of software (including web-based interfaces to databases) in conjunction with NIH-sponsored projects. Common test platforms and evaluation metrics should be defined, including clinically relevant performance measures.
- Encourage and facilitates reproducibility, technology transfer, and reusability.
- Provide or support data-driven benchmarks to support reproducibility.
- Consider creating archives of data used for evaluation or certification and testing.
- Require that software designs to be "what-oriented" (what we need) rather than "how-oriented."
- Support development of tools and interfaces supporting ad-hoc (vs. hypothesis-driven) queries.
- Support a study to gather user requirements from clinicians and use to define future directions.
- Require that software requirements documents be made public as part of deliverables.
- Consider adopting a set of guidelines for software requirements documents so that they could be mutually intelligible to other developers.
- Support "publishing" software and data as well as results.
- Require that deliverables include reproducible "use cases" for validating the tools.
- Support and encourage submitting data and application as publication requirement.
- Archive the supplementary information associated with journals – in an effort to move toward requiring fuller information about published studies.
- Provide incentives for contributing data into the public domain, possibly as one of the deliverables from funded research projects.
- Provide guidelines for management layers to encourage willingness to put data in the public domain.
- Create institutional arrangements for data repositories with controlled access for validation of research software is an important possibility to consider.
- Establish a website where various people can put links to their data. NIH will put disclaimers on there saying they are not responsible for the data.

3.3.3 *Discussion point*: *Data and Software Integration: Initial issues concerning interoperability, including software/system/application interoperability were discussed. These included*:

➢ Data portability, translation, and open transfer

➢ Software interoperability, including the combination of programming elements into parallelizable and/or sequential processing stages for large-scale computing

➢ Languages and ontologies for managing software as well as data; develop new conceptual abstractions for combining research efforts across disciplines

The recommendations made with regard to the above issues were as follows: NIH should promote:

- The development and dissemination of interoperable software and portable data.

- Fundamental research and software-engineered development of imaging, visualization, image-based databases and data mining.

3.3.4 *Discussion point:* *Shared, Distributed Resources: identify the need for operation and distribution of resources for shared models, tools, and datasets, including research on their continued development. Types of shared and distributed resources, issues related to collaborative exploration environments and the role of government in building sustainable infrastructure were discussed. It was discussed that whether tools (i.e. software) with general*

purpose should be preferred to paradigm driven ones. In former case putting them into use for a specific need the issues of evolutionary development and a central pursuit were addressed. In general purpose tools, you need evolutionary and progressive annotated databases where data is tested and reported in the same central database. A good example is treating images as multivariate functions which could be used for visualization and diagnosis.

Recommendations:

- Develop specifications for each of the stages so the collective force of the community can be unleashed.
- Be involved in making a national infrastructure for health maintenance and for sharing clinical data needs to be built and maintained.
- The operation and management of shared, distributed resources (community models, tools, datasets, computational servers), including ongoing development and enhancement activities.
- Consider the full life cycle of software tools and how to encourage development of a national infrastructure for sharing clinical (not just research) data.

3.3.5 Discussion Point: *Visual/Imaging-based Biomedical Informatics: identify the need for attracting research and development for visual/image-based data and integration into data mining development to include promotion of sophisticated image/data comparisons based on statistics, topology, semantics, and other abstract measures made of the non-textual data. Many advances still need to be made in fundamental computer science research, including multi-scale, multi-function visualization/imaging. Issues related to interactive access to multi-modal, multi-field data Error/uncertainty simulation and visual representation were discussed. Although, Imaging and visualization are ubiquitous, they are not coordinated. It's difficult for tools developers without access to representative groups of images. Data mining and data comparisons based on topology, statistics, model comparisons, and semantics are of interest and, therefore, integration of images with determining and network structuring would facilitate reaching these goals.*

Recommendations:

- Promote the visibility of image-based data and attract research and development for imaging, visualization, and image-based determining techniques and tools. Specific goals should include:

  - ➢ promotion of sophisticated image/data comparisons based on statistics, topology, semantics, and other abstract measures made of the non-textual data
  - ➢ Integration of image-based/non-text information with existing data mining methods.

