Evaluation of the Biospecimen Research Database as a Web-based Tool for Data-mining

Kelly B. Engel Ph.D.[1], Helen M. Moore Ph.D.[2]

[1] Preferred Staffing Group, [2] National Cancer Institute, Office of Biorepositories and Biospecimen Research

RATIONALE

The Biospecimen Research Database (BRD), under development by the National Cancer Institute's Office of Biorepositories and Biospecimen Research, is a searchable web-based literature database populated with peer-reviewed published research and review articles relevant to human biospecimen science. In order to assess the utility of the BRD as a vehicle for meta-analysis, representative variables relevant to formalin fixation and paraffin embedding of biospecimens were investigated.

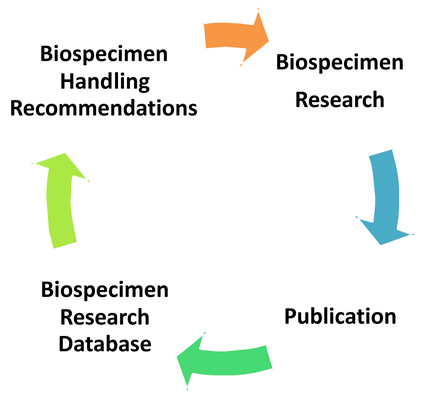
The goals of the data mining exercise were to:

(1) obtain a functional assessment of the BRD in terms of existing infrastructure and curation platforms;

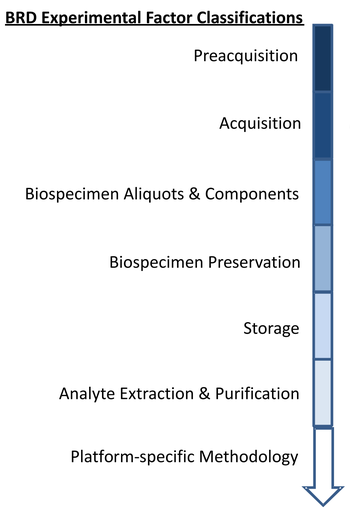
(2) appraise the current BRD paper population and identify specific and underrepresented variables;

(3) summarize consensuses in reported biospecimen fixation and handling when applicable;

(4) ascertain whether the present literature base is accurately capturing the state of the science.



METHODS



Fixation Parameters Investigated

Room temperature delay pre-fixation

Size of biospecimen

Temperature of fixation
Method of fixative delivery
Time in fixative

Duration of biospecimen archival

The influence of biospecimen handling and fixation parameters on molecular, histological, and morphological results was investigated using existing BRD content and infrastructure. The following search criteria was constant for all data-mining exercises:

Biospecimen type:	Cell or Tissue
Preservative type:	Formalin
Paper type:	Nonreview

The experimental factor, which was selected in conjunction with the above search criteria, was dependent on the variable of interest. Experimental factors are organized by classification, corresponding to stages in the lifecycle of a biospecimen.

RESULTS

Consensus

Conclusions supported by 3 or more papers populating the BRD.

Analyte	No. of BRD Papers	Consensus
Fixation Parameter: Biospecimen size		
DNA	3	PCR results of small biospecimens (2-10 mm diameter) were favorable to larger biospecimens.
Fixation parameter: Time in fixative		
DNA	6	PCR analysis was optimal in biospecimens fixed for 2-48 h, with adverse effects reported after fixation for =72 h.
RNA	3	Evidence of RNA degradation was observed in specimens fixed for 1 - 72 h.
	3	mRNA transcript stability was analyte -specific, with fixation thresholds potentially influenced by platform sensitivity and amplicon length. Quantified analytes included: COX-1, beta-actin, MART, MMP -1, MMP-1, VEGF, p21, EGFR, C -BCR.
Protein	5	Protein immunoreactivity was stable in biospecimens fixed for 6 h - 8 d. The antigens investigated included: p27Kip1, ER, PR, AR, c -erb-B2, HER/neu, EGFR, MMR-1, VEGF, p53, PCNA, Ki -67.
Fixation parameter: Archival of formalin -fixed paraffin tissue blocks		
DNA	4	PCR success and efficiency (of 90-435 bp fragments) were not impacted by paraffin block archival at room temperature for 1 wk - 8 y.
RNA	3	RNA degradation was more extensive in blocks stored for 3.5 -17 y compared to those stored for 1 y or less.
	3	RT-PCR success rate decreased by 0 -20% after 1 -10 y, 30 -50% after 10-30 y, and 60% after 40 y of paraffin block archival compared to fresh blocks, although amplicon length also influenced RT-PCR success. The analytes investigated included hepaptitis C, beta -actin, C-BCR.
	3	Real-time qRT-PCR analysis of paraffin blocks archived for 1 -8 y was successful; while analysis was impaired for blocks stored for 11 y or longer. The analytes investigated included LDHA, RPL32, beta -actin, RPL13A, RPLO, CYP1, GUS, TBP, TFRC
Fixation parameter: Archival of formalin -fixed, paraffin-embedded, slide -mounted sections		
Protein	3	Immunostaining was altered in slides stored for 3 mon -3 y at room temperature compared to freshly cut sections. Alterations in immunostaining intensity and duration threshold were antigen -specific. The antigens investigated included ER, PR, HER -2, Chromagranin, CD3, Vimentin, EGFR.

