Published in final edited form as:

Arch Pathol Lab Med. 2014 April ; 138(4): 526–537. doi:10.5858/arpa.2013-0250-CP.

A Call to Standardize Preanalytic Data Elements for Biospecimens

James A. Robb, MD, Margaret L. Gulley, MD, Patrick L. Fitzgibbons, MD, Mary F. Kennedy, CT(ASCP), MPH, L. Mark Cosentino, DPM, Kay Washington, MD, Rajesh C. Dash, MD, Philip A. Branton, MD, Scott D. Jewell, PhD, and Rosanna L. Lapham, MD

Department of Pathology, University of North Carolina, Chapel Hill (Dr Gulley); the Department of Pathology, St Jude Medical Center, Fullerton, California (Dr Fitzgibbons); the College of American Pathologists, Northfield, Illinois (Ms Kennedy); the Department of Translational Medicine, Biogen Idec, Cambridge, Massachusetts (Dr Cosentino); the Department of Pathology, Vanderbilt Medical Center, Nashville, Tennessee (Dr Washington); the Department of Pathology, Duke Medical Center, Durham, North Carolina (Dr Dash); the Biorepositories and Biospecimens Research Branch, National Cancer Institute, National Institutes of Health, Rockville, Maryland (Dr Branton); the Department of Biospecimen Science, Van Andel Institute, Grand Rapids, Michigan (Dr Jewell); and the Department of Pathology, Spartanburg Regional Medical Center, Spartanburg, South Carolina (Dr Lapham)

Abstract

Context.—Biospecimens must have appropriate clinical annotation (data) to ensure optimal quality for both patient care and research. Clinical preanalytic variables are the focus of this study.

Objectives.—To define the essential preanalytic variables (data fields) that should be attached to every collected biospecimen and to provide a complete list of such variables, along with their relative importance, which can vary, depending on downstream use, institutional needs, and information technology capabilities.

Design.—The College of American Pathologists Diagnostic Intelligence and Health Information Technology Committee sponsored a Biorepository Working Group to develop a ranked list of the preanalytic variables for annotating biospecimens. Members of the working group were experts in anatomic, clinical, and molecular pathology; biobanking; medical informatics; and accreditation. Several members had experience with federal government programs, such as the National Cancer Institute's Biospecimens and Biorepository Branch and the National Cancer Institute's Community Cancer Center Program. Potential preanalytic variables were identified and ranked along with available supporting evidence, definitions, and potential negative effects if the variable was not attached to the biospecimen. Additional national and international stakeholders reviewed the draft manuscript.

Results.—The ranked listing of 170 preanalytic variables produced can be used as a guide for site-specific implementation into patient care and/or research biorepository processes.

Corresponding author: James A. Robb, MD, 11613 Kensington Court, Boca Raton, FL 33428-2415 (jarobb33@yahoo.com).

Dr Robb is a consulting pathologist in Boca Raton, Florida. The authors have no relevant financial interest in the products or companies described in this article.

Conclusions.—In our collective experience, it is often difficult to choose which of the many preanalytic variables to attach to any specific set of biospecimens used for patient care and/or research. The provided ranked list should aid in the selection of preanalytic variables for a given biospecimen collection.

The College of American Pathologists (CAP) is committed to helping pathologists and the medical specialty of pathology successfully implement genomically informed, personalized, precision health care. As part of that mission, the CAP published a practical resource¹ for community pathologists. That publication promotes the creation of high-quality biospecimens and biorepositories and supports the emerging opportunities in the biorepository field.¹ Definitions of a biospecimen and a biorepository are given in that article. Discussion of the appropriate preanalytic variables needed for the creation of high-quality biospecimens was not within the scope of the previous study; however, it is the subject of the current study.

Successful implementation of genomically informed, precision, personalized health care requires the availability of high-quality biospecimens. These biospecimens are fundamentally necessary for the production of accurate downstream test results in clinical or research settings. These results, in turn, form the basis for determining optimal patient management and research outcomes. A critical ingredient in creating high-quality and high-value biospecimens is the clinical and processing information attached (annotated) and accrued to each biospecimen. Currently, only limited and/or relatively general recommendations are available as to what preanalytic information should be collected and attached to each biospecimen for clinical care and/or research.^{2–16} As a result, there is great variability in the collected clinical information, as well as in the quality of that information. A dynamic list of evidence-based and/or expert opinion-defined, preanalytic, clinical data fields should be useful for both patient care and research biospecimen annotation. The ability to retrospectively reassess or filter the overall quality of the biospecimen from well-annotated data is an additional benefit that the list provides. That use will be necessary when new discoveries emerge for determining how key parameters (variables) relate to or affect the quality of the previously collected biospecimen.

Although not every biorepository is engaged in the direct performance of laboratory testing, every organization that collects patient and/or research biospecimens must be the custodian of their biospecimens, whether the custodianship is short or long. Examples of the periods associated with custodianship include hours to days for biospecimens that are being transferred to another site for testing and/or storage; 1 to 2 weeks for clinical pathology biospecimens, such as blood, body fluids, or microbiology material; weeks to years for biospecimen-derived products, such as DNA and RNA; and years for formalin-fixed, paraffin-embedded tissues and their associated glass and digital slides, including immunohistochemistry, in situ hybridization, and genomic test results. That custodianship is crucial for the creation and maintenance of high-quality biospecimens for both patient care and research. Biospecimens collected today may need to be used for clinical care and/or research in the future. Recapturing preanalytic data elements not recorded at the time of collection, if possible, would be difficult.

MATERIALS AND METHODS

The CAP Diagnostic Intelligence and Health Information Technology Biorepository Work Group members were selected for their domain expertise in the areas of anatomic, clinical, and molecular pathology; biospecimen consenting, collecting, transporting, and processing; biorepository operations; biorepository accreditation; and information technology infrastructure support in both academic and community hospitals (see Appendix A). Their input was voluntary, and no conflicts of interest were identified that were specific to the topic of this activity.

We defined *preanalytic variables* as those variables that may affect the quality and/or value of a biospecimen from the time of consenting until the biospecimen is used for testing or biobanked for future testing, which included all biospecimens used either in direct patient care or in research or both. The *quality of a biospecimen* was determined by its suitability for the purpose or purposes for which it was collected. Currently, messenger RNA (mRNA), phosphorylated proteins, and peptides, such as cytokines, require the most stringent quality parameters because the technologies used to evaluate those targets are finicky (not robust) for biospecimen preparation variables. The *value of a biospecimen* has at least 3 components: (1) *the intrinsic value* of the biospecimen itself for testing; without the correct, properly collected and preserved biospecimen, accurate and appropriate testing cannot be done; (2) *the information value* associated with the biospecimen, especially for research because the more information about the “life history” of the biospecimen and the patient’s outcomes in response to particular interventions, the more valuable is the biospecimen; and (3) *the commercial value* of the biospecimen, which depends on the quality of its intrinsic and information values. Other biospecimen-specific values may be attached or accrue to particular biospecimens (eg, rare-disease biospecimens).