- Invest in fundamental (not just applied) imaging and visualization research.
- Create inter-institute programs to develop general (as well as area-specific) software tools.
- Create an open database of images to use as "gold standards" for researchers.

3.3.6 Discussion Point: *Evaluate the Review Process for Development of Software and Tools: discuss the potential need for separate scientific and engineering evaluation groups, policies regarding evaluation of non-hypothesis driven research and development, and interdisciplinary team-based research.*

Recommendations:

- NIH should clarify and articulate policies regarding the review and funding of software/tool development programs including the formal study of program announcements, and requests for applications for software development initiatives. This

should include the potential for separate scientific and engineering evaluation groups as well as reviewer familiarity with interdisciplinary team-based research.

3.3.7 *Discussion Point*: *Education & Curriculum Development: The need to move toward including all disciplines (biologists, computer scientists, clinicians) on PhD committees was discussed. This is due to the fact that whole approach of many disciplines has been dramatically changed by new technology and the clinical practice should not be an exception to this rule.*

Recommendations:

- Support training and development of new kinds of curricula to facilitate penetration and best practice application of information technology into clinical research and care.

- Support grants programs supporting interdisciplinary training involving biologists, computer scientists, and clinicians.

- Expand software packaging projects to include training clinicians in its use.

- Be more proactive in determining how new technology can and should fundamentally transform clinical practice – and in developing appropriate training.

3.3.8 *Discussion Point*: *Long-Range Approaches: Previous directions have focused on data vs. tools vs. research vs. practice, rather than a more holistic approach. Also, Biologists/bioengineers/computer scientists and doctors need to function as true peers. Focus has been on very large scale data, but ultimate impact is measured in terms of clinical decisions. Clinicians should always be involved as Computer Scientists should be in average study section.*

Recommendations:

- Adopt a "solution architectures" approach (problem-driven view of entire complex of tools /data /computation needed to solve similar type of problems, including validation requirements)
- Support research to integrate image data-base information with existing data-mining methods.
- Create study sections that better support computer science, bioengineering, biology, and practitioners' peer review.
- NIH should support software exploration and early development intended to benefit relatively few (pioneering breakthroughs). But, in this effort it needs procedural distinctions between this kind of software and clinical support (real-world) software. This should include:
  - ➢ Separate review bodies
  - ➢ Distinct evaluation criteria
  - ➢ Distinct meaning for "impact" and "results"
  - ➢ Distinct criteria for assessing outcomes.
  - ➢ Non-hypothesis-driven R&D

- Adopt policies/procedures that clarify when alternate approaches are acceptable, evaluation criteria.
- Adopt special procedures/criteria needed for interdisciplinary projects (pre-funding and project evaluation). This procedure should address:
  - ➢ Reviewers with relevant personal experience
  - ➢ Participants in large, interdisciplinary projects
  - ➢ Leaders of distributed collaborations
  - ➢ Criteria reflecting key barriers to success
  - ➢ Evidence of previous interdisciplinary collaborations
  - ➢ Sound coordination/management plan

- Support more computer-science oriented study sections with focus on key NIH-relevant directions.

3.3.9 Discussion point: *Work flow: Tools will only have an impact on health if people use them, and they have to be within the context of people's workplace. One of important concepts should be that programming for health care and clinical decision making support cannot be traditional in the sense that it assumes regularity and determinism. Biomedical data is variable, ambiguous, and noisy: biological sequences, images, experimental observations, written descriptions. Problem-driven curriculum should be created. Rich domain provides motivation and problem sets for central areas of computer science (algorithms, databases, machine learning, and software engineering). Problem-driven courses need basic modeling and computational concepts learn relevant software tools, connect tools to create analysis/model, and manage computational resources (time, space, processors).*

Recommendations:

- Take on early deployment of a concept to demonstrate the usefulness of the concept and encourage industry to pick it up.
- Searching an archive for software development and issues is an incredible resource, and the NIH can assist with this so people can communicate with others in the field addressing the same software problems; this could spark collaboration.
- Consider funding specific programs to allow clinicians to build software. The clinician should specify what they want to do. The system should provide constant feedback about how far along it is. NIH could fund a study to interview clinicians and find out what it is that they want.
- Adopt a set of guidelines for specification to get the software engineers together on this topic. More high-level building blocks are important.

