Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.988.7800 F.508.988.7956 www.lifescience-insights.com

LifeScience Insights™ An IDC Company

SPECIAL STUDY

Biobanks: Strategies to Address Next-Generation Challenges for Biobanking

Sponsored by: IBM

LIFE SCIENCE INSIGHTS: Alan Louie, Ph.D.
Brock Reeve

Special Contributors

Carolyn Compton, M.D., Ph.D., Director, Office of Biorepositories and Biospecimen Research, National Cancer Institute

Mylène Deschênes, BCL, LLB, LLM, Executive Director, P3G Consortium, University of Montreal, Canada

Benedikt Furrer, IT Architect, IBM Global Services

Simone Gauch, Associate Director Corporate Strategy, Qiagen GmbH

William Grizzle, M.D., Ph.D., Professor of Pathology, UAB Comprehensive Cancer Center

John D. Potter, M.D., Ph.D., Senior Vice President and Director, Fred Hutchinson Cancer Research Center

INTRODUCTION

A Vision for Information Based Medicine: by Brett Davis, Global Solutions Executive, Basic Research and Discovery Information Based Medicine, IBM Healthcare and Life Sciences

"The future is already here. It's just not evenly distributed yet."

-William Gibson

The healthcare and pharmaceutical industries have been buzzing with the promise of personalized medicine since the inception of the human genome project and other "big science" projects. Around the world we are seeing governments, employers, providers, and insurers driving a shift toward electronic health records and pervasive health technologies capable of delivering critical information at the point of care. These advances and investments in technologies to capture, deliver, integrate, and exchange biomedical information more effectively, coupled with advances in our molecular understanding of disease further strengthens this promise. Increasingly — almost daily it seems — we are seeing new announcements that indicate that the era of "personalized" medicine may be upon us.

Filing Information: October 2005, Life Science Insights #05LSIC4619, Volume: 1
Special Study

IBM®

As this decade continues to unfold, we will certainly see continued progress in science, biotechnology, and information technology that will generate even more excitement. However, many ethical, financial, intellectual property, political, and organizational challenges still stand in the way of seeing this progress translated into real improvements in patient outcomes. To ensure that this does not occur will require unprecedented collaboration and leadership between the public and private sectors. Synergies and connections also need to be created between the different stages of scientific endeavor and clinical practice, where historically there has been little collaboration and information sharing, resulting in information "silos" and lack of timely knowledge sharing. Enabling these synergies will not only require investment in interoperable IT systems, but a rethinking of business models, operational processes, funding mechanisms, and current incentive structures.

The need for proactive leadership and collaboration by the extended community to achieve this vision is perhaps no more apparent than in the area of biorepositories or "biobanks." Thought leaders from industry, government, and academia recognize that researchers and clinicians increasingly need access to molecular information but also the corresponding phenotypic information often contained in medical records or collected in the course of clinical trials. Biobanks — sometimes called biorepositories or tissue banks — provide both types of critical information, and thus can serve as an important translational bridge between research and clinical practice to accelerate the discovery and development of more personalized medicines.

IBM Healthcare and Life Sciences recognized the important need for collaboration across sectors and hosted two World Wide Biobank Summits in 2004 and one in May 2005 to bring the global "biobanking community" together to discuss and address the critical scientific, policy and technology challenges associated with biobanking. This community includes leaders from biopharmaceutical companies, technology providers, patient advocacy groups, government policy makers, and academic research centers.

After each summit, IBM teamed with IDC Life Science Insights to develop white papers based on the discussions and proposals made at each summit. Two white papers, *Biobanks: Accelerating Molecular Medicine* and *Biobanks: Collaborating for Cures* have been published detailing the challenges identified and debated at each summit. Once again, IBM has teamed with IDC Life Science Insights to develop this study to share more broadly the challenges and possible resolutions to these challenges that were discussed and debated at the third global summit held in Stockholm, Sweden at the Karolinska Institutet in May 2005.

Over the next several decades, personalized medicine is poised to transform medicine. Perhaps in no disease area does this promise seem more real than in cancer, a deadly disease that is caused by caused by complex interactions between genes, environment and lifestyle. Already, new diagnostic and prognostic tools are increasing our ability to predict the likely outcomes of drug therapy, and investment in pharmacogenomics is resulting in more focus on the development of targeted therapeutics.

Biobanks are one of the most critical resources required to help accelerate this transformation. They are a key resource in gaining an understanding of the interaction

between genes, the environment, lifestyle and disease, and then translate that knowledge into clinical practice quickly through innovative diagnostics, therapeutics and preventative treatment strategies. But once again, leadership from the community is needed to ensure that the samples and, more important, the data, that are captured as part of these initiatives can be used to maximum benefit in an ethical way. Recognizing this need, IBM Healthcare and Life Sciences is convening a fourth World Wide Biobank Summit focused specifically on harmonizing efforts across international biobanking efforts supporting cancer research and medicine.

Information technology will play an increasingly important role in this transformation. By providing the infrastructure and analytical tools needed to both integrate and understand the requisite genotypic and phenotypic information, IT is critical to revealing the complex underlying causes of disease and developing potential solutions. IBM Healthcare and Life Sciences is committed to playing a leadership role in accelerating this transformation, and has committed the people, technologies and resources to help make information-based medicine a reality.

TABLE OF CONTENTS

	P
Special Contributors	1
Introduction	**1**
A Vision for Information Based Medicine: by Brett Davis, Global Solutions Executive, Basic Research and Discovery Information Based Medicine, IBM Healthcare and Life Sciences	1
The Convergence of Societal Pressures, Business Drivers, and Technological Advances Is Driving Acceptance of Information-Based Medicine	1
Increased Recognition of the Value of Biorepositories	3
New and Improved Testing Capabilities: Technology Improvements Pacing Biorepository Development	8
Standardization: Moving from Samples to Quality Biorepository Biospecimens	11
Maturing of Biobank Infrastructure	14
Summary: Recap of Issues	**19**
Life Science Insights Perspective	**20**

The Convergence of Societal Pressures, Business Drivers, and Technological Advances Is Driving Acceptance of Information-Based Medicine

As evidenced by the discussions from the World Wide Biobank Summits I, II, and III, the biobanking community is actively moving toward building a sound technical and organizational foundation for biorepositories. Meanwhile, the world is moving forward independently in many directions that support their development as well. Societal needs and pressures, business drivers, and scientific and technological advances all are accelerating the adoption of biobanking as a critical resource on the path to information-based medicine.

