Banking of Biological Fluids for Studies of Disease-associated Protein Biomarkers*

Anne-Sofie Schroh‡§, Sidse Würtz‡§, Elise Kohn¶, Rosamonde E. Banks||, Hans Jørgen Nielsen**, Fred C. G. J. Sweep‡‡, and Nils Brünner‡§§

With the increasing demand of providing personalized medicine the need for biobanking of biological material from individual patients has increased. Such samples are essential for molecular research aimed at characterizing diseases at several levels ranging from epidemiology and diagnostic and prognostic classification to prediction of response to therapy. Clinically validated biomarkers may provide information to be used for diagnosis, screening, evaluation of risk/predisposition, assessment of prognosis, monitoring (recurrence of disease), and prediction of response to treatment and as a surrogate response marker. Many types of biological fluids or tissues can be collected and stored in biorepositories. Samples of blood can be further processed into plasma and serum, and tissue pieces can be either frozen or fixed in formalin and then embedded into paraffin. The present review focuses on biological fluids, especially serum and plasma, intended for study of protein biomarkers. In biomarker studies the process from the decision to take a sample from an individual to the moment the sample is safely placed in the biobank consists of several phases including collection of samples, transport of the samples, and handling and storage of samples. Critical points in each step important for high quality biomarker studies are described in this review. Failure to develop and adhere to robust standardized protocols may have significant consequences as the quality of the material stored in the biobank as well as conclusions and clinical recommendations based on analysis of such material may be severely affected. Molecular & Cellular Proteomics 7: 2061–2066, 2008.

The expression of proteins, either in amount or form, is often altered in disease. These altered proteins may be produced by diseased tissue or by normal adjacent tissues as a response to disease in a range of diseases types such as inflammatory, endocrine, cardiovascular, and cancer. The proteins may reach the peripheral blood circulation from where they can be sampled by a simple venipuncture. Disease-related proteins may also be present in other bodily fluids such as urine, saliva, cerebrospinal fluid, sweat, cystic fluid, and effusions; samples of these bodily fluids can be recovered for subsequent analyses in the laboratory.

A protein biomarker can be defined as a protein or peptide that can be objectively quantitatively measured and evaluated as an indicator of normal biological processes, pathological processes, exposure to environmental factors, lifestyle, presence of pathogens, or responses to a therapeutic intervention. In this regard, fluid samples from different body compartments may be used to measure biomarkers for several purposes; i.e. diagnosis, screening, evaluation of risk/predisposition, assessment of prognosis, monitoring (recurrence of disease), prediction of response to treatment, and as a surrogate response marker. For these purposes, the availability of high quality biobanks of body fluids using state-of-the-art methodologies with respect to sample collection, processing, and storage becomes extremely important. The quality of the operating procedures with which the biospecimens are handled directly determines the quality of the biomarker result and its ultimate applicability to the patient. Therefore, careful and thorough decision making before starting a collection is instrumental for getting the desired quality of samples.

The process of biomarker development and application, from the decision to take a sample from a patient to dispatch of the final assay result to the physician or researcher, consists of three phases: preanalytical, analytical, and postanalytical. Although the focus of many laboratory investigators is specifically on the analytical phase the pre- and postanalytical phases require similar robustness to have confidence in the final output. Analysis of errors occurring in routine clinical chemistry laboratories shows that at least 40% of the mistakes in the process are made in the preanalytical phase (1–4), whereas most attention goes to the laboratory process, i.e. the analytical phase (5). Lack of adhesion to simple guidelines and protocols regarding the preanalytical phase in particular may be the key reason for the fact that only a very limited number of biomarkers have emerged as clinically useful despite thousands of reports on promising candidates. Methodological problems in sampling, processing, and stor-
Regarding the age of samples in the preanalytical phase may result in inappropriate conclusions from any studies using such samples for biomarker discovery and additionally may hinder the velocity of subsequent analyses of a parameter in the laboratory even though the laboratory itself is accredited and has an outstanding performance.

The goal of this review is to focus on preanalytical issues and to increase the readers' awareness of potential problems, hazards, and biases that can arise when collecting body fluids for biobanks and protein biomarker studies with a particular focus on plasma and serum and specific named analytes. Similar considerations with other fluids and additional complementary information on serum and plasma and proteomics analysis have additionally been reviewed elsewhere (6).