The following staff organized the symposium:

Anna Retzke (NIH/NIBIB), Bruce Hamilton (NSF), Terry Yoo (NLM/OHPCC), Michael J Ackerman (NLM), Tom Aigner (NIH/NIDA), Mohammad Al-Ubaydli (NLM/NCBI), Tim Baldwin (NIH/NHLBI), Carol Bean (NCRR), Laurence Clarke (NIH/NCI), Milton Corn (NLM/EP), Peter Covitz (NIH/NCI), Mrinal Dewanjee (NIH/CC/LDRR), Sue Dubman (NIH/NCI), Gregory Farber (NIH/NCRR), Nancy Freeman (NIH/NIDCD), Peter Good (NIH/NHGRI), Alexander Gorbach (NIH/CC/DRD), Stephen Green (NIH/NIBIB), John Haller (NIH/NIBIB), Mervi Heiskanen (NIH/NCI), Michael Huerta (NIH/NIMH), Carl Jaffe (NIH/NCI), Saleet Jafri (George Mason U), Eric Jakobsson (NIH/NIGMS), Peter F Johnson (NCI CCR LP), Vipul Kashyap (NIH/NLM/LHC), King LI (NIH/CC/DRD), Martha Lundberg (NIH/NHLBI), Peter Lyster (NIH/NIBIB), Michael Marron (NIH/NCRR), Reza Momenan (NIH/NIAAA), Grace Peng (NIH/NIBIB), Karen Peterson (NIH/NIAAA), Marc Rigas (NIH/CSR), Karen Skinner (NIH/NIDA), Theresa Smith (NIH/NIBIB), Mollie Sourwine (NIH/NIBIB), Ram Sriram (NIST), Daniel Sullivan (NIH/NCI), Ronald Summers (NIH/CC/DRD), Richard Swaja (NIH/NIBIB), John Whitmarsh (NIH/NIGMS), and Chipper Whalan (Capconcorp)

The organizers would like to acknowledge the contribution of the following members of the external advisory committee

Committee Co-Chairs:
Ruzena Bajcsy, PhD, Director for Information Technology in the Interest of Society, University of California Berkeley
Mike Vannier, MD, University of Chicago Medical Center

Committee Members:
Ricardo Avila, MS, GE Global Research Center, Niskayuna, NY
David Brown, PhD, Food & Drug Administration, Rockville, MD
Reed Gardner, PhD, Medical Informatics, University of Utah School of Medicine
Mary Ellen Giger, PhD, Department of Radiology, University of Chicago
John Glaser, PhD, Partners HealthCare System
Ary Goldberger, MD, Harvard Medical School
H. K. Huang, PhD, Department of Radiology, University of Southern California
Issac Kohane, MD, PhD, Department of Pediatrics, Children's Hospital Boston
Harold Kundel, MD, Department of Radiology, University of Pennsylvania
Sarah Nelson, PhD, Division of Bioengineering, University of California, San Francisco
Joel Saltz, MD, PhD, Biomedical Informatics Faculty, Ohio State University
Ben Schneiderman, PhD, Department of Computer Science, University of Maryland
Edward Shortliffe, MD, PhD, Department of Biomedical Informatics, Columbia University
Ruth Taylor, IBM Life Sciences
Ralph Weissleder, MD, PhD, Center for Molecular Imaging Research, Harvard Medical School
Michael Welge, MS, National Center for Supercomputing Applications, University of Illinois