Potential variables were collected from the listed references and from personal experience. Five types of preanalytic variables were listed: (1) common variables; (2) tissue-specific variables; (3) blood-, serum-, and plasma-specific variables; (4) body fluid-specific variables; and (5) nucleic acid (DNA, RNA)-specific variables (see Tables 1–5). Although many of the listed variables pertain to other types of biomolecules, such as proteins, peptides, and metabolites, those biomolecules were considered to be outside the scope of this project.

Each member who scored the data variables used a semiquantitative scoring process to derive the priority for each variable. The categories were 1, required field; 2, clinically important, but the importance for biospecimen quality is not yet validated; 3, not necessary; and 4, parking lot for future assessment. The scores were added for each variable and divided by the number of scoring members to provide the average priority value in the spreadsheet. The variables were sorted by their priority score for each category.

Definitions and notes were supplied for all variables to reduce the variability in interpretations and to clarify the nature of each variable. The work group discussed negative effects, and 11 categories were chosen (Table 6). These negative-impact categories describe

what could happen if the variable is not collected and attached to the biospecimen. Specific examples are provided in Table 6.

Members of the CAP Diagnostic Intelligence and Health Information Technology Committee and other national and international domain experts reviewed the manuscript before submission (see the “Acknowledgements”).

RESULTS

By priority of their importance, 170 preanalytic variables were ranked in 5 categories: (1) common variables; (2) tissue-specific variables; (3) blood-, serum-, and plasma-specific variables; (4) body fluid variables; and (5) nucleic acid variables. Definitions and notes were given for all variables to help unify the meaning of the variables. One or more negative-impact categories were provided for each variable as a guide to what might occur if the specific variable was not collected and attached to the biospecimen.

COMMENT

Very little field-validated research is available about exactly what clinical, preanalytic variables should be attached (annotated) to a particular biospecimen. Although we initially tried to identify specific evidence for the recommended variables, there was too little literature support to be useful. Some efforts with a limited scope have been attempted, perhaps, most notably among them, the National Cancer Institute’s Common Biorepository Model.¹⁷ With high-level administration in mind, the Common Biorepository Model focuses on enumerating those data elements that help to identify broad categories of samples that are likely to be most relevant for general oversight (eg, answering how many samples of a certain type are present). Data elements that span anatomic source, diagnosis type, preservation types, biospecimen types, race, ethnicity, and sex are defined. Furthermore, coding of more than 3000 Common Biorepository Model concepts using SNOMED CT (Systemized Nomenclature of Medicine Clinical Terms, reference terminology for medical concepts; <http://ihtsdo.org>, Accessed July 26, 2013) was performed where applicable. Such coding is important for information system interoperability and will be the subject of ongoing efforts for the data elements proposed in this article (particularly because many of the proposed data elements currently have no direct mapping to an existing standard terminology). Specific value sets for each data element are listed. The authors propose that our current effort builds on existing efforts, such as the Common Biorepository Model, whenever possible. Therefore, we used national and internationally available guidelines and recommendations in this area,^{2–16} as well as an early implementation project.¹⁸ This information was then combined with the expert opinions of our working group and reviewers. Through this process, we created this first-of-a-kind, semiquantitative listing of recommended preanalytic variables for recording and annotating biospecimens. Clearly, this list is dynamic and will need periodic review and revision as further experience and supporting evidence become available.

The complete list is not suitable for all institutions. We emphasize that no single, universally applicable list can be generated. The choice and implementation of the variables will depend

on the needs of the institution and the capability of the institution's information technology infrastructure to support the collection of chosen variables, as well as on the intended uses of each biospecimen. Currently, much of the recommended annotation is being captured and entered manually, a process that is often difficult and time consuming. The future use of biospecimens in research, however, as well as the use of biospecimens in follow-up clinical testing (eg, genomic assays), will be more dependent on the need for quick searching and thorough analysis for associated preanalytic variables. We strongly support the extensive use of information technology implementation to automate the collection of clinical variables and subsequent annotation of the biospecimens with those variables as well as with biospecimen processing variables (eg, quality indicators). In most instances, manual collection with subsequent annotation is simply not an alternative.

As potential negative effects were discussed, it became clear that the risk of potential adverse events occurring downstream would increase if a variable was not attached to the biospecimen. Table 6 lists the 11 negative-impact categories and includes some specific examples.

The contemplated next steps include developing functional specifications for how these enumerated variables might be implemented to achieve measurable improvements in biospecimen quality and patient care. These specifications are the subject of follow-up studies. The CAP Diagnostic Intelligence and Health Information Technology project team will begin practical implementation discussions with the larger scientific community, with information technology vendors, and with other stakeholders to help communicate what the functional specifications of the preanalytic variables represent and why input is needed from a wide range of interested parties. Measurement units for the preanalytic variables, where appropriate, must be part of that effort.

Feedback is welcomed on the enumerated variables as well as the prioritization of variables for various clinical and research settings. This effort represents a beginning for a longer, more-involved process. The views and recommendations are the opinions of the authors and are not to be construed as official CAP policy, guidelines, or recommendations. Any comments or questions can be sent to informatics@cap.org.

There is tremendous benefit from explicitly proposing a set of standard data elements relevant to biospecimen quality in an era of personalized medicine. Although much of the current effort is based on expert opinion, our hope is that the evolution of this effort will crystallize those variables that have the most effect on health care outcomes. The authors hope that relevant stakeholders will join and engage in an effort to collate hard data on the utility of the proposed elements to concretely identify the most relevant (and practically implementable) variables. Finally, significantly more biospecimen research is needed to develop supportive evidence for the relative importance of the various preanalytic variables.

Acknowledgments

We thank the following domain experts for their valuable review of the manuscript. These experts have experience with activities similar to those described in this article. Their input was based on personal experience and does not necessarily represent the view of the organizations with whom they are affiliated: Victoria Blanc, PhD—experience with the College of American Pathologists' Biorepository Accreditation Committee, Office of Research, University

of Michigan Medical School, Ann Arbor; Sarah Dry, MD—experience with the College of American Pathologists' Biorepository Accreditation Committee, Department of Pathology and Laboratory Medicine, University of California, Los Angeles; Latarsha Carithers, PhD, and Angela Britton, MS—experience with the National Cancer Institute's Biorepositories and Biospecimen Research Branch, Rockville, Maryland; and Fay Betsou, PhD, HDR—experience with the International Society for Biological and Environmental Repositories and Integrated BioBank of Luxembourg.

APPENDIX A

Members of the College of American Pathologists (CAP) Diagnostic Intelligence and Health Information Technology (DIHIT) Biorepository Work Group (WG): Jewell - "Director of Core Technologies, Professor and Director, Program for Biospecimen Science"; Washington - "Professor of Pathology"; Fitzgibbons - "Staff Pathologist"; Kennedy - "Manager, Strategic Health IT Initiatives"; Branton - "Consulting Surgical Pathologist (Contractor)"; Lapham - "Staff Pathologist"; Gulley - "Professor of Pathology and Laboratory Medicine."