Many different protocols are being used worldwide for these fluid collections. In our view it is not possible to recommend one uniform standard operating procedure (SOP) for fluid collection for biomarker studies because the purpose of the sampling may differ between the clinical studies, particularly if a specific analyte is being measured, as different biochemical and biophysical properties may require tailor-made processing. Therefore, we will not provide a generic SOP for biofluid collection but instead give some general recommendations and suggestions on how to achieve adequate sampling, processing, and storage to build a high quality biorepository, in this case of serum and plasma as an example (key issues are listed in Table I). The different items discussed can easily be modified for other specific purposes, but it is important that a consistent approach is used within any biobank. We encourage the investigators to include detailed descriptions of all relevant information relating to the preanalytical phase whenever including data from banked samples in their scientific publications. Only a transparent and complete reporting of the materials and methods will aid other investigators in understanding the presented data and in judging whether the data are reliable and clinically useful. Moreover it will help others in designing a protocol from which they can independently reproduce the data.

**PREANALYTICAL PARAMETERS IMPORTANT FOR HIGH QUALITY BIOBANKING**

The fluid sample has been subjected to several crucial steps prior to and following entry to the analytical laboratory before the step of specific analyte quantification or profiling analysis if being used in a biomarker discovery experiment. Outside the laboratory, these steps include subject consent followed by sample collection, handling, processing, storage, and transport to the laboratory. Further processing and subsequent storage may occur in the laboratory. Important preanalytical parameters, exemplified for blood collection, will be described below, and a checklist is provided in Table II. It should be emphasized that the decisions made on the basis of the checklist depend strongly on the scientific question and the expected assay, and therefore not all the items raised might be applicable in specific situations.

**Collection of Samples**—Project design with a defined clinical relevance (hypothesis/objective) is required to determine the correct biospecimens and methods for collection and handling. This information is a required part of a petition for approval to collect samples from an Institutional Review Board or Institutional Ethics Committee. Application to use existing samples in a biorepository may not require independent informed consent from subjects. Prospective collection for a biorepository or for a specific project will require an approval and application of the informed consent prior to ascertainment of the biospecimen.

When building a biobank for future research it is preferable that as many sample types as possible from an individual are collected and stored. For blood, the minimum suggested is a serum and a plasma sample. Although venous blood is most commonly used, if indicated and approved arterial blood sampling can be included. For example, during surgery one can obtain arterial and venous blood, such as from afferent and

---

**Table I**

Specific aspects that should be included and described in any standard operating procedure for blood collection and handling (see text for details)

| Description of method of how the sample should be obtained from the patient, e.g. site of venipuncture, whether tourniquet should be applied. Additional information such as whether patient should be fasted, any specific posture such as supine or sitting, | Identify tube types to be used, e.g. manufacturer, any anticoagulant or other additives used, whether gel-based tubes, volume of sample to be obtained, any sample mixing that is needed. |
| Any specific aspects of sample collection, e.g. order of tubes filled, discard of initial 5–10 ml to clear the line of dwelling heparin or saline if indwelling catheter used. | Any specific aspects of sample handling, e.g. place on ice or leave at room temperature and period of time elapsed before further processing of samples. |
| Describe aliquoting process including avoidance of buffy coat, use of pooling tubes prior to aliquoting, volumes of aliquots, type of storage tubes to be used. | Describe the coding labeling of any samples, what the storage temperature should be and the location, and how sample locations are to be recorded. |
| Recording of all times of the various stages and any deviations from the SOP or noted abnormalities with the sample, e.g. hemolysis. |

---

1 The abbreviations used are: SOP, standard operating procedure; TIMP-1, tissue inhibitor of metalloproteinases-1.
efferent blood vessels of a tumor (7, 8). Such studies may be used to directly study differential secretion of biomarkers across a tumor bed.

