**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 quality management of the samples 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 permit 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 (NBGPS)**

The Cancer (clinical) NBGPS 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**

1. **Definition**
   - Evaluate various data standards and determine standards that align with VCB operational goals
2. **Assessment**
   - Assess existing data items against standards
   - Review all data values and apply appropriate standards
3. **Preparation**
   - Update data dictionary
   - Provide staff training on adopted standards
   - Define mappings to new specification
4. **Implementation**
   - Data transformation into BIMS
   - Verify data transformation is completed successfully

**Evaluation of Biobanking Data Standards**

**VRG3** Biobank Reporting for Improved Study Quality

**SPREC** Standard Preanalytical Code

**MIABIS 2.0 Core** Minimum Information About Biospecimen Data Sharing

**Conclusion**

- Harmonisation of VCB data items and values to national standard (Cancer clinical NBGPS) 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.