The Scottish Government Health & Social Care Data Strategy

Cancer Research UK's Consultation Response, August 2022

Introduction

Cancer Research UK (CRUK) is the largest independent funder of cancer research in the world, supporting research into over 200 types of cancer. Our ambition is to see 3 in 4 people survive their cancer by 2034.

Cancer Research UK have been working with people affected by cancer, public and patient data for almost two decades, and are one of only two health charities with "safe haven" status, allowing us enhanced access to certain types of data. Dedicated to saving lives through research, our long-term investment in state-of-the-art facilities has helped to create a thriving network of research at 90 laboratories and institutions in more than 40 towns and cities across the UK supporting the work of over 4,000 scientists, doctors and nurses. In 2020/21, we invested over £30 million in research in Scotland.

Part 1: Empowering People

A key aim of the strategy is to empower individuals to have greater access to, and greater control over, their own health and social care data.

1. We all have different perceptions of what our health and social care data may be:

1A. When considering the term 'your health and social care data' what does this mean to you and what do you consider it to be?

Data is fundamental to our progress against cancer. Data and data-driven technologies underpin every effort we make to understand cancer, prioritise resources and identify disparities. However, much of this potential is untapped due to fundamental issues right across the data lifecycle – from collection to its use. It is crucial that the Scottish Government implements a health and social care data strategy that will unleash the power of data and we are fully supportive of this ambition.

An essential prerequisite for harnessing the power of data is ensuring that patients and the public are willing for their data to be shared and used for purposes beyond direct care, including lifesaving research. Confidence in the measures in place to safeguard their data, clear communication about the purposes it will be used for and a good understanding of why their data is so important and valuable is critical to building trust and confidence. Public trust is essential to ensure that datasets are representative and mitigate issues with large numbers of people choosing to opt out, which can often lead to biases in the datasets.

There are many ways in which the Scottish Government and the health and social care system organisations can demonstrate trustworthiness and thereby increase the likelihood that patients and the public are comfortable with their data being used for purposes beyond direct care. We are pleased to see the Scottish Government commit to improving patient access to data regarding their health and social care, being transparent with the use of this data, and ensuring that individuals are aware of their choices regarding sharing options and can act upon those choices. However, alongside these positive steps, we strongly recommend that more is done to communicate and educate the public about the value of their data. Clear and compelling communication about the benefits of data sharing with different entities as well as increasing awareness of concepts of privacy such as pseudonymisation, data minimisation and the implications of different access models for privacy and data security on the types of insights that can be gained are essential.

To shape our response to part 1 of this consultation we have drawn from a broad evidence base. This includes our own evidence generated from our networks of people affected by cancer as well as the large base of literature around patient and public attitudes to data[i][ii][iii][iv].

At Cancer Research UK we have a diverse group of 2,000 people affected by cancer from across the United Kingdom (UK) within our Patient Involvement Network that we consult, collaborate and partner with to inform and shape our research, policy, fundraising, HR and technology work. The individuals within the network want to make a real difference to the lives of other people facing cancer now and in the future. They share their experiences of diagnosis, treatment and care through a range of involvement opportunities. Included in this network is a dedicated panel that specialises in data, which we have regularly involved and sought advice from over the past few years, including in the development of our guiding principles for commercial access to research data, and the development of our Research Data Strategy. We have also engaged the wider network with surveys related to the use of health and social care data, so are in a good position to speak to the experiences and opinions of people affected by cancer.

To inform this consultation response, we conducted a survey with our network of people affected by cancer in Scotland about their views on health and social care data. We designed 11 questions, asking our respondents about their opinions on health and social care and data collection, access and sharing. It is worth noting that the survey received just 15 responses, due to a short deadline and respondents likely self-selected themselves based on interest in data and willingness to voice their opinions. With this in mind, the survey results, whilst beneficial, are limited in representing the overall viewpoints of people affected by cancer in Scotland.

2. Our ambition is to give everyone greater access to and a greater say over their health and social care data. Health and social care data examples include results from a blood test, a diagnosed condition or interaction with specific health and social care services.

2A. When thinking about accessing your own health and social care data, what data about you would be your priority for having access to and greater control over?

