Published in final edited form as:
Syst Biol Reprod Med. 2011 October ; 57(5): . doi:10.3109/19396368.2011.604818.

Proactively Establishing a Biologic Specimens Repository for Large Clinical Trials: An Idea Whose Time has Come

Peter R. Casson, **Stephen A. Krawetz**, **Michael P. Diamond**, **Heping Zhang**, **Richard S. Legro**, **William D. Schlaff**, **Christos Coutifaris**, **Robert G. Brzyski**, **Gregory M. Christman**, and **Nanette Santoro on behalf of Esther Eisenberg for the Reproductive Medicine Network**
Department of Medicine and Obstetrics and Gynecology, University of Vermont, Burlington, VT (Casson); Department of Obstetrics and Gynecology, Wayne State University, Detroit, MI (Diamond; Krawetz); Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT (H Zhang); Department of Obstetrics and Gynecology, Department of Obstetrics and Gynecology, Pennsylvania State University, Hershey, PA (Legro); Department of Obstetrics and Gynecology, University of Colorado, Denver (Schlaff and Santoro), University of Pennsylvania, Philadelphia, PA (Coutifaris); Department of Obstetrics and Gynecology, The University of Texas Health Science Center at San Antonio (Brzyski); Department of Obstetrics and Gynecology, Reproductive Sciences Program, University of Michigan, Ann Arbor, Michigan (Christman); Reproductive Sciences Branch, National Institute of Child Health and Human Development, Bethesda, MD (Eisenberg).

Abstract

Large randomized clinical trials are becoming more costly, and resources to support them increasingly scarce. Biologic materials, such as blood, DNA, body fluids, or tissue samples collected and stored as a component of these studies represent an invaluable resource, to answer immediate secondary hypotheses, but also as archived material, linked to the study data, for the use of investigators long into the future. The regulatory climate surrounding the storage and future unconstrained utilization of biologic materials is evolving quickly. It is no longer acceptable simply to store samples and use them in an unbridled and unregulated fashion. Thus, to fully utilize the tremendous potential of biologic samples generated from large clinical trials and their related databases, investigators should consider proactively creating a biologic specimen repository, or biorepository. A repository likely assures appropriate subject consent, sample provenance, secure storage, and codified procedures for sample and data retrieval and sharing that protect the subject's confidentiality, the investigator's need for accurate data, and the limited resource. Importantly, the biorepository specimens/samples are typically collected in addition to local and core specimens obtained for the parent study that provide baseline assessments for safety and efficacy outcomes.

Keywords

Biorepository; Repository; Reproductive Medicine Network

Corresponding author:, Peter R Casson, Professor and Director, Division of Reproductive Endocrinology and Infertility, Department of Obstetrics, Gynecology, and Reproductive Sciences, University of Vermont, Smith 401, MCHV, Fletcher Allen health Care, 111 Colchester Avenue, Burlington, VT 05401, Fax: (802)-847-9243, Peter.casson@vtmednet.org.

Introduction

This paper outlines the experience of the NICHD-funded Reproductive Medicine Network (RMN) in developing a biorepository and the associated regulatory infrastructure, with the aim of amplifying the scientific impact of the network work product. It is expected that the long term storage of well-characterized specimens will facilitate full scientific exploration of RMN-generated hypotheses by both internal and external investigators. In the case of the RMN, 20 years, 5 iterations and 5 randomized clinical trials have led to a disparate set of samples and data at multiple centers, collected under widely differing conditions. As the findings of the RMN studies became known, increasing numbers of investigators, both internal and external to the network, approached the RMN with worthy requests for access to stored data and materials. Four such requests have occurred in the last two years. A need was thus seen to codify the process of sample and data acquisition into a repository. An accompanying paper (Krawetz et al, 2010) documents the technical details of DNA, tissue and blood storage of the RMN repository. Prospective consideration of the issues outlined in both these papers has led to the establishment of a well-functioning biospecimen repository for the RMN, and it is the hope of the RMN investigators to provide a resource for colleagues who might have an interest in setting up repositories, and to encourage the practice of standardized sample collection, storage, and sample dissemination.