Societal Pressures

As society becomes more comfortable with the ability of science and technology to understand, predict, and cure major diseases, there is increasing support by the general community to participate directly in finding solutions. Whether motivated by their own self-interest or the needs of future generations, people increasingly understand that they can be involved by participating in genetic testing, consenting to scientific research in both clinical trials and after medical interventions, and through directed political activism in search of cures for diseases that have affected friends and family. As the population at large becomes more informed about the underlying causes of diseases, they are more likely to support advanced technological approaches that offer the potential to prevent and cure diseases, especially in those areas that have previously proven untreatable. Furthermore, such activism is causing funding and regulatory organizations to take notice and provide support toward accelerating development of cures. This level of focus and attention benefits the biobanking community in a multitude of ways — from improved access to samples to increased funding to regulatory support in establishing common standards.

Additionally, there is increased broad awareness of the value of preventative care and the impact of personal and environmental factors on health. This awareness provides substantial support to efforts to reduce and eliminate damaging lifestyle factors that cause or trigger disease. At the same time, increased scientific understanding of the biological mechanisms underlying disease is also driving improved nutrition and lifestyle choices by the public. From the research perspective, more well-informed patients are acting as disease research advocates in addition to their direct contributions of biospecimens supporting scientific studies.

Business Drivers

There is widespread agreement that a major shift in the fundamental "blockbuster"-based business paradigm of the pharmaceutical industry is under way. With the number of potential blockbusters in development diminished, more focused advances based on information-based medicine are only expected to increase. Knowledge of the molecular basis of disease is growing and, when combined with population-based genetic data, offers the potential for more "personalized" medicine to be realized in the near to mid-term. For any given product, personalized medicine inherently creates a reduced total market size, but may allow for multiple versions of safe and effective

pharmaceutical solutions that can be approved in a timely, directed fashion. As stated by Dr. Anna Barker of the NCI at Biobank Summit II, "Unless substantial action is taken with biobanking and biospecimens, we'll delay personalized medicine by 30 to 40 years."

Biobanks' store of biospecimens and their associated clinical data is a key enabler of improved genetic analysis and screening. Improved genetic screening will, in turn, result in improved effectiveness and efficiency of clinical trials and should allow for shorter directed trials. For a specific disease, personalized medicine may result in the development of multiple personalized therapeutic options. These cumulative therapeutic approaches, combined with supporting genetic testing, may begin to replace the traditional blockbuster model, assuming combined markets begin to approach traditional market sizes.

Movement towards this more information-based medicine will also allow for the potential to recover failed drug candidates through elimination of at-risk patients and improved identification of target patients. Directed clinical trials to validate efficacy and safety issues should allow for potential recapture of previously written-off investments and will provide for new revenue opportunities. Biospecimens and associated data collected from clinical trial participants will be critical to identifying the key efficacy and safety factors that can drive product recovery.

Advances in Technology

Technology is advancing medical research at a rapid rate. From ever-increasing genomic data to noninvasive molecular imaging, technological advances are offering the potential of information-based medicine to mitigate disease through knowledge-based solutions. As the understanding of disease is increasing at the molecular level, both predictive and personalized medicine are considered to be within our reach. Advances in technology continue to reduce the hurdles to extraction of valuable data from biospecimens and will improve research efficiencies at increasing rates. In addition, advances in IT infrastructure (i.e., email, global networks, common database platforms, etc.) are enabling the fluid, almost transparent, ability of organizations to connect together, which will provide the biobanking community with the communication infrastructure needed to facilitate collaboration.

Key Learnings from Biobank IT implementations

Benedikt Furrer, IBM Global Services

Just in time for the third Worldwide Biobank Summit hosted at Karolinska Institutet (KI) in Stockholm, the Biobank Information Management System (BIMS) was launched into production. During approximately nine months (from September 2004 until May 2005), IBM and Karolinska Institutet architected, developed, and deployed the BIMS solution in a collaborative approach. The project benefited from a wide range of perspectives on the problem to be solved as the project team consisted of both IT professionals and researchers/academics from KI and IBM Sweden, backed up by global IBM Healthcare & Life Sciences skills and resources.

One of the critical concerns for the development of the BIMS solution was the design and implementation of necessary data security measures. On one hand, unauthorized access had to be prevented with the level of security of an ebanking solution; on the other hand, access to data for authorized users had to be limited flexibly to subsets of data stored in BIMS. Those requirements resulted in an advanced security architecture that was developed and deployed in order to satisfy these requirements.

Another challenge was to align the data models originating from the disparate data sources integrated to the BIMS solution. Data sources were not only to be integrated, but data models originating form these data sources had to be aligned in order to enable researchers to benefit from data aggregations of data originating from the integrated data sources. Data elements from integrated data sources were divided into a shared data model and into various source-specific data models. This required time-consuming studying of all data sources together with data owners. The deployed solution provides a first level of data alignment but the extension of the shared data model is an ongoing process that will provide more and more powerful data aggregation features in future BIMS releases.

The key success factor for the project was the flexible project delivery approach featuring a small project core team based in Sweden that was supported with leading-edge expertise in various areas, such as security, data federation, data abstractions, utilizing resources from IBM's worldwide organization.

Increased Recognition of the Value of Biorepositories

Biobank repositories are actively contributing to advances in medical research today. As the body of knowledge built on biobank specimens grows, recognition of the value of biobanks is becoming increasingly apparent to researchers, pharmaceutical companies, government and other funding organizations, and the general public. Researchers look to biobanks to advance their scientific efforts. Pharmaceutical companies are finding biorepositories as tremendous resources in advancing drug discovery and development. Government and other funding organizations look to advance long-term healthcare goals. The general public is beginning to see the potential for disease solutions in their lifetimes. As biobank-enabled medical research yields knowledge, forward momentum is accelerating. Several biobank-enabled research programs were described at Biobank Summit III, including:

- Gene expression in breast cancer (Karolinska Institutet and Genomics Institute of Singapore)
- Origins and outcomes of rheumatoid arthritis (Karolinska Institutet)
- Genetic variation present in CETP (Pfizer)
- Molecular profiling in oncology (Pfizer)

- The PREVENT [prospective randomized evaluation of the vascular effects of Norvasc Trial] study (Pfizer)
- The Genographic project (National Geographic and IBM)
- Atherosclerosis risk in communities (ARIC) study (National Heart, Lung, and Blood Institute [NHLBI], and the National Institutes of Health [NIH])
- Genome-wide mapping of cancer pertinent TFs and NHR (Genomic Institute of Singapore)
- The HapMap Project (International Research Consortium)

New advances in medicine will increasingly require biological samples that specifically provide insights into individual disease expression. As we study diseases, it is becoming increasingly apparent that diseases once considered to be simple results of a genetic flaw (e.g., sickle-cell anemia) are actually much more complex manifestations of impacts on different parts of multiple pathways, resulting in the same effect (e.g. cancer). As we begin to study the molecular basis for disease, there is an increasing need to access significant numbers of disease-specific biospecimens to find direct molecular insights into the underlying biology. Since specific disease biospecimens are relatively rare within most research institutions, the potential for access for uniformly high-quality disease biospecimens from biobank repositories across the globe offers the potential to greatly accelerate disease research.