CRUK supports Public having access to as much of their data as possible. Improving public access to their health and social care data could be beneficial for empowering individuals to be more engaged in their health. It would also demonstrate transparency and ensure that individuals are able to make informed choices around their data sharing preferences as they will know what data exists about them.

This is a key requirement for building the trust and confidence needed for people to feel comfortable with data sharing for purposes beyond direct care. Indeed, our recent survey with people affected by cancer in Scotland, respondents emphasised their desire to have access to all health and social care data pertaining to their direct care. This is likely to be driven by a desire to know what data exists about them.

Moreover, whilst we strongly support the opt-out models in place for many health and social care data flows, including disease registries, it is nevertheless important to ensure that this model is clearly communicated, and that people are aware of the choices they can make about their data sharing preferences. Respondents to our survey for people affected by cancer in Scotland all expressed that they would like to have greater control of their data, including their data sharing preferences.

We strongly advise that as well as providing clear options for individuals to access and control their data, the Scottish Government and health system should place emphasis on public education about the benefits of using health and social care data[v].

2B. When considering the rights of individuals who are unable to interact with their own health and social care data, do you feel that delegating access to a guardian/carer/trusted individual would be appropriate? - If yes, what safeguards need to be in place?

Yes
Appropriate safeguards and privacy measures are essential to protect everyone's personal data. There should be measures in place to validate the guardian/care/trusted individual and controls surrounding eligibility criteria. These measures should be directed towards validating the guardian/ carer/ trusted individuals' identity, their relationship to the individual and ability to handle delegated data appropriately and safely.

3. We are committed to providing clarity over how your data is used and the need for this to be built on ethical principles. When thinking about the ethical principles (read our ethical principles on www.gov.scot) **that must be maintained when gathering, storing, and using health and social care data:**

3A. What information would you find most useful in providing clarity over how your data is used in a consistent and ethical manner?

There are two key areas where clarity over data must be provided for the Public.

1. Firstly, priority should be placed on increasing data transparency by clearly communicating to the public on why their data is being used, by who and the ethical principles underpinning this.
2. Secondly, it is important to clearly communicate complex concepts associated with privacy (e.g., pseudonymisation, anonymisation, re-identification) and the reasons healthcare data can rarely be fully anonymised.

We are supportive of the ethical principles set out in the Scottish Government Data and Intelligence Network: ethics framework. Transparency around how health and social care data is used, for what purposes and how individual choices can be made around data sharing is essential for building public trust in data sharing. Concerningly, the 2022 Centre for Data Ethics and Innovation report on Public

Attitudes to Data & AI (PADAI)[vi] found that 52% of 4,200 respondents knew little or nothing about how data about themselves was collected and used. This highlights the need for proper involvement and engagement alongside transparency of what, how and why data is collected and shared.

Respondents to our survey for people affected by cancer in Scotland all expressed that they would like to have more information about how their data is used. This included having access to information about when their data had been accessed, who by and for what purpose.

In general, there is a broad range of views and diversity of the level of stipulations about what is considered acceptable when data is used beyond the purposes of direct care. There are individuals who are comfortable and willing to share everything and are not concerned about the collection, use or sharing of their data – as long as they feel assured that it is for public good. In contrast, others are much more cautious about their data and need more assurances and clarity on how their data is being collected, used, accessed and shared.

The spectrum of positions is exemplified in our most recent survey, where one respondent fully supported the use of data in health and social care, stating:

"*I have no problem whatsoever that health and social care data be used to accelerate understanding of diseases and to improve efficiency and effectiveness of healthcare services.*"

In contrast, another respondent said that they were *unsure* of their data being utilised for research and expressed the need to see *"detailed explanations"* of how and why their data would be used.

Moreover, many of the respondents stipulated that their data should always be anonymised if it is to be used for purposes beyond direct care. It is important to note that survey respondents were not given any information about the benefits of data use or the reasons why data relating to individuals' health and care can rarely be truly anonymous, particularly when used for research and planning purposes. The concepts of privacy and safeguards for data use – for instance, anonymisation, pseudonymisation, de- and re-identification – are very complex. It can therefore be challenging to ensure good public understanding and clarity on how data is used, however it remains important. We should be careful in drawing conclusions from this survey, which had a low response number and was not completely in line with the wider literature on such issues, nor our previous work engaging patients and the public. However, the results perhaps do highlight the need to clearly communicate the value of health and social care data. Specifically, the necessity to explain how personal data is safeguarded, can be made less identifiable and that to unleash the power of data, it often needs to be linked to other datasets and therefore can no longer be considered truly anonymous.

