Considerations for uniform and accurate biospecimen labelling in a biorepository and research environment

Andrew B Kay,1,2 Daniel K Estrada,1,2 Sergey Mareninov,1,2 Steven S Silver1,2 Clara E Magyar,2,3 Sarah M Dry,2,3 Timothy F Cloughesy,4,5 William H Yong1,2,5,6

ABSTRACT
Correct labelling of specimens in a biorepository or research laboratory is vital, especially for translational or clinical studies linking clinical data with biospecimens. While patient privacy must be carefully protected, confusing or inadequate labelling can potentially result in the study of the wrong biospecimens with detrimental effects to the accuracy of published findings or a requirement for invaluable biospecimens to be discarded. Labelling guidelines are described in the biorepository of the University of California—Los Angeles Brain Tumour Translational Resource, and in recipient neuro-oncology laboratories to which biospecimens and derivatives are provided. This approach includes specifying identifier types, types of dates and institutions on the biospecimen labels; using multiple identifiers on each specimen when feasible; and developing a three to four-letter alphanumeric code to aid in label recognition. In addition, steps are being taken to educate recipient laboratories on best practices in labelling.

METHODS
Routine specification of type of identifier
Undesignated identifiers can easily lead to confusion, especially if two different identifiers have the same number of digits or similar numbering systems. Rarely, we have encountered identical case numbers from different institutions for different patients. Therefore, if surgical or autopsy numbers are used, the institution of origin must be specified.

Our laboratory has instituted a standard practice of specifying the identifier type and institution on the biorepository specimen label (see figure 1A). In addition to the surgical or autopsy pathology accession number, the block should also be specified for paraffin blocks and slides (eg, block B3 in figure 1A). As multiple biospecimens from different patients may be requested, we also use a job number to identify biospecimens corresponding to a specific request or job (see table 1). A proposed universal nomenclature for identifier types is shown in table 1, along with abbreviations and suggested situations for usage.

Routine use of two identifiers when possible
One possible source of misidentification is research personnel incorrectly transcribing identifiers onto samples. We believe it is optimal for all samples to have at least two different identifiers so one may act as a back-up should the other be incorrect (see figure 1B). The use of two identifiers is a safety practice that has been advocated for clinical settings. Researchers have a better probability to back-track to patient information in a database. Dates can aid in identification but must be specified as to what date they reflect. A date nomenclature is proposed (see table 2).
Incorporation of an alphanumeric labelling system

Purely numerical identifiers are prone to transcription or recognition errors. Recently, a technician started working on a specimen coded 1211 when 12111 was the desired specimen—a situation that might be avoided with an extra letter string (see figure 1C). We have thus created a 4841-word library of unique three and four-letter words attached sequentially to a series of numbers, although not in alphabetical order. The words on the list are chosen from a public list of words, and randomly paired with numbers in a text file. When exhausted, the word library is attached in the same order to the next series of numbers, until 100,000 numbers are paired. In figure 1C, the alphanumeric string is far easier to recognise than the highly similar numbers alone. This alphanumeric list is available upon request. Bar-coding software could be easily programmed to include the alphanumeric string.

A proposed ideal label

In our opinion, an ideal sample label would have at least two identifiers and a date labelled as to type (see figure 1D). In this case, a surgical pathology number and patient research number are used. A specimen identification number can be substituted for the surgical pathology number. We routinely include a job number because it facilitates gathering materials relevant to a specific job request and, to a lesser degree, can help narrow possibilities in the event of mislabelling. In situations in which the institution review board status of a study requires a higher degree of anonymisation, this labelling schema will not be possible. In that setting, at the time of distribution it may be appropriate to use a patient research number or unlinked number known only to the biorepository as well as an appropriately identified date for redundancy.

Placement of the label

As lids or caps may be separated from the cryovial body, labels with two or more identifiers should always be placed on the body of the container wherein the biospecimen resides. In addition to the more detailed label on the cryovial body, we also affix a small label to the cap of the cryovial with one identifier to minimise the possibility of confusion. The use of a label printer is strongly preferred over handwritten labels and can achieve legibility even when lettering is small. Cryovial-specific labels and solvent-resistant labels are commercially available and should be used when relevant. The placement of slide labels at