Several factors related to the individual and the environment should be considered before starting the collection of samples. Concentrations of proteins may be affected by normal physiologic factors, e.g., race, gender, age, menopausal status, stage in the menstrual cycle, fasting, exercise, smoking, alcohol/caffeine consumption, timing (day/year), co-morbidity, medication. Sampling (use of tourniquet, resting time, posture, stress, nurse experience, timing relative to other interventions, e.g. anesthesia) Collection of samples (blood or biofluid collection vessel with choice of anticoagulants or other additives such as proteinase inhibitor cocktails) Concentrations of proteins may be affected by normal physiologic factors, e.g., race, gender, age, menopausal status, stage in the menstrual cycle, concomitant infection, etc. (9). In addition, various stages of pregnancy may influence the levels of some biomarkers, including levels of angiogenesis-regulating proteins and hormone-binding and transport proteins (9). Where allowable, data collection should include these parameters. A vast number of other factors such as fasting, exercise, smoking, consumption of alcohol or caffeine-containing beverages, time of day and year, use of a tourniquet, time of rest before sampling, posture (sitting versus supine), stress, co-morbidity, medication, and experience of the nurse who collects the sample could also be of importance (10–14). For instance, we have shown that tissue plasminogen activator concentrations in blood taken after use of a tourniquet are significantly higher than samples taken while blood vessels were not occluded. In practice, not all of these parameters can be controlled, but any available information about these parameters should be carefully collected, registered, and stored whenever possible along with the samples (15).

The project design may require samples from control or reference individuals. It should be emphasized that “healthy” or “normal” blood donors or other volunteers, although often used, may not be the most appropriate reference population, but this depends on the specific question being asked. Blood donors are persons generally less than 60 years of age with no known diseases and receiving no medication, and accordingly they represent neither the general background population nor a disease control state. Other volunteers may, by requirement, have an underlying condition that provides a control to the project in question; for example they may have a benign condition that provides an important source of samples for comparison with patients with a malignant illness. Another source of control or reference samples could be from a case-control study. For example, we performed a large prospective study in which serum and plasma were collected from individuals who underwent endoscopy for symptoms that could be related to colorectal cancer (15). Among the ~5,000 individuals who were enrolled in the study 305 had colorectal cancer. However, the study provided 4,700 blood samples from individuals who were comparable to the colorectal cancer patients with regard to all demographic and

<table>
<thead>
<tr>
<th>TABLE II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Checklist for preanalytical considerations associated with banking of biological fluids (see text for details)</strong></td>
</tr>
<tr>
<td>It should be emphasized that the decisions taken on the basis of the checklist depend strongly on the scientific question and the expected assay, and therefore not all the items raised might be applicable in specific situations.</td>
</tr>
<tr>
<td><strong>Phase/purpose</strong></td>
</tr>
<tr>
<td><strong>Collection of samples</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Transport of samples to the laboratory</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Handling and processing of samples</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Storage of samples</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Bioinformation</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
clinical variables but without colorectal cancer. Similarly we have performed a prospective sampling of serum and plasma from more than 1,000 women who were scheduled for a breast surgical procedure. Approximately one-third of these women had benign diseases, and because they reflect the demographic and clinical features of the breast cancer patients, they can be used as reference group for the cancer patients.

Different technical aspects apply to collection of different types of samples. Plasma is collected in a tube with anticoagulant added (commonly EDTA, citrate, or heparin), whereas serum is collected into a tube that contains no additives or in many cases a form of micronized silica to promote clotting. For some types of collections it might be feasible to add a mixture of protease inhibitors to avoid extensive degradation of the proteins in the sample; the parameters of the proteins in these tubes may differ substantially from samples collected in tubes without protease inhibitors.2 Tubes can be made of different materials and may adsorb analytes differentially. We therefore recommend using standardized and internationally available tubes and those predefined by the project objectives where prospective collections are done. It is especially important to use identical types of tubes for all patients and control subjects included into one particular (clinical) study. Often if the biobank is being established for biomarker discovery, the type of tube will not be determined by the requirements of a specific analyte but by more pragmatic considerations of what is practically possible and likely to be of most use generally. Obviously if the analyte of interest is known and available in purified form, recovery and parallelism studies should be performed before starting collection and banking, thus enabling selection of the most appropriate tube material for this specific type of analyte. If samples to be analyzed have been obtained from external banks it is recommended to perform recovery studies for the particular protein in the tube type that this other bank(s) used when they collected their samples.

When obtained from the donor, blood samples are often kept at room temperature until processed. However, some analytes may require that the blood sample for plasma sampling is immediately put on melting ice. Serum samples are usually kept at room temperature for 30–60 min when coagulation should be complete, and again in some instances it may be required that the serum tube is put on ice before transportation to the laboratory for further processing. Although many samples and proteins may be stable for prolonged periods of time, subsequent processing is generally recommended to be as rapid as possible, and examples of analytes such as vascular endothelial growth factor illustrate the need for this (16).