The Rationale for and Utility of a Repository

Simply stated, the rationale for establishing a repository is to add value to the parent trials by banking tissue, serum, and DNA, long after such trials have been completed, in order to answer both secondary hypotheses that may arise and also other related scientific questions raised as the results become disseminated. As such, the repository acts to amplify the scientific impact of the parent studies. In order to serve such a function over a substantial period of time, in the face of a constantly evolving regulatory and cultural climate, a robust set of policies and procedures must be prospectively instituted that codify and regulate access to the samples and related data, and assure maintenance of full informed consent for the participating subjects that will survive scrutiny for years to come. Central to this informed consent are the institution of systems that assure both the subjects and future IRB committees that adequate and complete subject de-identification has occurred, and that stewardship of the data and samples will be ethical and responsible.

As a result of setting up a repository and its related policies to satisfy the imperatives noted above, several other benefits accrue. A transparent path is provided to data and sample sharing, an increasingly important concern to funding agencies. By codifying collection transport and storage procedures, sample integrity and uniformity is assured. Further, the accuracy of the link between the biologic samples and the clinical data is maintained with a high degree of stringency.

The Legal and Regulatory Environment Existing Around Repositories

For many years, the utilization of samples or data generated from patients and clinical trials were not carefully regulated after the trial was completed. Investigators would retain samples without explicit subject permission, and use them for unrelated purposes, often moving them from institution to institution, or transferring them between investigators. These practices often occurred without IRB oversight. Such practices led to situations that would now be considered abusive, such as the use of Henrietta Lack’s cervical cancer cells 60 years ago to establish a pervasive immortalized cell line (Skloot, 2010). With establishment of the Belmont commission in the 1970s (The Belmont Report, 1979) and the subsequent development of regulations and IRBs designed to protect the rights of subjects participating in clinical trials, egregious violations have become largely a thing of the past.

Informality is clearly no longer acceptable and subjects' rights must be adequately protected. IRBs now require evidence of subject informed consent for unconstrained use of their biologic materials with processes in place to assure de-identification. Additionally, peer review journals now require assurance of such approval prior to publication.

Concurrent with the maturation of the scientific ethical environment, a legal framework has developed that impacts the storage and future use of samples and data generated from clinical studies. Independent of the Belmont report, much of this impetus arose from the institution of the Health Information Portability and Accountability Act (HIPAA) in 1995, which requires certificates of confidentiality from institutions and researchers performing clinical research, and also governs the maintenance of confidentiality when samples are transferred from institution to institution. Regulation aside, subject concerns regarding unrestrained future use of their biologic samples, especially genetic material, has led to a body of case law in this area that should give investigators pause for thought. This is best illustrated in the area of newborn screening in Texas, where a settlement led to the destruction of 5.3 million newborn DNA samples, and has led several states to regulate the collection of newborn blood for this purpose, either with an opt in or an opt out clause (Beleno v. Texas Department of States Health and Human Services, 2009). In response to this controversy, draft recommendations have been issued regarding the retention and use of neonatal DNA by the Department of Health and Human Services (The Advisory Committee on Heritable Disorders in Newborns and Children, 2010). Additionally, the Havasupai Indian tribe has recently won a lawsuit against Arizona State University regarding the retention and unauthorized use of their research samples, in a fashion directly impinging on their cultural belief structure, resulting in a substantial settlement and return of the material (Mello, 2010). It is clear from a legal point of view that society is becoming increasingly sensitized to the long term ramifications of unregulated storage of data, genetic material, or biologic specimens in the course of patient care or scientific investigation.

For these reasons, central biospecimen repositories have evolved, as well as their certification within a regulatory framework. They standardize specimen collection, protect sample and data integrity, and assure a level of informed consent that will endure ethical and legal scrutiny well into the future. The legal framework for repositories is provided by HIPAA and federal human subjects regulations (45CFR46). The development of standardized repository procedures furthers these aims. Such standards have been put into place by the International Society of Biological and Environmental Repositories (ISBER; Best Practices for Repositories)(ISBER, 2005), the National Cancer Institute (NCI; Best Practices for Biospecimen Resources)(NCI, 2007) and the NHLBI (Guidelines for Human Tissue Repository, National Heart, Lung, and Blood Institute, National Institutes of Health) (NHLBI, 2000). These documents act to codify minimum standards of repositories for sample storage (inventory, storage, backup in the case of storage failure, release, and shipping), and should provide an invaluable resource for individuals wishing to establish a repository.