Delivering Value Today

Through the genetic assessment of both healthy and disease-specific biospecimens obtained from biobanks, the potential for personalized medicine is becoming realized. It has long been recognized that individual genetic variation is likely to be a major factor in disease manifestation, both in terms of predisposition and in reference to therapy. Advances in technology and efficient application of these technologies to both normal and disease biospecimens is beginning to uncover the genetic basis for the differential manifestation of specific diseases. Significant efforts are underway at pharmaceutical companies to identify biomarkers that can explain these differential effects. For some cancer patients, personalized medicine is already a reality. Novartis' leukemia drug, Gleevec, and Genentech's breast cancer drug, Herceptin, are well-known current examples. Continued development of such targeted, effective products will require the availability of significant quantities of high-quality biospecimens from biobanks.

As biological knowledge grows as a result of advances in technology, predictive medicine will become possible. The ability to identify early presymptomatic manifestation of disease offers the potential to exploit current disease-static therapeutic solutions to delay the onset of disease (e.g. Biogen-Idec's Avonex and Serono's Rebif therapeutics for multiple sclerosis). Biorepositories offer unique opportunities to directly assist in the identification of the molecular basis of genetic diseases, especially where biospecimens from related family members and/or longitudinal biospecimens (samples obtained from the same individual over time) are available. As with all research studies, the standardized collection of families of samples and longitudinal biospecimens will be critical to ensuring that results

obtained are representative of the patient's biology and not artifacts of the sampling procedures and sample-handling variations.

Recruiting People, Not Just Collecting Specimens

John D. Potter, MD, Ph.D., Fred Hutchinson Cancer Research Center

Chronic-disease epidemiologic methods grew out of infectious-disease epidemiology, which involved both field and laboratory methods. In diseases where intermediate biology was, for a long time, not observable (particularly cancer), record- and interview-based epidemiology revealed some key exposures (e.g., smoking, radiation). With measurable biologic intermediates (e.g., blood lipids), cardiovascular epidemiology also yielded inferences on causal pathways. Major changes are remaking epidemiology in ways that will ultimately influence all aspects of medical practice. These include, particularly, high-throughput genotyping, allowing genetic and gene-environment causes of disease to be identified; high-throughput proteomics, allowing the development of early-detection markers; new tools for the measurement of exposures, improving precision; and a molecular disease taxonomy. These will allow a much better understanding of both the causes and intermediate stages of disease. However, new methods do not obviate the necessity for good epidemiologic study design; the greatest opportunities to informed medical practice come from the application of new methods to large-scale human cohort studies.

The essential ingredients of the standard epidemiologic cohort study are people and time. Individuals (e.g., from the general population, from a specifically exposed group such as smokers or asbestos workers, cancer survivors, or those at high risk of a specific disease) are recruited; specific data on behavior, exposures, family history, medical history, etc. are collected from participants; blood is collected for a variety of biologic measures; and the participants are followed over time to disease endpoints of interest or death. Inferences are made about possible causal relationships by comparing the rates of disease and mortality in the exposed, or the genetically predisposed, with the rates in the unexposed (or nonpredisposed). The new tools will allow exploration of a wider range of biology in this cohort-study setting and the cohort design itself can be used to expand the range of questions that can be asked. In short, what is essential, in addition to assembling collections of well characterized, but anonymous, tissues in general and specific biobanks, is to recruit a living population laboratory.

Until recently, cohort studies would establish immortalized white cell lines or amplified DNA for etiologic studies (focusing on environmental exposures, genetic variation, and acquired host characteristics) that such a cohort allows. High-throughput genotyping and a good draft of the human genome, in concert with appropriate bioinformatics tools, will ensure that, within the foreseeable future, such a study is not limited in this way. Soon, a genome sequence on each study participant will be available as an electronic datum in exactly the same way epidemiologists store data on smoking, sexual behavior, and family history.

Second (already the norm in some cohort studies), recontact with participants allows updating of, for example, behaviors, medical history, and family history (family history changes with time). More importantly, recontact with participants allows a scheduled collection of further samples: tissues and fluids (e.g., serum, white cell mRNA, urine) that may be used as biologic exposure dosimeters (smoking, diet, etc). Moreover, in nested case-control studies, these biologic samples can provide markers of early detection: sequential blood samples, collected from healthy cohort members, can be interrogated using tools to identify both specific patterns of, say, proteins, at a moment in time and changes in those patterns over time. This approach may identify, at the earliest possible time, otherwise undetectable disorders such as occult cancer, early neurodegenerative disease, or low-grade chronic infection.

Third, cancer, particularly, is classified largely using techniques completely recognizable to a 19th century pathologist. There are already attempts to produce molecular classification systems using cDNA microarrays, proteomics, mutation spectra, and DNA methylation patterns. In the setting of a human population laboratory of the sort under discussion here (with data on exposure, germline genetics, early detection markers), a variety of new opportunities present themselves, allowing connections to be made among the various classes of data in hand (see below) and appropriate inferences to be made about much more homogenous subsets of disease.

Recruiting People, Not Just Collecting Specimens (Continued)

Fourth, in this population-laboratory setting, there is an opportunity to breach the artificial barrier between etiologic epidemiology, where follow-up of the study participant is usually censored at the time of diagnosis, and clinical epidemiology, which often begins with the assembling of a cohort of individuals who have been recently diagnosed with the disease of interest. If cohort members are followed through (molecular) diagnosis and treatment, and treatment data added to the database, it becomes possible to link therapeutic regimen to molecular diagnosis and to outcome (e.g. therapeutic response, survival). These data can then inform the design and conduct of relatively small independent clinical trials, focused on specific homogenous molecularly defined disease entities. Initiation of new clinical trials will make it more likely that agents are effective, but only with a subset of the cases that would normally be diagnosed using current conventions (e.g., Iressa, aka Gefitinib). Finally, continued observation of the population laboratory and further characterization of outcomes — imaging, molecular studies, clinical observation — will add to knowledge of progression, treatment response, and survival.

Thus, making the best use of the new biologic tools (e.g., those that are mature, like high-throughput genomics, and those that are still in development, such as high-throughput proteomics and new imaging agents) as well as enlarging the scope of the standard etiologic cohort study to engage with its clinical counterpart, can ensure the establishment of one or more cohorts as population laboratories. Ideally, these laboratories will use standardized instruments and protocols and be based in settings (geographic, cultural) with contrasting environmental exposures and genetic backgrounds. Such collaboration allows comparison of findings, between centers, on common outcomes and accumulation of cases of rarer ones across centers.

