Towards all-Island sharing of Irish lymphoid blood cancers data: landscape and gap analysis

Kluivert Boakye Duah[1,9], Michael Quinn[2], Eva Szegezdi[3,9], Lisa Crawford[1], Aedin C. Culhane[4,9], Mark Lawler[1,9], Siobhan Glavey[5,6,9], Ruth Clifford[4,7,9], and Ian M. Overton[1,8,9,*]

[1]Patrick G. Johnson Centre for Cancer Research, Queen's University Belfast, UK, [2]Belfast Trust, Health and Social Care Northern Ireland, UK, [3]School of Biological and Chemical Sciences, University of Galway, Ireland, [4]Limerick Digital Cancer Research Centre, Health Research Institute, School of Medicine, University of Limerick, Ireland, [5]Department of Haematology, Beaumont Hospital, Ireland, [6]Department of Pathology, Royal College of Surgeons in Ireland, Ireland, [7]Department of Haematology, University Hospital Limerick, Ireland, [9]Health Data Research Wales and Northern Ireland, Queen's University Belfast, UK, [9]eHealth–Hub for Cancer.

*Correspondence address: E-mail: i.overton@qub.ac.uk

Background: Sharing health data significantly improves public health, clinical care, personal care and associated research. However, privacy laws, limited data standardisation and interoperability, and insufficient data integration have made it difficult for health data to be shared across health institutions and borders. We examine the health data landscape in Northern Ireland (NI) and Ireland (IE). Also, a federated approach with multiparty homomorphic encryption is proposed to analyse and share Chronic Lymphocytic Leukaemia (CLL) and Multiple Myeloma (MM) data on the Island.
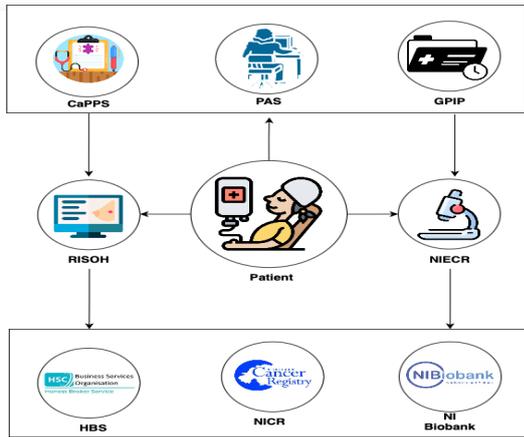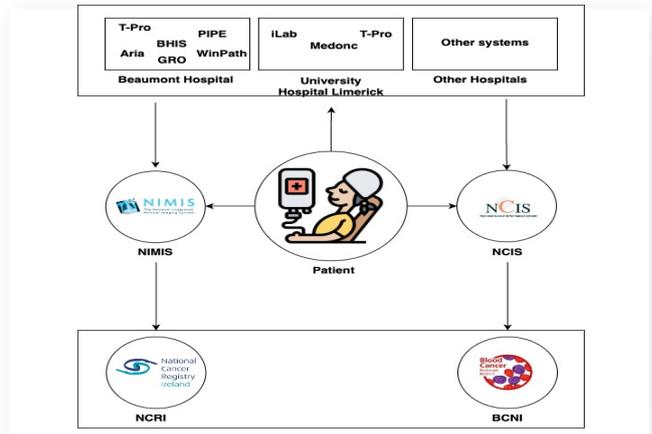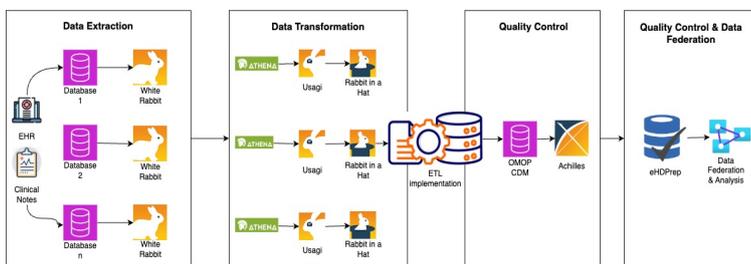
Figure 1: Cancer–related data ecosystem in NI.
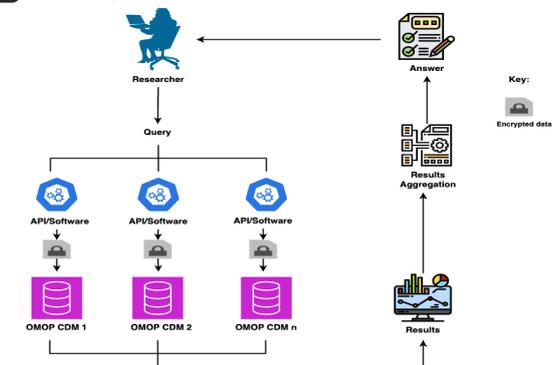

Figure 2: Cancer–related data ecosystem in IE.

Future work

1. Mapping of Irish lymphoid blood cancers data workflow.

2. Proposed data federation and analysis workflow.

Method: Data managers and haematologists of selected health institutions were interviewed in a semi– structured manner. The interviews were done to assess the current state of Irish cancer data including CLL and MM. The nine key areas covered were a) Tools/systems/software for collecting data, b) Data sources, c) Data quality, d) Variables collected, e) FAIR principle, f) Data sharing, g) Conditions inhibiting data sharing, h) Data ownership, and i) Data governance.