

# Improving Interoperability and Quality of Data within the Victorian Cancer Biobank

Maggie Ling<sup>1</sup>, Wayne Ng<sup>1</sup>, Vicky Thursfield<sup>2</sup>, Catherine Shang<sup>2</sup>

<sup>1</sup>Victorian Cancer Biobank, Cancer Council Victoria, Melbourne, Australia, <sup>2</sup>Victorian Cancer Registry, Cancer Council Victoria, Melbourne, Australia.

Corresponding author: [maggie.ling@cancervic.org.au](mailto:maggie.ling@cancervic.org.au)



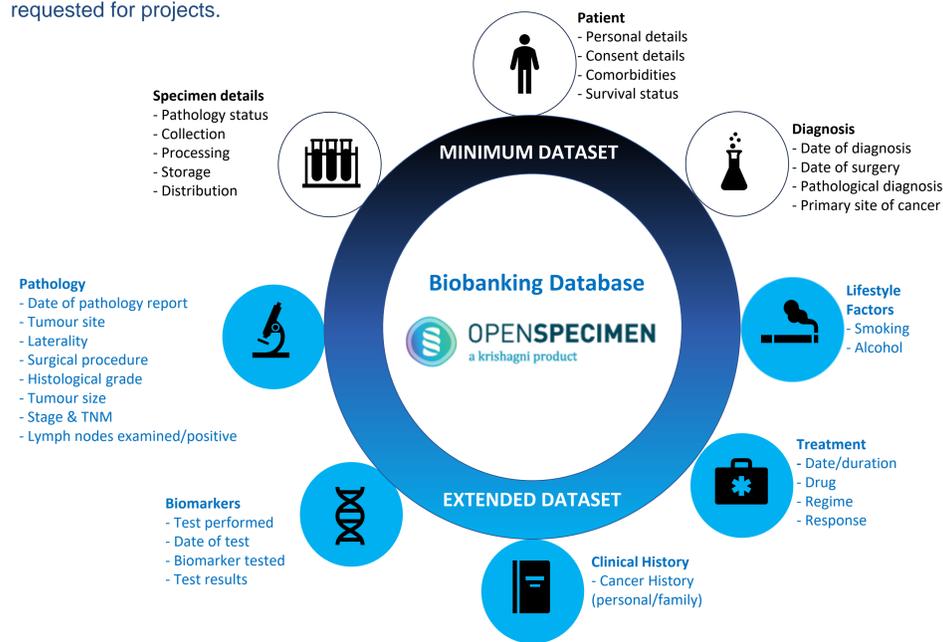
## Introduction

The Victorian Cancer Biobank (VCB) is a not-for-profit consortium with members comprising of five major Victorian health precincts and the Cancer Council Victoria (CCV) as the lead agency. Since VCB was established in 2006, it has recruited more than 34,000 donors and serviced more than 220 research projects. VCB has collected more than 600,000 biospecimens from a broad range of tumours, and approximately 423,000 biospecimens are currently in our inventory. OpenSpecimen was implemented as the biobank information management system (BIMS) for VCB in 2016 to improve the management of the specimens and the associated data. In 2018, VCB integrated into the Victorian Cancer Registry (VCR) division at CCV.

The ability to correlate data and biospecimens from different biobanks is becoming integral for accelerating the pace of translational research. In order to facilitate the exchange and pooling of data and biospecimens between biobanks, it is essential to harmonise biobanking operational procedures and best practices. In light of this, VCB is working towards improving the interoperability and quality of data across the entire network.

### VCB Dataset

Currently VCB collects an extensive dataset via laborious manual annotation of cases from medical records. For all specimens there is a mandatory minimum dataset that must be entered into the BIMS at collection, the extended dataset (additional clinical data) is entered when time permits or when requested for projects.