David L. Booker, MD, Department of Pathology, Augusta Hospital, Augusta, Georgia

Philip A. Branton, MD, Biorepositories and Biospecimen Research Branch (BBRB), National Cancer Institute, National Institutes of Health, Rockville, Maryland

Mark Cosentino, DPM, Biorepository Project Manager, SAIC-Frederick, Frederick, Maryland; as of January 2013, Dr Cosentino is the Biobank Director, Department of Translational Medicine, Biogen Idec, Cambridge, Massachusetts

Rajesh C. Dash, MD, Department of Pathology, Duke Medical Center, Durham, North Carolina

Denise Driscoll, MS, MT(ASCP), SBB, College of American Pathologists, Northfield, Illinois

Patrick L. Fitzgibbons, MD, Department of Pathology, St Jude Medical Center, Fullerton, California

Margaret L. Gulley, MD, Department of Pathology, University of North Carolina, Chapel Hill

Scott D. Jewell, PhD, Director of Biospecimen Science, Van Andel Institute, Grand Rapids, Michigan

Mary F. Kennedy, CT(ASCP), MPH, College of American Pathologists

Rosanna L. Lapham, MD, PhD, Department of Pathology, Spartanburg Regional Medical Center, Spartanburg, South Carolina

Sandi Larsen, MT(ASCP), MBA, College of American Pathologists

James A. Robb, MD, Chair, Consulting Pathologist, Boca Raton, Florida

James Sorace, MD, Senior Medical Officer, Office of Science and Data Policy, Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services, Washington, DC

Kay Washington, MD, Department of Pathology, Vanderbilt Medical Center, Nashville, Tennessee

References

1. Dash RC, Robb JA, Booker DL, Foo WC, Witte DL, Bry L. Biospecimens and biorepositories for the community pathologist. *Arch Pathol Lab Med*. 2012; 136(6):668–678. [PubMed: 22646276]
2. Wolff AC, Hammond ME, Schwartz JN, et al.; American Society of Clinical Oncology; College of American Pathologists. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for epidermal growth factor receptor 2 testing in breast cancer. *Arch Pathol Lab Med*. 2007;131(1):18–43. [PubMed: 19548375]
3. Hammond ME, Hayes DF, Wolff AC. Clinical notice for American Society of Clinical Oncology-College of American Pathologists Guideline Recommendations on ER/PgR and HER2 testing in breast cancer [update letter to Reference 2]. *J Clin Oncol*. 2011;29(15):e458. doi:10.1200/JCO.2011.35.2245. [PubMed: 21502545]
4. Hammond MEH, Hayes DF, Allred DC, et al.; American Society of Clinical Oncology; College of American Pathologists. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *Arch Pathol Lab Med*. 2010;134(7):e48–e72. doi: 10.1043/1543-2165-134.7.e48. [PubMed: 20586616]
5. Gelber RD, Gelber S; for the International Breast Cancer Study Group (IBCSG) and the Breast International Group (BIG). Facilitating consensus by examining patterns of treatment effects. *Breast*. 2009;18(suppl 3):S2–S8. [PubMed: 19914537]
6. Hewitt SM, Robinowitz M, Bogen SA, et al. Quality Assurance for Design Control and Implementation of Immunohistochemistry Assays: Approved Guideline. 2nd ed. Wayne, PA: Clinical and Laboratory Standards Institute; CLSI document I/LA28-A2; 2011; vol 3, no. 4.
7. Rainen L, Arbiq JC, Asthana D, et al. Collection, Transport, Preparation, and Storage of Specimens for Molecular Methods; Approved Guideline. Wayne, PA: Clinical and Laboratory Standards Institute; CLSI document MM13-A; 2005; vol 25, no. 31.
8. Kiechle FL, Betsou F, Blakeney J, et al. Procedures for the Handling and Processing of Blood Specimens for Common Laboratory Tests; Approved Guideline. 4th ed. Wayne, PA: Clinical and Laboratory Standards Institute; CLSI document H18-A4; 2010; vol 30, no. 10.
9. Rabinovitch A, Arzoumanian L, Curcio KM, Dougherty B, Halim A-B. Urinalysis; Approved Guideline. 3rd ed. Wayne, PA: Clinical and Laboratory Standards Institute; CLSI document GP16-A3; 2009; vol 29, no. 4.
10. Ernst DJ, Balance LO, Calam RR, et al. Procedures for the Collection of Diagnostic Blood Specimens by Venipuncture; Approved Standard. 6th ed. Wayne, PA: Clinical and Laboratory Standards Institute; CLSI document H03-A6; 2007; vol 27, no. 26.
11. Berte LM, Ball JE, Charity KS, et al. Quality Management System: A Model for Laboratory Services; Approved Guideline. 4th ed. Wayne, PA: Clinical and Laboratory Standards Institute; CLSI document QMSO1-A4, formerly GP26-A4; 2011; vol 31, no. 15.
12. Lehmann S, Guadagni F, Moore H, et al.; International Society for Biological and Environmental Repositories Working Group on Biospecimen Science. Standard preanalytical coding for biospecimens: review and implementation of the sample preanalytical code (SPREC). *Biopreserv Biobank*. 2012;10(4): 366–374. [PubMed: 24849886]
13. International Society for Biological and Environmental Repositories. 2012 Best practices for repositories: collection, storage, retrieval, and distribution of biological materials for research. *Biopreserv Biobank*. 2012;10(2):79–161. [PubMed: 24844904]

14. Mohanty SK, Mistry AT, Amin W, et al. The development and deployment of common data elements for tissue banks for translational research in cancer—an emerging standard based approach for the Mesothelioma Virtual Tissue Bank. *BMC Cancer*. 2008;8(1):91–109. doi: 10.1186/1471-2407-8-91. [PubMed: 18397527]
15. [NCI] National Cancer Institute. NCI best practices for biospecimen resources. Office of Biorepositories and Biospecimen Research <http://biospecimens.cancer.gov/practices/2011bp.asp>. Updated 2011. Accessed May 14, 2013.
16. [NCI] National Cancer Institute. Common biorepository model (CBM). [https://wiki.nci.nih.gov/display/TBPT/Common+Biorepository+Model+\(CBM\)](https://wiki.nci.nih.gov/display/TBPT/Common+Biorepository+Model+(CBM)). Updated November 21, 2012. Accessed May 14, 2013.
17. [NCI] National Cancer Institute. Common biorepository model vocabulary [Excel file]. https://ncisvn.nci.nih.gov/WebSVN/filedetails.php?repname=common_biorepository_model&path=%2Ftrunk%2Fdatabase%2Fdata%2FCBM_VocabularyTerms.xls. Updated March 9, 2011. Accessed May 14, 2013.
18. Brooks E, Wease K, Wiles K, Washington K. Configuring and implementing an error reporting feature in a donor repository procurement application. In: *Proceedings of the ISBER 2013 Annual Meeting: Turning the World Upside Down—Emerging Perspectives on Biorepositories*; May 5–9, 2013; Sydney, NSW, Australia Abstract RM16.
19. [NCI] National Cancer Institute. GTEx donor eligibility criteria form. PM-0003-F4. [http://biospecimens.cancer.gov/resources/sops/GTex_SOPs/I/GTex%20Donor%20Eligibility%20Criteria%20Form%20\(PM-0003-F4%20v1.1.0\).pdf](http://biospecimens.cancer.gov/resources/sops/GTex_SOPs/I/GTex%20Donor%20Eligibility%20Criteria%20Form%20(PM-0003-F4%20v1.1.0).pdf). Updated April 19, 2012. Accessed May 14, 2013.
20. [NCI] National Cancer Institute. NCI thesaurus. <http://ncit.nci.nih.gov>. Version 13.05d. Updated May 14, 2013. Accessed May 14, 2013.

