

A data standard for facilitating the sharing of data about clinical samples

Philip R Quinlan[†], Confederation of Cancer Biobanks Working Group 3

[†] Breast Cancer Campaign Tissue Bank, University of Dundee, Dundee, DD1 9SY

Introduction

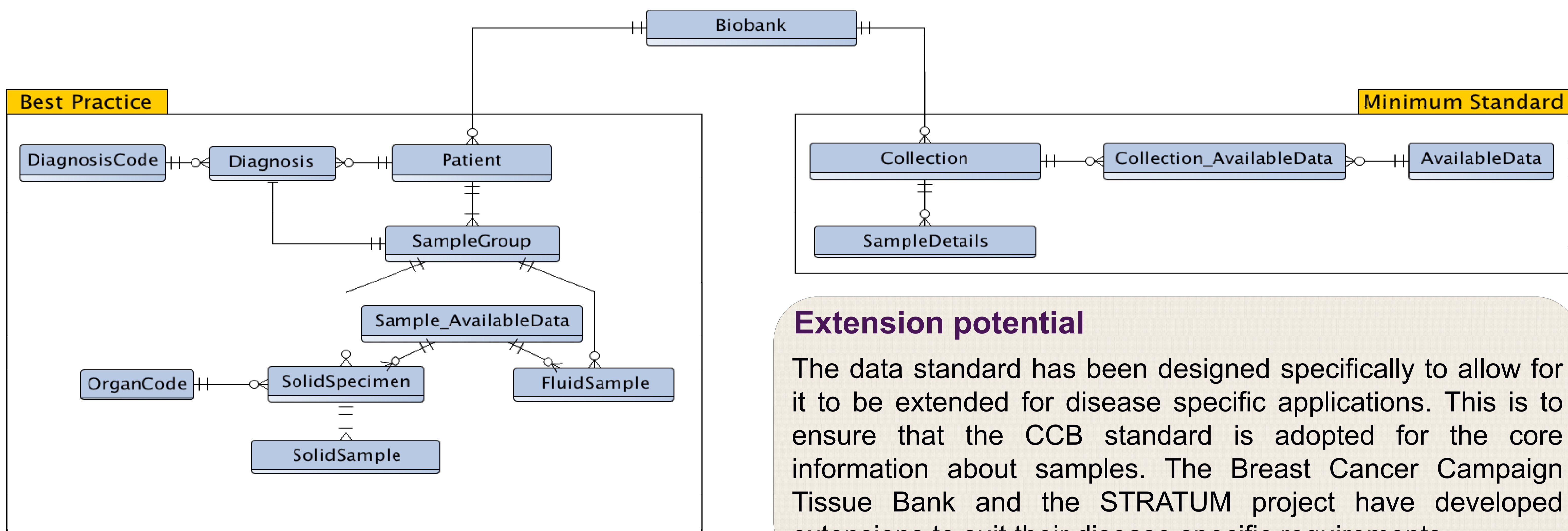
There currently is no infrastructure in the UK to integrate tissue banks and with the advent of stratified medicine, cancer treatment and diagnosis is becoming ever more personalised. A single biobank cannot provide sufficient samples to capture the full spectrum of any disease. The Confederation of Cancer Biobanks (CCB) is a networked consortium of independent organisations in the UK involved in the development, management and use of biosamples resources for cancer research. During 2012 four working groups were established to work on standardisation and harmonisation across all biobanking activities; consent, ethics and public engagement, sample quality, biobank governance and IT and quality assurance.

As part of their remit, CCB Working Group 3 looked to establish a data standard to enable tissue banks to communicate about their holdings and facilitate development of an integrated national network of tissue banks. The standard developed needed to:

- use existing standards where applicable
- cater for different types of collections
- allow for multiple disease types
- not dictate data terms

The long term goal is to develop a national search system to allow any researcher to find samples available across the UK.

Confederation of Cancer Biobanks data standard



Extension potential

The data standard has been designed specifically to allow for it to be extended for disease specific applications. This is to ensure that the CCB standard is adopted for the core information about samples. The Breast Cancer Campaign Tissue Bank and the STRATUM project have developed extensions to suit their disease specific requirements.

Minimum vs best practice

There is an array of different types of biobanks in the UK that collect clinical samples. The main priority of the standard was to ensure that any biobank could communicate about the samples they hold. Therefore the minimum standard, borrowed heavily from the MIABIS¹ standard, is designed to allow a biobank to broadly describe the samples it holds. However, at the same time, there is clear benefit to allow the researcher a greater search capability – such as to find the number of patients from which there is both tissue and blood and that is the purpose of the best practice component.

The unknown

A data standard should not be an all or nothing initiative. Just because a biobank cannot provide one field, should not mean that bank cannot be part of a network. Therefore, the standard allows a biobank to specify information as 'Unknown'. This allows the researcher searching the network for samples to determine how important that information is to their research.

Data term flexibility

One the barriers to the adoption of a data standard is the use of data terms that are not universally used. Therefore this standard has not specified any data terms. Tables such as 'DiagnosisCode' and 'Organ' allow the biobank to provide their own data terms. The aim is to encourage the use of standards without dictating the exact terms that should be used.

Acknowledgements

The data standard was developed by CCB Working Group 3, consisting of Anne Carter, Bill Greenhalf, Elwyn Shing, Gita Mistry, Helen Bulbeck, James Flanagan, John Brinsley, Kwok Pang, Mairead MacKenzie, Martin Groves, Stuart Griffiths. Chaired by Ian Forgie and Philip Quinlan, with help from Daniel Naeh – Wales Cancer Biobank STRATUM Project

(1) Norlin, Loreana, et al. "A Minimum Data Set for Sharing Biobank Samples, Information, and Data: MIABIS." *Biopreservation and Biobanking* 10.4 (2012): 343-348.



Review it in full at: <http://bit.ly/CCBdata>