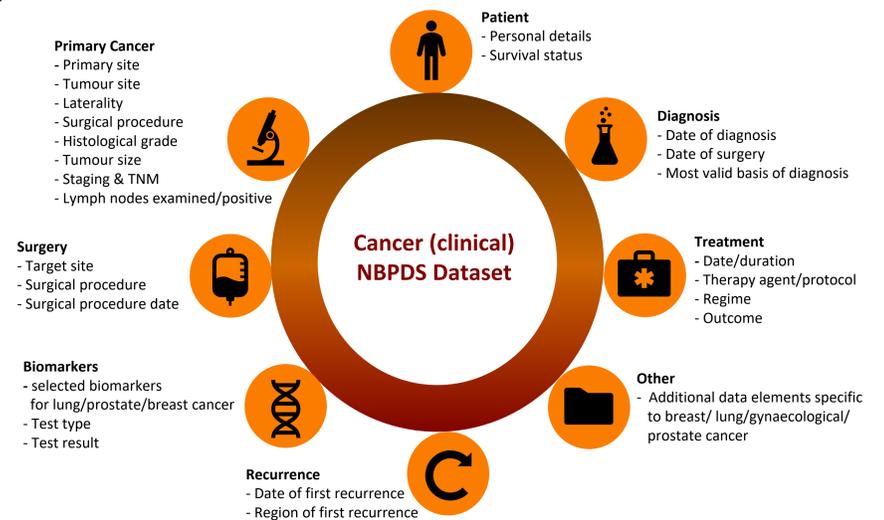


## Evaluation outcomes:

- Some of the VCB data items align with BRISQ and MIABIS recommended standards, whilst the remaining data items are VCB specific (such as additional clinical data for tumour specific cases).
- BRISQ and MIABIS provide recommended datasets with specific data elements, but does not list permissible values or guidelines on the usage of values. For VCB to improve interoperability and quality of data, a better defined framework is required.
- SPREC provides a standardised method for annotating the biospecimen quality based on defined pre-analytical variables that relate to collection, processing, and storage. The quality of biospecimens obtained from different sources can be evaluated and this is essential for facilitating the reproducibility of experimental results. The VCB currently collects data which can be mapped to SPREC.

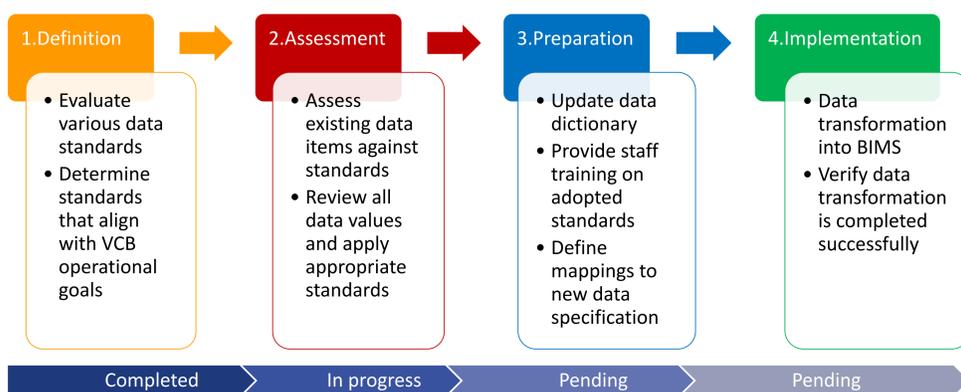
### Cancer (clinical) National Best Practice Dataset (NBPDS)

The Cancer (clinical) NBPDS developed by a working group under the stewardship of Cancer Australia, defines data standards for the national collection of clinical cancer data. The metadata is stored on Australia's Metadata Online Repository METeOR which is administered by the Australian Institute of Health and Welfare (AIHW). Permissible values and guidelines for collection and usage are clearly defined for each data item. Most of this dataset is collected by the VCR according to METeOR data element definitions. This data standard provides a robust framework for VCB to annotate cancer clinical data to national standards, hence improving interoperability and quality of data.

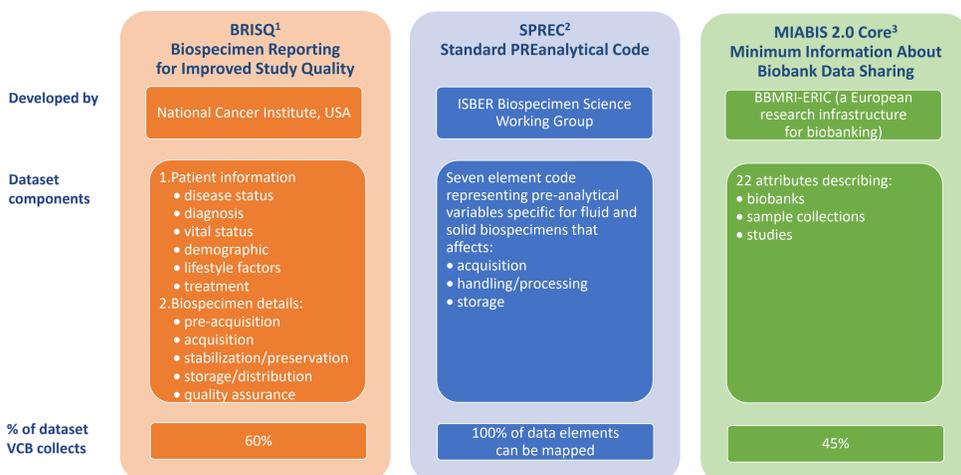


## Improving Interoperability via Harmonisation of Data

### Roadmap to Data Harmonisation



### Evaluation of Biobanking Data Standards

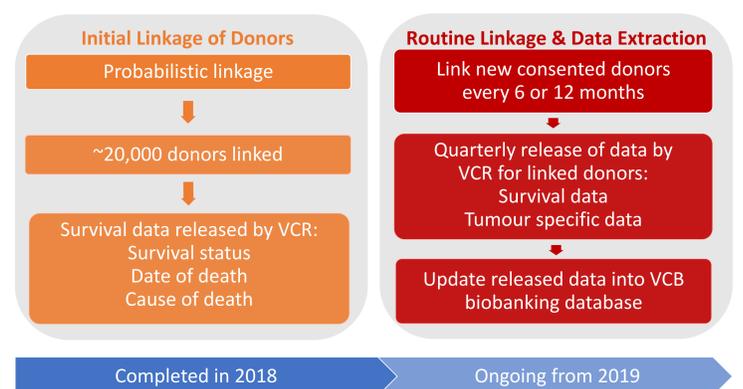


BRISQ<sup>1</sup>: Moore, H. M. et al (2011) Biospecimen Reporting for Improved Study Quality (BRISQ) *J Proteome Res.* 10(8):3429-3438.  
 SPREC<sup>2</sup>: Betsou, F. et al (2018) Standard PREanalytical Code Version 3.0. *Biopreserv Biobank* 16(1): 9-12.  
 MIABIS 2.0 Core<sup>3</sup>: Merino-Martinez, R. et al (2016) Toward Global Biobank Integration by Implementation of the Minimum Information About Biobank Data Sharing (MIABIS 2.0 Core) *Biopreserv Biobank* 14(4): 298-306.

## Leveraging Cancer Registry Data To Improve Data Quality

- Recent integration of VCB into the VCR has provided the opportunity to leverage the VCR expertise in data management and to use the VCR quality assured dataset to improve clinical data annotation. In 2018, initial linkage of VCB donors with VCR was performed.
- The VCB through agreement with VCR can now access tumour specific data and follow-up data, this will be followed with routine electronic extracts and linkage. This ongoing collaboration will improve the overall quality of the annotation through data validation of the current VCB specimen annotation and completion of missing data in legacy cases using the VCR dataset.

### Data Linkage & Extraction with VCR



## Conclusion

- Harmonisation of VCB data items and values to national standard (Cancer clinical NBPDS) will facilitate extraction of data for linkage with VCR and other administrative and disease registry datasets. This will improve the interoperability and quality of VCB data.
- Adoption of SPREC will promote interoperability between biobanks and allow researchers to make informed decisions with respect to the suitability of biospecimens and provide useful information for comparing the biospecimens obtained from different biobanks.
- To further enhance data quality and streamline the quality assurance process, an enhanced version of OpenSpecimen is currently being developed to validate data in real time.