Subject Informed Consent for a Repository

Critical to the formation of a biospecimen repository is the prospective execution of a informed consent. Indeed, if separate repository consent is not obtained, either explicit permission for future use of specimens needs to be documented in the informed consent study from which the samples originated, or the criteria for waiver of consent need to be applicable (Common Rule, 45CFR46.116d). The repository informed consents obtained from subjects must outline fully the plans to store additional or extra biologic materials obtained for future as of yet undefined research purposes, regardless of whether it is in a separate biorepository consent or in the body of the original study consent.

Irrespective of how it is obtained, the repository informed consent must outline processes implemented to permanently deidentify the repository samples and the linked related data, the procedures for withdrawal of consent, and the ultimate disposition of samples and data from those subjects who choose to withdraw, or when the repository ends. Additional protection for the subjects can be obtained if the repository consent acknowledges and describes the existence of a Certificate of Confidentiality, a document signed by everyone involved in the study and the repository. It protects the subject's data and specimens from inappropriate release, for instance, because of a court order. Perhaps the best way to do this from an IRB's point of view is with an appended certificate of confidentiality. Blank IRB-approved consents from each site should be maintained in conjunction with the samples to assure future IRBs that appropriate informed consent for the use of these samples was obtained. For all these reasons, a separate IRB approved repository protocol consent is encouraged. In the case of the RMN, with multiple sites and IRBs, with differing local requirements, such a process became somewhat unwieldy and repository consents within the parent study consents were in many cases adopted. In the future, as the RMN and other cooperative clinical networks move towards a centralized IRB process, some of obstacles that local IRBs presented to a separate repository protocol and consent will subside.

Establishing the Physical Plant of the Repository

Once the decision has been made to create the repository, the next step is to decide to establish the physical plant in house or to subcontract the responsibility to an external repository service. For single site studies, or for studies with only a few sites, an in house repository is feasible. However, as in the case with the RMN repository where multiple sites are involved, an external physical site is preferred. This is required if the repository is to maintain ISBER standards. Among others, examples of such external repository vendors include for example SeraCare (http://www.seracare.com/), and Fisher biosciences (http://www.fisherbioservices.com/).

Regardless of where the repository is situated there are several considerations. Various biologic specimens have differing storage requirements. Blood spots for DNA, slides, or paraffin-fixed tissues can be stored in ambient environmental conditions. Although not optimal, serum or plasma for future steroid assays, DNA, or whole blood for DNA extraction may be stored at −20°C. Ideally serum or plasma samples should be stored at −80°C. If cellular viability is required (leukocytes, sperm) a liquid nitrogen environment (liquid or vapor phase) is needed. In all cases, storage systems require alarms, backup capacity and power in case of catastrophic failure, and an explicit management structure for immediate round-the-clock sample supervision. For large collections the dispersion of samples to several sites may be considered to minimize the impact of natural disasters on the sample set.

Sample reception, inspection, and inventory are also important parts of an effective biospecimen repository. Samples should be inspected for damage, and entered into the repository data base as soon as they are received, ideally within a few hours. Consider the situation if the samples arrive on a weekend. Clearly, shipping methods must be robust enough to allow the samples to survive for a considerable period of time (dry shippers for liquid nitrogen, dry ice for −20°C or −80°C). For these reasons, the RMN established a Monday-Wednesday shipping window from the clinical sites, on a quarterly basis. Many external repository subcontractors provide explicit shipping instructions and shipping materials to minimize sample risk in this regard.

A sample release protocol must be thoroughly defined. The criteria as well as the mechanics of release are important. Appropriate shipping methods, as noted above, must be used to

maintain sample integrity in the face of uncertain receipt procedures and if possible the procedure should be tested in house prior to shipment to a receiving institution. Ideally, a materials transfer agreement (MTA) should be executed between the repository and the receiving institution. An example of a MTA is presented in supplemental file 1.

When setting up the repository, considerable thought must be given to disposition of the biospecimens under several scenarios. First, a subject may request to withdraw the samples from the repository. Provision must then be made for sample identification and destruction, (or return to the subject). If this is required then sample retrieval may put the integrity of the repository de-identification process at risk. In the case of the RMN, it was decided that the risks of allowing the subjects to withdraw samples and data from the repository, in terms of the anonymity of the process, were too great. However, the RMN has provided a window for each participant to withdraw his/her samples by holding the samples on site after collection for at least three months prior to shipping. While the samples remain at the collection site, they are identifiable, and the subjects have the right to withdraw them from the repository if they wish.