The organizers would like to acknowledge the contribution of the following presenters and panel members:

Presenters:
Stephen Katz, MD, PhD, National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH
Pedro Del Nido, MD, Children's Hospital, Boston
Laura Esserman, MD, MBA, University of California at San Francisco
Edward Shortliffe, MD, PhD, Columbia University
Michael Hehenberger PhD, IBM Life Sciences
Maryellen Giger, PhD, University of Chicago
Jim Bassingthwaighte, PhD, University of Washington
Ricardo Avila, MS, GE Global Research Center

Isaac Kohane, MD, PhD, Harvard Medical School
Ken Buetow, PHD, National Cancer Institute
Aravinda Chakravarti, PhD, Johns Hopkins University

Breakout Session 1: Clinical Challenges and Related Software/Informatics Requirements
Moderator: Edward Shortliffe, MD, PhD

Panelists:
Andrea Bradford, MD, Georgia Department of Human Resources
Lawrence Friedman, MD, National Heart, Lung, and Blood Institute, NIH
Reed Gardner, PhD, University of Utah School of Medicine
Gary Kelloff, MD, Cancer Imaging Program, National Cancer Institute, NIH
Michael Knopp MD, Ohio State University
Emanuel Petricoin, PhD, Co-Director, NCI-FDA Clinical Proteomics Program, Food and Drug Administration

NIH Coordinators:
Milt Corn, PhD, National Library of Medicine
Carl Jaffe, MD, National Cancer Institute

Breakout Session 2: Databases as Required for Assessment and Application of Software Tools
Moderator: Maryellen Giger, PhD, University of Chicago

Panelists:
David Brown, PhD, Director, Division of Imaging and Applied Math., OSEL, Center for Devices and Radiological Health
Peter A. Covitz, PhD, Director, Bioinformatics Core Infrastructure, National Cancer Institute,
Ary Goldberger, MD, Harvard Medical School
Ron Kikinis, MD, Brigham & Women's Hospital
Geoffrey McLennan, MBBS, FRACP, PhD University of Iowa
Thomas Wittenberg, PhD, Fraunhofer-Institute for Integrated Circuits (IIS)

NIH Coordinators:
Peter Good, PhD, National Human Genome Research Institute
Michael Ackerman, PhD, National Library of Medicine
Mervi Heiskanen, PhD, National Cancer Institute

Breakout Session 3: Software Tools for Modeling, Data Analysis, Data Integration, and Work Flow
Moderator: James Bassingthwaighte, PhD, University of Washington

Panelists:
Narendra Ahuja, PhD, University of Illinois
Christopher Johnson, PhD, University of Utah
Wayne Kubick, MBA, Lincoln Technologies, Inc.
Cherri Pancake, PhD, Oregon State University
Fernando Pereira, PhD, University of Pennsylvania
Stephen Rosenfeld, MD, MBA, Chief Information Officer, Associate Director for Clinical Research Information Systems, Clinical Center, NIH
Ricardo Avila, MS, GE Global Research Center

NIH Coordinators:
Michael Marron, PhD, National Center for Research Resources
Terry Yoo, PhD, National Library of Medicine
Reza Momenan, PhD, National Institute on Alcohol Abuse and Alcoholism

Appendix A: Final Report on Satellite Meeting: Using Standards to Integrate Biomedical Imaging into Clinical Decision Making.

Carol Bean (cabean@mail.nih.gov), NIH and Ram D. Sriram, NIST (sriram@nist.gov)

Introduction

Advances in molecular and genomic imaging technologies in combination with anatomical and functional medical imaging technologies are poised to have a tremendous impact on healthcare. Hence, it is important to effectively integrate various types of images generated by these technologies into clinical decision making. Standards have enormous potential to facilitate and coordinate such integration and interoperability of biomedical imaging applications and practices for clinical decision-making. However, the development and implementation of standards is extremely resource-intensive and can take years to accomplish; further, standards evolve so must be tracked and maintained through on-going efforts.