When collecting cerebrospinal fluid, saliva, urine, cystic fluid, or effusion biospecimens, similar precautions apply. In addition, biomarker levels in these fluids may be dependent on flow (volume produced/time) or total volume, and a marker to adjust for differences in flow rate or volume may be needed (e.g. creatinine in urine). Note that standard procedures for collection of biospecimens for clinical purposes, such as diagnosis or therapeutic benefit, may include additives that might adversely affect planned biomarker studies. Additives may include heparin to prevent clotting of proteinaceous ascites or pleural effusions or preservatives normally found in some liter collection bottles. Biospecimens taken for both research banking and medical purposes should be addressed prior to collection to avoid these confounding issues.

Transport of Samples to the Laboratory—Samples destined for a biobank are preferably directly brought to the laboratory by dedicated transport or personnel wherever practically possible (15). This will ensure increased awareness of transport time and transport conditions and immediate receipt and processing. In some institutions, processing occurs prior to transport to the biobank whereas in other situations the processing laboratory may be within the same institute but not in close proximity to the site of collection. If it is not possible for processing to occur at the parent institution, it is recommended that a transport SOP is defined that is considered optimal for most if not all analytes of interest. As an example of the importance of determining this type of information before accessing and using samples from a biobank, we received access to blood samples obtained in an ovarian cancer biomarker study of more than 600 patient samples with full clinical follow-up. However, the biobank also received samples shipped from many hospitals as whole blood by regular surface mail without cooling. Before analyzing this valuable sample set, we tested whether the concentration of our analyte of interest (tissue inhibitor of metalloproteinases-1 (TIMP-1)) would be affected by such a transport procedure. The results showed that our analyte was completely degraded 24 h of storage of whole blood at room temperature, precluding our analysis of this putative biomarker in this set of samples.

Handling and Processing of Samples—Time is an important factor throughout the process of biospecimen collection and processing as protein modifications, proteins, and peptides may be subject to alteration and/or spurious cellular release in vitro after sampling. This occurs among other things from ex vivo cellular injury, disintegration and cellular granule release, and action of proteases. The sooner the separation of plasma from cells occurs, the higher the quality of the resultant sample. This is particularly important for biomarkers that are both circulating and stored in platelets (e.g. vascular endothelial growth factor and TIMP-1).

Blood samples should be centrifuged upon arrival or at completion of clotting for serum within the set gravitational force (g) and time given in the relevant SOP (Table I). The supernatant should be collected, on ice if required, from each of the tubes to obtain serum and plasma as appropri-
ate. The gravitational force, time, and temperature may influence the content of the sample. Different gravitational force yields plasma with different contents of platelets (platelet-rich or platelet-poor) and may thus affect the biomarker level in the sample. We therefore recommend that the influence of gravitational force (g), time, and temperature on a particular analyte (where known) be tested before starting the biosampling. This validation should be performed for any new analyte.

The effect of different anticoagulants on the analyte should be examined. When comparing TIMP-1 levels between different plasma types, we obtained very consistent results when comparing EDTA and citrate plasma samples obtained simultaneously from the same individual, whereas heparin plasma gave variable results (17). When analyzing the level of TIMP-1 in plasma and serum from healthy individuals and from patients, we have observed that serum consistently contains higher values and that these values fluctuate as compared with plasma, which has very stable and reproducible TIMP-1 levels (18). The main explanation for this observation most probably is that because TIMP-1 is stored in the α-granules of platelets it will be released from these cells during ex vivo coagulation when platelets are activated (19). This particular phenomenon may be universal for proteins stored in granules of platelets and/or granulocytes (19–23), although it does not appear for proteins synthesized by specific organs such as liver-derived proteins (24).

