Harnessing data for better cancer care: a policy report from All.Can International

Research methodology

May 2021

Contents

The All.Can initiative is made possible with financial support from Bristol Myers Squibb (main sponsor), Roche (major sponsor), Amgen, MSD and Johnson & Johnson (sponsors), Baxter and Illumina (contributors), with additional non-financial (in kind) support from Helpsy, The Health Value Alliance and Goings-On.

1. Introduction

The aim of the All.Can report was to develop an engaging, evidence-based policy report and peer reviewed publication that assess the current evidence showing how data can improve, or their potential to improve, efficiency in cancer care. In doing so All.Can hopes to make a significant contribution towards achieving the optimal use of data to help improve efficiency in cancer care. This document outlines the research methodology used to develop the report.

2. Defining efficiency and health data

2.1. Efficiency

All.Can's definition of efficiency takes a patient-centred approach – focusing on what matters to patients throughout their cancer care.

Efficient cancer care should:

- **Improve outcomes for patients** – through the delivery of accessible, patient-centric, evidence-based and high-quality cancer care that achieves best possible outcomes for all cancer patients individually and collectively with the resources at hand.
- **Optimise allocation of resources** – use available resources in such a way as to achieve optimal outcomes across the system. Resources should be distributed equitably across the population.
- **Use data to continuously learn** – newly available data should be used to contribute to an adaptive and learning healthcare system that strives for continuous improvement to benefit cancer patients and their families.

2.2. Health data

Health data is a broad term that can be defined in a number of ways. For the purpose of this report, All.Can used the definition proposed by the Data Saves Lives initiative:

> *[Health data are] any data describing a person's health, their healthcare or anything affecting any health issues or diseases they may have. This includes information created by health and care professionals, as well as information generated by patients; from illnesses monitored through mobile applications and smart devices, to screening tests and nutritional data*.

The most common types of health data were identified and are outlined in **Box 1**. They formed a starting point for the search terms used in the in-depth literature review.

Box 1. Health data types

- Administrative
- Behavioural
- Biomarker
- Claims
- Economic
- Electronic health records (EHRs)
- Genomic
- Laboratory
- Medical imaging
- Medical reports
- Patient-reported experience measures (PREMs)
- Patient-reported outcome measures (PROMs)
- Records: hospital records + medical records + nursing records
- Registries
- Surveys and questionnaires
- Vital statistics

Based on the literature review findings, the data types included within the scope of the report were further narrowed down to include those deemed most relevant to cancer care. They included:

- cancer registry data
- electronic health records data
- genomic data
- patient-generated health data.

In addition to these health data types, the report also investigated topics related to analytical tools and techniques for extraction of insights from data sources such as big data analytics, artificial intelligence, and machine learning. The report does not attempt to cover all facets of the complex ecosystem of data in healthcare. For reasons of feasibility, it focuses on data generated during the provision of routine clinical care.

3. Research questions

The literature review aimed to answer the following questions:

- What are the predominant types of health data generated as part of routine oncology care?
- How can the use of data, as defined for the purposes of this report, improve or potentially improve the efficiency of cancer care?

- What are the key barriers and facilitators surrounding the use of data in cancer care?
- How can the barriers (identified above) be overcome in practice?

We examined the research questions with a focus on the four areas of opportunity for improving efficiency in cancer care that were identified in the All.Can Patient Survey.

These are:

- Ensure early, speedy, accurate and appropriately delivered diagnosis
- Improve information-sharing, support and shared decision-making
- Make integrated multidisciplinary care a reality for all cancer patients
- Address the financial burden of cancer.

Recognising that there are a number of ways in which data could improve efficiency in healthcare beyond the four points listed above, we allowed for some level of flexibility in exploring topics across the cancer care pathway. For reasons of feasibility, we looked at the impact of data on patient outcomes (i.e. patient experience, survival, quality of life etc.) in the literature review, and did not look at impact on costs.

4. Structured literature review

A structured literature review was conducted to compile existing evidence from peer-reviewed (using PubMed and ScienceDirect) and grey sources (using Google search, news outlets and stakeholder websites) which formed the basis of the policy report.

4.1. Peer-reviewed publications

We conducted a structured literature review of peer-reviewed articles available on the PubMed and ScienceDirect academic databases to answer the research questions. Search terms and searches performed are outlined in **Section 4.2** and were adapted as needed. The literature review was performed until a saturation point was reached.

4.2. Search terms and searches performed

The search combinations outlined below were used to identify published literature. The search strings and terms outlined were used as a starting point, and a snowball approach was used, i.e. interrogating the reference lists of identified sources to identify further papers.

An initial broad search was conducted (Search 1), which yielded too many publications, which were broad and not relevant. As a result, the search was narrowed down (Search 2). Findings from **Search 2** were used to inform the report.