The repository may eventually be closed. On one hand, perhaps the samples have been utilized so much that a significant number of them are exhausted, rendering the remainder scientifically useless. On the other hand, perhaps demand for stored materials is small and it is no longer cost effective, or the funding to maintain the collection is not available. As part of the repository protocol, there may also be a predefined date beyond which the repository closes (as the RMN secured funding for the repository through the next iteration, we elected not to enact such a provision). In all such cases, a clear plan for sample destruction must be in place and the existence of such provisions made known to the subjects in the informed consent. Succession plans must also be considered if the repository vendor closes or merges. Usually these provisions are negotiated in the initial vendor contract. Finally, repositories may need to perpetuate over multiple funding cycles of a network, and multiple iterations of investigators, as in the case of the RMN. In such cases, the governance of the repository must be passed on seamlessly. Alternatively, the repository could merge with another repository, in which case the governance would be transferred to that entity.

Linking Data to Samples While Preserving Anonymity (‘De-identification’)

It is absolutely critical to develop a robust system for irreversible de-identification of the research subjects’ samples when they are entered into the repository. Potential research subjects and IRBs need to be aware of this subject protection. It must be stringent enough to survive the scrutiny of future IRBs. In the RMNs case, data linking the subject to the biological samples remain at the clinical sites, subject to HIPAA and institutional regulations regarding the maintenance of data after the termination of the study. The Central Data Coordinating Center (DCC) generates a unique subject identifier code at randomization, and creates the data sets with this code. Only the sites, and not the DCC or the repository, can link this code to the subject. The biologic samples are stored at the repository, linked to this data set with this unique second identifier. When appropriate, the samples are released to requesting investigators, again linked to the data that was requested by the investigator with this second identifier. Thus, while the investigator gets the complete dataset, the samples remain anonymous to the DCC, repository, and requesting investigator. As noted above, in order to protect this essential process, once the sites have sent samples to the central repository, or data to the DCC, the subjects will no longer be able to withdraw samples. As a consequence of this de-identification policy, once the specimens are sent to the repository, and later used, and an abnormal result is obtained, it will not be possible to communicate that abnormal result back to the subject, and subjects need to be carefully appraised of this possibility beforehand, during the informed consent process. It would be

possible however, and even desirable, to add the data from repository generated samples back into the de-identified database at the DCC.

Governance of the Repository

A biologic repository existing on its own, without governance, is greatly limited in its utility. Codified and transparent governance assures data and sample integrity, safeguards the valuable resource, maximizes scientific impact, and assures fair access by both internal and external investigators. Governance should be prospectively established and in the public domain. The RMN repository is under the control of the Research Access and Data Acquisition (RADA) Committee, a subcommittee of principal investigators that reports to the RMN Steering Committee. Interested investigators approach the RMN committee with a request and provide a detailed hypothesis-driven protocol, evidence of adequate funding to perform the protocol, and written intra- or extra-departmental peer review. While it is not the intent of the RADA Committee to provide exhaustive peer review of the scientific worthiness of the proposed research, inevitably some degree of prioritization must occur to safeguard the resource. If the RADA Committee judges the request worthy, it recommends release to the Steering Committee with a priority score.

After final approval is obtained from these bodies for sample release, and following documentation of IRB approval for the proposed research and receipt of a uniform Materials Transfer Agreement (MTA) with the DCC, the DCC provides a list of the samples, type and the unique identifiers to the repository site, and sends the requisite data to the investigator. In the case of an investigator's request for DNA, the DCC will also notify the repository of the required amount of DNA to be released, and in what form (native DNA, transformed leukocytes, or blood spots). The repository then ships the requested samples or DNA directly to the investigator, while continuing to store and keep track of the remaining samples. The investigator is responsible for the fees of the biorepository and the DCC for handling and shipping of the specimens and data.