Table 1.**Prenalytic Variables—Common to All Biospecimens**

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact, If Data Missing ^b
Aliquot (sample) fresh: time/date	1.0	Documentation of the month/day/year and time that a fresh aliquot (sample) is prepared from a biospecimen: tissue, plasma, serum, urine, pleural fluid, etc.	8, 9, 2
Was donor alive, dead, or an infused transplant donor when the biospecimen was collected? Or "Status of donor at time of biospecimen collection (alive, dead, infused transplant donor)"	1.0	The oxygenation status of the donor is very influential on the molecular integrity of a biospecimen.	8, 9, 2
Biospecimen source site	1.0	The body site from which the biospecimen is collected: body location, organ, arm from which blood drawn, cavity from which fluid is drawn, etc.	1, 3
Biospecimen type	1.0	Tissue, blood, urine, or other human-derived biological material. A single biopsy may generate several biospecimens, including multiple paraffin blocks or frozen biospecimens. A biospecimen can include subcellular structures (DNA), cells, tissue (eg, bone, muscle, connective tissue, and skin), organs (eg, liver, bladder, heart, and kidney), blood, gametes (sperm and ova), embryos, fetal tissue, and waste (urine, feces, sweat, hair and nail clippings, shed epithelial cells, and placenta). Portions or aliquot (samples) of a biospecimen are referred to as samples, eg, fine-needle aspiration, tissue, whole blood, urine. Do not confuse with the diagnosis used for the collection of the biospecimen (eg, cancer, normal).	1, 3
Biospecimen-derived sample type	1.0	The type of sample derived from the primary biospecimen; eg, an excisional biopsy of a breast cancer will typically be used to create the formalin-fixed, paraffin-embedded blocks of tissue, which are samples of the biospecimen. These samples may be used to create histologic, immunohistologic, or in situ hybridization slides, and possibly, cytologic touch preparations. Nucleic acids, such as DNA and RNA, may be extracted and are additional samples of the original "biospecimen." Frozen tissue samples may also be created.	1, 3
Centrifugation: time/date	1.0	Documentation of length in minutes and the month/day/year that a biospecimen is centrifuged.	8, 9
Collection: time/date biospecimen removed from patient	1.0	Documentation of the hour/minute and month/day/year that a biospecimen is taken from the body.	8, 9
Field(s) for linking other types of pertinent information	1.0	An identified area (eg, database column, URI, etc) with the ability to store or link accessible information (eg, diagnostic reports and applicable test results, such as, but not limited to, APand CP, IHC, ISH, genomic, radiologic, etc).	5, 4
Sex	1.0	The state of being biologically male or female. If this information is not available, social and/or cultural "gender (sex)" may be used.	5
Handling/storage history, if not in an SOP	1.0	The documentation of the tracking of the biospecimen from collection time through storage history if SOP documentation is not available. If this information is in an SOP, provide link to the SOP.	8, 9, 2
Infectious disease present	1.0	Use GTEx exclusion list. ¹⁹	5
Pathology accession no.	1.0	The identifier used for each separate pathology biospecimen, sample, or case.	7, 6
Preservation agent	1.0	Use of chemical agents, alterations in environmental conditions, or other means during processing to prevent or retard biologic or physical deterioration of a biospecimen [Source: ISBER, ¹³ 2012] (eg, formalin, liquid nitrogen, proprietary preservative, etc).	8, 9
Preservation: time/date biospecimen into preservative	1.0	Documentation of the month/day/year and hour/minute that a biospecimen is placed into a preservative.	8, 9

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact, If Data Missing ^b
Receipt at intermediate storage site: time/date	1.0	Documentation of the month/day/year and hour/minute that a biospecimen is received at an intermediate location between collection and final receipt at the biobank.	6, 8, 9
Release from intermediate storage site: time/date	1.0	Documentation of the month/day/year and hour/minute that a biospecimen is released from an intermediate storage site between collection and final receipt at the biobank (eg, collection of blood and tissue that is preserved in vapor-phase LN ₂ for 2 wk before being sent to the biobank).	6, 8, 9
Shipping: time/date of preparation	1.0	Documentation of the month/day/year and hour(s)/minute(s) needed to prepare a biospecimen for shipment.	8, 9
Submitting facility name	1.0	Information necessary to identify the origin of the facility submitting the biospecimen.	6
Temperature deviations during storage	1.0	Documentation of fluctuations in the storage temperature of biospecimens that deviate from the expected norm or specific SOP.	8, 9
Cold ischemia time	1.1	The documented period in hour(s)/minute(s) between stoppage of the blood supply to the biospecimen, or removal from the patient (eg, core biopsy, blood, urine), and the preservation of the biospecimen, including dissection of the biospecimen if necessary. This period can begin while the organ is still in the patient or after it has been removed from the patient, if the organ is to be used for transplantation. [Source: NCI Thesaurus, ²⁰ 2013]	8, 9
Diagnosis	1.1	The primary pathology diagnosis used for the decision to collect the biospecimen is preferred to the clinical diagnosis.	5
Biobank receipt of biospecimen time/date	1.2	Documentation of the month/day/year and hour/minute that the biospecimen is received at the biobank.	4
Comment: deviation from SOP	1.2	Documentation of a process performed that is deviant to an SOP.	8, 9
Receipt at destination: condition	1.2	The condition of the biospecimen at the time of opening the shipping container and examining the biospecimen (eg, frozen biospecimen is thawed, blood tube is broken, shipping temperature within the acceptable range).	8, 9, 2
Aliquot (sample) fresh: no. of aliquots	1.3	Record of the number of fresh, separate samples derived from the same biospecimen (eg, tubes, vials). This information is necessary to ensure all prepared aliquots are properly stored, in order to reduce the risk that one or more of the aliquots is improperly stored.	2, 10
Aliquot (sample) fresh: volume	1.3	Volume of the aliquot (sample[s]) prepared.	2, 10
Aliquot (sample) frozen: time/date of refreezing	1.3	The minute/hour and month/day/year that a biospecimen is refrozen.	8, 9
Biospecimen disposal request	1.3	Solicitation by an authorized entity to discard a biospecimen. Link to the request, if possible.	6
Collection site: name and location for national or international site	1.3	The institution that collected the biospecimen. If the biospecimen is sent to an offsite institution for temporary storage, the institution's information should be included, using a separate field.	5, 6
Ethnicity	1.3	Pertaining to a race (<i>Oxford English Dictionary</i>).	5
Label (all primary biospecimens and derived products must have both a machine-readable and human-readable label if physically possible)	1.3	Any biospecimen (primary and derived) should have a human-readable ID, and if possible, the same ID in a bar-coded version (1-D or 2-D). Whichever ID is used, it needs to be random and never contain any patient-relevant identifiable information. The human-readable version is to assist staff in quickly confirming the ID of a biospecimen if a scanner is not readily available or if a receiving facility does not have/use bar-coding capabilities. Some facilities can only read a 1-D barcode, even though 2-D barcodes are better for redundancy. A repository must always consider the technical capabilities of the site that is going to receive the biospecimen. Repositories need to be forward thinking but need to realize that not everyone may be able to afford barcode scanning. Some facilities now place both 1-D and 2-D barcodes, as well as a human-readable ID, on the label.	3

