Sample PREanalytical Code for labeling of biospecimens: an analysis of specimen labeling protocols

Silvia Riondino1,*
Umberto Nanni2,*
Fay Betsou3
Luisa Rossetti2
Corrado Campisi4
Raffaele Palmirotta1
Patrizia Ferroni1
Mario Roselli5
Fiorella Guadagni1

1BioBIM-Interinstitutional Multidisciplinary BioBank, IRCCS San Raffaele Pisana – Research Center, Rome, Italy; 2Department of Computer, Control and Management Engineering “Antonio Ruberti”, “Sapienza” University, Rome, Italy; 3Integrated Biobank of Luxembourg, Luxembourg City, Luxembourg; 4LC Campisi Laboratory Ltd, Avola (SR), Italy; 5Department of Systems Medicine, University of Rome “Tor Vergata”, Rome, Italy

*These authors contributed equally to this work

Abstract: The suitability of samples for a given research depends on many variables, among which, the best preanalytical conditions for the type of test that should be carried out. Thus, standardization and harmonization of processing conditions for samples entering a particular study should be highly pursued to facilitate research collaborations among different institutions and biological banks, allowing specimen comparison both for sample sharing and for the development of large-scale clinical trials. In this study, we review some issues and scenarios related to the adoption of the Sample PREanalytical Code, which deserves attention for a proper management of the samples and, ultimately, for maximizing the efficiency of the consistent investments required to set up biobanking resources.

Keywords: biobanking, preanalytical variables, sample labeling

Introduction

Biobanks are defined as operational units that provide a service for the storage and management of biological material (organs, tissues, blood, cells, and fluids having an amount of DNA or RNA that allows genetic analysis) and associated personal data, in accordance with the codes of good practice, privacy law, and ethics guidelines as defined by the European Oviedo Convention.1

The increasing frequency of collaboration between research institutions and biobanks renders the knowledge of a biosample “quality” mandatory. The availability of biological samples has, in fact, assumed a crucial role in the field of biomedical research, from a basic to translational one, to assess predisposition to complex diseases, and identification and validation of new diagnostic biomarkers, drug targets, or the improvement of monitoring strategies. All these tasks critically depend upon the availability of a large number of standardized biological specimens, and many research activities are seriously invalidated by the different methodological approaches employed during sample preparation.2 At present, in fact, the requirements for a specific intended use of the stored biomaterials are often lacking, as well as information on origin and preparation (preanalytical phase), and there are no standardized methods that can be used in comparative approaches between different institutions.3

In this review, we will consider some issues and scenarios related to the adoption of the Sample PREanalytical Code (SPREC), which deserves attention for a proper management of biospecimens and, ultimately, for maximizing the efficiency of the consistent investments required to set up biobanking resources. Moreover, we will describe some information technology (IT) solutions developed by joint efforts within our institutions, aimed at facilitating the adoption of SPREC by all the organizations for managing
biosamples. These functionalities enable interested biological resource centers to establish shared searchable collections of SPRECs among a network of biobanking facilities.

A systematic literature review was performed by searching in PubMed, Scopus, and Web of Science. The inclusion criteria regarded relevant research studies published in English. Furthermore, we extended our search on health care systems due to their similarity with the subject reported. For the selection of the search terms, we referred to previous literature reviews and the keywords of leading papers on the topic of biobanking as provided by the International Society for Biological and Environmental Repositories (ISBER) guidelines.

**Sample quality and methodological pitfalls**

The National Cancer Institute (NCI)/Office of Biorepositories and Biospecimen Research (OBBR) guidelines represent, together with the Biospecimen Research Network Symposium that took place in 2009, the first evidence of the urge to address the issue of biosample quality and of standardized procedures for biospecimen collection, processing, and storage. Members of the previous working groups participated in Biospecimen Reporting for the Improved Study Quality (BRISQ) in 2011, which released recommendations intended for the reporting of data elements of human biospecimens (solid tissues and bodily fluids), running through the whole life cycle of a sample (from collection to analysis). The recommendations gave birth to the BRISQ reporting item list, documenting sample preanalytical variables aimed at harmonizing the methods to trace information and facilitate effective inter- and intralaboratory specimen sharing/use. Furthermore, BRISQ efforts were focused on the standardization and improvement of procedures for tissue processing to assess quality and suitability for research use, to identify quality control tools for tissues stored in biorepositories, and to interpret and validate data from archived specimens. Tissue processing is of the utmost importance for the preservation of both molecular and morphological integrity, and several preanalytical factors have been shown to significantly affect sample quality (eg, postmortem interval; cold ischemia time; specimen size; fixative buffer, delivery method; temperature, and duration; and section thickness and storage) (for an extensive review see).