Numerous standards exist in the healthcare arena, most of which focus on a specific biomedical or clinical subject domain at some level of granularity or a specific type of healthcare encounter. Within each area there exist standards to address variously the meanings of terms and the relationships between them, the overall conceptual model for the domain or enterprise, or the ways in which information must be formatted in order to be transmitted electronically and used by machines.

The DICOM family was designed primarily as a data interchange standard to deal with the acquisition and transportation of digital diagnostic images and associated information over networks. Today virtually all imaging modalities in radiology (e.g., CT, MRI, ultrasound, RF, PET, etc.) support the DICOM standard. An interest in linking all images associated with a single patient into some form of integrated master patient record is spurring other specialty domains (e.g., pathology, ophthalmology, dermatology, gastroenterology, etc.) to develop and implement extensions to the DICOM standard. Similarly, the HL7 series comprises a set of data interchange formats for use in other applications of the healthcare enterprise.

Additional dimensions of standardization beyond data interchange formats are important for using interpretations of these images in the context of clinical decision-making. These semantic dimensions correspond to the meaning of terms used to express concepts, the relationships that obtain between and among concepts and the knowledge structures used to represent them, and various means of modifying the attributes of both concepts and relationships.

The primary purpose of this workshop (held on June 22nd 2004 as a satellite workshop of the joint BECON/BISTIC Biomedical Informatics for Clinical Decision Support symposium) was to explore the potential of semantic standards for enhancing integration of biomedical imaging into the clinical decision-making process.

We offered the following themes for guiding the discussions:

1. Clinical decision-making relies increasingly on information gathered from a variety of sources. Because of differences in perspective and tradition, health-care providers in different specialty domains that use images (such as radiologists, surgeons, internists, etc.) tend to use different terms for description of the same anatomical features and landmarks as well as for clinical observations and interpretations. How can standards be used to integrate and coordinate the annotation of medical images with clinical observations from different specialty domains?

2. The same image object corresponding to a set of measurable observations may be interpreted in a different manner, yielding variation or conflict in subsequent treatment planning. Further,

these interpretations are likely to be highly contextual, depending on the patient data, anthropological gene pool, clinical objectives, etc. Often these criteria address degree or severity; for example, the same tumor image object may be interpreted as mild by an oncologist and as severe by a radiologist. To what extent do such criteria that influence decision-making need to be standardized? How can ontologies be used for standardization of decision-making criteria? Are existing standards sufficient for these tasks?

Standards Workshop Summary

The workshop chairs, Ram Sriram and Carol Bean, gave an overview of the field. Their presentation can be found at http://www.becon.nih.gov/symposium2004.htm. The following questions were posed to the audience.

A. Adequacy of existing and planned semantic standards

1. Are *current approaches and standards* sufficient for integrating imaging into clinical decision-making tasks?

2. If current approaches and standards are inadequate or insufficient, where are the important *gaps and deficiencies*?

3. In what specific *domains* are standards lacking?

4. What *enhancements or extensions* are needed for effective use of standards in the domains of interest?

5. What *additional areas* (e.g. degree and severity in different specialty domains) might exist for semantic standards and integration?

6. To what extent do *contextual criteria* that influence decision-making (e.g., clinical objectives, anthropological gene pool) need to be standardized?

7. Can reliable standards be established for *qualitative classes* or categories (e.g., severity), especially across domains?

B. Implementation

1. How can semantic standards be incorporated into *imaging practices* in a realistic and efficient way?

2. How can ontologies and other structured terminology resources be used for standardization of *decision-making criteria* using imaging?

3. How can semantic standards be used to integrate and coordinate the *annotation of medical images* with clinical observations from different specialty domains?

4. What are the *obstacles to implementation* and use and how can they be overcome?

5. What *obstacles to interoperability* inhibit sharing, coordinating, and integrating biomedical image data across modalities and among sites for support of clinical decision-making?