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact, If Data Missing ^b
Prepreservation transportation time	1.3	Documentation of the transport hour(s)/minute(s) of a biospecimen before placement into a preservative, if not immediately preserved, eg, transport of a breast excisional biopsy between the surgicenter and the pathology department. Although this period is part of the cold ischemia time, it may affect the quality of the biospecimen, eg, a significant and/or unexpected delay in the delivery of the breast biopsy above.	8, 9
Provide link to any variables in the biospecimen collection and processing SOPs that are not listed in the already defined variables	1.3	The ability to retrieve information and definitions of data elements not included in the provided SOPs.	8, 9
Race	1.3	A group of persons related by common descent or heredity.	5
Shipping: maximum temperature during preparation	1.3	The maximum temperature of the biospecimen or sample during one or more of the steps necessary to prepare the biospecimen or sample for shipment (eg, the maximum temperature reached by an aliquot [sample] of DNA that has been stored in the vapor phase of LN ₂ and is prepared to ship in a LN ₂ dry shipper). Ensure the proper scale of measurement is recorded.	8, 9
Shipping: maximum temperature	1.3	If a continuous temperature-recording device is not available or practical, an inexpensive maximum temperature-recording device should be included in every shipment. Ensure the proper scale of measurement is recorded.	8, 9
Shipping: method	1.3	The method in which a biospecimen is shipped (eg, vapor-phase LN ₂ dry shipping, dry ice, ice pack, ambient, etc).	8, 9
Storage temperature	1.3	The temperature at which the biospecimen is stored, including the duration of any periods above or below the desired temperature. Ensure the proper scale of measurement is recorded.	8, 9
If the biospecimen was removed from storage because of a freezer malfunction, what was the maximum temperature it attained?	1.3	During a freezer malfunction, when the biospecimens have to be removed and temporarily stored, the biospecimens may undergo some thawing, which, in turn, may reduce the quality of the biospecimen.	8, 9
If the biospecimen was removed from storage because of a freezer malfunction, what was the duration of the maximum temperature it attained?	1.3	During a freezer malfunction, when the biospecimens have to be removed and temporarily stored, the biospecimens may undergo some thawing, which, in turn, may reduce the quality of the biospecimen.	8, 9
Thaw-freeze events: combined duration at maximum temperature	1.3	The time during which the biospecimen is at its maximum (highest) temperature during the thawing and refreezing process. If multiple events, this figure should be the total of the combined durations.	8, 9
Thaw-freeze events: maximum temperature	1.3	The maximum temperature that is attained by the biospecimen during the thawing and refreezing process. Ensure the proper scale of measurement is recorded.	8, 9
Age	1.4	The age of the patient/donor at the time the biospecimen was collected. The optimal recording is birth date (month/day/year).	5, 6
Pathology department—receipt of biospecimen by pathology: time/date	1.4	The time (minute/hour) and date (month/day/year) of the receipt of the biospecimen in the pathology department.	6, 8, 9
Patient medical record no.	1.4	The institutional identification no. that was assigned to the patient/donor at the time of the biospecimen collection.	3, 7
Pretransportation state	1.4	The condition of the biospecimen as it is in transportation from the patient/donor to pathology (eg, fresh, in formalin, in LN ₂ liquid or vapor, in dry ice/butanol, etc).	8, 9
Timeline: first cancer progression after first prescription	1.4	Date cancer first reappeared after primary treatment—local recurrence or metastasis.	5
Consent linked to biospecimen in biobank's inventory system	1.5	Is the patient/donor consent form for the biospecimen linked to the biospecimen in the biobank's inventory system?	6, 10
Receipt-at-destination temperature	1.5	The temperature of the biospecimen(s) upon receipt and inspection at the receiving site, such as a biobank or researcher's laboratory. Ensure the proper scale of measurement is recorded.	8, 9

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact, If Data Missing ^b
Return: time/date	1.5	The date and time when a distributed biospecimen is returned and accepted at the distributing biobank.	8, 9
Transport total time in transit to pathology	1.5	The time between removal of the biospecimen from the patient/donor to receipt in the pathology department.	8, 9
Unrestricted consent versus tiered consent	1.5	A tiered consent gives the patient/donor multiple choices regarding the use of the biospecimen(s), both currently and in the future, including specifying specific tests to be included or excluded. The single tier (unrestricted) consent does not allow patient choices: patients either choose to donate or they do not.	6
Centrifugation: duration	1.6	The time the fluid biospecimen undergoes centrifugation.	8, 9
Consenting outcome	1.6	Was the proper consent obtained or not?	6
Date of birth	1.6	The date of the patient/donor's birth (month/day/year).	5
Medications, before death if biospecimen collected after death	1.6	This list should include insulin, steroids, and NSAIDs as specific categories because they are known to have significant influence on the genomic expression of almost all biospecimens. Consider over-the-counter supplements as well.	5
Medications, before biospecimen collection	1.6	This list should include insulin, steroids, and NSAIDs as specific categories because they are known to have significant influence on the genomic expression of almost all biospecimens. Consider over-the-counter supplements as well.	5
Medications (during surgery, surgery case only)	1.6	This list should include insulin, steroids, and NSAIDs as specific categories because they are known to have significant influence on the genomic expression of almost all biospecimens. Consider over-the-counter supplements as well.	5
Postpreservation transport maximum temperature	1.6	The maximum temperature of a biospecimen during transport after it had been preserved. For example, core biopsies in formalin sent to pathology at another facility, fresh tissue taken to pathology for dissection, blood or body fluid sent fresh to another facility, LN ₂ snap-frozen tissue sent to an offsite biobank, etc.	8, 9
Postpreservation transportation state	1.6	The state of the biospecimen between the time it has been removed from the body until it has been accessioned (eg, fresh in transport medium, preserved in formalin, frozen in dry ice/butanol or LN ₂ vapor, in a Vacutainer tube, etc).	8, 9
Pretransportation maximum temperature	1.6	The maximum temperature of the biospecimen before it was transported either within the collecting institution or before it was sent to an offsite biobank.	8, 9
Study	1.6	Link the biospecimen to a specific research protocol.	8, 9, 10
Treatment prior: chemotherapy	1.6	The date and type of all chemotherapy given to the patient/donor before the biospecimen was collected.	5
Treatment prior: radiotherapy	1.6	The date and type of all radiotherapy given to the patient/donor before the biospecimen was collected.	5
Treatment, prior: biological (eg, immunotherapy, vaccination for disease treatment or infectious disease)	1.6	The date and type of all biological therapy given to the patient/donor before the biospecimen was collected, eg, immunotherapy, vaccination for disease treatment or infectious disease, etc.	5
Aliquot (sample) frozen: duration before refreezing	1.7	Time that a frozen biospecimen was out of the storage compartment before it was refrozen.	5, 4
Consent obtained date	1.7	The date that the patient/donor signed the appropriate consent.	8, 9, 6
Hematologic disease history	1.7	If blood is used as the "normal" control, all biospecimens from this patient/donor should have this information.	5
Label: (Prelabeled biospecimen accession no.)	1.7	When kits/supplies are prepared for clinical trials, all contents are labeled with a random ID, which is useful during the biospecimen collection process. This ID will change/evolve after receipt at the kit-producing facility and/or the site where biospecimen processing occurs. Considerable labor efficiencies can be achieved if the biospecimen containers are prelabeled, and that information is in the	5, 4