- ***Search 1:*** *[Field 1] AND [Field 2] AND [Field 3] AND [Field 4] [all fields]*
- ***Search 2:*** *[Field 2] AND [Field 3] [title/abstract only]*

Findings from Search 2

Search combination	**[Field 2] AND [Field 3] [title/abstract only]**		
	Included (n)	Excluded (n)	Total (n)
Total records identified: Filters applied: - Published in 2015 and onwards - Humans - English language - Full text available *Publications were sorted using the 'best match' filter*	8,431	0	8,431
Filtered for exclusions: NOT Clinical trial + Clinical research + Market research	4,581	3,850	4,581
Titles screened for inclusion criteria and relevance	428	2,572	Scanned 3,000 out of 4,581 total – until reached saturation point
Abstracts and/or full-text articles assessed for eligibility	100	328	428
Final studies identified for inclusion in literature review findings report	100	0	100

Search terms

Field 1: Patient outcome terms
"Patient outcome" OR "Patient goals" OR "Patient priorities" OR "Patient perspective" OR Survival OR "Quality of life" OR "Clinical outcomes" OR "Clinical indicators" OR "Quality of health care" OR "Patient satisfaction" OR Hospitalization OR Progression OR "Resource use" OR "Return to work" OR Productivity OR Employment OR Morbidity OR Mortality OR Incidence [all text] Yields 6,018,140 results

Field 2: Disease area term
Cancer OR Oncology [all text] Yields 4,173,872 results

Field 3: Health data term

Data OR "Health data" OR "Patient-generated health data" OR "People-generated health data" OR "Patient-reported outcome measures" OR "Patient-reported outcomes" OR PROMs OR "Patient-reported experience measures" OR "Data collection" OR "Data analysis" OR "Data insights" OR Records OR "Medical records" OR "Hospital records" OR "Nursing records" OR "Big data" OR "Information technology" OR "Digital health" OR "Cloud computing" OR "Cancer Registry" OR "Medical informatics" OR "Health surveys" OR "Surveys and Questionnaires" OR "Learning health systems" OR "Administrative data" OR "Behavioural data" OR "Observational data" OR "Claims data" OR "Economic data" OR "Electronic medical records" OR "Electronic health records" OR Genomics OR "Precision Medicine" OR "Personalized medicine" OR "Laboratory data" OR "Medical imaging" OR "Real-world data" OR "Vital statistics" OR "Machine learning" OR "Artificial intelligence" OR "Text mining" OR "Blockchain" OR "Clinical decision support" OR VBHC OR "Value-based healthcare" OR "Staging data" OR "Cause of death data" Yields 5,796,331 results

Field 4: Areas of inefficiency/areas of interest terms
Screening OR Prevention OR Diagnosis OR Misdiagnosis OR "Diagnosis delay" OR "Delayed diagnosis" OR Information OR Support OR Communication OR "Shared decision-making" OR "Side-effects" OR "Late effects" OR Monitoring OR "Care pathway" OR "Pain management" OR "Coordination of care" OR "Care manager" OR "Integrated care" OR "Continuity of care" OR "Key contact" OR Discrimination OR Palliative OR "Multidisciplinary care" OR Nurse OR "Cancer specialist nurse" OR Caregivers OR "Holistic needs assessments" OR HNAs OR "Peer-support" OR "Allied health professionals" OR "Complementary care" OR "Complementary therapies" OR "Distress screening" OR "Psychological support" OR Costs OR "Financial burden" OR "Out-of-pocket costs" OR OOP OR "Follow-up" OR Insurance OR Survivorship Yields 18,937, 510 results

4.3. Inclusion and exclusion criteria

Academic publications, grey literature, and case studies selected for inclusion in the findings report met these criteria:

- Disease area: Oncology
- Language: English
- Date of publication: In 2015 and onwards
- Age group: Any
- Study type: There are no limitations on study design (e.g. data will be accepted from multiple study types)
- Geographic location: Europe mainly, North America and Australia also considered
- Outcomes of interest: Studies must show positive impact of data-related intervention on patient outcomes

- Duration: Initiatives may be completed or ongoing. Ongoing initiatives should show promise of efficiency gains for patients and/or the healthcare system. Completed initiatives must show concrete data demonstrating improvements in patient outcomes to be considered for inclusion in the findings.

Other important criteria should be noted, related to the non-promotional nature of All.Can:

- Data generation, processing or analysis initiatives including commercial products launched by commercial entities may be included if they met **both** of the following conditions:
  - They offered improvements to the standard of care (marked improvement for healthcare systems or patients and their families).
  - Findings published in the literature and available in the public domain (i.e. not available only in company promotional materials).
- Private sector approaches can be included; however, care was taken to ensure examples are never promotional of a given company or product. Case studies were thoroughly validated by the data working group and the external advisory committee before being included in the final report.