Evaluation of BRD Search Terms

Fixation Parameter	Corresponding Experimental Factor	Studies Returned (No.)	Relevant Studies (%)
Room temperature delay prior to fixation	*Time at room temperature*	5	100%
Size of the fixed specimen	*Aliquot size/volume* *Biospecimen collection method*	11	45% *
Temperature of fixation	*Temperature of fixative*	3	100%
Method of fixative delivery	*Method of fixative delivery*	1	100%
Time in fixative	*Time in fixative*	34	85% **
Embedding Reagents	*Embedding media* *Embedding reagents*	4	100%
Duration of specimen archival	*Storage duration*	26	73% ***

Unrelated topics include:
* Thickness or quantity of paraffin sections, Specimens procured during autopsy
** Multiple fixatives of a single duration
*** Storage of extracted analytes

.Experimental factors relevant to the subject of interest were selected using the Advanced Search option. Search restrictions included nonreview articles and the preservative formalin.

CONCLUSIONS

Variables Excluded from Meta-analysis

Fixation Parameters Under-represented in the BRD

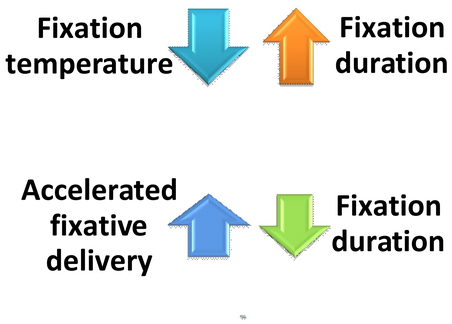
- Room temperature delay pre-fixation
- Temperature of fixation
- Embedding reagents
- Method of fixative delivery
- Post fixation of frozen specimen sections

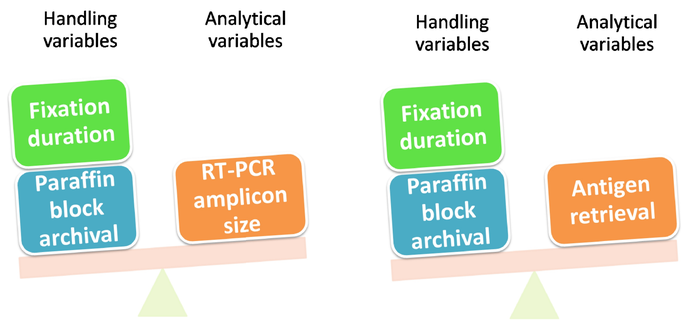
Fixation Parameters Absent from the BRD

- Refrigerated delay prior to fixation
- Fixative pH
- Fixative source
- Fixative age
- Biospecimen size / fixation volume ratio

Relationships Among Preanalytical and Analytical Variables

Although definitive conclusions were not possible for many of the experimental variables investigated, several relationships among variables became evident. Fixation temperature directly affected the duration of fixation required, as did accelerated fixation (injection, immersion, ultrasound acceleration). Further, RT-PCR success was influenced by the size of the mRNA fragment of interest and its corresponding amplicon, as RNA fragmentation was observed with prolonged fixation or storage.





Conclusions and Suggested Improvements

The BRD was successfully used as a data-mining tool. While conclusions were limited, principally restricted by paper abundance, the data-mining exercise successfully identified

(1) consensuses for five subjects of interest

(2) specific biospecimen handling and fixation variables for future literature searches; and,

(3) potential relationships among biospecimen handling and analytical variables.