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact, If Data Missing ^b
		kit-producer's IT system. This ensures that the kit producers know what the receiving institution should be getting before it arrives, who sent it, and what it is called. It adds an additional layer of QC to the operation. If the information in the kit-producing/biospecimen-receiving site system does not agree with what arrives, then the receiving site knows it has a problem with that biospecimen collection site and that biospecimen.	
Newly added biospecimen accession no.	1.7	Some biospecimens may have an additional ID no. attached other than the primary patient/donor ID no.	3, 4
Postpreservation transportation time	1.7	The time (minutes/hours) between the preservation of the biospecimen and its receipt in the pathology department.	4, 8, 9
Return: condition	1.7	The condition of the biospecimen when it is returned to the biobank after distribution, eg, thawed, biospecimen container broken, etc.	8, 9, 4
Return: refreeze time/date	1.7	The time and date that a distributed biospecimen is returned and refrozen.	8, 9, 4
Return: temperature	1.7	The temperature at which a biospecimen is received upon its return.	8, 9
Allergies	1.8	The allergies that are recorded in the medical record for the patient/donor.	8, 9
Centrifugation: No. of times	1.8	The no. of times that a biospecimen has been centrifuged (eg, whole blood, plasma, serum, body fluids, urine, etc).	5
Centrifugation: speed (<i>g</i> force)	1.8	The speed of each centrifugation using <i>g</i> force is preferable to providing a link to the appropriate section of an SOP.	8, 9
EtOH history	1.9	The ethanol abuse history in the patient/donor medical record.	6, 8, 9
Medications, immediately preoperatively, surgical case only	1.9	List of medications given to patient within 24 h of biospecimen removal.	5
Nonneoplastic history	1.9	The patient/donor's diagnoses that are not neoplastic.	5
Smoking history	1.9	The smoking history in the patient/donor medical record.	5
Weight	1.9	The weight of the patient at the time of biospecimen collection.	5
Consent collection comment	2.0	Any comment that has been added to the consent document or about the consenting process.	5, 6
Patient/donor first name	2.0	The first (given) name of the patient/donor.	6, 3
Patient/donor last name	2.0	The last name of the patient/donor (surname).	3, 6
Illicit drug use	2.0	Criminal behavior should be protected information and not shared unless permission is included in the consent.	3, 6
Method of transportation from biospecimen collection site to biobank	2.0	The method of transportation used to transport the biospecimen to the biobank, eg, walking, taxi, institutional vehicle, bus, airplane, etc.	8, 9
Method of transportation used for distributing biospecimen	2.0	The method of transportation used to distribute the biospecimen from the biobank (eg, walking, taxi, institutional vehicle, bus, airplane, etc).	8, 9
Shipping: duration of preparation	2.0	The time (minutes/hours) between retrieval of the biospecimen and its hand-off to the person responsible for its delivery process (eg, the time between retrieval of the biospecimen and the hand-off of the biospecimen to the person who will take it to the taxi, courier, airplane, etc).	8, 9
BMI (calculated)	2.1	The Basal Metabolic Index as calculated from the patient/donor's height and weight.	8, 9
Consent scanned	2.1	Has the appropriate consent been scanned into the patient/donor's medical record and is it available if needed by collecting and/or processing staff?	5, 6
Consent type (precollection versus postcollection)	2.1	Was the appropriate consent obtained before or after the collection of the biospecimen(s)?	6

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact, If Data Missing ^b
Patient/donor middle initial	2.1	The patient/donor's middle initial, as defined in the medical record.	6
Recreational drug history	2.1	If allowed by local, state, or federal law, any history of recreational drug use from the medical record.	6
Centrifugation: temperature	2.2	The internal temperature of the centrifuge at the time of biospecimen centrifugation.	8, 9
Consenting personnel name	2.3	The first name, middle initial, and last name of the person obtaining the approved donor consent.	6
Height	2.3	The height of the patient/donor at the time of biospecimen collection. Specify unit of measurement.	5
Pathologist name	2.3	The first name, middle initial, and last name of the pathologist, or appropriate designee, responsible for obtaining and/or processing the biospecimen for the biobank.	5, 4
Consent mailed date	2.4	The date (month/day/year) that a copy of the appropriately signed consent was mailed to the patient/donor.	6, 8
Consent sent to medical record	2.4	The date (month/day/year) that the appropriately signed consent was sent to medical records or attached to the medical record.	6, 8
Submitting physician name	2.4	The first name, middle initial, and last name of the physician or other appropriate collector/orderer of the biospecimen.	6, 5, 4
Biospecimen (if the biospecimens are of the same material type, they should be stored in consecutive spaces if possible).	2.5	This procedure allows for convenience in retrieval later.	7, 8, 9
Patient Karnofsky (performance) score	2.5	The Karnofsky (patient/donor performance score) obtained from the medical record.	None

Abbreviations: 1-D, 1-dimensional; 2-D, 2-dimensional; AP, anatomic pathology; BMI, Basal Metabolic Index; CP, clinical pathology; EtOH, ethanol; GTEx, genotype-tissue expression; ID, identification; IHC, immunohistochemistry; ISH, in situ hybridization; IT, information technology; LN₂, liquid nitrogen; NCI, National Cancer Institute; NSAIDs, nonsteroidal anti-inflammatory drugs; PR, priority; QC, quality control; SOP, standard operating procedure; URI, uniform resource identifier; no, number.

^aPriority numbers: 1, required field; 2, clinically important, but not yet validated; 3, not necessary; and 4, parking lot for future assessment (No. of variables = 102).

^bNegative-impact definitions are linked to the numbers in Table 6.

Table 2.