Several points are important to consider in this process. Once samples leave the repository, they cannot be returned, to ensure repository sample integrity. Once they are used for their stated purpose, the requesting investigator must destroy any remaining portion to protect the rights of the research subject and the parent network. However, this destruction needs to be considered in terms of requirements for publication and the ability to retest the exact same sample set if the data are called into question. Additionally, the parent network needs to be acknowledged once the results of the research are published. Finally, rules regarding when external investigators can gain access to samples and data from the repository need to be explicitly stated. For example, in the case of the RMN, the repository is open to external investigators after completion of the parent study plus one year from publication.

Releasing Samples from the Repository

The samples are released after vetting the sample request. Consideration is given to release the minimal aliquot to adequately complete the task. Thus, careful prospective consideration must be given to sample aliquoting at the outset of the repository, in order to optimize the utilization of the resource and avoid the need for multiple freeze-thaw cycles that can compromise the sample. The ultimate decision about how to aliquot samples must reflect a thoughtful balance between the expense of multiple aliquots, the volume of specimen available, and the anticipated future use needs of the specimens. An example of the RMNs sample aliquoting protocol is outlined in the accompanying paper.

In the RMN repository, the requesting investigators are charged for the shipping and handling of the samples from the repository. The amount of these charges are at cost and

serve simply to defray costs, so as to minimize obstacles to access to the samples, congruent with data and specimen sharing policies of federal funding agencies.

Conclusions

In this age of scarce research resources and increasing societal skepticism of the motives of clinical investigators, the establishment of a sample and data repository is an idea whose time has truly come. One can no longer store samples from clinical studies indefinitely and use them for unrelated research purposes without expecting objections from research subjects, IRBs and the legal profession. Given past abuses, such objections are justified.

In response to these concerns, the concept of a biospecimen repository has evolved, as outlined above, with standards, a legal framework, and governance, that protects the right of research subjects, the needs of the investigators, and the valuable resource itself. The establishment of such a repository has several other benefits. It ensures data and sample integrity, prioritizes access to the resource on the basis of scientific worth, and amplifies manyfold the impact of the parent study, well into the future. Given the expense and rigor of clinical studies today, particularly large multicenter randomized trials in the context of clinical trial networks like the RMN, the establishment of a repository can no longer be considered a scientific luxury. While the repository setup issues raised in this paper are not inconsequential, the RMN found that considering them beforehand made the eventual establishment of the repository quite seamless. Further, the overall cost of the repository was only about 4% of the total cost of the three parent trials. It is the hope of the RMN investigators, that by outlining the steps toward developing a repository in this paper, the path towards repository establishment by other groups of investigators is shortened and facilitated.

Acknowledgments

The authors wish to acknowledge the help of Louis De Paolo, PhD in the preparation of this manuscript.

List of abbreviations

RMN	Reproductive Medicine Network
IRB	Institutional Review Board
NICHD	National Institute for Child Health & Human Development
HIPPA	Health Insurance Portability and Accountability Act
ISBER	International Society for Biological and Environmental Repositories
MTA	materials transfer agreement
DCC	Data Coordinating Center
RADA	Research Access and Data Acquisition.

References

Beleno v. Texas. Department of States Health and Human Services, Case No. 09-cv-00188, First Amended Complaint (filed 9/29/2009).

ISBER. Best practices for repositories I: collection, storage and retrieval of human biological materials for research. Cell Preserv Technol. 2005; 3(1):5–48. 2005.

Mello MM, Wolf LE. The Havasupai Indian Tribe Case — Lessons for Research Involving Stored Biologic Samples. N Engl J Med. 2010; 363:204–207. [PubMed: 20538622]

National Institutes of Health. Guidelines for Human Tissue Repository, National Heart, Lung, and Blood Institute. 2000. April 14/2000.

NCI. Best Practices for Biospecimen Resources, 2007. 2007. (http://biospecimens.cancer.gov/global/pdfs/NCI_Best_Practices_060507.pdf)

Skloot R. The Immortal Life of Henrietta Lacks. Crown Publishers (Random House 2010).

The Advisory Committee on Heritable Disorders in Newborns and Children. Considerations and Recommendations for National Guidance Regarding the Retention and Use of Residual Dried Blood Spot Specimens after Newborn Screening. Federal register 75(79). 2010 Apr 26. (http://www.thefederalregister.com/d.p/2010-04-26-2010-9625).

The Belmont Report. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979 Apr. (http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm).