HHS Public Access

Author manuscript

Clin Lab Med. Author manuscript; available in PMC 2024 July 09.

Published in final edited form as:

Clin Lab Med. 2018 March ; 38(1): 183–207. doi:10.1016/j.cll.2017.10.015.

Biospecimens and Biobanking in Global Health

Maimuna Mendy, PhD [Group Head],
Laboratory Services and Biobank Group, International Agency for Research on Cancer, Lyon, France

Rita T Lawlor, PhD [Research Centre Coordinator],
ARC-Net Applied Research on Cancer Centre, University of Verona, Verona, Italy

Marianne K Henderson, MS,
Senior Advisor on Biobanking and Division Resources, National Cancer Institute, NIH, DHHS, Bethesda, MD USA

Anne Linda van Kappel, MSc, PhD [Managing Director],
Netherlands, IMV Technologies, 61300 l'Aigle, France

Peter HJ Riegman, PhD [Director],
Tissue Bank, Erasmus MC, Department of Pathology, Rotterdam, The Netherlands

Fontini Betsou, PhD, HDR [Chief Scientific Officer],
Integrated BioBank of Luxembourg, Luxembourg

Oliver D Cohen, MD, PhD [Professor]
Medical School of Grenoble, Joseph Fourier University, La Tronche, France

Keywords

biobanks; pathology; data; biospecimens; tissue; utilization; catalogues; information technology

Corresponding Author: Mailing address and email address for one author who will receive article proofs, Marianne K. Henderson, MS, Senior Advisor on Biobanking and Division Resources, National Cancer Institute, NIH, DHHS, 9609 Medical Center Drive, Rm 3W534, Bethesda, MD USA 20892, hendersm@mail.nih.gov.

author Contact information

Maimuna Mendy, PhD, International Agency for Research on Cancer, 150 Cours Albert Thomas, Lyon 69372, France

Rita T Lawlor PhD, Research Centre Coordinator, ARC-Net Applied Research on Cancer Centre, University of Verona, Piazzale LA Scuro 10, Verona 37134, Italy

Marianne K. Henderson, MS, Senior Advisor on Biobanking and Division Resources, National Cancer Institute, NIH, DHHS, 9609 Medical Center Drive, Rm 3W534, Bethesda, MD USA 20892

Anne Linda van Kappel MSc, PhD, Managing Director, Netherlands, IMV Technologies, 61300 l'Aigle, France

Peter HJ Riegman, PhD, Director, Tissue Bank, Erasmus MC, Department of Pathology, Dr Molewaterplein 40, Rotterdam, The Netherlands

Fontini Betsou, PhD, HDR, Chief Scientific Officer, Integrated BioBank of Luxembourg, 6 rue Nicolas Ernest Barble, L-1210 Luxembourg

Oliver D Cohen, MD, PhD, Professor, AGEIS EA 7407 Laboratory, Medical School of Grenoble, Joseph Fourier University, Domaine de la Merci, 38700 La Tronche, France

disclosure statement

The authors of this chapter have no conflicts of interest associated with the material presented.

1. Background and introduction

Infectious diseases continue to be a major burden of disease globally. According to the Global Burden of Disease Study 2015, while the epidemic of HIV/AIDS deaths peaked in 2005 and have annually decreased since 2015, with the scale-up of ART and PMTCT particularly in Sub-Saharan Africa, there continues to be large-scale HIV/AIDS epidemics in many low and middle income countries (LMICs). (1) While life expectancy in many regions has risen due to the investments in interventions for infectious disease, many countries are seeing an increase in rates of non-communicable disease (NCD) burdens as their populations age. In fact, the global total of new cancer cases is projected to increase by 75% to 22.2 million annually by 2030, with an estimated 13.1 million deaths from cancer yearly. About half of these cancer deaths will occur in low-income countries and more than 80% of these in African countries. (2, 3)

It is crucial that appropriate interventions and infrastructure be implemented to confront this disease crisis. Biobanks play an important role in the study of infectious and non-communicable disease etiology and identification of new potential diagnostic markers, and are central to the development of personalized drug treatment and translational research. (4–6) Investments in biobank infrastructure will enable scientific progress, on which effective disease control measures depend.

The aim of this chapter is to provide information on the collection, processing and storage of biospecimens and the management of biobanks as a valuable tool for global health research in low and middle income countries (LMICs). (7) A biobank, defined as a facility for the long-term storage of biospecimens, is a key resource providing for access to high-quality human biospecimens. The combination of infrastructure, facilities and resources is referred to as a biological resource centre (BRC). Tumor banks are BRCs; they have been defined by the Organisation for Economic Cooperation and Development (OECD) as service providers and repositories of living cells, of genomes of organisms, of cells and tissues, and of information relating to these materials.

Technological advances in molecular biology and genetics have greatly enhanced our ability to investigate the interactions among genetics, the environment, lifestyle and health. Biobanks consisting of biospecimens from clinical and epidemiological studies provide the opportunity to more effectively study disease causation and prognosis. At the present, analytical methods have developed to a level where they can be applied to large numbers of biospecimens, so biobanks play a cornerstone role in genetic and molecular epidemiology studies. The management of BRCs requires comprehensive quality management systems with appropriate controls. These are necessary to ensure that biospecimens collected for clinical or research purposes are of consistently high quality and are appropriate for the intended analyses and study goals. (8)

Despite advances in biobanking activities in high-income countries, populations in LMICs are underrepresented in sharing of these resources owing to their economic constraints and related issues. This means that studies are conducted without adequate representation of the populations that are mostly affected by the life- threatening diseases. Many research studies

have been conducted in LMICs, but apart from the biobanks created in HIV treatment facilities for HIV research involving large numbers of individuals, (9–12) very few other research studies have found it necessary to establish a biobank, mainly because the sample sizes for many non-HIV studies are small, and the studies very rarely collect and store frozen plasma or DNA for further biochemical and genetic studies. (13)

When such biospecimens are collected, their collection and storage are not often planned or organized in any systematic way. Noting the absence of biobank studies in sub-Saharan Africa, Campbell and Rudan (14) conducted a systematic review of birth cohort studies to assess the resources available to support genetic epidemiological studies. Their results showed that less than 40% of the 28 studies included in the review collected biological material and less than 20% collected and stored DNA.

In the absence of adequate funding for and awareness of the benefits of biobanks, the challenge for LMICs to establish and maintain suitable BRC infrastructure, consisting of biospecimen cold storage facilities, databases, reliable electricity supply to maintain the equipment, and quality assurance tools continues to be a challenge. Nevertheless, the first national DNA bank in Africa was established in 2000 in the Gambia as one of 14 such collection sites created by the Medical Research Council to study the genetics of the complicated diseases of malaria, HIV/AIDS and tuberculosis. (15) The facility in the Gambia has expanded over the years and has continued to support research activities.

Although there is a plethora of guidelines and protocols for biospecimen processing (9, 16–19), the tools are not easily accessible in LMICs, which makes it difficult to adhere to principles defined in international protocols and best practices. It is important therefore that alternative and cheaper options of evidence-based protocols be developed for LMICs.

The aim of this article is to promote good practices in human sample biobanking in LMICs to facilitate the appropriate collection of samples for the development of local biomedical research and international collaboration. Underpinning all this is the need to have well-trained staff to operate the facilities and manage the different processes involved in providing high-quality biospecimens, to develop appropriate technologies applicable to local settings, to handle the day-to-day activities, and to deal with issues relating to sample access and patient confidentiality. Information is provided on requirements for cold storage facilities and on the development and management of databases. The emphasis here is on presenting the basic requirements for BRCs or biobanks to store and maintain high-quality biospecimens and on providing guidelines to ensure that research is conducted with integrity and in adherence to the highest ethical standards according to international regulations and rules governing ethical, legal and social issues (ELSI).

This article provides information on the value of studies on pre-analytical variability of biospecimens which are crucial in ensuring the integrity of downstream analytical results (20) and how biospecimen science research can offer the opportunity to develop and validate appropriate technology and tools for LMICs. In particular, research to help identify quality-control biomarkers for assessing the quality of samples before they are included in expensive research platforms, would reduce costs and free up the limited funds to be spent elsewhere.

Quality control biomarkers can also contribute to higher reproducibility of data in biomarker research.

2. Collection and processing of biospecimens

Collection, annotation and use of human biospecimens are essential activities in biomedical care and research. Biobanking is also becoming a critical process in allowing patients access to molecular- based diagnosis, prognosis and precision medicine. Tumor banks need to comply with strict technical requirements. The definition of a tumor bank includes not only the infrastructure for collecting, archiving and storing biospecimens and data, but includes the entire biobank continuum beginning with the procedures and services for informing patients; obtaining consent; collecting and processing specimens for secure, potential long-term storage; appropriately accessing and retrieving specimens for analysis; and processing specimens for preparation of molecular derivatives such as DNA, RNA and proteins for quality control and for distribution to researchers.

Two types of methods must be distinguished: (a) biospecimen processing methods, which include different types of biospecimen handling, such as snap-freezing, paraffin embedding, plasma and buffy coat preparation, nucleic acid extraction, and establishment of cell lines, and (b) biospecimen quality control methods, which enable characterization of the biospecimens, and include such elements as the minimal sample characterization data set.

Biospecimen processing

The types of samples collected during clinical practice include bodily secretions, tissue and fluids. Using samples left over from clinical diagnostic procedures for biobanking purposes is generally not an optimal practice. Where possible, dedicated specimens for biobanking and research should be collected at the same time as diagnostic specimens, but in separate containers and processed through separate standardized workflows. However, tissue sampling should only occur once diagnostic procedures have been completed, such as the evaluation of margins and sufficient tissue has been sampled for diagnostic tests. The critical steps in each processing method should be acknowledged and controlled. The biospecimen processing method may depend on the anticipated end use. It is difficult for a biobank to anticipate all the different future uses for the samples, therefore, the most stringent processing requirements should be followed where possible by the biobank to maximize the lifespan and potential uses of the samples. As the impact of freezing and thawing on future target molecules is unknown for all sample types, the number of freeze-thaw cycles should be kept at a minimum. For that reason, small volumes of aliquots should be prepared before cryostorage, about ≤200μℓ for serum and plasma, and ≤0.5 cm3 for frozen tissue.

Prospective biospecimen collections generally have the most added value. Longitudinal follow-up of patients allows the establishment of causal links instead of simple associations between observed clinical end-points and candidate surrogate biomarkers. Improving such studies requires coding of samples instead of irreversible anonymization and, most importantly, adequate human resources such as clinical research nurses for follow-up.

Furthermore, inclusion of coded family links adds value to the collection. For biological fluids or solid tissue samples intended for immunological, molecular biology or proteomic analyses, critical in vitro pre-analytical details should be accurately recorded. For fluids, this information includes the type of primary collection tube, pre- centrifugation time delay and temperature, centrifugation conditions, post-centrifugation time delay and temperature, and long-term storage duration and temperature. For solid tissues, this information includes warm and cold ischemia times, type and duration of fixation, and long-term storage duration and conditions. (21)

Annotation of refrigeration and short processing delays are crucial, especially for urine collected without preservatives. If the samples are intended for proteomic downstream applications, high-speed centrifugation should be used, ensuring removal of white blood cells and platelets. The “as soon as possible” recommendation for pre-analytical processing is not precise enough. A simple way of tracking pre-analytical information is through the Standard PREanalytical Code (SPREC), a simple seven-element code that enables all pre-analytical information to be captured for the different types of specimens. (22)

This information can be added as a searchable data element in the biobank database. Finally, if metabolomic applications are anticipated, in vivo pre-analytical data—including the time of day when the blood or urine samples were collected, medications taken by the patient and food intake— should also be recorded in appropriate databases.

Best Practices and Standard Operating Procedures (SOPs) for different types of biospecimens being collected and processed in tumor banks can be found on the web sites of the International Society for Biological and Environmental Repositories (ISBER) Best Practices for Repositories (16), the Canadian Tumor Repository Network (CTRNet) (18), the US National Cancer Institute, Biorepositories and Biospecimen Research Branch (BBRB) (23), and the International Agency for Research on Cancer (IARC) Common Minimum Technical Standards and Protocols for Biobanks dedicated to Cancer Research. (19)

Other types of samples that can be stored include human DNA for genetic susceptibility testing, human RNA from peripheral blood for gene expression signatures such as prognostic or predictive biomarkers of treatment outcome, and peripheral blood mononuclear cells for cell sorting and cell immunophenotyping. Coordination of project-specific prospective collections is also possible, for example collecting urine with protease inhibitors or plasma for peptidomics analyses.

As a general rule, tubes and kits for collecting and processing biospecimens should be obtained from commercial suppliers. The advantage of using such devices is that they have already undergone significant validation by the suppliers. However, since this validation often focuses on specific quality attributes or specific molecules in the sample, the tumor bank still must validate the collection or processing device for other target molecules as these become known.

All human specimens regardless of the known disease state of the patient should be treated as potential biohazards. This is because the patient may have a known or potentially undiagnosed contagious disease. Appropriate measures should be taken to protect laboratory

workers who handle specimens and to prevent others from being exposed to the specimens during transportation. This is good laboratory practice. The most commonly collected and processed biospecimens include blood, urine and tissue.

Blood Specimens

Whole blood samples.—Whole blood does not require any special processing for storage and can be stored at −80 °C or room temperature or as dried blood spots on filter paper. Storage of whole blood samples is necessary if the end use is DNA analysis. DNA is a very stable molecule that is robust to a range of storage conditions. Whole-genome sequencing requires higher-quality DNA samples than do single-target PCR assays. However, whole-genome amplification can be performed to obtain large quantities of DNA from minute amounts of initial material. (24) When anticoagulated blood is centrifuged it separates into the red blood cell (RBC) fraction and the buffy coat layer containing white blood cells (WBC), platelets and plasma. When coagulated blood is centrifuged it separates into the clot (RBC, WBC and platelets) and serum.

Dried blood spots.—In the low resource setting context, storage of whole blood as dried blood spots (DBS) makes sense, as this avoids technical problems related to cryostorage and logistical arrangements. DBS can be used effectively for DNA sequence analysis for up to three decades of storage (25) and for cytokine measurements for up two decades of storage at −20°C. DBS testing is a powerful tool for screening programs and large population-based surveys, for detection of biomarkers such as hepatocellular carcinoma (26) and for large-scale testing for HIV infection. (27) Special attention should be paid to card selection, collection method and storage. The ethical concerns of long-term storage and use of DBS should be considered when consenting for the biospecimens. (28)

Plasma is the liquid fraction of anticoagulated blood.—Different anticoagulants may be used, such as ethylenediamine tetra-acetic acid (EDTA), heparin and acid citrate dextrose (ACD). The end use of the blood influences the choice of the anticoagulant. ACD is the preferred anticoagulant when lymphocytes from the cellular fraction of the blood are to be used to establish lymphoblastoid cell lines, but citrate interferes with future metabolomic analyses in the plasma. Heparin may inhibit nucleic acid amplification and future molecular biology analyses in the nucleic acid samples obtained either from the cellular blood fraction or plasma itself (circulating nucleic acids). Therefore, EDTA is preferred as it allows proteomic, genomic and metabolomic analyses to be performed in the future. (29) Platelet-poor plasma can be obtained after blood centrifugation at high speeds (>3000 g) and is more suitable for proteomic analyses because it allows less interference by circulating platelets and other coagulation factors. (30)

Time delays and temperatures to which the blood is exposed between collection and centrifugation and between centrifugation and plasma storage must be carefully documented. (22) There is no consensus about time delays and temperatures that the collected blood can tolerate, but it has been shown that pre-centrifugation delays of up to eight hours at room temperature do not significantly alter proteomic profiles. (31) For longer pre-centrifugation delays, storage at 4 °C is preferred. For specific target analytes in

the context of biomedical assays, validation should be performed to deal with the impact of the time delays. (32) Storage of plasma should be at −80°C, which is the temperature that has been found to ensure stability of the vast majority of molecules examined to date. (24) Validation has not yet been performed of lyophilized plasma and its storage at different temperatures, including room temperature.

Serum is the liquid fraction of clotted blood.—Preparing and storing serum instead of plasma offers two advantages for a tumor bank: (a) serum does not contain platelets and coagulation factors and therefore it allows proteomic analysis of a greater number of proteins, including those that cannot be identified in plasma because they are bound to plasma coagulation factors (31), and (b) the absence of additives in serum ensures there will be no interference from such elements in downstream spectrometric analyses. (29) As with plasma, for serum, it is very important to document the time delays and storage temperatures from blood collection to centrifugation (clotting time) and from centrifugation to storage. The same considerations for storage temperatures are observed for serum as for plasma. For both types of samples, the inflammatory status of the donor is an important confounder, if the anticipated use is proteomic analysis; therefore, normalization of the samples relative to the inflammatory status may need to be performed.

Peripheral blood mononuclear cells (PBMCs) include lymphocytes and monocytes.—These cells can be isolated from the buffy coat layer of centrifuged anticoagulated blood through Ficoll gradient centrifugation. PBMCs should either be stored at −80°C, preferably in lysis buffer if they are intended for gene expression analyses, or cryopreserved as viable cells in liquid nitrogen (LN2) using the cryopreservation medium dimethyl sulfoxide (DMSO) for future cell sorting, cell immunophenotyping, immortalization and establishment of lymphoblastoid cell lines. For transcriptional analysis, commercially available RNA-stabilization blood collection tubes, such as PAXgene RNA tubes (Qiagen) or Tempus tubes (Thermofisher), are preferred due to the potential influence of pre- analytical conditions on gene-expression profiles. (33) For viable lymphocyte isolation, Acid Citrate Dextrose (ACD) blood collection tubes are preferable. If Ficoll gradient centrifugation is not possible, whole blood can be cryopreserved and used for viable lymphocyte processing or analysis in the future. Viable lymphocytes may be recovered from whole blood held at 4°C for several hours before progressive-rate freezing in DMSO. (34)

Urine specimens

Urine samples have become quite valuable to measure metabolites associated with human disease state. (35) Urine contains DNA, RNA, proteins and metabolites, and it is easy to collect and store for analysis of all of these molecules. However, because urine composition lacks tight homeostasis and depends on disease status, the time of the day it was collected and donor hydration status, needs to be normalized for proteomic analyses. For proteomic or metabolomic analysis, urine is centrifuged and the supernatant is aliquoted and stored at −80°C. Centrifugation is necessary to avoid interference by cell components. For DNA or RNA analysis, the pellet is stored preferably in a nucleic acid stabilization solution such as a cell protect reagent. Filtration of the supernatant must also be performed for proteomic analysis, (36) but this step can occur after thawing and immediately before

analysis. Collecting midstream, first or second morning urine is preferred, and in all cases, collection time and delays should be documented. (22) Refrigeration is preferred to avoid bacterial proliferation.

Different urine preservatives such as boric acid or sodium azide may also be used to prolong processing delays if the urine is stored at either room temperature or 4°C. (37) Metabolomic analysis is affected by the use of preservatives so for this use, urine should be kept at 4 °C or frozen within one hour of collection. (29) While EDTA has been reported as a DNA stabilizer in urine, its effciency has not been reproducible in Africa. (38) Validation of standards for urine collection, processing and analysis have been the focus of active research and publications. (36)

Tissue specimens

As 90% of cancers are tissue, the collection of tissue samples provides the basic material for most cancer research. Tumors are made up of cancer cells and stroma and this must be evaluated during sampling and quality control of a cancer tissue sample. Furthermore, most cancers are heterogeneous in nature, differing not only from one to another, but also within different areas of the same tumor. This heterogeneity is not always identifiable by morphological evaluation and therefore sampling of cancer tissue for molecular classification must take this into consideration. Where possible, attempts to sample multiple areas of the cancer tissue should be done.

Dissection.—An absolute rule in tumor banking is that tissue sampling locations and tissue amounts for research must not interfere with routine diagnostics, resection margins and staging. Some samples may be sampled directly in the operating suite but pathology must be notified. Once the specimen has been excised, it should be transported to the grossing room and the time from excision to processing should be noted. Grizzle et al. reviewed the time to processing and the effects of cold ischemia on gene expression in human surgical tissues. They noted that variability in molecular results were more affected by specimen and patient characteristics than on cold ischemia; that ischemia differed with tissue organ and tissue size; and recommended that the timing be matched with the fit for purpose of specific research questions to be examined. (39)

The specimen should be weighed, measured and photographed by the pathologist, areas with normal and abnormal tissue portions should be sampled by gross examination following diagnostic sampling and procedures. Dissection of the specimen should be performed in a way to ensure its sterility and avoid cross- contamination during dissection, which are critical factors for downstream molecular biology analyses. For each tumor sample collected, a frozen section should be performed for quality control during sampling.

Freezing.—Tissue sections are immersed into either an isopentane bath previously cooled in LN2 or directly into an LN2 dewar. Tissue should be 0.5cm3 or thinner for quick freezing. A minimum of two to three minutes is needed for complete freezing. Frozen samples are then transported to the tumor bank in dry ice (−80°C) or colder. An optimal cutting temperature (OCT) compound can be used to embed and freeze the tissue, to allow for future cryosections and morphological examination, or molecular extractions and analyses.

However, as OCT can make frozen samples unusable for mass spectometry measurements, removal techniques need to be applied for this use. Frozen tissue specimens should never be allowed to thaw, which would not only destroy the sample's morphology but also cause severe RNA degradation.

Stabilization and fixation.—Different types of tissue fixatives are available. The standard fixative for preserving morphology is 10% neutral buffered formalin (NBF). However, be aware that NBF causes molecular cross-linking and undermines the quality of DNA, RNA and proteins that can be extracted from a formalin-fixed, paraffin-embedded (FFPE) block. For complete fixation, samples should not exceed 0.5cm3 in size. Samples should be fixed in 10 volumes or more of NBF. Routine fixation is typically done for approximately 12 hours overnight. After fixation, the tissue specimen is removed and placed in 70% ethanol for shipping or further processing in paraffin. PAXgene tissue fixative (Qiagen) allows for morphology preservation and at the same time ensures high quality of DNA, RNA and proteins. The only drawback seems to be the eventual necessity to revalidate the immunohistochemical (IHC) assay parameters for PAXgene-fixed, paraffin-embedded (PFPE) tissue. (40) Alcohol-based fixatives also are available, including Omnifix (Omni, Xenetics Biomed) and other proprietary fixatives. A section should always be made from fixed tissue for immediate histopathological quality control.

Tissue stabilizers exist that allow stabilization of molecules but do not preserve tissue morphology. RNALater (Ambion) is an aqueous non-toxic tissue storage reagent that rapidly permeates tissue and stabilizes all nucleic acids. AllProtect (Qiagen) reagent allows stabilization and subsequent extraction of both nucleic acids and proteins. These stabilizers eliminate the need to process tissue samples immediately or to freeze them. However, such methods usually do not preserve tissue morphology. Finally, heat stabilization under vacuum (Denator) conditions with the subsequent storage of the heat- stabilized tissue at −80°C has been shown to effectively preserve tissue phosphoproteome, although it does not allow preservation of morphology or use in nucleic acid analysis. (41)

Laser capture microdissection (LCM) is a technique that allows isolation of pure cell populations from heterogeneous tissue sections through direct visualization of the cells. Automated LCM platforms combine a graphical-user interface and annotation software together with staining reagents for visualization of the tissue of interest and robotically controlled microdissection. Microdissected cells for protein analysis may be stored at −80°C before extraction, while for DNA analysis they may be stored desiccated at room temperature for up to one week before extraction. For RNA analysis, the RNA extraction should be performed immediately after microdissection and the RNA samples then stored at −80°C.

Tissue microarrays (TMAs) are paraffin blocks that contain an array of minute specimen cores taken from different FFPE blocks or different areas of the same FFPE block. They are usually composed of several cores for each "donor" and usually contain many donors on the same block. TMAs are ideal for efficient screening of prospective biomarkers in multiple samples by IHC, fluorescence in situ hybridization (FISH), and RNA in situ hybridization methods. They are prepared by transferring paraffin tissue cores from many "donor" blocks

to one “recipient” block. TMAs should have positive controls for the anticipated IHC assays. Each slide cut from this recipient block is called a TMA slide. “Frozen” TMAs may also be prepared using frozen donor tissues embedded in the OCT compound. These samples are arrayed into a recipient OCT block. This allows high throughput evaluation of frozen tissue with corresponding visualization of tissue morphology. To preserve antigenicity, a fresh section should be cut at the time of evaluation or the TMA sections should be stored in a vacuum, in a nitrogen gas environment or at −80°C, to avoid antigen degradation due to oxidation.

Biospecimen quality control

Quality control (QC) procedures are important to ensure data and sample quality. For data, this includes clinical data accuracy, whereas for biospecimens it includes assays on the authenticity, integrity and identity of the samples. (21) Biospecimen QC is required to ensure accurate sample characterization and classification and to avoid introducing bias in downstream research due to intrinsic heterogeneity of the sample. This bias was shown in a specific African breast cancer classification study. (42) The type of QC depends on the intended end use of the sample. For example, samples to be used as reference samples in commercialized diagnostic kits must undergo mandatory testing for HIV, hepatitis B virus (HBC) and hepatitis C virus (HCV.) A central QC laboratory can undertake this testing, and the critical steps in each QC assay should be acknowledged and controlled by the laboratory. Accurate characterization of the samples supplied by a biobank focuses on both the authenticity and the integrity of the biomaterial.

Authenticity.—Phenotypic QC methods generally used for authentication of tissue specimens involve histopathological assessment. Trained pathologists should perform the histopathological evaluation of tissue samples (fixed and/or frozen sections), to confirm the tissue type, whether it is from a tumor or normal tissue, and the basic histopathological diagnosis and classification, based on standard hematoxylin-eosin staining. The evaluation includes assessment of cellular composition, which is of critical importance in any downstream molecular analysis. A sample with highly heterogeneous cellular composition makes any definitive molecular analysis more challenging. A general rule is cellularity higher that the detection rate of the instrument, the minimum cellularity of tumor is generally set at 40% for whole genome sequencing. (43) For lower cellularities, enrichment processes such as microdissection may be used. The standard histological control also includes assessment of specimen morphological degradation. Histopathological evaluation allows identification and marking of the block areas that are the most suitable for TMA cores. There may be special advantages in developing and implementing histopathological QC by telepathology in LMICs. In October 2015, the American Society for Clinical Pathologists (ASCP) and a coalition of world partners launched a project to provide patients in underserved areas of Sub-Saharan Africa and Haiti with access to rapid disease diagnostics and resource-appropriate treatment. (44) Static telepathology, or microscopic photographs, is based on offline imaging without interaction between operators. Virtual microscopy and dynamic telepathology allow production of “virtual slides” using navigation tools on the Internet, allowing for experts around the world to support diagnosis of diseases in LMICs that have such systems.

Integrity.—Very few data are available on the use of QC tools in assessing collection procedures, shipping and storage conditions. However, homogeneity in these steps is key for quality multi-center research studies. Recently, the H3Africa project team published on their experiences and solutions to the challenges of intra-country biospecimen shipment logistics associated with their multisite and multi-biobank network. (45)

For effective QC of prospective collections, biobank managers can proceed in several ways. QC may be performed on every specimen received at the biobank. In some instances, this may be highly recommended and cost-effective, such as haemocytometry for all blood samples. In other instances, generalized quality control for samples received at the biobank is not cost-effective, such as for specimens for DNA extraction and analysis. In that case, QC may be performed before distribution of samples to researchers, such as for verification of DNA concentration, purity or Taq amplifiability, if such the QC does not destroy the sample. (46) Retrospective QC is always an option and two alternatives are available: testing either a randomly selected percentage of the collected specimens or samples considered to have undergone the most “inconsistent” processing. The first approach allows comparisons of collection sites, while the second allows targeted assessment of “highest risk” samples.

There are a variety of QC tests for assessment of sample integrity and purity, shipping temperature variations, processing and storage conditions, to assess the fitness for purpose of biospecimens in a biobank (45,46,47,48).,

Plasma and serum QC.—Biospecimen research is in progress to identify appropriate QC tools for serum, plasma and urine. (47, 48) Such QC markers may be serum sCD40L to assess the duration of exposure of the sample to room temperature, protein S activity in plasma and matrix metalloproteases in serum or plasma. Serum sCD40L assays are particularly relevant in Africa, where high ambient temperatures are more often observed.

Biospecimen research in LMICs

Variability in acquisition, processing and storage of samples may contribute to experimental variability, particularly in high-throughput analyses, and may result in false research conclusions. This is especially true for the most labile bioanalytes like RNA, functional proteins and metabolites. In this respect, biospecimen research is linked to biospecimen QC.

High-quality nucleic acids and adequate antigen preservation can be obtained from formalin-, PAXgene- or ethanol-fixed, paraffin-embedded tissues or from DBS on filter paper. These types of biospecimen processing techniques suit the logistical conditions in low resource settings because they do not need cryostorage. Biobanks in LMICs may choose to develop expertise in these approaches of maintaining biospecimen stability and robustness in the context of novel, dried blood, room-temperature technologies, including lyophilization, and in room-temperature storage devices for biofluids, nucleic acids, or even for whole blood.

Choosing the best collection method.

Choosing the right procedures for collecting and storing samples and data is very important, and such decisions need to be carefully considered even before the samples and data are

collected. A review of the optimum conditions for collection and storage of many human sample types provides evidence for specific choices. (49) The focus should be to maximize the level of quality to ensure that the sample is at least fit for its intended use(s). However, if possible and affordable, the highest possible standards should be pursued to permit wide use of the samples in future tests, applications and collaborations. (11) Collaboration must consider not only sample sharing, where not only sample quality must be comparable, but also donor consent and access where approval by local stakeholders has been considered up front. (50)

All samples and data used within a study must at least be collected to be “fit for purpose” for the planned (and anticipated) experiments and all samples must be of comparable quality. Comparable quality demands that pre-analytical conditions be kept as constant as possible for every sample. Therefore, every sample should be collected in analogous manner following a standard operating procedure (SOP) for the collection. The object for SOPs should be focused on the highest quality possible in the specific collection setting. The QC process should be designed to reject samples from studies that are not fit for purpose.

Changing collection or storage procedures while the collection is underway can have consequences for the results from the experiment in which these samples are used and should therefore be carefully considered. Changes implemented during sample collection should be documented. Protocols should be established for the processes to be used for the entire collection period, prior to initiating collection. To achieve a level of standardization of processes, it is very important to have written SOPs for at least the collection, processing, storage and distribution of samples. In addition, it is essential to ensure that the SOPs are followed and interpreted in the same way at all collection sites for a study. Regular audits and QC are needed to check compliance with the SOPs. Any variance to the collection procedures should be documented and this documentation can be used during analyses to correlate outliers in experimental results or during cohort selection for exclusion of samples that are not fit for purpose.

It is important to include an SOP for transportation between collection site and the biobank including chain of custody procedure. Often in LMIC settings, there may be challenges in shipment of samples from remote collection locations to the biobank. Transportation of samples in warm cars before processing and storage, can be detrimental for many methods, and thus cooling is often required and should be addressed in a shipment SOP, to maintain sample integrity and quality.

Quality assurance

Certification of a biobank to international standards such as ISO 9001 (quality management systems— requirements) by an independent body is proof that the biobank is effectively organized and managed. Furthermore, subcontracting testing to laboratories that are themselves accredited to international standards such as ISO 17025 (general requirements for the competence of testing and calibration laboratories) or ISO 15189 (medical laboratories particular requirements for quality and competence) by a national accreditation body is proof of reliability of the sample characterization processes. A new ISO standard specifically for biobanks is in development (ISO/DIS 20387). Although compliance with

these standards is important, it essentially remains voluntary for the biobanks and can be expensive.

3. Biospecimen storage facilities and equipment

Biospecimen storage facilities are the most visible part of biobanks, and storage systems are important factors in maintaining sample quality. The variety of storage systems available for specimen collection increases as technologies advance. FFPE materials must be stored at room temperature, meaning that in most countries the storage room would need air conditioning to maintain the temperature below 25°C. Mechanical freezers need to be kept in clean, cool rooms since they produce a lot of heat and must have dust removed from their filters to work efficiently.

Storage equipment should be selected based on the type of specimens to be stored, the anticipated length of storage time for the specimens, the use intended for the specimens, and the resources available for purchasing the equipment. (16) In selecting equipment, quality issues should be considered, but for a local setting with limited access to tools, the primary considerations should be the available resources, staffing requirements and equipment support and maintenance. For the sample storage equipment, such as freezers, and for infrastructure equipment, such as electrical power and backup systems, LN2 bulk tanks, and transport pipes, compatibility with local conditions and the capacity of the vendor to provide on-site support and maintenance for the time they are used by the biobank, should be verified. Two types of storage systems are described here: ultra- or low-temperature storage systems and ambient temperature storage systems.

Sample storage containers

In selecting storage containers for biospecimens, consideration should be given to:

- **a.** sample volume;
- **b.** necessary cooling and warming rates for both the individual container and the racks, boxes, or goblets;
- **c.** potential risk of contamination of the sample or the environment;
- **d.** storage temperature and conditions;
- **e.** space available for sample storage;
- **f.** frequency of access;
- **g.** specimen identification requirements;
- **h.** specimen preparation and after-storage processing techniques;
- **i.** economic aspects.

Containers used in cryogenic temperatures should be rated for these temperatures. Containers used for storage in LN2 should be hermetically sealed to avoid penetration of LN2 into the container and consequent risk of contamination and explosion when the container is removed from the freezer. All human specimens should be treated as potential

biohazards, and the choice of storage container should integrate minimizing the risk of contamination of laboratory workers who handle the specimens and preventing others from being exposed to the samples in the laboratory or during transportation. This is good laboratory practice. Identification labels, that do not contain personal identifiers, should be compatible with the storage temperature and medium and should always include eye-readable codes when access to scanners for barcodes, 2D codes or RFID codes cannot be guaranteed by the sample processing institution or end-user.

Liquid nitrogen freezers

Cryogenic storage using LN2 is an effective long- term storage platform because its extreme cold temperatures slow down most chemical and physical reactions preventing biospecimens deterioration. LN2 vapor-phase containers can maintain samples below Tg (glass transition temperature, i.e., −132°C) while submersion in LN2 guarantees a stable −196°C temperature. Where a regular supply of LN2 is available, LN2 freezers reduce reliance on mechanical freezers and electrical power and guarantee sample integrity under critical temperatures during power cuts as closed LN2 freezers can maintain samples at below −130°C for longer periods without refilling. Initial investment together with availability and cost of LN2 can be major drawbacks. When LN2 freezers are used, oxygen level sensors should be used and calibrated regularly. Use of protective goggles and gloves should be mandatory, and easily accessible. Appropriate training in the safe handling of cryogenics and samples stored in cryogenics should be provided and part of a SOP for health hazards and safety precautions.

Mechanical freezers

Mechanical freezers are used for a variety of storage temperature ranges and come in a wide range of sizes, configurations and electrical voltages. Ice crystals may form in biological samples at temperatures of about −70°C, therefore freezer temperatures should preferably be below −80°C. Cascade compressor technologies may produce temperatures as low as −140°C. Mechanical freezers, generally require a lower initial investment than LN2 freezers and provide easier and safer access to samples, can be installed if electrical power is available. However, compressor technology requires constant electrical power to maintain subzero temperatures, so backup power and an emergency response plan are needed. Ambient temperature and humidity influence temperature stability so they should be set apart in rooms that are air- conditioned and/or have equipment for extraction of hot air generated by the compressors. Internal temperature is easily affected by open doors during sample loading. Freezers need to be periodically cleaned to remove frost and filters should be cleaned.

Refrigerators

Refrigerators are commonly used where the longevity of the material being stored is enhanced by storage below ambient temperature. Storage at 4°C can also be an intermediate step before preparation for ultralow-temperature storage. For refrigerators, as for mechanical freezers, it is important to maintain and monitor the temperature in the required operating range and to organize for a backup power plan.

Ambient-temperature storage

In the absence of mechanical or cryogenic equipment owing to practical or financial reasons, specific biological storage matrices may be used for long-term maintenance of some biological components at room temperature. Formalin-, PAXgene- (51) or ethanol-fixed, paraffin-embedded tissues and lyophilized samples can be stored at such temperatures. The storage matrices should be evaluated before use to ensure that they are appropriate for downstream applications. Temperature, humidity and oxygen levels should be controlled at the biobank, to avoid mold growth and microbial contamination.

Information technology in biobanking

Information Technology (IT) has a fundamental role in biobank organization. Indeed, software applications known as Laboratory Information Management Systems (LIMS) have been adapted and developed to address biobank processes, optimize workflow efficiency, drive quality assurance, and maximize the use of collections. IT systems have also been developed to manage the large amounts of resulting molecular research data and provide controlled access to the data by the research community.

IT tools are particularly important as part of quality management systems of biological resources. Documenting actions (metadata) associated sample collection, processing and preservation times, storage locations and temperatures (traceability), provide important quality indicators for both fluid and tissue samples. This sample metadata, together with sample-associated clinical annotation and consent documentation, create enriched, valuable information about the quality of the sample collections. The ISO 9001:2015 standard sets out the criteria for a quality management system. The ISO/TC276 standards in development will include specific requirements for a QMS for biobanks and biological resources. The standards aim to harmonize business practices, IT requirements and optimizes exchanges between biobanks. As part of a QMS, LIMS tools facilitate auditing of biobank procedures, to allow for evaluation of conformity to SOPs, and continued quality process improvement.

Use of information technology in tumor biobanks

Large numbers of primary and derived samples can be easily managed with most LIMS software. Every process and procedure should time-stamped and recorded with information on operator, equipment used, and reagents and other consumables used. Digital versions of forms documenting sample collection, sample reception forms, sample processing and anonymized/de-identified sample may be designed and completed electronically to reduce manual input of data to reduce user errors, increase workflow efficiency and improve quality assurance.

Most LIMS software provides the ability to trace the movement of samples over time in a biobank. Storage information shows the current position and movement of samples in storage units (freezers). A storage location tool within a LIMS can suggest sample storage locations for new collections, based on the total number of samples to be stored and the collection to which it belongs. Free space in storage units can be identified to optimize the use of space in the biobank’s units.

All processing and storage-related incidents should also be recorded. Each incident should be described by 1) event type, e.g a missing subject consent or deviation from protocol, and 2) event status, e.g. whether corrective action is pending, is ongoing or has been taken.

LIMS can also be used to track and manage donor consent, if scanned copies of consent documents (or copy of the informed consent template) are saved, consent information is registered with the IT system and this data is permanently linked to patient and/or sample data. It is important to only release samples for the uses contained in the approved scope of the donor's consent.

When utilizing samples, the LIMS should track the distribution of samples and therefore often permits the registration of dates such as sample requests and transfers, parties involved, sample recipients, and sample quantities requested, transferred or returned. Some LIMS may also store the scanned material transfer agreements or MTA number associated with a transfer requisition. Criteria may be set in some biobank LIMS that can alert staff when a pre-established minimum sample aliquot number threshold has been reached, to ensure adequate volumes are maintained and stored for future use.

Many biobanks now include a "return of research information" policy that requires data from experiments or publication citations be returned to the biobank. This data if uploaded to the biobank database permits for continuous enrichment of the collection and increases the collections value to the research community. However, the format of the data and whether the biobank has the capacity to store and analyse raw data needs careful consideration.

Biospecimen catalogues and directories

A catalogue is an electronic list of the biobank's collections and associated metadata. Catalogues are established as institutional, national or regional infrastructures promote harmonization and can provide standardized resources for enhanced utilization of collections in international collaborative studies (e.g. NCI's Specimen Resource Locator, BBMRI-ERIC catalogue, CTRNet and BCNet). Catalogues may be updated manually or periodically by a formatted program issued from the biobank's database. Biobank catalogues or directories are developing at a rapid pace to provide visibility for biological resources and data for secondary access to collections available once studies have completed their primary objectives and end points. A new study collection can be built up from samples (biological resources) selected for their common biological, clinical, and preanalytical characteristics, such as those in SPREC (22), from across several collections within or amongst biobank catalogues. IT solutions can facilitate the selection of homogeneous or comparable resources and for research on diagnostic or therapeutic biomarkers.

For example, the BBMRI-ERIC directory (52) represents the European biobanking infrastructure for collections and biobank networks and provides information on the what sample sets and data sets are available for sharing. The US National Cancer Institute's funded extramural collections are listed in the Specimen Resource Locator catalogue. (17) The Canadian Tissue Resource Network (CTRNet) has a biobank locator created to increased collaboration and use of Canadian specimen collections. The locator individually

lists the networked biobanks, their major collection type and the contact information for the biobanks, across Canada. (53) The Biobank and Cohort Building Network global catalogue (54) provides the ability to register, store and track biological collections held by BCNet LMIC member organizations. It is being developed to enhance research sample and data sharing to support public health research within the member's countries/regions.

Commercially available catalogue solutions are becoming more widely available, as well as those integrated into biobank LIMS software.

Types of information technology

Biobank LIMS software must be selected or designed for its security, robustness, interoperability and configuration features. In making the decision on software, attention should also be paid to the data items to be recorded and the quality of the data. The IT solution must ensure the security and accuracy of donor identification. Donors must be anonymized at all times. Hospital patients must be identified by their permanent ID numbers and their hospital visit numbers so that the samples can be linked to the respective hospital visits. Many systems are developed that protect and maintain patient information confidentiality per the standards of the Title 21 Code of Federal Regulations (21 CFR Part 11) of the United States Food and Drug Administration (FDA). Specific security attention should be paid to systems that hold genetic data. Software solutions based on web informatics systems are preferred for biobanks as they offer high security. They also have the advantage of interoperability with other clinical informatics systems. Selection of a LIMS should consider the ease of configuration and ability of the local administrator to customize content and functionality, extensibility and maintenance. The amount of pre-existing specimen metadata to be imported to a new LIMS solution, will help drive specific software selection, based on the ease of data migration.

Off-the-shelf LIMS solutions include open-source software or tables. While open-source software solutions may appear to be free to the biobank, they may have secondary costs such as those for installation, "training" or "support" to set up the application and adapt the solution to the specific requirements of the biobank. Some open source LIMS require expert local administration at the biobank.

OpenOffice or Excel tables are other forms of off-the-shelf IT solutions, but are most likely not the best option for biobanks with multiple collections and users. Furthermore, tables and spreadsheets do not offer data traceability, and do not have the functionality to alert or log changes in data content.

The marketplace for off the shelf, commercially available LIMS for small and large-scale biobanking, regardless of sample type, has exponentially increased over the past several years.

Annotation of biospecimens

Data items that are recorded in biobank LIMS likely include:

a. Patient identification and demographic data, with sex and date of birth, as well as state of the vital signs.

b. Diagnostic data, with principal diagnostic end- point and date, and clinical tumor-lymph node- metastasis (cTNM) classification.

c. Specimen data, including sample identifiers, sampling date, sample nature, organ from which sample was collected, collection method, stabilization process and preservation details.

d. Lesion data, including histological type, size, event nature, whether primary tumor or metastasis, lesion advancement and invasion and pathological tumor-lymph node metastasis classification (pTNM), grading and other important data such as gleason score for prostate cancers.

e. Sample data, type, number, size, characterization (for tissue, whether the sample is tumor (neoplastic), non-neoplastic, tumor, normal adjacent, normal distant, node and a quality control evaluation based on the percentage of tumor, necrosis, stroma), and sample pre-analytical data such as SPREC.

f. Derivative data, including type, number, quantity, SOP used to produce derivative (e.g. DNA or RNA) and its characterization (e.g. concentration, purity, integrity).

g. Storage data, including temperature, location (freezer, shelf, rack, box, position in box) and events.

All data items should use standard terminology to optimize their use and to allow their future export into a specific IT system. While many histopathology reports use free text, this does not promote good practice for the biobank and, where possible, should be translated into standardized formats, for easier comparison. All data should be entered through drop-down lists with standardized terms based on dictionaries, ontologies and international nomenclatures like the International Classification of Diseases for Oncology (ICD-O) (55) or the Systematized Nomenclature of Medicine (SNOMED). (56)

Recent publications have emphasized the importance of a common terminology for improved data sharing (57) and ISO 11179–3 conformant metadata repositories. (58)

Data Backup and Disaster Recovery

IT systems must include a required process that backs up data up to an alternate database, housed in a physically distinct location (and if possible geographically distinct) for safety and disaster recovery. (59) Backup intervals could be daily, weekly or monthly depending on the volume and nature of activity. If no IT system is available and the use of electronic tables and spreadsheets are used, backups of these should save time-stamped versions. This allows for minimal tracking of changes to the database over time as well as safety against loss.

Regulation in Biobanking

The regulations governing biobanking need to address legal and ethical issues concerning the use of biological materials and data in research. These regulations must deal with the rights and responsibilities of donors, biobank managers and researchers. Biobank

governance must respect individual donors and guarantee their privacy and confidentiality. At the same time, it must not inhibit the provision of samples for potentially beneficial research. Regulations can be found dispersed in different declarations, such as acts governing the use of human tissue and data. (60–63) Many African countries have regulatory bodies with guidelines for the use of biological samples and data for research. (64, 65) Where national guidelines do not exist, international guidelines can be used to address legal and ethical aspects concerning the collection and use of biological material and data for research. However, traditional cultural values placed on human biological materials and data by local communities must be considered.

International guidelines

The WMA Declaration of Helsinki (66) provides guidelines for medical research on human beings. It aims to promote the ethical conduct of research and to protect human subjects from associated risks. The Declaration of Helsinki was the first set of international research guidelines that required research participants to provide informed consent. Since its adoption, it was amended by the addition of the WMA Declaration of Taipei on Ethical Considerations regarding Health Databases, Big Data and Biobanks. (67) Large collections of data and human specimens allow for the development of new research strategies and models, as well as new predictive types of research and analysis. The promise of large sample sets and genomic data create potential hazards for identifiability of donors, privacy worries and a need for strong governance of biobanks and resulting data. The United Nations Educational, Scientific and Cultural Organization (UNESCO)(68) emphasizes the protection of human genome-derived genetic data. (41) The use of human biological resources and data for genetic research is addressed in the OECD Guidelines on human biobanks and genetic research databases (69) and the Belmont Report and US Code of Federal Regulations, 45 CFR 46 (Common Rule). (60)

Governance

The key aspects relating to biobank governance are the policies, processes and procedures in place to ensure correct operation of the biobank. Governance should take into account the biobank ethical and legal issues of informed consent of samples, benefit sharing by donors' countries, confidentiality, ownership and public participation. (70) These should include oversight mechanisms for the development, implementation and use of the biobank collections; stakeholder support and accountability; and sustainability of the biobank. The responsibilities of funders, biobank developers, researchers and the various institutions involved must be clearly spelt out. For samples and data, there must be well- defined and documented processes for initiating collections, acquiring specimens, sharing samples and long term support, as appropriate. Oversight mechanisms should include ethics policies to ensure the ethical collection and use of samples and data, consistent with the consent granted by the subjects. Scientific policies should control the scientific validity of sample requests and consider the availability of samples and their rarity or scarcity.

Data access policies should guide researchers' access to data and define the conditions for access and the review process for access. Access policies also guide the deposition of research data from the use of biospecimens, back to resources to provide wider use by

approved users.(71) Governance processes must address the possible or eventual closure of the biobank and how samples and data will be disposed of or transferred to a third party. These processes must respect the initial consent granted by the donors, and disposal must comply with local regulations. Two documents set out the conditions for collection and use of samples: the informed consent and the material transfer agreement.

Informed consent

Informed consent is a fundamental legal and ethical principle in sample donation and biobanking. It underlines the basic rights of autonomy, liberty and dignity. It outlines the agreement between the donor and the custodian biobank on the provision of samples and data for research. Any deviation from a donor's consent must be authorized by a supervisory board, such as an institutional ethics review board.

The basis of informed consent is that donors understand the request being made for collection, storage and use of their samples and data. The consent form should be simple, clear and in the colloquial language of the donor.

Consent must be voluntary and should indicate the purpose of the biobank; possible physical risks associated with collection of the sample and risks associated with collection of personal data; procedures for maintaining privacy; methods of protecting donor identity; future use and access to the samples and data; the right to withdraw consent and request destruction of remaining samples or render them anonymous; the possibility of sharing samples and data with other institutions, exporting them across borders or using them commercially; and the right to refuse to provide samples, with clarification that such refusal will not affect the care to the donor/patient. It is also necessary to indicate the possibility that the donor might be re-contacted for follow-up or more information or for further consent. It may be unrealistic to expect researchers to re-contact individual participants to obtain specific informed consent for each access to their samples in a new research project. In addition to being expensive and impracticable, such requests may also be against the wishes of the donor.

Types of informed consent

Several types of informed consent exist, defined by the level of permission from donors for use of their samples. Specific consent limits the use of the sample and data to a specific research project whose details are made aware to the donor. It is used when samples and data are identifiable. Partially-restricted consent is used in a specific research project but allows future unspecified use directly or indirectly related to the research. Broad consent allows unspecified future use of the sample and data and the donor is provided with general information about possible future research, but which should comply with applicable national or local regulations and policies. Layered or tiered consent permits the donor to consent to particular aspects of the research but not others. Specific consent is advisable for identifiable samples and data, whereas broad consent may be used where samples and data are anonymized and the research is approved by an ethics committee or other body. (72, 73).

Community engagement prior to the launching of a project that involve the use of samples and data is a useful method of informing, consulting and actively involving relevant

communities that have a legitimate interest in the research process. Individuals often take decisions in consultation with family, friends and community members. Frequently, there are also clear authority structures that must be respected in the engagement process such as permission from village chiefs and elders. Adhering to these principles is important to build respect and trust between research teams and the respective communities.(74)

Vulnerable subjects

Safeguards should be put into place for the use of tissue from vulnerable donors, such as patients with mental incapacity, for example, heavily sedated people, people with dementia or impaired consciousness, and children. In the case of deceased donors, consent should be based on the views of the deceased person or of the family, if known. When research includes an ethnic minority, single community or cultural group, a representative from that group should be involved in the consent process. (75)

Exceptions to informed consent

The requirement of consent may be waived by an ethics committee, in accordance with applicable laws and regulations, in cases where the researcher will not come into possession of identifying information and the specific research has been approved by a recognized research ethics committee.

Material transfer agreements

A material transfer agreement (MTA) is a contractual document governing the conditions under which samples and data may be used in research. It defines the rights and obligations of both the biobank and the receiving researcher. Provision of samples and data must be consistent with the given consent of the donor. Cross-border collaboration and sample export must be governed by the permission from the donor's informed consent and the local legislation, which will indicate whether biological material may be exported and what necessary permits are required.

An MTA should specify (a) the purpose of the transfer of the material and its intended use; (b) restrictions on the use of the samples, such as their redistribution to third parties or sale for commercial purposes; (c) restrictions on re-identification, where de-identified specimens are provided; (d) requirements for handling biosafety hazards; (e) disposal or return procedures for unused samples; (f) ownership of intellectual property rights; (g) acknowledgement arrangements and publication rights; (h) provision of aggregated or raw research data; (i) guarantees and waivers; and (j) other factors that may govern sample transfer and the applicable regulations and law. The specific protections for data may be included in the same MTA or a separate agreement. Such an agreement must deal with the future use of the data, including their possible redistribution, requirements for maintaining privacy and confidentiality, governance for access to data and protection of data against unauthorized access. Most MTAs provide a time limitation and should provide guidance on the final disposition of the samples and data – whether the residual samples should be returned or destroyed. Based on the funding for the research or the regulations of the sample provider, data resulting from the use of the samples may be required to be deposited in centralized databases for open access to the research community. (71)

Data privacy and data protection

Data protection is a key principle in protecting donor privacy and confidentiality. (76) The types of data collected by the biobank are (a) sample-related data on quantity, quality and methods of collection and storage, and (b) donor-related data on clinical, pathological and lifestyle aspects. These data are considered sensitive because they provide information for potential identification of the donor and possible familial indicators. Therefore, procedures must be put in place to store data in a manner that protects the identity of the donor, including de-identifying or coding any identifying data, storing samples without associated identifying data, and ensuring that data are stored securely with access restricted to authorized personnel, including access to coding keys that may re-identify data or associate them with other data sets.

Sample identification

Identifying information should not be provided to researchers unless the research specifically requires it and approval from either the donor or the ethics committee has been received. If identifiable samples are used for research, donors should be informed about any implications, for example if they will be re-contacted by researchers or receive feedback or additional requests for access to medical records.

Genetic data

In disease research, genetic data can be defined as somatic or germline. Somatic alterations are genomic anomalies that are confined to diseased cells and have the potential to provide valuable diagnostic, prognostic and treatment information.

Germline variations are gene variants that increase the risk of disease either related to hereditary predisposition or promoted by lifestyle and environmental exposures. Sharing germline data requires specific consent, because they are considered as sensitive identifying material with consequences not only for donors but also for others family members.

Particular attention needs to be paid to the protection and release of these data. (77) A dedicated data access committee should make the decision about providing genetic data to the general research community. Complex ethical issues such as whether to provide genetic and genomic research information to donors about heritable factors or disease risks mean that ethic boards require that biobanks should choose to err on the side of caution, and these data should remain in the domain of research unless clinically validated and actionable in a clinical context. There is ongoing debate on the topic of release of genetic data directly to individuals or in research projects. (78)

Sustainability of biobanking

Creating a successful biomolecular research program requires the creation of a quality biobank infrastructure that spans the continuum from consent of donors, through collection, processing, storage, and utilization of biospecimens. Quality biobanking infrastructure is inherently expensive and requires a long-term financial commitment. Currently, the funding of research biobanks is not uniformly regulated, mainly because of the different focuses, services and concepts. (79) On the other hand, the samples and associated

data of biobanks gain significant value over time, and therefore their long- term storage should be planned from the start. (80)The biobank must have the staffing, the equipment, processes and quality control plans to provide fit for purpose specimens to the research community. Sustainability of biobanking requires attention to operational, social and financial dimensions of the process. (81) Securing the organization's commitment to support a biobank; creating a business plan that is reviewed and modified as the biobank matures and asthe collections grow; planning for utilization of the collections and providing services to the research community in an ethically-robust framework; are key components to long-term sustainability. (82–84)

Conclusion

Biobanking has become a key infrastructure to study the molecular basis of disease across the world. The biobanking continuum includes the initiation of the informed consent for the sample donation, to collection, processing, storing and retrieving the samples for research. Each of these important steps should be controlled by SOPs and documented in a robust information technology system. The SOPs should be developed to be appropriate for the available resources, without forfeiting the quality needed to result in meaningful molecular data for the research population being studied. The lack of biobanks and quality biological resources from LMIC populations should be addressed to make the largest impact on the global burden of disease.

References

1. Wang H, Naghavi M, Allen C, Barber RM, Bhutta ZA, Carter A, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet.388(10053):1459–544.
2. WHO. GLOBOCAN 2012. (http://globocan.iarc.fr/). 2012.
3. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer. 2013;132(5):1133–45. [PubMed: 22752881]
4. Arbyn M, Ronco G, Cuzick J, Wentzensen N, Castle PE. How to evaluate emerging technologies in cervical cancer screening? Int J Cancer. 2009;125(11):2489–96. [PubMed: 19626591]
5. Hainaut P, Vozar B, Rinaldi S, Riboli E, Caboux E. The European Prospective Investigation into Cancer and Nutrition biobank. Methods Mol Biol. 2011;675:179–91. [PubMed: 20949388]
6. Pukkala E, Andersen A, Berglund G, Gislefoss R, Gudnason V, Hallmans G, et al. Nordic biological specimen banks as basis for studies of cancer causes and control--more than 2 million sample donors, 25 million person years and 100,000 prospective cancers. Acta Oncol. 2007;46(3):286–307. [PubMed: 17450464]
7. World Bank Country and Lending Groups: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.
8. Vaught JB, Henderson MK. Biological sample collection, processing, storage and information management. IARC Sci Publ. 2011(163):23–42.
9. Biobank Quality Standard: Collecting, storing and providing human biological material and data for research National Cancer Research Institute UK, 13 March 2014 http://ccb.ncri.org.uk/wp-content/uploads/2014/03/Biobank-quality-standard-Version-1.pdf
10. Abdur Rab M, Afzal M, Abou-Zeid A, Silverman H. Ethical practices for health research in the Eastern Mediterranean region of the World Health Organization: a retrospective data analysis. PLoS One. 2008;3(5):e2094. [PubMed: 18461183]

11. Grizzle WE, Gunter EW, Sexton KC, Bell WC. Quality management of biorepositories. Biopreserv Biobank. 2015;13(3):183–94. [PubMed: 26035008]
12. Thio CL, Smeaton L, Saulynas M, Hwang H, Saravanan S, Kulkarni S, et al. Characterization of HIV-HBV coinfection in a multinational HIV-infected cohort. AIDS. 2013;27(2):191–201. [PubMed: 23032418]
13. Abimiku A, Mayne ES, Joloba M, Beiswanger CM, Troyer J, Wideroff L, et al. H3Africa Biorepository Program: Supporting Genomics Research on African Populations by Sharing High-Quality Biospecimens. Biopreserv Biobank 2017;15(2):4.
14. Campbell A, Rudan I. Systematic review of birth cohort studies in Africa. J Glob Health. 2011;1(1):46–58. [PubMed: 23198102]
15. Sirugo G, Schim van der Loeff M, Sam O, Nyan O, Pinder M, Hill AV, et al. A national DNA bank in The Gambia, West Africa, and genomic research in developing countries. Nat Genet. 2004;36(8):785–6. [PubMed: 15284842]
16. ISBER Best Practices for Repositories; http://www.isber.org/bpr.
17. https://specimens.cancer.gov/.
18. http://www.ctrnet.ca/operating-procedures.
19. Mendy M, Caboux M, Lawlor RT, Wright J, Wild C. Common Minimal Technical Standards and Protocols for Biobanks dedicated to Cancer Research.. IARC Technical Publication No44 ISBN-13 (PDF) 978-92-832-2464-8
20. Moore HM, Compton CC, Alper J, Vaught JB. International approaches to advancing biospecimen science. Cancer Epidemiol Biomarkers Prev. 2011;20(5):729–32. [PubMed: 21430299]
21. Betsou F, Lehmann S, Ashton G, Barnes M, Benson EE, Coppola D, et al. Standard preanalytical coding for biospecimens: defining the sample PREanalytical code. Cancer Epidemiol Biomarkers Prev. 2010;19(4):1004–11. [PubMed: 20332280]
22. Lehmann S, Guadagni F, Moore H, Ashton G, Barnes M, Benson E, et al. Standard preanalytical coding for biospecimens: review and implementation of the Sample PREanalytical Code (SPREC). Biopreserv Biobank. 2012;10(4):366–74. [PubMed: 24849886]
23. http://biospecimens.cancer.gov/rsources/sops/library.asp.
24. Lasken RS, Egholm M. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. Trends Biotechnol. 2003;21(12):531–5. [PubMed: 14624861]
25. Sjoholm MI, Dillner J, Carlson J. Assessing quality and functionality of DNA from fresh and archival dried blood spots and recommendations for quality control guidelines. Clin Chem. 2007;53(8):1401–7. [PubMed: 17586590]
26. Mendy M, Kirk GD, van der Sande M, Jeng-Barry A, Lesi OA, Hainaut P, et al. Hepatitis B surface antigenaemia and alpha-foetoprotein detection from dried blood spots: applications to field-based studies and to clinical care in hepatitis B virus endemic areas. J Viral Hepat. 2005;12(6):642–7. [PubMed: 16255766]
27. Schim van der Loeff MF, Sarge-Njie R, Ceesay S, Awasana AA, Jaye P, Sam O, et al. Regional differences in HIV trends in The Gambia: results from sentinel surveillance among pregnant women. AIDS. 2003;15;17(12):6.
28. Tarini BA. Storage and use of residual newborn screening blood spots: A public policy emergency. Genet Med. 2011;13(7):3.
29. Bernini P, Bertini I, Luchinat C, Nincheri P, Staderini S, Turano P. Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. J Biomol NMR. 2011;49(3–4):231–43. [PubMed: 21380509]
30. Rai AJ, Gelfand CA, Haywood BC, Warunek DJ, Yi J, Schuchard MD, et al. HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. Proteomics. 2005;5(13):3262–77. [PubMed: 16052621]
31. Hsieh SY, Chen RK, Pan YH, Lee HL. Systematical evaluation of the effects of sample collection procedures on low-molecular-weight serum/plasma proteome profiling. Proteomics. 2006;6(10):3189–98. [PubMed: 16586434]
32. Clark S, Youngman LD, Palmer A, Parish S, Peto R, Collins R. Stability of plasma analytes after delayed separation of whole blood: implications for epidemiological studies. Int J Epidemiol. 2003;32(1):125–30. [PubMed: 12690023]

33. Debey S, Schoenbeck U, Hellmich M, Gathof BS, Pillai R, Zander T, et al. Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. Pharmacogenomics J. 2004;4(3):193–207. [PubMed: 15037859]
34. Stevens VL, Patel AV, Feigelson HS, Rodriguez C, Thun MJ, Calle EE. Cryopreservation of whole blood samples collected in the field for a large epidemiologic study. Cancer Epidemiol Biomarkers Prev. 2007;16(10):2160–3. [PubMed: 17932366]
35. Playdon MC, Sampson JN, Cross AJ, Sinha R, Guertin KA, Moy KA, et al. Comparing metabolite profiles of habitual diet in serum and urine. Am J Clin Nutr. 2016;104(3):776–89. [PubMed: 27510537]
36. Thomas CE, Sexton W, Benson K, Sutphen R, Koomen J. Urine collection and processing for protein biomarker discovery and quantification. Cancer Epidemiol Biomarkers Prev. 2010;19(4):953–9. [PubMed: 20332277]
37. Vu NT, Chaturvedi AK, Canfield DV. Genotyping for DQA1 and PM loci in urine using PCR-based amplification: effects of sample volume, storage temperature, preservatives, and aging on DNA extraction and typing. Forensic Sci Int. 1999;102(1):23–34. [PubMed: 10423850]
38. Cannas A, Kalunga G, Green C, Calvo L, Katemangwe P, Reither K, et al. Implications of storing urinary DNA from different populations for molecular analyses. PLoS One. 2009;4(9):e6985. [PubMed: 19746164]
39. Grizzle WE, Otali D, Sexton KC, Atherton DS. Effects of Cold Ischemia on Gene Expression: A Review and Commentary. Biopreserv Biobank. 2016;14(6):548–58. [PubMed: 27551929]
40. Belloni B, Lambertini C, Nuciforo P, Phillips J, Bruening E, Wong S, et al. Will PAXgene substitute formalin? A morphological and molecular comparative study using a new fixative system. J Clin Pathol. 2013;66(2):124–35. [PubMed: 23125305]
41. Methodology Boren M. and technology for stabilization of specific states of signal transduction proteins. Methods Mol Biol. 2011;717:91–100. [PubMed: 21370026]
42. Awadelkarim KD, Arizzi C, Elamin EO, Hamad HM, De Blasio P, Mekki SO, et al. Pathological, clinical and prognostic characteristics of breast cancer in Central Sudan versus Northern Italy: implications for breast cancer in Africa. Histopathology. 2008;52(4):445–56. [PubMed: 18315597]
43. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagn. 2017;19(3):341–65. [PubMed: 28341590]
44. www.ascp.org/globalimpact.#sthash.8Q5SQZ3f.dpuf.
45. Croxton T, Swanepoel C, Musinguzi H, Kader M, Ozumba P, Pillay A-D, et al. Lessons Learned from Biospecimen Shipping Among the Human Heredity and Health in Africa Biorepositories. Biopreserv Biobank. 2017;15(2):8.
46. Simbolo M, Gottardi M, Corbo V, Fassan M, Mafficini A, Malpeli G, et al. DNA qualification workflow for next generation sequencing of histopathological samples. PLoS One. 2013;8(6):e62692. [PubMed: 23762227]
47. Betsou F, Barnes R, Burke T, Coppola D, Desouza Y, Eliason J, et al. Human biospecimen research: experimental protocol and quality control tools. Cancer Epidemiol Biomarkers Prev. 2009;18(4):1017–25. [PubMed: 19336543]
48. Betsou F, Gunter E, Clements J, DeSouza Y, Goddard KA, Guadagni F, et al. Identification of evidence-based biospecimen quality-control tools: a report of the International Society for Biological and Environmental Repositories (ISBER) Biospecimen Science Working Group. J Mol Diagn. 2013;15(1):3–16. [PubMed: 23195791]
49. Hubel A, Spindler R, Skubitz AP. Storage of human biospecimens: selection of the optimal storage temperature. Biopreserv Biobank. 2014;12(3):165–75. [PubMed: 24918763]
50. Riegman PH, de Jong B, Daidone MG, Soderstrom T, Thompson J, Hall JA, et al. Optimizing sharing of hospital biobank samples. Sci Transl Med. 2015;7(297):297fs31.
51. Kap M, Smedts F, Oosterhuis W, Winther R, Christensen N, Reischauer B, et al. Histological assessment of PAXgene tissue fixation and stabilization reagents. PLoS One. 2011;6(11):e27704. [PubMed: 22110732]

52. http://old.bbmri-eric.eu/bbmri-eric-directory.
53. http://biobanking.org/locator.
54. http://bcnet.iarc.fr/projects/biobank_catalogue.php.
55. http://codes.iarc.fr/.
56. http://www.snomed.org/snomed-ct.
57. Ellis H, Joshi MB, Lynn AJ, Walden A. Consensus-Driven Development of a Terminology for Biobanking, the Duke Experience. Biopreserv Biobank. 2017;15(2):126–33. [PubMed: 28338350]
58. Ulrich H, Kock AK, Duhm-Harbeck P, Habermann JK, Ingenerf J. Metadata Repository for Improved Data Sharing and Reuse Based on HL7 FHIR. Stud Health Technol Inform. 2016;228:162–6. [PubMed: 27577363]
59. Henderson MK, Simeon-Dubach D, Zaayenga A. When bad things happen: lessons learned from effective and not so effective disaster and recovery planning for biobanks. Biopreserv Biobank. 2013;11(4):193. [PubMed: 24845583]
60. US HHS Office of Human Research Protections Regulations and Policies. (Belmont Report and 45 CFR 46) https://www.hhs.gov/ohrp/regulations-and-policy/index.html.
61. US Code of Federal Regulations, Title 45 Part 46 (45 CFR Part 46): https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html
62. EU General Data Protections Rule: http://www.eugdpr.org/the-regulation.html.
63. Single European Code for Tissues and Cells : http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015L0565&from=EN
64. Rothstein MA, Knoppers BM. Harmonizing Privacy Laws to Enable International Biobank Research. The Journal of Law, Medicine & Ethics. 2015;43:2.
65. International Compilation of Human Research Standards in Africa: https://bioethicsresearchreview.tghn.org/site_media/media/medialibrary/2013/11/2014intlcomp_Africa.pdf
66. Declaration of Helsinki. 2008.
67. WMA Declaration of Taipei on Ethical Considerations regarding Health Databases and Biobanks: https://www.wma.net/what-we-do/medical-ethics/declaration-of-taipei/
68. UNESCO. Universal Declaration on Bioethics and Human Rights. 2005.
69. OECD guidelines on human biobanks and genetic research databases. Eur J Health Law. 2010;17:14.
70. Chen h, Pang t. A call for global governance of biobanks. Bulletin of the World Health Organization 2015;93:113–117 doi: http://dxdoiorg/102471/BLT14138420 [PubMed: 25883404]
71. US HHS Public Access Policy for Research Data (https://www.hhs.gov/open/plan/hhs-public-access-policy-for-research-data.html).
72. Elger BS, Caplan AL. Consent and anonymization in research involving biobanks: differing terms and norms present serious barriers to an international framework. EMBO Rep. 2006;7(7):661–6. [PubMed: 16819458]
73. Salvaterra E, Lecchi L, Giovanelli S, Butti B, Bardella MT, Bertazzi PA, et al. Banking together. A unified model of informed consent for biobanking. EMBO Rep. 2008;9(4):307–13. [PubMed: 18379580]
74. Jao I, Kombe F, Mwalukore S, Bull S, Parker M, Kamuya D, et al. Involving Research Stakeholders in Developing Policy on Sharing Public Health Research Data in Kenya: Views on Fair Process for Informed Consent, Access Oversight, and Community Engagement. J Empir Res Hum Res Ethics. 2015;10(3):264–77. [PubMed: 26297748]
75. Andanda PA. Human-tissue-related inventions: ownership and intellectual property rights in international collaborative research in developing countries. J Med Ethics. 2008;34:9.
76. Malin B, Karp D, Scheuermann RH. Technical and policy approaches to balancing patient privacy and data sharing in clinical and translational research. J Investig Med. 2010;58(1):11–8.
77. Kaye J, Heeney C, Hawkins N, de Vries J, Boddington P. Data sharing in genomics--re-shaping scientific practice. Nat Rev Genet. 2009;10(5):331–5. [PubMed: 19308065]

78. Knoppers BM. Consent to ‘personal’ genomics and privacy. Direct-to-consumer genetic tests and population genome research challenge traditional notions of privacy and consent. EMBO Rep. 2010;11(6):416–9. [PubMed: 20448662]

79. Kirsten R, Hummel M. [Securing the sustainability of biobanks]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2016;59(3):390–5. [PubMed: 26769112]

80. Watson PH, Nussbeck SY, Carter C, O’Donoghue S, Cheah S, Matzke LA, et al. A framework for biobank sustainability. Biopreserv Biobank. 2014;12(1):60–8. [PubMed: 24620771]

81. Henderson MK, Goldring K, Simeon-Dubach D. Achieving and Maintaining Sustainability in Biobanking Through Business Planning, Marketing, and Access. Biopreserv Biobank. 2017;15(1):1–2. [PubMed: 27860502]

82. Bledsoe MJ, Henderson MK, Sawyer SJ, Otto JM, Suh KS. Ensuring Biobank Value Through Effective Utilization. Biopreserv Biobank. 2015;13(6):5.

83. Henderson MK, Simeon-Dubach D, Albert M. Finding the Path to Biobank Sustainability Through Sound Business Planning. Biopreserv Biobank. 2015;13(6):2. [PubMed: 25686039]

84. Simeon-Dubach D, Henderson MK. Sustainability in biobanking. Biopreserv Biobank. 2014;12(5):287–91. [PubMed: 25314050]

Key Points

- Biobanks or Biological Resource Centres provide critical infrastructure for clinical research and biomarker discovery when samples are well annotated with pre-analytic data.
- Sample collections should be encouraged from geographically- and genetically-diverse regions to assure relevant clinical health data for all populations, including those in low and middle income countries.
- Biospecimen (tissue, and biofluids) collection, processing, storing and retrieval should be carried out with strict standard operating procedures (SOPs) to assure sample quality and fit for purpose use.
- The SOPs should be developed that are appropriate for the available resources, without forfeiting the quality needed to result in meaningful molecular data.
- Documentation about samples, including pre-analytic data, donor consent for use and linkage to clinical data from the donor should be kept in a robust laboratory information system (LIMS) to protect the data and ensure privacy and encourage ethical use of the samples and data.

Synopsis

Biobanks provide a critical infrastructure to support research in human health. Biospecimens, including tissue, biofluids (blood, serum, plasma, urine, saliva) and purified DNA and RNA and their accompanying data are increasingly needed to support biomedical research and clinical care. Insights from basic research on human biospecimens include biomarker discovery and can be translated back to the clinic and patient care. The original text was initially published in the Handbook for Cancer Research in Africa. The value of this publication is great as it underlines the importance of biobanks in Africa as a key resource to increase quality scientific research and participate in global health research. Therefore, a revision to extend these principles to other low resource contexts, to include updated material and references and the add the topic of biobank sustainability were relevant.