Given the benefits of simplifying data and appropriately informing the Public, organisations such as Understanding Patient Data (UPD) has produced public guides that explain large-scale health and social care data sets and pseudonymised patient data in a clear and balanced way[vii]. Set up in 2016, UPD aims to support conversations with the public, patients and healthcare professionals about uses of health and care data. This sort of work is essential to ensuring meaningful public discussions. However, they will not be funded from 2022 onwards, resulting in a significant gap and halt in progressing public and patient knowledge and engagement in health and social care data.

The Scottish Government should utilise these resources to clearly communicating concepts of privacy and the safeguards in place for data usage. They should also consider prioritising how to bridge the communication and engagement gap regarding health and social care data between the public and Government/healthcare organisations without the support of organisations such as UPD.

3B. To what extent do you believe it is important to collect data to enable our health and social care services to understand how they are serving those with protected characteristics?
(Very important / fairly important / neutral / not important)

Very Important - Health and social care should be accessible to all individuals, provide equitable outcomes, and not exacerbate existing inequalities.

To achieve this goal, it is important that health and social care data-sets are representative of the diverse Scottish population. Representative datasets can drive innovation, positively effecting research, specifically into rare diseases where incidence is already low and further shrinking of data could compromise analyses.

Through our most recent survey with people affected by cancer, a commonality amongst all participants was the recognition on the importance of addressing healthcare inequalities and willingness to share data to understand these inequalities, and drive innovation in research, treatment and care for those with protected characteristics. Many recognised the importance and value of their data to improving the experience and care of others:

"Using available data to identify health care inequalities is a very useful way of identifying under - treated communities […] to help improve adequate care and outcomes for the patients who live in these areas."

However, there are significant limitations to using some of the key datasets due to issues with completeness and timeliness[viii]. As an immediate priority, the strategy should prioritise providing timely, high-quality, transparent, and integrated data to enable a better understanding of the health landscape in Scotland. In our answer for question 7a, we highlight key priorities and steps to identify healthcare landscape including inequalities and drive progress using data. It is crucial to understand where improvements need to be made, and to measure and drive forward the progress within health and social care data.

3C. When thinking about health and social care professionals accessing and using your health and social care data, what more could be done to improve your trust?

There two key areas that we believe will improve public trust in health and social care professionals accessing and using patient data:

- Ensuring that data is stored, managed and used appropriately and the safeguards that exist are well understood and clearly communicated
- There is a need to prioritise authentic and meaningful public involvement and engagement throughout the data life cycle, including through the process of collection, usage, sharing and analysing the data.

Appropriate safeguards and privacy measures are essential to protect everyone's personal data, and indeed many of these are already in place. More should be done to communicate what safeguards already exist and how they are being adhered to and implemented.

CRUK strongly supports the development and implementation of secure data environments, including trusted research environments (TREs). These secure environments, where data users go to the data rather than data being shared into multiple places, are a new paradigm for how personal data can be accessed and used. They not only provide technical benefits, such as the ability to store and manage large volumes of data, it is also hoped they will build public and patient confidence in how their data will be accessed, even when it is not anonymised. It is important that the Scottish Government and data controllers consider how to effectively communicate what TREs are and how they mitigate risks associated with data sharing to the public.

In addition to findings from our most recent survey, through our engagement with our networks of people affected by cancer and wider research[ix], a common theme is that they want more information about the benefit and impact of data sharing and use.

The Scottish Government and health and social care organisations should ensure they can clearly articulate the importance of data sharing, including demonstrating the tangible impacts on individuals and population health. This should be via a range of methods but could include the development of case studies that highlight real life impacts of data sharing and taking opportunities to communicate how data drives insights have led to certain aspects care.

4. **When considering sharing of your data across the health and social care sector:**

4B. Under Data Protection legislation, your health and social care data can be shared in order to administer care. For what other purposes would you be comfortable with your health and social care data being shared within the health and social care sector?