Concomitant to the development of BRISQ recommendation, the International Society for Biological and Environmental Repositories Biospecimen Science Working Group endorsing the need of a deeper knowledge of the preanalytical conditions of specimens used for research activities generated the SPREC. The SPREC identifies the main preanalytical factors of clinical fluid and solid biospecimens and their simple derivatives, which, together with its accompanying data, represent an essential part of its clinical and scientific value. Accordingly, quality of specimens becomes a multidimensional concept that determines the scientific value and exchangeability potential for health care research. Recording information about a specimen and its processing (eg, defining a set of features that are to be explicitly traced), in fact, gives the opportunity of choosing the right options for the right specimens and, at the same time, the right specimens for the target study (“fitness for purpose”). This concept finds its highest applications in the field of oncology, and applies both for sera and tissue samples.

Thereby, biospecimen processing methods are recognized as a critical issue whose inappropriate application causes detrimental effects which are widely acknowledged. Among the many examples of methodological pitfalls, one may cite the important confounders in the interpretation of soluble CD40 ligand (sCD40L) level measurements due to preanalytical and analytical interferences. Indeed, platelet preactivation during the course of improper sample handling may substantially increase sCD40L levels, as well as keeping samples at room temperature for prolonged periods before processing or freezing. Measurement of serum sCD40L levels is sensitive to all storage-associated preanalytic conditions to such an extent that it has been proposed not only as a decay marker but also for the establishment of standard operating procedures in biobanking.

Other methodological examples come from studies on vascular endothelial growth factor (VEGF), whose levels might be concealed in the presence of thrombocytopenia. To overcome this bias, VEGF levels should always be normalized by platelet counts. Another interesting issue is represented by the use of ethylenediaminetetraacetic acid anticoagulated samples for mean platelet volume measurement, which are biased by ethylenediaminetetraacetic acid-induced platelet changes over time. This effect must be controlled by standardizing the duration between sampling and analysis, thus ensuring a relative homogeneity among the samples used and minimizing the differences in analyses. Furthermore, particular attention should be paid to the choice of samples to be analyzed for the research of lupus anticoagulant in studies on coagulation alterations occurring in conditions of acquired thrombophilia.

Numerous sources of variability have been described for the nonobservance of the basic requirements for an optimal
sample, such as the contamination of the plasma with platelets, the presence of activated platelets, or the concentration of calcium ions.\(^\text{21}\) It is evident that even a nonoptimal pre-treatment (as might occur during the normal steps of plasma preparation), rather than a real error, would render the sample not fit for the purpose of the study.

Finally, several authors emphasize the importance of knowing the status of the subject who has agreed to provide the sample for the study, as some conditions can significantly interfere with the use of certain methods. For example, inconsistent results may be obtained with the use of functional tests of coagulation in carriers of the mutation R506Q FV Leiden and in case with the presence of lupus anticoagulant, increase of FVIII, or presence of factor VIIa.\(^\text{22}\)

However, beside properly defined errors, the suitability of samples for a given research also depends greatly on the type of test that should be carried out. Thus, standardization and harmonization of processing conditions for samples entering a particular study should be highly pursued to facilitate research collaborations among different laboratories, both allowing specimen comparison for sample sharing, and for the development of large-scale clinical trials. In this context, comprehensive guidelines for the best practices for managing biosamples in repositories have been provided by the ISBER guidance on best practices for collection, storage retrieval, and distribution of biological materials for research.\(^\text{23}\)

### Proper sample labeling

The urge to use “talking” codes rather than simple numeric ones stems from the need to mitigate errors in sample identifiers during biosample mislabeling in research laboratories or biorepositories.\(^\text{24–26}\) The biospecimen publication guidelines of the NCI/OBBR\(^\text{1}\) first envisaged the need for more detailed information used for research activities.

As a possible support for this issue, ISBER has proposed and refined the SPREC, a labeling methodology which provides detailed information on the preanalytical conditions that a single stored specimen has encountered during its manipulation.\(^\text{11,27}\) SPREC is a seven-element-long code, where each element corresponds to a punctual preanalytical variable and contains a string of letters (different for fluid or solid samples).\(^\text{11}\) This code can be integrated into the local institutional IT system (such as Biobank Information System [BIS]) and tracking databases so that the SPREC can share the same record and the information contained can be applied to all aliquots of the parent sample. The stand-alone application SPRECbase ensures SPREC output-coding and input-decoding of each specimen. Basic information concerning preanalytical data can be selected from SPREC tables by drop-down menus; the resulting SPREC will be stored in the local database, and can be easily retrieved.\(^\text{28,29}\) Also, 2D barcodes or quick response codes can provide plain text information combined with multiple links to online content (videos, photos, hypertexts). This last coding system, although highly envisaged, would require the purchase of specialized software/application, which might represent a substantial investment both in time and resources.\(^\text{30}\) Such implementation of the SPREC system is already under study in our institutions, and is going to be delivered together with its package and validated in a biological bank environment.

SPREC can be applied both to primary samples (samples directly collected from the donor) and to their derivative tubes, and provides different information on sample withdrawal, processing, and storage for tissues and bodily fluids, all summarized in one single string on a sample label.\(^\text{11}\) Obviously, this will be of particular help in the case of samples resulting from extensive technical manipulations that introduce preanalytical variables related to each laboratory’s standard operating procedures and operators. Indeed, the prevalence of errors in the selection of a sample by a laboratory that processes specimens manually is difficult to estimate. On the other hand, if the biological samples are SPREC-labeled by means of a barcode (as in the case of SPRECbase), any personnel handling the sample (even nontrained operators) can easily and explicitly obtain all the information associated with that preanalytical encoding by means of any reading device.

Since not all research laboratories are equipped with a management system shared with the central “core lab”, an application has been developed (SPRECware), and a server has been set up (SPRECbase), allowing for barcode (or quick response code) reading by means of any code-scanning device: this provides the immediate decoding of the SPREC into the corresponding preanalytical information (Figure 1).\(^\text{28}\)

In this scenario, checking whether a biological sample has characteristics that prevent its inclusion in a given study can be done by reading/scanning the SPREC. Moreover, if the information related to the nature of a biological sample is paired with the exact location within the biological bank, the construction of new studies and national and international collaborations is facilitated, with a remarkable saving of time otherwise spent for manual sample searching, and there is a significant increase of availability in terms of numbers and quality of samples as well as data related to them.
Besides preanalytical variables, the whole biospecimen chain of custody, ie, the sample life cycle, should also be carefully monitored and checked. Indeed, this should occur even when minimizing the amount of time during which samples of a given batch undergo a drastic change in temperature (ie, when the batch is taken from the freezers at −80°C to insert/pick up a cryovial). Tracking temperature observations can undoubtedly help to rapidly identify artefacts during downstream sample assay.  

All samples (eg, sera, plasma, whole blood, urine, tissues) are stored in freezers at −80°C equipped with systems that allow the temporal marking of the stages of insertion/withdrawal of the samples themselves. These systems, which may include radio-frequency identification (RFID)31 or bluechip® technology (a passive wireless technology based on micro-electromechanical systems),32 can overcome the possible problems stemming from the need to maintain the cold chain. RFID is a technology that uses communication via radio waves to exchange data between a reader (interrogator) and an electronic tag attached to an object (label), for the purpose of identification and tracking.31 The cryotag uses a transmission contactless passive type and can be placed on the cryovial, on the wall of plates, on cryoboxes, and on the freezer door. Electronic information is then written on a portable remote writer/reader device that allows multiple and contemporary readings of RFID supports, and can associate a time record to every operation performed. Moreover, if samples are also mapped by software, this will allow their immediate localization and recognition, and will minimize the sample temperature-related alterations. Indeed, as an example of the detrimental effects of repeated freeze–thaw cycles,33,34 it has been reported that even one freeze–thaw cycle is capable of leading to substantial changes in the concentrations of serum analytes.35 The use of a RIFD-implemented SPREC will allow the operators to find and collect in the shortest possible time the required samples that might otherwise be exposed to sudden changes in temperature and consequent decay. It is obvious that this could ultimately lead to the acquisition of erroneous results due to no-longer-suitable samples. This is the case of plasma samples which, once frozen at −80°C, should be rapidly thawed at 37°C to prevent denaturating fibrinogen.36 The analysis of a sample not properly thawed would provide unreliable results.21

In general, the availability of samples of different quality levels in a biobank – if known and properly managed – can result in the optimization of costs and in the maximization of the outcome. Several procedures in sample preparation and maintenance have an intrinsic relevant cost. A crucial point is to avoid expensive procedures for specimens, which, at some point in their early life cycle, have been managed out of the best practices. In other words, the resulting final quality of the sample usually does not only depend on the last step, but rather on the “worst stage” over the whole life cycle.

In this scenario, a SPREC label which allows one to easily obtain most of the information associated with that preanalytical encoding can be very useful for:

- avoiding degradation of high-quality specimens due to a poor management;
- avoiding expensive handling procedures for specimens that do not fulfil the required quality criteria;
- selecting specimens which are suitable for inclusion in specific biomarker discovery projects.

For an easier selection of the suitable specimens, once the inclusion criteria have been defined, sample searching can be performed by partial “pattern-matching”. In other words, the user may specify only some of the fields (eg, type of sample, pre-centrifugation conditions, or elapsed time from sample acquisition), and the SPRECbase application would filter and list all the samples complying to such partial specification.28

### Transfer of SPREC-labeled samples between laboratories/institutions

As part of the collaboration between research institutions, it is very frequent that a laboratory should send a large amount of samples to another laboratory. This may be the case in retrospective studies involving the analysis of a large segment of the population with special characteristics or belonging to a temporal situation or to a specific geographic area.37

---

**Figure 1** Generating sample Sample PRAnalytical Code, with the corresponding barcode and quick response code.
or of multicenter studies designed to determine the incidence and clinical relevance of a given condition.

In the event of such large enlistments, it is not sufficient that the biological samples are provided with identifiers and paper documents containing the demographic information, medical history, and methodological description, since the amounts of data having to be consulted would be so elevated that it would expose the personnel of the “receiving” laboratory to manual mistakes during the compilation of the database with sample-related information.

Conversely, if the laboratory that sends samples uses SPREC coding, then the laboratory which receives the samples can easily decode the SPREC of the sample batch and subsequently store this code on the local installation of SPRECbase. In this way, the transfer of samples and that of an integrated database with all the necessary pre-analytics information can be achieved simultaneously, and the event of accidental mistakes due to manual handling of the information can be minimized. It is also very important to emphasize that the information related to a particular batch of samples will always be available so that the sample data can be retrieved at any time, facilitating the use of the samples for future scientific collaborations and allowing the update of the existing ones according to the new results.

If the facility receiving the SPREC-labeled samples does not adopt SPREC for internal purposes, it can still get interesting information about the samples by scanning the SPREC labels and decoding them, restoring the original preanalytical information in its own laboratory information system (Figure 2).

**SPREC-based sample retrieval within a network of laboratories**

This scenario applies well to all those trials that require the involvement of several research institutes due to either the objective difficulties of recruitment or the need for a multicenter approach for external data validation. Both situations, requiring the exchange of large amounts of samples, allow for the translation of relevant findings into clinical practice. For this reason, it is desirable to establish a network of institutions sharing a labeling coding system in which context the availability and location of even few samples with given characteristics could be easily identified, thanks to the ability to consult a privileged common database (Figure 3).

The purpose of this desirable scenario would be the establishment of a SPREC database shared among a network of biorepositories. We can assume that SPREC-labeled samples are imported from every laboratory pertaining to the network where biological samples are made available to the entitled operators. This would allow for certain advantages in the planning of clinical trials: to avoid unnecessary duplication of research, to not wasting precious samples, to foster collaboration between researchers, and to optimize the distribution of research funds.

**Future perspectives**

The SPREC-coding system can be applied at various levels from handwritten to “high-tech” facilities to summarize both preanalytical and relevant post-analytical information. Moreover, SPREC has infinite applications, and in many
different, not strictly clinical, environments. Presently, it has been successfully applied to algal culture collections, proving optional coding from the moment of sample collection to post-cryostorage manipulations. Future applications include its exploitation for a microbial culture bank within the Multidisciplinary Interinstitutional Biobank (BioBIM) facilities.

Thus, the “SPRECware architecture” previously described and here reported, does not represent an end point, but is susceptible to further implementation to increase its easiness to use and widespread diffusion among researchers. Preliminary results have been obtained in a pivotal study specifically designed to facilitate the exchangeability of information using a web-based platform that does not require any software installation, and tests are being performed between the BioBIM of the IRCCS San Raffaele Pisana in Rome (the “sender” laboratory) and the LC Campisi Laboratory Ltd in Sicily (the “receiver” laboratory). This has been made possible by the establishment of a permanent centralized web server, which is being implemented with the following features:

- SPREC-coding and -decoding servers with no installation requirements;
- SPREC-sharing server allowing enabled users to formulate queries with search/filtering capabilities;
- development of a downloadable version of SPRECbase.

These novel functionalities will ultimately provide the users with easy-to-use solutions for upload/download capabilities to/from shared SPREC collections on a centralized server. Implementation with novel IT tools will also allow adjustments and timely updates to keep the SPREC’s optimal use in line with future technological developments.

Author contributions
All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

Acknowledgments
This study was partially supported by research funding from the European Regional Development Fund PO FESR 2007/2013 Linea di Intervento 4.1.1.1 – SIASOP (CUP G83F11000290004), the European Social Fund, under the Italian Ministry of Education, University and Research PON03PE_00146_1/10 BIBIOFAR(CUP B88F12000730005), and Lazio Region, with financial resources from the Italian Ministry of Economic Development, Decree May 7, 2010.

Disclosure
The authors report no conflicts of interest in this work.

References