6. How can various standards organizations, including government agencies, *help* in the implementation process?

The responses to above are summarized below.

Adequacy of Standards

Current standards need to be extended. It was pointed by one participant that in the clinical arena, physicians refer patients for imaging studies when it has been proven for diagnostic value—e.g., for an ultrasound to diagnose appendicitis because it has been shown to identify patients with and without appendicitis. People (such as radiologists) who see images, do not see patients, they see a shadow of patients. They use an ontology of shadows: they see masses. It would be unrealistic to expect that the radiologist may be able to suggest interpretations. Hence, it was it is the responsibility of the physician to make sure which of those best fits the situation. However, the majority of the participants felt that the current scenario for radiological interpretations was untenable. It was asserted that when radiologists may see a shadow it is important for them to report it in their reports to physicians in as definitive a manner as possible, i.e., to communicate it consistently. There was a feeling that there were limitations in terminologies like SNOMED, deficits, including limitations such as describing what was actually done. e.g., "How a contrast (agent) was administered during a CT scan." The need for global assessments – a measure of characterizing the extent of cancer in a patient within the limitations of imaging, so that what the radiologist says in the report (about the risk of the cancer) can be expressed in a manner that can be unambiguously understood by others was stressed. This clearly underscores the case for semantics.

Standards should incorporate semantics. It was observed that nowadays one can DICOM any image. But we were to have a semantic standard where the precise definition is explicit, then we could expand it with new terms that enable connection between the terminology and the old terms. As an illustration it was pointed out that radiologists provide a verbal interpretation of the image using standards like SNOMED. However, if information could be incorporated into the image, then image processing would enable a semantic interpretation of the image.

Standards for interoperability are needed. If we can represent what we see in an effective way we can make it interoperable. Standards and terminologies are important. Without them we cannot make systems interoperable. If there is a need for a terminology to be frozen, then it may place limitations in future advancements. It was felt that we should allow enhancements on top of existing standards. While it is important to encourage standardization and adoption to enable queries, dissatisfactions with the terminology should be voiced in order to enable improvements.

Inadequacy of DICOM for pathological slides. Although there was a portion of the DICOM standard -- called visible light standard --- which could accommodate the imaging of pathology slides, in general there seems to a lack of standards for representing pathological slides. There were comments from the workshop members about DICOM compliant imaging systems, non-adoption by pathology vendor community, difficulties encountered with JPEG 2000 images, and challenges to be overcome due to the size of images.

Lack of computerized reporting tools. A major problem is a lack of computerized reporting tools for annotating images in a structured manner. It seems that radiologists generally produce only a written report, not with annotated images. Standardized tools for structured reporting would be a paradigm shift for radiologists, and would start them on the road to using more sophisticated semantic tools.

Standards are needed for proper data usage. Optimal exploitation of data from radiological/imaging examinations can be achieved only if certain minimum standards are adhered to in terms of: a) data acquisition procedures; b) data analysis (computer-based) and quantification, and c) data translation, i.e., what does an enhancement pattern or uptake pattern correlate to in physiological, morphological and functional terms.

Standards should be free. There was a comment to the effect that there should not be licensing of industry standards; in effect the standards themselves should not be impediments for the practice of medicine and should therefore be above licensing. One example that was cited: "No one pays

a licensing fee for buying a [standards-compliant] screw from a hardware store." While in principle this is a valid notion it was pointed out that organizations like ANSI (American National Standards Institute) or others may sell standards documents for a fee, but that was for the publication not the standard itself. In other countries, standards are given away for free (to facilitate commerce and streamline operations) and that standards were not licensed.

Implementation Issues

Indexing and retrieval of images. One of the participants enquired how images could be indexed using parameters to enable searching database systems, or text retrieval systems. A response to this enquiry was as follows: "We saw an example earlier, showing an image on the left and XML on the right (from the overview slides). The standard is at a low level (DICOM) that we cannot get that kind of information right now. Integrated Health Enterprise solutions (IHE) is using DICOM-SR which will enable images and report structure annotations."