Population laboratories (recruiting people, not just collecting specimens) will allow standard epidemiologic questions to be asked (what is the relationship between a particular exposure or a particular genotype and disease outcome?) More critically, they will also allow the relationship of genes/allelic variants and exposures to specific molecularly defined disease to be identified with greater precision. Further, they will help develop, or confirm the utility of, plasma protein profiles as markers of early detection. Finally, molecularly defined disease entities and treatment data, in conjunction with pharmacogenomics to account for variation in drug metabolism, will point toward subsets of disease against defined genetic backgrounds that are particularly successfully treated with specific agents.

The Role of the Public

The general public has become increasingly savvy regarding diseases. This is especially true with disease-specific special interest groups (SIGs), where most members possess a personal interest in finding disease interventions (i.e. family members or friends have the disease). While some disease-specific SIGs are relatively small, many are large, global groups with thousands of interested members. Most of these types of SIGs are medically well read and are aware of the ongoing research in the area. As individuals with a vested interest in finding disease cures, most of these groups recognize the value of biospecimens in enabling disease solutions and often actively solicit the public to contribute biospecimens in hope of accelerating progress or otherwise facilitating solution development.

Along with the recognized potential for genetic testing to predict an individual's susceptibility to a given disease, the general public has recognized the potential for abuse by employers, managed care organizations, and insurance providers. Concerns over policy rejections or exclusions for preexisting disease conditions (e.g., HIV, HPV) and extending to the "potential" for the disease sometime in the future (e.g., family histories of heart attacks or diabetes) are real. Governments and other regulatory bodies are developing laws to prevent abuse. As these laws are put in place, people are beginning to allay their concerns over improper use of biodata

generated from biospecimen donations. These controls are expected to result in increased donations to biorepositories and the potential for accelerated research. As a caution, however, recent studies show that the public still does not yet fully trust government or industry in this area and that their attitude has not changed much over the last year.

An Emerging Imperfect Resource

Core to the value of biobank repositories are the patient, clinical, and molecular data that accompany the biospecimen. This information is critical to maximizing value from both healthy and diseased tissues. The wealth of information associated with research samples can allow researchers to apply genetic relevance, correlate with specific environmental factors, and otherwise gain the "big picture" relevance of the biological material being evaluated. In the absence of this associated data, as much as 90% of the relevant value may be missing, severely limiting the amount of useful information that can be derived from the biospecimen.

In many of the situations where disease tissues are generated (and especially within the surgical suite), biospecimens are an afterthought of the procedure. The physician's primary responsibilities are focused on achieving a successful outcome for the patient. Several different fates can await biospecimens, including:

- Excised disease tissues may lie in the physician's office or surgical suite for extended periods of time at room temperature.
- Samples going to different laboratories for analysis may be labeled using different naming conventions, preventing later consolidation.
- Samples may be transferred to the pathology laboratory for analysis (again at room temperature) for extended periods and only then stored as an afterthought. During these prolonged periods, biospecimens will degrade in a number of ways, rendering them of limited value for future analyses.
- Data associated with the sample (including handling conditions, excision data, etc.) are often not captured.

Subsequent analysis of samples can produce results that do not reflect the condition of the original sample and may provide the researcher with false data. False data then leads to incorrect conclusions and delays accurate understanding of the disease impact. In some cases, reported data produce results that are impossible *in vivo.* As standards continue to be established by the biobanking community, recognition of the potential impact of improperly handled samples is increasing, and procedures to ensure the proper handling of samples are being put in place.

Unless testing of biospecimens has been performed exclusively on samples entirely under the control of the researcher, it may be possible that samples have degraded or otherwise do not reflect the original condition of the sample immediately after it was excised from the patient. Under these circumstances, it is likely that the experimental data obtained from these samples may be flawed. Historically, specific lack of quality assurance procedures have resulted in improperly prepared samples, which have been used to generate potentially erroneous results. As a problem that is only

recently receiving expanded attention, specific discussions are needed to determine whether it will be possible to filter out bad data. It is clear, however, that some data collected are completely unaffected by storage or handling conditions (e.g., genomic or SNP data) while other data (e.g., mRNA or transient metabolite levels) are highly affected. As described previously, the standard procedures anticipated for biobank repositories are expected to minimize the potential handling and degradation issues, especially when awareness of specific targeted measurements is considered at the time of collection.

New and Improved Testing Capabilities: Technology Improvements Pacing Biorepository Development

The fundamental purpose for the existence of biobank repositories is to provide medical research with material and data that will allow researchers to understand the workings of biological systems, with particular interest in understanding the disease; the impacts of genetic, nutrition, environmental, and lifestyle factors on disease; and potential interventions that can result in cures.

What Biorepositories Offer Research

From a practical and ideal perspective, biobanks offer a resource of biological samples that are:

- Available in limited quantity
- Potentially part of a collection of different tissues and fluids cocollected from the patient, including possible longitudinal samples
- Collected under standard conditions designed to retain maximum biological fidelity for specific analytical testing
- Fully annotated with associated and relevant patient, laboratory, and disease data, but stripped of identifiers that could be misused
- Released by the patient for broad research purposes

Recognizing the limited availability of important biospecimens, it is important to take care to ensure that the resource is not wasted. With numerous opportunities to consume biospecimens in creation of research data, many of the advances in technology are aimed at more efficient sample use while generating ever-increasing amounts of data. In addition, new technologies are improving sample throughput to reduce data acquisition time, which becomes especially important in assessing multiple biospecimens in a timely fashion.

Several examples of applications where improvements in sample utilization and throughput is being applied to biobank samples are provided below:

- Breast cancer studies using the Swedish cancer registry in a collaborative effort between the Karolinska Institutet and Genomics Institute of Singapore

- RA studies at the Karolinska Institutet using twins data, including antibody levels and the impact of smoking

- Nested case setting used for sophisticated analysis of potential effects of drugs on rare events at the Karolinska Institutet

- Rapid sequencing technologies that allow for characterization of cancer genome variations based on work at the Genomics Institute of Singapore

New Technologies are Coming

The storing of biospecimens in biobank repositories offers the potential for future testing using powerful analytical research tools that may not have been available when the samples were first collected. While some caution must be exercised to ensure that the data being collected accurately reflect the presence and levels of targeted analytes preserved in the biospecimen, the biobank offers a unique, high-value resource capable of gleaning new knowledge from archived biospecimens and directly leveraging prior research efforts. Several examples highlight the potential research opportunities for biobank specimens, including:

- Anticipated development of high throughput analytical technologies, including rapid sequencing technologies leading ultimately to the "$1,000 genome" and whole proteome proteomic analysis methods to fully characterize biospecimens in a timely and cost-effective manner.