4.4. Grey literature

Grey literature includes position papers, guidelines, opinion pieces, and policy documents, etc. Google search was used as the primary tool to identify grey literature. The search terms described in **Section 4.2** were used to find grey literature.

In addition, we used these search methods:

- interrogated and regularly monitored websites of key stakeholders identified through an initial stakeholder mapping exercise
- signed up for key newsletters, regularly browsed websites, and scanned social media
- searched relevant global, European, Australian and North American institutions and their websites to uncover any government publications, guidelines, policy documents, and white papers
- monitored web-based news sources including Politico, CancerWorld, EurActiv and The Parliament Magazine for relevant opinion pieces from key stakeholders.

Peer-reviewed publications were used, where possible, over grey literature sources as this type of evidence is deemed superior in quality.

5. Case study identification and selection

We collected case studies of projects, programmes or initiatives where data have shown to improve patient outcomes (including survival and quality of life) and efficiency in cancer care. Case studies served the function of exploring data solutions in practice, key challenges faced, lessons learned from their implementation, including how any hurdles and challenges in the implementation have been overcome. Case studies were identified through:

- the structured literature review
- submissions by members of All.Can
- submission by the external advisory committee
- the All.Can efficiency hub.

5.1. Shortlisting case studies

Case studies were shortlisted according to the inclusion and exclusion criteria outlined in **Section 4.3** and shared with the data working group for approval and sign-off before including them in the final report.

5.3. Drafting process

Case studies were drafted based on desk research. Each written example aimed to follow a similar format – with some flexibility:

- **Initiative:** a description of the initiative itself.
- **Problem:** a description of the inefficiencies that the initiative aimed to address through the systematic collection and use of data.
- **What it has achieved:** an overview of the initiative's impact on cancer care (e.g. cost reduction, shorter waiting times, improved patient outcomes). Where initiatives were ongoing, their potential impact was outlined. An overview of resources used in order to implement the case study was welcomed, if available.
- **Any challenges/facilitators:** challenges that the initiative needed to overcome (e.g. data protection laws, interoperability, access to data, working with others) and how they have done so in practice, or facilitators of an initiative's success.
- **Key lessons learned:** anything they would recommend others know and consider before trying to start a similar data initiative.
- **Next steps:** any future aims and projects related to the initiative.

6. Expert consultation

In addition to the structured literature review, we consulted several groups of experts throughout the different stages of project work to ensure we covered all relevant ongoing work in the field and to cross-check the validity of our findings. The consultations took place virtually via videocalls, through two main forums:

1) All.Can Data Working Group
2) External Advisory Committee

6.1. All.Can Data Working Group

The All.Can Data Working Group is a subset of the international membership of All.Can, drawn from members with a research interest in data. The group was responsible for overseeing the entire project, including the research methodology and final report development, and held full editorial control. Members of the Data Working Group met regularly to discuss project progress.

The members of the Data Working Group were:

Name	Affiliation
Sangeeta Agrawal	Helpsy Health
Antonella Cardone	European Cancer Patient Coalition
Ivana Cattaneo	Novartis
Dave Duplay	Vital Options International
Caroline Falciola	F. Hoffmann-La Roche
Alex Filicevas	World Bladder Cancer Patient Coalition
Stefan Gijssels	Digestive Cancers Europe
Matthew Hickey	The Health Value Alliance
Petra Hoogendoorn	Goings-On
Agnieszka Krukowska	Johnson & Johnson
Laura McDonald	Bristol Myers Squibb
Jan van Meerbeeck	Antwerp University Hospital
Matthijs Van Meerveld	Merck Sharp & Dohme
Borna Mueller	F. Hoffmann-La Roche
Vivek Muthu	Marivek Consulting
Kathy Oliver	International Brain Tumour Alliance
Titta Rosvall-Puplett	Bristol Myers Squibb
Christobel Saunders	University of Western Australia
Puneet Singhal	Merck Sharp & Dohme
Henriette Thole	Novartis
Julian Shepelev	Baxter Healthcare
Veronica Zilli	Johnson & Johnson

6.2. External Advisory Committee

An External Advisory Committee was assembled to provide validation and external guidance to the data report. The committee took part in the following activities:

- individual introductory calls
- virtual advisory committee meetings to discuss feedback on iterative drafts of the policy report

- written thoughts and advice on final policy recommendations and report contents.

The External Advisory Committee members were:

Name	Affiliation
Fatima Cardoso	Advanced Breast Cancer Global Alliance
Sybo Dijkstra	DigitalEurope
Nigel Hughes	Janssen and European Health Data and Evidence Network
Adrian Jonas	National Institute for Health and Care Excellence
Sabrina Montante	Instituto Superiore di Sanità
Francesco Pignatti (observer role)	European Medicines Agency
Ray Pinto	DigitalEurope
David Roder	University of South Australia
Abdullahi Sheriff	GE Healthcare Europe