There are diverse views regarding data use for purposes beyond direct care, but we believe that transparency, and meaningful involvement, engagement and communication are key to building trust in how data will be used so that the full potential of data can be harnessed. As mentioned in the previous answer, it must be authentic and meaningful and happen throughout the data life cycle. People must be assured that their data will be used ethically with proper safeguards and steps taken to mitigate risks, such as against misuse.

Our extensive experience consulting with people affected by cancer highlights strong support for data to be used to improve research, patient outcomes and patient experience. The advocacy group, UseMyData[x], exemplifies this general attitude. The group is a strong proponent of using health and social care data for research and improving care. Their vision is of every patient willingly giving their data to help others, knowing that effective safeguards to maintain the confidentiality and anonymity of their data are applied consistently, transparently and rigorously. This group was originally founded by people affected by cancer but has since expanded to include anyone who believes in their vision.

Additionally, studies find the public to be accepting of the idea of data related to their health and social care being used for purposes beyond direct care as long as it is for public benefit, and that this benefit outweighs risk. It's also important to them that the benefits are spread across the country and doesn't exacerbate inequalities[xi]. In addition, we have found in consultation with people affected by cancer, it's important to them that benefits of commercial access to data are returned to the health service in some way.

Overall, it is recognised that used to its full potential, information routinely collected by the NHS, in national datasets and medical records, is an invaluable resource driving research into the causes of disease, improving the effectiveness of diagnosis and treatments, and optimising NHS services. Previously, we conducted a survey with our patient panels during our response to the Scotland Cancer strategy consultation. Of the 25 responses, 94% of those who responded to our patient involvement network consultation said they support the use of data related to their care being used to improve cancer services and research.

Many recognised the importance and value of their data to improving the experience and care of others:

"Any data which can be used to help prevent cancers or improve the care for others should absolutely be used"

"I feel that anything that can be gained and used to help others based on my own experience means I didn't suffer for no reason."

Building on previous engagement results, in our most recent survey this trend has continued. However, many of them emphasised the prerequisite that their data should be anonymised due to concerns, including that data will be used for discriminatory purposes, it will not be safely managed, or won't be accurate. Thus, as mentioned above, this suggests there is work to be done in communicating what safeguards are in place, the value of data and how data can be used to drive lifesaving research.

Many public attitudes studies demonstrate concern about private companies' use of health and social care data, and some show low awareness of how patient data is used or why companies need access to patient data[5]. However, research by University of Sheffield and HRA[xii] found that when they provided public with information and clarity on why commercial entities need access to and use health and social care data through informative workshops, acceptance to share anonymised patient data with commercial organisations increased from 18 - 45%[xiii].

In order to build public trust in data uses beyond direct care, the Scottish Government, health and social care services and those who access and use health and social care data must ensure meaningful, transparent and ongoing public involvement and engagement. They must ensure that involvement and engagement activities reach all demographics of the population. Communications should also address key areas of concern for patients for different types of data use and different types of researcher access, what controls and sanctions are in place, and the options people have to opt-out, as well as the benefits of different types of data use.

Part 2: Empowering Those Delivering Health and Social Care Services

Empowering the people who deliver health and social care services by giving them the confidence and ability to securely access, gather and share relevant information to make timely decisions and deliver better outcomes.

6. Considering skills and training opportunities for those delivering health and social care services:

6A. What are the top skills and training gaps relating to data in Scotland's health and social care sector?

It is important for the Scottish health and social care services to identify and understand what data exists and where to find it.
Followed by:

- Knowledge of how to access data
- Confidence in using data
- Other – Data linkage and interoperability

We agree with the gaps highlighted but would suggest the following addition regarding data linkage:
From: Understanding of what data exists and where to find it
To: Understanding of what data exists, where to find it and how (and what) data can be linked

In addition to the specific training needs mentioned above, provision of foundational training on data should be prioritised. This should focus on supporting understanding the importance and value of data utilisation and its role in improving patient outcomes, service delivery and data driven innovation.

7. Thinking about improving the quality of data that is used by health and social care services:

7A. What three things are needed to improve quality and accessibility?