- Future translational research allowing for development of new research data from biospecimens with known clinical outcomes.

As an example of how current technologies have successfully added to completed collections, retrospective analysis of atherosclerosis risk In communities (ARIC) biospecimens collected between 1987 and 1989 has allowed for the supplementing of original research with new information, with continued efforts planned through 2007. Many of the analyses of ARIC biospecimens have been performed using technologies that were not available when the samples were collected (e.g., comprehensive genetic analysis methods, including SNP analysis). As a result, the ARIC study provides a definitive example of how biobank repositories can expand the scope and value of research programs by continuing to provide new results and insights over time that could not have been predicted during initial research study planning.

Recognized Business Imperatives

The pharmaceutical industry has recognized that biobank repositories will be a critical resource in the development of new innovative therapeutic solutions. In addition to providing biospecimens for pharmaceutical testing, significant added value is brought by biobanks to the effort in the form of standardized sample collection, informed consent, associated phenotypic data, and correlation of phenotype and genotype relationships.

The Need for Standardization in Biobanking — Focusing on Preanalytical Steps for Molecular Analysis

Simone Gauch, Qiagen GmbH

Biobanks are key to the future advancement of personalized medicine. However, despite their great opportunities, there are a number of diverse challenges, including technical, ethical, and financial that must be overcome.

Even though it is important to look at the "big picture" of strategies and open questions in biobanking, there is also the need to focus on specific challenges to reduce complexity. This will show which questions can already be answered today with practical solutions, and from there identify the path for future progress. Here we describe how this approach can be applied to the preanalytical biobanking steps for molecular analysis.

Progress to address the necessity to standardize biospecimen collection, storage, and processing for biobanking has been made in a number of areas. For example, comprehensive best practice guidelines for biospecimen management, including the critical topic of preanalytical steps, are currently being produced by the National Cancer Institute (NCI) and the International Society for Biological and Environmental Repositories (ISBER). In addition, initiatives for the standardization of technologies — such as the External RNA Control Consortium (ERCC), which produces external RNA controls for microarray platforms and RT-PCR — have been established by the public and private sectors. Meanwhile new enabling technologies have greatly facilitated sample management by offering integrated solutions such as blood collection tubes containing RNA stabilizing agents that immediately freeze the gene expression profile and whole genome amplification technologies that enable unlimited DNA amplification from precious samples.

Although significant progress has been made, a number of questions still need to be addressed in order to take the standardization of biospecimen management to the next level. For certain issues (e.g., the impact of ischemia on expression profiling results), there is still not enough scientific data available to derive best practice guidelines, and more research is required in order to address these open questions. Other issues (e.g., the fixation of tissue samples for parallel analysis of RNA and proteins) still represent technical challenges, and better tools are required. Also, an increase in the availability and use of automated systems for sample management would significantly streamline workflows, ensure reproducibility, and facilitate documentation.

In some instances, even though best practice is known, it cannot always be incorporated into the daily workflow of a biobank since patient management has highest priority. Therefore, it is also desirable to have guidelines on what to do if optimal conditions cannot be maintained due to practical constraints and the potential use of suboptimally managed samples. Therefore, more information regarding the success and failure of molecular analysis techniques when using suboptimally managed samples is required. Here the challenge is that unsuccessful experimentation is typically not reported.

In order to take the next steps, public sector funding is necessary to address the open scientific questions around sample management. Forums within the biobanking community for open sharing of successful and unsuccessful techniques and results, including unpublished data, should be encouraged. There is also a need for more advanced tools in order to overcome current technical challenges. In addition, more tools are required to assess sample quality, enabling accurate predictions for the success or failure of downstream molecular analysis. For these developments to take place, close interaction between the public sector and companies dedicated to offering solutions for sample management, is crucial.

We envision the ultimate goal to be a searchable, Web-based central repository of guidelines, SOPs, tools, and references for molecular analysis of biospecimen. While we are still facing a number of challenges toward this goal, we should not forget that we have already taken many positive steps toward the standardization of biobanking.

Biobank repositories are a major enabling factor in advanced pharmaceutical research. At Pfizer, biorepositories are critical to the company's focused research efforts in the areas of pharmacogenomics and biomarker research. As part of their discovery strategy, Pfizer is using DNA from biorepositories to determine how genetic variation impacts disease (e.g., by identifying all common CETP gene variants, they were able to identify associations with HDL cholesterol and a reduced risk for CHD).

From a business perspective, increased knowledge (especially from biomarker research) enabled by biorepositories has allowed pharmaceutical companies to terminate unproductive research early (saving significant resources), add confidence to promising efforts, aid in dose selection, and provide the basis for product differentiation and future research (ideally resulting in new indications).

As a success story, Pfizer's Prospective Randomized Evaluation of the Vascular Effects of Norvasc® Trial (PREVENT) study of 825 patients with angiographic evidence of coronary artery disease produced more than 3,000 biospecimens that were used to identify nine potentially valuable biomarkers. These biomarkers are undergoing further validation to determine their value in supporting Pfizer's research programs. As a caution, Pfizer's original expectations for the rapid development of validated biomarkers (i.e., several validated biomarkers within the initial 3- to 4-year period) has not been realized. However, the difficulty can likely be attributed to the complexity of disease causality and is unrelated to the contributions from biobank specimens.

Standardization: Moving from Samples to Quality Biorepository Biospecimens

"Big Picture" Benefits of Standardization

Biological discovery research requires significant quantities of biospecimens to ensure the statistical validity of research experimentation. With the exception of studies where all biospecimens have been directly collected by the researcher, samples will come from multiple sources with varying degrees of quality assurance. Associated with these multisourced samples comes the potential for new sources of experimental error that can greatly impact the researcher's ability to extract valid results and draw appropriate conclusions. The move toward uniformity in sample collection and annotation will allow for larger high-quality data sets to be compiled across organizational boundaries, enabling research at the scale needed for effective biological discovery.

The adoption of standard validated procedures across the scientific and business communities is an important step toward enabling uniformity and transparency between organizations. Within the biobanking community, the establishment of standard validated procedures is a critical early requirement to ensure that samples obtained from different biorepositories are comparable. These high-quality samples provide the basis for productive discovery research without compromise.

In addition to experimental data derived from biospecimen testing, it is important that other biospecimen-associated information (e.g., patient data, disease information, collection and storage conditions, etc.) be collected using standard processes and terminologies to ensure maximum relevance of the data. This is important not only in the short term for research studies already anticipated but is especially important for unplanned future studies (see the ARIC study described previously). As new analytical technologies and analysis tools (e.g., systems biology modeling and simulation) are developed, associated data provide valuable context and relationships that will contribute significant value above and beyond the experimental data provided by the biospecimen alone.