Legislation may accelerate terminology adoption. An example of the development of a very useful and well-defined terminology emerged in the arena of testing for cervical cancer was noted. After the emergence of the PAP test there was widespread use in the testing of the Human Papiloma Virus (HPV). Due to abuses in cytology practices including the use of technologists to interpret excessive numbers of slides, Congress passed Public Law 100-578, the Clinical Laboratory Improvement Amendments of 1988 (CLIA '88). Afterwards, the Bethesda System of Classification was introduced (under the sponsorship of NIH). The ontology in the Bethesda classification (since it is more than just a classification or terminology) has had a significant impact in the reporting of HPV testing results and interpretations in a uniform, unambiguous, and readable format and better follow up. It was predicted recent advances in a DNA probe for HPV would soon result in a YES/NO test. Another example is the Mammography Quality Standards Act of 1992 (MQSA). Since data fed back to a radiologist would help to improve quality of mammography, there was a formalization of terms used in mammogram reports in conjunction with the Breast Imaging Reporting and Data System (BI-RADS) to unambiguously state the risk of malignancy when interpreting mammograms and also helps to make follow ups for patients to be much easier.

Lack of compliance measures hamper use of standard terminologies in radiological reports. There was an enquiry whether the above approach cannot be adopted for spine X-rays, and others. One response, which appeared to resonate with the rest of the attendees was that there were no pressing requirements to develop or adopt such a standard, due to the cost of compliance. Mammography was forced by legislation — MQSA. But no similar provisions exist that can be enforced. We need to encourage the adoption of similar measures for all other diagnostic imaging.

Role of medical imaging vendors and societies. A question was asked whether vendors of medical imaging systems could be asked to create extra parameters for the image (e.g., resolution of image, and others) that could be automatically generated. The response to this question was that RSNA (Radiological Society of North America, http://www.rsna.org) is providing a framework for formalization of terms. RADLEX is a pilot project for the thoracic arena, in which a lexicon is being developed using terms in UMLS and other terminologies and filling in gaps in terminologies.

Planning for a Future Workshop

The workshop chairs indicated plans to conduct a two day workshop on the role of standards for integration of multi-modalities in clinical decision making which was favorably received by the participants. The participants felt that use cases and thinking of practical solutions would be necessary and should involve ontologists, clinicians, radiologists, decision support folks. It was acknowledged that we needed more knowledge representation experts to participate in the meetings.

There was a suggestion that one venue for the proposed workshop could be the annual AMIA symposium. A workshop either at or before or after would be suitable. And importantly, it would

not have allegiances that may come in the way. It could be a half day workshop followed by in-depth workshops on specific topics that would be identified.

Another member of the workshop suggested that we gather input from the semiconductor industry which had successfully addressed issues that were similar to those in imaging. He added that they have several tools based on the position of a component on the wafer, similar to anatomy being defined in relationship to the body, instead of saying it in words. The chairs concurred, indicating that tools and approaches in topography and Geographic Information Systems (GIS) may also be appropriate and should be considered

Workshop Participants

The participants in the workshop include the following: Larry Reeker (NIST), Alan Rowberg (Rowberg and Associates), Mehmet Kayaalp (NLM/NIH), Steve Ray (NIST), Richard Liu (SAIC), Kurt Augustine (Mayo Clinic), Tony Pan (Ohio State University), Barbara Beekerman (Y12 National Security Complex), Michael Galdzicki (Fujitsu Labs of America), Andres Kriete (Drexel University), Curtis Langlotz (University of Pennsylvania), William Hayes (Astra Zeneca R&D), Craig Liddell, Charles Sneiderma (NIH), Thomas Wittenberg (Franhaufer IIS, Germany), Bret Peterson (NCRR/NIH), Alex Wang (NIH), Seza Orcun (Purdue University), Amie Fitchard, Karina Tulipano (Columbia University). Additional comments through e-mail were provided by Ron Summers NIH) and Willy Eidesaunet (GE Healthcare, Olso). The workshop minutes were scribed by Ravi Raman (NIST).