Storage of Samples—The method of storage is important for the stability of the sample, ability to share resources, and optimization of storage space. The vials used for aliquot storage should be sealed tightly to prevent lyophilization of the samples; for example cryovials with screw caps, storage straws, or appropriate Eppendorf-style tubes may be used with sizes appropriate to the volume of sample being aliquoted. For practical reasons, some biobanks may acquire samples that have been frozen in large aliquots prior to shipping with subsequent thawing, aliquoting, and refreezing being introduced once at the biobank. This introduces the issue of possible freeze/thaw damage, and aliquoting to useful volumes at the time of the first handling minimizes this occurrence. The importance of this procedure is illustrated by a study on plasma TIMP-1 levels where we found that more than six cycles of repeated freezing and thawing significantly changed the TIMP-1 concentrations (25). Some analytes will not be sensitive to repeated cycles of freezing and thawing (steroids), although other analytes may degrade or precipitate out with as few as one cycle. Such information is needed for each analyte to be tested from biobank material, and an SOP that precludes or minimizes freeze/thaw cycles must be in place to protect these important resources. Information on the effect of storage temperature and duration is also required. Our recommendation is that storage at −70/−80 °C is adopted as a minimum, but it is recognized that some biomarkers may degrade spontaneously during storage even at −80 °C over a prolonged time (11, 26), and storage in liquid nitrogen (−120 °C) may be required for these biomarkers.

Protein degradation during prolonged storage represents a unique problem that may introduce bias when existing biobank resources are applied to future putative biomarker analyses. Often biomarker data are produced from repository materials collected over years. It is mandatory that the reference/control group is derived from a cohort that is not only matched in terms of age, gender, etc. but also source, storage duration, and temperature. If not, this may introduce unanticipated and unrecognized bias unless the differences in handling are systematically investigated and shown not to affect the specific analyte measurements.

Even within a well controlled environment accidents can happen. For instance, the freezer can break down or the current can be interrupted. To guard against the detrimental effects of such events, we suggest that each freezer or liquid nitrogen container containing samples from a biorepository has an electronic device that registers the temperature and reports on any significant change in temperature. This device should be able to alarm the responsible persons 24 h a day, 7 days a week, and all year-round. Surveillance systems that call mobile phones of technicians if the temperature rises are on the market. Additionally mirror banking of samples ensures that if samples are compromised for whatever reason, replicate aliquots of good integrity will still be available. Storage of biosamples in a commercial repository is an alternative possibility, and companies offering such services mostly have well trained people contracted, are International Organization for Standardization-certified, and have all facilities (fireward, temperature registration and checks, and automatic sample retrieval) available. However, the cost of this may be prohibitive for many relatively small biobanks of university hospitals.

Moreover a state-of-the-art laboratory follows the best practices and/or guidelines for biorepositories from the International Society for Biological and Environmental Repositories and Organisation for Economic Co-operation and Development. The suggested measures for safeguarding the collection mentioned in this review are an absolute minimum for those not able to follow the guidelines.

Finally it should be mentioned that an important aspect of constructing a high quality biobank is the availability of a dedicated software package that allows safe, secure, and reliable registration of information regarding the stored samples and process. This system should also enable dependable data retrieval.

CONCLUSIONS

The use of biomarkers is aimed at improving health care of the individual by supporting clinical decision making related to screening, risk evaluation, prognostic stratification, personalized treatment, and monitoring. Inappropriate sampling, transport, handling, and storage of samples as well as errors in data entry, recording, and withdrawal may all cause results

Molecular & Cellular Proteomics 7.10  2065
that are irreproducible and, more importantly, unreliable. Scientific reporting of false biomarker results with linkage to a specific disease stage may not only cause false hopes among patients but may also harm any future research in this particular field. It is therefore recommended that strict guidelines, including SOPs and adherence to good laboratory practice, and assessment by external audits are introduced in all laboratories that bank human biological fluids as a major activity. In addition, it is recommended that whenever data obtained from materials from such a bank are published the study should contain a detailed description of all the parameters that potentially could have influenced the overall results.

Present and future improvements of health care are for a large part based on molecular characterization of disease-related changes and exploitation of these changes in tailored treatments. Both steps, i.e. the characterization and the development of related treatments, have a large dependence on analyses of banked tissue and bodily fluids. Thus, the strict adherence to standardized protocols, an investigation of the stability of molecules of interest (once known) under different conditions, and a thorough reporting of results and possible sources of bias are crucial.

* This work was authored, in whole or in part, by National Institutes of Health staff. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Both authors contributed equally to this work.

 §§ To whom correspondence should be addressed: Dept. of Veterinary Pathobiology, Faculty of Life Sciences, University of Copenhagen, Ribe Banevej 9, 1870 Frederiksberg C, Denmark. E-mail: nbr@life.ku.dk.

REFERENCES