Because biospecimens are fundamentally the property of the individual from whom they were obtained, the ability to study and analyze individual samples is limited by the informed consent conditions approved by the donor. At present, the range of informed consent is highly variable and may be open to all scientific research or only for specific targeted studies. Recognizing the rights of the donor as well as the research opportunities (both those currently recognized and those yet to be determined), the establishment of standardized informed consent options will be important to ensure that the broadest opportunities to gain knowledge from biospecimens will be possible.

Standard Methods Enabling Quality Results

The preservation of biospecimens is not as simple as aliquoting materials into cryotubes and placing them in a -80°C freezer. While DNA would be effectively preserved under these and less-severe conditions, proteins, mRNA, and small molecule metabolites may continue to degrade over time or may be directly affected by the freezing conditions. Because biological samples are complex (and potentially self-degrading), it is important to anticipate the types of testing that are to be performed on biospecimen samples.

Basic research has well characterized the appropriate storage conditions required for most, if not all, of the different types of biological materials that can be present in biospecimens. In cases, where specific research requirements are known, it is easy to identify proper collection and storage conditions for specific biospecimens. Alternatively, where biospecimens are broadly collected with the potential for future testing of unidentified biological analytes, multiple storage conditions may be required. Based on the major categories of biological analytes of interest (i.e., DNA, RNA, proteins, lipids, small molecule metabolites, etc.), identification of standard storage options should be possible, allowing for maximum flexibility for current and future testing.

In addition to standardization of storage conditions, standardization of extraction and testing procedures will also be needed, based on quality assurance, and data documentation requirements. Identified areas requiring specific standards processes and procedures include, genetics (especially DNA), gene expression, metabolomics, lipidomics, and proteomics. These category-specific standards recognize the inherent differences among categories, including analyte stability, nomenclature, and associated materials (e.g., primers for DNA testing, probes and splice variants for gene expression, and protein isoforms for proteomics).

The ability to differentiate specific variations caused by specific disease conditions requires comparative analysis with appropriate normal reference samples to ensure that the variations are really linked to the disease. As part of their general approach, biobank repositories are a significant resource in providing standardized normal reference samples. With standardized collection procedures ensuring comparability between different samples and appropriate relevance to disease tissues, biobank repositories can ensure that comparative analysis is possible. For example, in a simple comparison of cancerous and noncancerous tissues, other potential biases (e.g., age, sex, fasting versus nonfasting, etc.) may be just as important in analytical testing, resulting in potentially incorrect conclusions.

Defining Standards: Precedents and Hurdles

There is general consensus that standards are required to ensure that quality, comparability, interoperability, and communication are possible among biobanks, research institutions (both academic and industrial), and individual researchers. Unfortunately, in many cases, this is where the consensus ends. Different interest groups, competing standards organizations, professional reputations, and divergent goals all offer standards that are often irreconcilable with each other, based on different expectations and goals. In many cases, it may be possible to resolve such conflicts by establishing standards-setting authority at the governmental level supported by financial incentives. However, as biobank efforts become increasingly global, regulatory authority becomes unclear. Several options for establishment of standards are currently competing in the community and a unified vision has yet to be determined. These options include:

- As one of the largest global biobanking efforts, Sweden and specifically, the Karolinska Institutet, with its Biobank Information Management System (BIMS) offers a *de facto* standard in moving toward sample information conventions and standardization.

- The P3G Consortium has as its primary focus the harmonization of biobanks and population studies. In collaboration with international working groups, P3G is working to establish commonality among existing cardiovascular research databases, including identification of comparable variables and establishing the basis for inter-biobank correspondence criteria. Moving forward, P3G is also applying its approach to the collection of new data with the addition of a harmonized patient questionnaire to develop standardized associated data for collected biospecimens.

Standards for Population-Based Biobanks
Mylène Deschênes, BCL, LLB, LLM, P3G Consortium

The series of World Wide Biobank Summits has clearly demonstrated the shared frustration of the scientific community caused by the inability to communicate between different biobanks. Data and samples are not comparable, samples are collected and stored under different conditions, the datasets that emerge cannot be validated, and mostly, we simply don't have an inventory of the various biobanks.

The need for longitudinal studies and high-quality biobank resources for research has never been more critical. To unravel the complex genetic and environmental interactions responsible for most common diseases, we need large numbers of samples, and high-quality biobanks and datasets that testify to the complexity of human lives, including ongoing socio-demographic data. These types of biobanks take time to create and require significant financial investments.

To that end, several large-scale epidemiological research projects and resources have been created around the world and their number is likely to grow over the next decade. There is a need to maximize the potential benefits of these critical investments in public health and research. While it is important that each research resource keeps its own specificity, we need to think, at an international level, about the development of strategies (such as harmonization), that will ensure the highest standards of quality. These standards will pave the way for meaningful and sustainable international collaborations and are pivotal to understanding key issues involving the cause of human disease, and effective strategies for ensuring the health of our populations.

The Public Population Project in Genomics (P3G) is an international organization dedicated to the development, creation, and sharing of common research tools for the constitution and management of population biobanks so as to prospectively enable effective collaboration and ensure knowledge transfer for the benefit of the public.

P3G has created three "international working groups" to meet the scientific, information technology, and ethical challenges raised by the creation and management of population biobanks, in an environment of collaboration. As their first mandate, these groups are working on the identification of a minimum dataset to be collected by biobanks and the identification of key ethical barriers to international collaboration. We have also launched a P3G Internet Observatory that will provide free access to a standardized description of major research biobanks from around the world.

So far, P3G has gathered nine major international biobanks from Europe, North America, and Australia and other experts in the field of population genomics. With their collective expertise, we will develop common research tools and offer these tools to the international research community through the P3G Observatory.

Maturing of Biobank Infrastructure

As existing biobanks continue to grow and new biobanks are established, the opportunity to leverage positive experiences gained in successful biobank efforts and learn from negative experiences encountered in less successful efforts grows. In addition, collaborative discussions from conferences, including the prior three World Wide Biobank Summits, highlight best practices that can be shared to provide the basis for growth of both existing and new biobank initiatives.

Unambiguous Success

Sweden and the Karolinska Institutet have taken a major role in establishing a solid foundation for biobanking. With its eight different nationwide registers, including major cancer and twin registries, and between 50 and 100 million samples, Sweden has developed strong momentum in advancing the aspirations of biobanking. In support of their extensive biobank collections, Sweden has also made strong efforts to create

and develop the necessary infrastructure and standards to ensure that it will be able to access and fully benefit from its investments over the long term.