Appendix B: Final Report on Satellite Meeting: Public/Private Partnerships: Potential Means To Enhance Biomedical Informatics Resources.

Larry Clarke NIH (lclarke@mail.nih.gov), and Karen Skinner NIH (kskinner@mail.nih.gov)

This satellite meeting explored the feasibility of developing biomedical research resources such as databases for the standardized assessment of application specific informatics software tools for translational medical research and clinical trails. The first two presentations described a model for informational technology and informatics software tools to support clinical trials for drug research and clinical decision-making, both private sector initiatives. The next presentation described a success story for a public-private partnership that supports the development of informatics software research resources for image processing methods as applied to osteoarthritis, a public resource supported by a public-private partnership organized by the Foundation for NIH (FNIH). The next two presentations by the FNIH and NCI provided a report on progress for an emerging FNIH public-private partnership to support validated reference image databases for the standardized assessment of informatics software tools for cancer imaging, where lung cancer screening, diagnosis and therapy response data base resources are being developed demonstration projects as web accessible resources. The final presentations included case reports from two imaging companies that described their interest in the development and support for informatics databases as a useful resource for both the device and drug industry.

The panel members reviewed how these research resources could be developed as a trans-NIH effort, initially as image databases resources, but to include other databases such as molecular profiling or other biosensor data. It was agreed than an array of databases is required for the evaluation of different application specific software tools, data integration and eventually clinical decision methods. The panel members agreed that (a) the development of these database resources and (b) the related software tools for the creation of validated data bases, image annotation methods, and web query systems for benchmarking software performance; would be a very useful trans-NIH, FDA, NIST and federal government resource. These resources are important for translational research for the evaluation of informatics software tools as required for NIH and privately supported clinical trials as described in the BECON-BISTI 2004 symposium report.

Satellite Agenda

Extramural Chairs: Larry Clarke PhD, NCI;   Karen Skinner PhD, NIDA.
Lclarke@mail.nih.gov,   Kskinner@mail.nih.gov

External Chairs: Michael Vannier MD, University of Chicago.
Michael Knopp MD, Ohio State University.

1.50-2.15 PM   Shaping Government and Industry: Vision for Health Care
Information Technology. www.nahit.org
National Alliance for Health Information Technology.
Bill Head. Vice President for Policy and Government Affairs.

2.15-2.30 PM   Standards for Commercial Data Management Systems (CDISC).
Wayne Kubick PhD, http://www.cdisc.org/

2.30-2.45 PM   Public Private Partnerships: A success story.
Osteoarthritis Initiative. www.fnih.org
Gayle Lester PhD., NIAMS.

2.45-3.00 PM   Overview of the FNIH and the Imaging Data Base Initiative:
Wendy Sanhai PhD. www.fnih.org

3.00. -3.15 PM   Scientific Goals: FNIH-NCI Imaging Database Initiative.
Potential expansion as a trans-NIH Initiative.
Larry Clarke PhD. NCI.

3.15-3.30 PM   Device and Drug Industry Case Report:
Imaging database Initiative:
Rick Avila, GE Global Research, USA.

3.30-4.00 PM   Device Industry Case Report: Imaging database Initiative.
Informatics Resources Requirements for Clinical Trials.
Gudrun Zahlmann. Siemens: Erlangen, Germany

4.00-4.30 PM   Panel Discussion.
Co Chairs and presenters
NIH Road Map Representatives:
Dushanka Kleinman (NIDCR), Carl Roth (NHLBI)
FDA Representatives
David Brown (FDA CDRH),

4.30-5.00 PM   Open Discussion.

5.00 PM   End of Satellite Meeting.