Preanalytic Variables—Tissue Specific

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact ^b
Biospecimen type	1.0	Type of biospecimen (eg, fine-needle aspiration, core biopsy, resection, cytology preparation).	1, 3
Blood transfusion: before, during, and after surgery	1.0	Any blood products (eg, packed RBCs, FFP, platelets).	5
Tissue site of origin	1.0	Organ or tissue from which the biospecimen is obtained.	1, 3, 5, 8
Type of preservative (fixative)	1.0	Preservative used (eg, formalin, 70% EtOH, proprietary fixative, etc).	1, 4, 8
Histologic diagnosis	1.1	Histologic type. For metastatic tumors, include primary site of origin. Include grade and other relevant elements of the surgical pathology report. Link to pathology report if possible.	1, 5, 6
Location where aliquot was removed from the lesion of interest	1.1	Site from which the sample was excised (eg, from a cancer resection surgery, this biospecimen was taken from the center of the tumor mass, or it was taken from the tumor/normal border, or it was taken from the surgical margin).	1, 4
Pathologic stage of cancer (TNM)	1.1	http://www.cancerstaging.org , accessed May 14, 2013.	5
Total time in formalin	1.1	Must include tissue processor formalin time (minutes/hours).	2, 4, 8
Formalin fixation start time	1.3	Specify date (month/day/year) and time (minute/hour) for start of fixation.	2, 4
Type of quality control tissue if this is a "normal control" biospecimen	1.4	For example, whole blood or derived DNA, resection margin, peritumoral, etc. Peritumoral "normal" tissue often has premalignant "field" mutations.	8, 9
Warm ischemia time	1.4	Time the tissue remains at body temperature after its blood supply has been significantly reduced or stopped.	4
Primary, recurrent, or metastatic cancer	1.6	If the biospecimen is from a recurrent or metastatic cancer, specify the site from which biospecimen was obtained, and if the lesion was before or after therapy.	1, 3, 5, 6
Storage: humidity	1.6	Percentage average humidity. Maximum high and minimum low data if available.	2, 4
Tissue abnormalities other than the primary histologic diagnosis	1.6	Significant pathology detected at the time of morphologic review that is not part of the primary diagnosis used to identify the biospecimen for collection (eg, inflammation, fibrosis, infection consistent with <i>Helicobacter</i>).	1, 5, 8
Type of tissue: nonalcohol dehydration	1.6	The liquid used to dehydrate the alcohol-impregnated biospecimen during processing (eg, xylene).	1, 4, 8
Liquid nitrogen freezing method	1.7	The exact way the biospecimen is snap-frozen in liquid nitrogen (eg, immersion, float on raft, vapor, ±OCT, etc).	2, 4
Tissue FFPE block of choice	1.9	Identification of the best block(s) of the selected lesion and/or normal tissue to optimize downstream molecular analyses. It is not necessary to indicate the location of the preferred tissue in the block, only that a particular block will be the "best" for further testing on the tissue.	1, 3, 8
Formalin expiration date	2.0	Expiration date on formalin container if premade or on concentrated formaldehyde container if mixed on site.	2, 4
Warm ischemia time: first artery clamp time	2.0	Some resections have only one main arterial supply (eg, kidney).	4
Warm ischemia time: last artery clamp time	2.0	Some resections may have several significant arterial supplies (eg, colon).	4
Anesthesia start time	2.1	Obtained from anesthesia record.	5
Anesthesia stop time	2.1	Obtained from anesthesia record.	5
Anesthesia type	2.1	Obtained from anesthesia record.	5

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact ^b
Paraffin expiration date	2.1	The expiration date on the paraffin stock container.	2, 4
Paraffin temperature: embedding	2.1	The temperature of the paraffin used during embedding.	2, 4
Paraffin temperature: processing	2.1	The temperature of the paraffin bath in which the tissue biospecimen is submerged (should not be >60°C).	2, 4
Blood pressure significant variation during surgery	2.4	There is no standard definition for this parameter.	5
Fluid administration during surgery	2.4	Way in which fluid is administered during surgery (eg, whole blood, packed red blood cells, plasma, saline, dextrose, etc).	5
Formalin lot no.	2.4	The lot no. on the stock formalin (premixed) or formaldehyde (if mixed on site) container.	2, 4
Paraffin replacement date	2.4	The date when the paraffin bath used for the tissue biospecimen was last replaced.	2, 4
Paraffin lot no.	2.6	The lot no. on the paraffin stock container.	2, 4
Paraffin vendor	2.6	The name of the company that made the paraffin. If not available on the container or invoice, use the name of the vendor.	2, 4
Formalin manufacturer	2.7	The name of the company that made the formalin/formaldehyde. If not available on the container or invoice, use the name of the vendor. Include the catalog number.	2, 4
Percentage of formaldehyde used for the initial preservation in formalin	2.7	The percentage of formaldehyde listed on the container of the premixed formalin or of the final formalin if mixed on site.	2, 4
Percentage of formaldehyde in the formalin used in the tissue processor	2.7	The percentage of formaldehyde listed on the container of the premixed formalin or of the final formalin if mixed on site.	2, 4
Tool used to remove biospecimen	2.7	Tool or method for excising biospecimen (eg, scalpel, cautery, punch biopsy, core-needle biopsy, fine-needle aspirate).	2, 4

Abbreviations: EtOH, ethanol; FFP, fresh frozen plasma; FFPE, fresh frozen paraffin embedded; OCT, optimal cutting temperature; PR, priority; RBCs, red blood cells; TNM, tumor, node, metastasis.

^aPriority numbers: 1, required field; 2, clinically important, but not yet validated; 3, not necessary; and 4, parking lot for future assessment (No. of variables = 36).

^bNegative-impact definitions are linked to the numbers in Table 6.

Table 3.

Preanalytic Variables—Blood, Serum, Plasma Specific

Variable (Data Field)	PR ^a	Definition/Note	Negative Impact ^b
Anticoagulant used	1.0	Type of anticoagulant used (eg., EDTA, heparin, etc).	1, 2, 4, 8
Storage temperature	1.0	Temperature at which the whole blood was stored before processing.	2, 4
Type of biospecimen	1.0	Indicate whether the biospecimen being annotated is serum, plasma, whole blood, etc.	3, 4
Date/time processed	1.3	Date/time primary tube of whole blood was further processed (eg, to fractionate blood into serum or plasma).	4
Lipemia	1.3	Note the degree of lipemia if present: mild, moderate, severe.	2, 4
Method of blood separation	1.3	Note separation method (eg, record the centrifuge make and model number, g, time spun, tube type used in centrifuge, whether brake was used to assist the stop, how much serum or plasma was removed. Link to SOP if possible.)	2, 4
Volume of blood collected	1.3	The quantity of whole blood collected.	2, 4
Fasting	1.5	Was the donor fasting at the time the biospecimen was collected: yes or no.	4
Temperature at which blood was processed	1.5	If the whole blood was processed further, at what temperature did that processing take place?	2, 4
Collection tube type	1.7	The color of the blood tube's top (stopper) and/or its specific name or model number.	2, 4
Type of blood collection apparatus	1.8	What type of whole blood collection apparatus was used (eg, Vacutainer, butterfly needle, etc).	4
Collection position of donor	1.9	Note position at time of collection (eg, lying down, sitting, standing, etc).	8, 9
Order of tube type collection	1.9	List the order of tube types used during the blood collection.	2, 4
Did not clot	2.0	Did the biospecimen form a complete clot before it was processed: yes or no.	2, 4
Hemolysis	2.0	What was the degree of hemolysis if present: mild, moderate, severe.	2, 4
Bilirubinemia	2.0	What was the degree of bilirubinemia if present: mild, moderate, severe.	2, 4
Needle gauge	2.0	What was the gauge of the needle used to collect the biospecimen?	4
Partial clot	2.0	Did a partial clot occur in the blood before it was processed: yes or no.	2, 4
Name of collecting person	2.1	The first name, initial, and last name of the person collecting the biospecimen. If only initials are provided, use the initials.	4
Vacutainer tube holder type	3.0	The type of Vacutainer used, not the type of tube used.	4

Abbreviations: EDTA, ethylenediaminetetraacetic acid; *g*, g force or gravity for centrifugation; PR, priority; SOP, standard operating procedure.

^aPriority numbers: 1, required field; 2, clinically important, but not yet validated; 3, not necessary; and 4, parking lot for future assessment (No. of variables = 20).

^bNegative-impact definitions are linked to the numbers in Table 6.

Table 4.