While not necessarily the best option for other biobank efforts, Sweden has advanced specific approaches to address several biobank infrastructure issues. Specific approaches to infrastructure issues include:

- Recognizing the economies of scale, Sweden centralized many of its efforts to provide critical mass and to take advantage of economic and other efficiencies.
- Established in the 1960s, the Swedish twin registry is a strong example of effective information and biospecimen collection procedures.
- To optimize access, sample handling processes have been standardized to allow for efficient retrieval of biospecimens.
- In recognizing the importance of patient privacy rights with regard to its biospecimens, the Swedish Biobank law requires strict informed consent approval from patients.
- The Karolinska Institutet Biobank has developed the BIMS (Biobank Information Management System, mentioned previously) as part of its biobank informatics efforts to provide the Swedish biobanking system with informatics infrastructure allowing for key access to sample supporting information.
- Sweden's LifeGene population-based project provides a knowledge base of routine, baseline cohort, fully annotated with Swedish health registries.
- Sweden has taken efforts to harmonize with other biobank collections from the start, including the Singapore and U.K. biobank efforts.
- Recognizing that biobanking efforts are costly, Sweden has accepted the fact that anticipated returns from biobanking investments will be realized over decades.

A number of successful biobank efforts have also been achieved outside of Sweden. Canada's iCapture Center, the Estonian Genome project, the Genome Austria Biobank, the GenomeEUtwin project, deCode Genetics' commercial effort in Iceland, the Taiwan Biobank, the U.K. Biobank, the WA Genetic Epidemiology Resource in Australia, the U.S. National Children's Study, the U.S. Black Women's Cohort, the California Teacher's Study, and the ARIC study are all successful biobanking efforts, and new efforts from Japan and the Danubian Biobank appear promising.

Commercial Success

While the knowledge gained through using biobank specimens to understand the molecular basis for disease is inherently valuable, the benefits to society are best captured through the use of knowledge to develop practical medical interventions. In most cases, these interventions will require commercial pharmaceutical development. The pharmaceutical industry has recognized the value that biobanks offer and are actively working with biobanks to accelerate both medical research and direct

development of new therapeutic solutions. In addition, pharmaceutical companies are developing their own biorepositories using biospecimens produced during company-sponsored clinical trials. Significant experimentation with biobank specimens has focused on their use to help identify biomarkers related to either drug efficacy or drug safety. In addition, biomarker research promises to improve the efficiency of drug development and enable pharmacogenomics as a new medical approach.

Not Everyone Is On the Same Page

While the general arguments for biobanking appear compelling, specific projects to establish and develop biobanks have encountered hurdles and pitfalls that threaten to derail promising efforts. These efforts include:

- The UK National Translational Cancer Research Network (NTRAC) effort to develop a national cancer tissue resource (NCTR) was initially funded in 2003. Advancing what was perceived to be an effective approach to biobank development, the NCTR encountered significant difficulties in obtaining cancer samples. A number of issues were identified including societal acceptance questions, the appearance of business conflicts of interest, and logistical problems. In an effort to restart the NCTR effort after having addressed the key problems, onCore UK was formed in May 2005. This emerging effort is expected to quickly determine whether the hurdles to success have been overcome.

- The U.S. National Biospecimen Network (NBN) Blueprint is an extensive dialog on the pathway to developing a national biospecimen repository with the ultimate purpose of accelerating scientific discovery in the battle against cancer. The Blueprint addresses many of the fundamental issues associated with biospecimens, including patient privacy and informed consent, collection, and testing standards, and biospecimen associated data. While inputs from most involved constituencies were considered, it is yet to be determined whether any of the differing viewpoints will derail the effort. As with all large consensus-driven efforts, differing opinions as to the balance between patient-specific information and personal privacy issues may limit availability of some data for scientific research.

Proposed and new biobank initiatives should carefully consider prior historical biobanking experiences to ensure that these hurdles and pitfalls do not derail efforts or otherwise impact progress.

The NCI Views Biorepositories as a Critical Resource for Translational Research and Molecular Medicine

Carolyn Compton, MD, Ph.D., National Cancer Institute, USA

The increased use of high-throughput methodologies for molecular analysis (e.g., genomic and proteomic technologies) in translational research is focusing new attention on biorepositories with collections of human biospecimens that are annotated with pre-analytical parameters, molecular, demographic, and clinical information. Over the past few years, several biorepositories have been established domestically and internationally to support these new research applications. With the sequencing of the human genome, biorepositories are providing a critical platform for researching the genetic basis of disease and enabling the identification of new biomarkers to predict and monitor drug response at the level of a single individual, the vision of personalized medicine.

The success of the nationwide and international biomedical research communities supported by the National Cancer Institute (NCI) in developing the molecular medicine of the future is increasingly dependent on high-quality, well-annotated biospecimens. The NCI recognizes that there is a growing need for harmonized approaches and data-driven operating procedures for biorepositories to meet the needs and increasing rigor of translational research. Technologies that allow scientists to create, share, and aggregate genomic and proteomic information on an unprecedented scale are generating new demands to assemble powerful databases that facilitate the comparison of research results derived from biospecimens collected at different institutions. To enable resource sharing and team science through single-investigator studies and multi-institutional, high-throughput analyses, the NCI has undertaken a series of initiatives across the cancer research enterprise that will ultimately facilitate the harmonization and perhaps the eventual virtual integration of NCI-supported biorepositories.

The NCI has been carrying out an integrated effort across all its divisions to identify and implement quality practices in NCI-funded biorepositories that continues to date and is to be further developed through the newly established Office of Biorepositories and Biospecimen Research. This effort began with collaborative work on the National Biospecimen Network (NBN) Blueprint, a futuristic vision for a nationwide biorepository management system to support genomic and proteomic research. The requirements outlined for this system emphasized the use of best practices-based standard operating procedures (SOPs) for the collection, processing, storage, annotation, and distribution of biospecimens; harmonization of informed consent procedures; and development of a common informatics platform. The NBN Blueprint also detailed the importance of upholding the highest ethical and legal standards by establishing a "chain of trust" beginning with the patient who donates a biospecimen and extending throughout the network to each scientist who uses the biospecimen for research purposes.

IT Infrastructure Maturing

The information technology infrastructure supporting biobanking has been able to successfully leverage progress in the broader information management systems marketplace. Information technology systems are being called upon to link physical biospecimens with the large databases containing patient, sample, molecular, and clinical data. In addition, the systems must provide connectivity with software query and analysis programs to allow researchers to access both biospecimens and associated data. Both commercial and custom solutions have been developed to address these needs. Examples of major biobank IT management solutions are provided below:

- **The Ardais System.** Ardais Corporation, with its exclusive focus on biospecimen management, has developed a comprehensive Web-based system that provides a complete suite of biobanking software applications.