Table 1 Identifier types, abbreviations and usages

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Abbreviation</th>
<th>Usage</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autopsy pathology number</td>
<td>APN</td>
<td>Links a specimen to its autopsy case</td>
<td>APN: A11-12345 B1*</td>
</tr>
<tr>
<td>Surgical pathology number</td>
<td>SPN</td>
<td>Links a specimen to its surgical case</td>
<td>SPN: S11-54321 A1*</td>
</tr>
<tr>
<td>Patient research number</td>
<td>PRN</td>
<td>Unique patient identifier in research database</td>
<td>PRN: 9876-LOST</td>
</tr>
<tr>
<td>Specimen identification number</td>
<td>SID</td>
<td>Specifies particular specimen of a patient who may have multiple specimens</td>
<td>SID: 5318†</td>
</tr>
<tr>
<td>Medical record number</td>
<td>MRN</td>
<td>Unique patient identifier in clinical database</td>
<td>MRN: 111-11-11</td>
</tr>
<tr>
<td>Unlinked number</td>
<td>UN</td>
<td>Used when specimens are to be completely anonymous and untraceable</td>
<td>UN: 698</td>
</tr>
<tr>
<td>Job number</td>
<td>J</td>
<td>Used to group multiple specimens that are part of the same job requested of the biorepository</td>
<td>J10-113</td>
</tr>
</tbody>
</table>

*Note that blocks (B1, A1, respectively) are specified for the autopsy pathology number and surgical pathology number.
†The specimen identification number (SID) is a numeric only to distinguish from the patient research number (PRN). SID and PRN are generally used together on a label.
the top of a slide as customary is appropriate. While imprinting or writing an identifier on the top edge of a paraffin block cassette is common, additional adhesive labels can be attached to the side or back of the plastic cassette. This approach is needed to attach our biorepository identifiers to paraffin blocks received from outside institutions. Placement of the labelled cassette in a small plastic bag can provide containment for the label coming off and for the tissue if the paraffin should melt because of high ambient temperatures.

**Education of biorepository and research laboratory personnel**

We suggest that it is valuable to educate sample recipients of proper labelling practices, because specimens are no longer under the control of the biorepository once given to recipient research laboratories. Even within the same laboratory, labelling practices may differ. One person’s labelling nomenclature is potentially incomprehensible to other members of their own laboratory—a situation that is exacerbated upon their departure. The provision of a uniform labelling system and nomenclature would alleviate this situation.

**CONCLUSION**

By specifying types of identifiers, institution names and dates, using multiple identifiers, using an alphanumeric code, and educating recipient laboratories on these practices and standard nomenclature, we suggest that biorepositories will be able to reduce misidentication of specimens in translational research. One or more of these practices may apply depending on the situation. While different institutions may find other nomenclature more appropriate, similar practices to ours may still be applicable. Automated labelling, bar codes and RFID tags have roles to play in the biorepository, but the redundancy of a legible label is desirable. Furthermore, for individual researchers or for biorepositories without such resources, we suggest that uniform labelling practices are vital.

**REFERENCES**


### Table 2 Date types and abbreviations

<table>
<thead>
<tr>
<th>Date type</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of operation</td>
<td>DO</td>
</tr>
<tr>
<td>Date of autopsy</td>
<td>DA</td>
</tr>
<tr>
<td>Date received</td>
<td>DR</td>
</tr>
<tr>
<td>Date of experiment</td>
<td>DX</td>
</tr>
<tr>
<td>Date of culture</td>
<td>DC</td>
</tr>
<tr>
<td>Date frozen</td>
<td>DF</td>
</tr>
<tr>
<td>Date of procedure</td>
<td>DP</td>
</tr>
<tr>
<td>Date immunostained</td>
<td>DI</td>
</tr>
</tbody>
</table>

**Take-home messages**

- Mislabelling of biospecimens in the research laboratory or biorepository may impact the accuracy of research results including clinical trials.
- Labelling with two identifiers can mitigate an error in one of the identifiers.
- Specification of the type of identifier can minimise confusion related to identical numbers originating from different labelling systems.
- A three to four-letter alphanumeric code rather than a purely numerical identifier may aid accurate reading of a label.
Considerations for uniform and accurate biospecimen labelling in a biorepository and research environment

Andrew B Kay, Daniel K Estrada, Sergey Mareninov, et al.

J Clin Pathol 2011 64: 634-636 originally published online January 7, 2011
doi: 10.1136/jcp.2010.080655

Updated information and services can be found at:
http://jcp.bmj.com/content/64/7/634.full.html

These include:

References
This article cites 3 articles, 1 of which can be accessed free at:
http://jcp.bmj.com/content/64/7/634.full.html#ref-list-1

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/