Preanalytic Variables—Body Fluid Specific

Variable Data Field	PR ^a	Definition/Note	Negative Impact ^b
Collection method	1.3	Describe the way in which biospecimen was collected (eg, for urine, clean catch, midstream, 24 h, catheter/bag).	1, 2, 4
Volume of fluid collected	1.4	Urine collection is unique in its collection and the parameters of interest vary depending on the type of collection.	4
Cloudy	1.5	What is the degree of cloudiness if present: mild, moderate, severe?	2, 4
Color	1.6	Describe the color of the biospecimen (eg, amber, pink, red, etc).	4
Particulates observed	1.7	List the particulates observed if any (eg, crystals, cells, etc).	2, 4
Cell count	1.7	The no. and type of cells within the fluid, reported in international units.	2, 4

Abbreviation: PR, priority.

^a Priority numbers: 1, required field; 2, clinically important, but not yet validated; 3, not necessary; and 4, parking lot for future assessment (No. of variables = 6).

^b Negative-impact definitions are linked to the numbers in Table 6.

Table 5.**Preanalytic Variables — Nucleic Acid Specific**

Variable (Data Field)	PR^a	Definition/Note	Negative Impact^b
Concentration	1.0	The concentration of the biospecimen that is stored.	1, 4, 10
Concentration, determination method	1.0	The method used to measure the concentration of DNA, RNA, or total nucleic acid.	4
Contamination by protein	1.0	Note if it is present and the method used to assess it.	1, 4, 10
Contamination by salt	1.0	Note if it is present and the method used to assess it.	1, 4, 10
Derivative type	1.0	The exact type of sample that is derived from the primary biospecimen (eg, DNA, RNA, Total NA, microRNA, crude extract, etc).	1, 3, 4, 10
Freeze-thaw cycles	1.0	Record date/time of freezing and of each prior freeze/thaw.	2, 4
Spiked nucleic acid	1.0	If exogenous DNA or RNA was added, specify its name (eg, ERCC no. 134), concentration added, and timing of the spike (eg, before versus after extraction).	4
Isolation method: extraction kit name	1.0	Make and model no. of extraction kit and instrument.	4
Isolation method: link to SOP	1.0	Link to SOP for biospecimen extraction.	4
Original biospecimen (frozen tissue, FFPE, macrodissected or microdissected tissue, whole blood, plasma, serum, body fluid before or after centrifugation, etc)	1.0	List the type of biospecimen or link to SOP for biospecimen preparation.	1, 3, 4, 10
Q-rtPRC of housekeeping RNA	1.0	Specify the targeted RNAs and the unit of measurement; link to SOP.	4, 8
RIN and link to analyzer tracing	1.0	Link to RNA analyzer tracing, record the RIN value.	4, 8
Diluent used for nucleic acid	1.2	The diluent used for aliquoting and storing the nucleic acid.	1, 4, 10
No. of vials stored	1.2	The no. of vials of the biospecimen that are sent to storage.	1, 4
Volume of NA vial and amount/vial remaining if partially used	1.2	Provide the volume of NA/vial and amount/vial remaining if partially used.	1, 4
Bar-coded vials	1.4	Describe how the vials are labeled. May link to labeling SOP if available.	3, 4, 7

Abbreviations: ERCC, excision repair cross-complementing; FFPE, formalin-fixed, paraffin embedded; NA, nucleic acid; PR, priority; Q-rtPRC, quantitative real-time reverse transcription polymerase chain reaction; RIN, RNA integrity number; SOP, standard operating procedure.

^a Priority numbers: 1, required field; 2, clinically important, but not yet validated; 3, not necessary; and 4, parking lot for future assessment (No. of variables = 16).

^b Negative-impact definitions are linked to the numbers in Table 6.

Table 6.

Preanalytic Variables—Negative Impact Categories With Examples

Category	Tissue Example	Blood Example ^a
1. Biospecimen: insufficient for testing or incorrect	A thyroid FNA cell block containing only blood is sent for <i>BRAF</i> testing.	Whole blood previously processed, but only the plasma fraction of blood is now available in the repository, and plasma does not contain adequate leukemic cell DNA.
2. Biospecimen: degraded inventory	A breast tumor that was unfixed for 96 h is sent for hormone receptor testing.	Whole blood received, but after refrigerating the residual whole blood for 5 mo, there is insufficient intact DNA.
3. Biospecimen: source misidentified	A negative lymph node is sent for <i>BRAF</i> testing when a node with metastatic melanoma was the intended substrate. Wrong accession no. used on downstream aliquot (sample) and/or slides.	Whole blood processed, but all the patients with leukemia were in remission rather than at initial diagnosis.
4. Biospecimen: chain of custody and/or history	Information about the biospecimen source and clinical history is mistakenly deleted and no longer available to researchers. Consent comments not saved in database and/or not attached to biospecimen. Complete consent form not attached to biospecimen.	Whole blood received from patients with leukemia that had been processed at an outside facility although it was not clear how long the blood had been stored and at what temperature it was stored before processing at the outside facility: insufficient annotation.
5. History: insufficient or incorrect	A colon cancer sent for microsatellite instability testing is mistakenly described as having been obtained following neoadjuvant chemoradiation, confounding interpretation.	Whole blood previously processed, but the anticoagulant type (eg, EDTA, heparin) had not been recorded.
6. Legal risk	Genetic testing is done on a biospecimen for which the patient did not provide consent.	Whole blood previously processed, but it was not clear whether the subjects had consented to whole-genome sequencing. Another example would be if the receipt at a location is missing, and the biospecimen was actually lost in transit.
7. Patient: incorrect identification	Biopsy of tumor mistakenly labeled as belonging to a different patient.	Whole blood previously processed and one of the samples had the diagnosis "PCD" in accompanying paperwork. Because PCD was an ambiguous term (eg, plasma cell dyscrasia, primary ciliary dyskinesia), this sample did not meet the needs of the investigator.
8. Quality: testing result issue	False-positive <i>HER2</i> test results are recorded. Histologic slide is mislabeled. Incorrect weight is attached to biospecimen.	Whole blood previously processed with karyotyping done at site C, which was not a CLIA-certified laboratory. Because the annotated karyotype results from site C were not clinical-grade results, the samples did not meet the needs of the investigator.
9. Quality: missing quality indicator	Total time in fixative is not available.	Whole blood previously processed, but neither blood smears nor blast counts were checked before processing to determine whether the samples met study inclusion criteria regarding minimum proportion of leukemic cells.
10. Resources wasted	An incorrect panel of biomarker tests is done on a small biopsy sample, and the tissue is depleted as a result. Cancer aliquot (sample) is too large for shipping and has to be trimmed.	The samples had not been annotated with information on how the blood was processed.
11. Operational: investigator confidence in biobank	Billing for biospecimens not properly explained.	Although a database of available biospecimens can be made accessible, a contact person is needed to clarify data elements and to promote efficient use of tangible and information resources.

Abbreviations: *BRAF*, proto-oncogene B-Raf gene; CLIA, Clinical Laboratory Improvement Amendments; EDTA, ethylenediaminetetraacetic acid; FNA, fine-needle aspiration; *HER2*, human epidermal growth factor receptor 2.

^a100 Patients for a Study of Leukemia Cell DNA.