- **LabVantage Sapphire Biobanking Solution.** LabVantage Solutions Inc. has customized its Sapphire core technology platform to provide a browser-based biobanking module that, like Ardais, comprehensively addresses all fundamental biobanking requirements. In addition, in combination with LabVantage's Sapphire LIMS product, seamless access to molecular and clinical data is possible.

- **Karolinska Institutet's BIMS system.** The Karolinska Institutet has developed a custom Biobank Information Management System (BIMS) to manage and access biospecimen samples and data. The BIMS system was developed using IBM's WebSphere platform as the foundation together with the federated database Information Integrator (formerly known as Discovery Link) as the data handling backbone. In combination with Karolinska Institutet's LIMS systems and information databases, the BIMS system provides a mature IT infrastructure to support the Karolinska Institutet Biobank.

The Importance of Informatics in the Prevention of Bias in Tissue Resources

William E. Grizzle, M.D., Ph.D., University of Alabama at Birmingham

The collection, processing and storage of human tissues to support biomedical research requires great attention to details as well as the development and use of standard operating procedures for each step in the process from collection to storage of the tissues and ultimately through their analysis. Without attention to such details, biases may be introduced into tissue collections. For example, at one site, "A," tissues may be frozen and stored within 1 hour of removal from patients, while at a different site, "B," tissues may be frozen and stored within 24 hours after removal from patients. If most controls (e.g., patients without prostate cancer) are obtained from site "A" while most cases (e.g. patients with prostate cancer) are obtained from site "B," the differences between the specimens may not be due to the presence or absence of prostate cancer but rather to the time between collection and freezing of samples. Other potential sources of bias include patient demographics, comorbid conditions in patients (e.g. adult type diabetes), patient homeostasis (e.g., fed versus fasting), collection vagaries (e.g. type of collection container – red top versus tiger top for collection of serum), processing differences (e.g. time, temperature, and methods), and time and temperature of storage as well as freeze-thaw cycles of specimens. Many of the above factors may be beyond the control or knowledge of the tissue collection site. Whether such factors can be controlled or not, the factors should be documented, if known, to aid in the prevention of bias or the identification of potential sources of bias. An informatics system which includes fields to record important features which might affect the molecular features, quality or usefulness of solid or fluid tissues is very important in efforts to prevent or to identify sample bias.

Thus, in the design of the informatics system of a tissue resource, care should be taken to include fields that are needed to identify potential parameters which affect or may affect the quality of samples of tissue. Most important is to perform and to record a quality control examination of the specific samples provided or stored for research; thus, the exact diagnosis of the specific tissue to be used in research must be known and recorded in the database. Similarly, the times of collection and of processing should be included in the database as well as the demographics and co-morbid conditions of patients from whom tissue samples are obtained. Of special importance would be to record any changes and dates of changes to standard operating procedures.

When new issues are raised that may affect samples, such as time of warm ischemia (i.e., the time between vascular compromise and removal of a tissue from a patient) or freeze-thaw cycles of samples during aliquoting, the database may need to be modified to include new fields. Alternatively, a database may include enough "comment" fields so that additional fields are not necessary when new features are identified as being important in identifying the condition of human tissues used for research.

SUMMARY: RECAP OF ISSUES

The stated goals for Biobank Summit III were:

- Enable and facilitate many different international stakeholders to share best practices and discuss possible resolutions to the scientific, technology, and policy challenges facing the global biobanking community
- Raise the visibility and importance of biobanking to realize the vision of information-based medicine
- Facilitate and encourage collaboration and cooperation within the biobanking community and begin to address challenges raised at the second summit
- Continue the dialog started at the first two summits through facilitated roundtable discussions and "deep dive" sessions

These goals were achieved through discussion of specific progress in four areas. These areas were:

- Increased recognition of how biorepositories can build value and advance the goal of information-based medicine
- How advances in technology are supporting biobank specimen-based research
- The critical need for standardization in biobanking as the foundation for quality science, interoperability, and future growth
- Maturation of biobanking infrastructure as models for best practices

Key takeaways from the 2005 Worldwide Biobank Summit III were:

- Standardization, standardization, standardization
- Biorepositories with high-quality biospecimens and their associated data are needed.
- Biobank specimens are being used in research now to advance information-based medicine.
- Don't "reinvent the wheel" — we must learn and build off the body of knowledge the biobanking community has built.
- Best practice-based standard operating procedures (SOPs) are needed for:
  - Biospecimen collection, storage, and distribution
  - Data collection
  - Ethical, legal, and social issues

- Technology is advancing independent of progress in biobanking. Current and new technology applications are making extraction of knowledge from biobank specimens increasingly easier.
- Communication with all stakeholders is key — leadership must define and frame the debate.
- The biobanking community needs clear leadership — where will it come from? - the biobanks themselves, government, or industry?
- Now is the time to move forward. Discussions are not enough.

LIFE SCIENCE INSIGHTS PERSPECTIVE

Brock Reeve

It is clear that biobanking is here to stay and that biorepositories have a tremendous opportunity to develop resources that have a significant potential to accelerate research efforts to develop therapies for major diseases. It is also obvious that the biobanking community has much work to do to achieve this potential. Significant hurdles to progress exist, both internal and external. Addressing these issues will require unprecedented levels of cooperation and collaboration as well as collective focus on achieving targeted goals.

Some hurdles, including naming conventions, agreement on standards, ontologies, and rules on informed consent, must be addressed by all parties. Other hurdles apply only to larger efforts and include broad procedural standards and conventions for collaboration. Even in the absence of collaboration, current biobanking efforts can be expected to succeed through incremental achievement of goals. With every small step, progress is made, and some movement toward information-based medicine is achieved.

At the same time, technology continues to advance, especially in the area of genomic analysis and understanding. These (and other) advances can be and will be applied to biobanking, but will occur regardless of what progress is made by the biobanking community. The challenge will be for the biobanks to make sure they are able to take advantage of, contribute to, and build on this progress.

While some might disagree, despite discussions of national and international networks, the goal of the biobanking community should not be the formation of a global repository for all biobank specimens. This is not the goal of most current biobanks and should not be. Many biobank initiatives, recognizing practical limitations, have more targeted goals, limited resources, and focused expectations. Such biobanks will be successful based on their own goals and expectations, having achieved what they initially set out to do. Everything else is a bonus.

That said, such a bonus could be big since there is considerable value in biobanks being able to collaborate and share data, scientific learnings, and even specimens. Accomplishing that will require continued work on developing standards (scientific, clinical, and IT) and harmonizing practices such as these summits have begun and some of the national networks are trying to encourage and put into place.

Copyright Notice

G299-0806-00