In our view, the three key things needed to improve the quality and accessibility of data are:

1. consistency and completeness in data collection;
2. transparent, appropriate and consistently applied access routes;
3. and improving the quality and availability of data relating to health inequalities.

Consistency, achieved through comprehensive data standards, and completeness in data collection will provide richer datasets that will enable effective comparison between regions and between the four UK nations. The ability to provide meaningful service level comparisons and effective identification of gaps in service delivery and resourcing will, in turn, support the design and implementation of targeted interventions and service improvements, ensuring quality and equity in care provision.

There is considerable burden and uncertainty surrounding data access routes for external researchers as well as those within the health and social care system. Implementation of consistently applied and transparent access routes and processes that are appropriate for the intended use of data are essential. Clarity over access will enable a quicker and more streamlined approach to service level audits, support the identification of potential areas which would benefit from increased resourcing or enhanced processes to support the delivery of care and unleash the power of data for research.

It is essential that improvements to the quality and availability of data relating to health inequalities are prioritised and that this data is utilised to improve population health and patient outcomes. Data should be routinely collected on protected characteristics including age, gender, ethnicity and sexuality, and socio-economic group, and published in a timely and accessible manner. Within cancer

datasets, this should include breakdowns on disparities across the pathway (from incidence, mortality, cancer waiting times, early diagnosis, survival etc) and regular reporting on those datasets including an assessment of impact on inequalities.
The Scottish Government and healthcare organisations such as Public Health Scotland should prioritise data collection to better understand any inequity of access to cancer care and research for those living in Scotland's urban areas and rural communities.

7C. What data, that is generated outside of the health and social care sector, do you think could be made available to health and social care professionals to improve health and social care outcomes in Scotland?
Social and administrative datasets, especially those related to wider determinants of health such as education information, household income band and housing status. This is essential to provide a more holistic understanding of the wider determinants of disease risk, outcomes and experience on a population level and to ensure individual patients and families receive the best support possible.

Additionally, the ability to compare with other countries and regions within the UK can also enhance understanding of health and health service performance.

8. We have heard that a more consistent approach to data standards will help improve insight and outcomes for individuals:

8A. To what extent do you agree with the proposal that Scottish Government should mandate standards for gathering, storing, and accessing data at a national level?
[agree, disagree. Unsure]

Agree – As mentioned in our answer to 7a (above), consistency in data standards supports effective use by health and social care professionals and researchers through the creation of richer and more complete datasets, which leads to more valuable insight for both individuals and the health service as a whole. As discussed in part 1 of this consultation response, public trust is essential for the collection of high quality, representative data. If people do not trust that their data will be handled safely, there is a risk that high numbers of people will opt out of data sharing, as exemplified by the recent backlash against the General Practice Data for Planning and Research Programme, launched by NHS England last year. As such, adherence to clear and consistent standards for data collection, storage, and dissemination is integral to reassure individuals that their data is being handled safely, mitigating the impact of self-exclusion through opt-outs.

8B. What data standards should we introduce?
The priority should be to ensure that the standards help implement adherence to the FAIR Principles (https://www.go-fair.org/fair-principles/). There are multiple standards available that the Scottish Government should consider, such as the Fast Healthcare Interoperability Resources standards and open EHR. Importantly, whichever standards the Scottish Government decides to mandate, they should be applied across the system, and they need to support interoperability. They must also be mindful that adherence to standards will need technical, subject matter expertise and financial support.

9. When considering the sharing of data across Scotland's health and social care system:

9A. Do you agree with the idea that greater sharing of an individual's health and social care data between the organisations in the health and social care sector will lead to better quality services? [agree, disagree, unsure]

Agree - Yes, as this will help reduce the risk that the patient is asked to relay the same information multiple times and should ensure that the information is accurate and accessible when needed. This could also enable H&SC systems to identify where there are gaps in services and the resources they need to address this.

10. Thinking about the actions needed to improve the quality of management information and internal reporting data across health and social care:

10A. What are the priority pieces of management information needed (that are not currently available) to provide better health and social care services?
We would prioritise data from primary care, which is invaluable for assessing the cancer pathway, from symptom presentation to treatment monitoring to community-based care. At present, data from primary care is not easily extracted from GP clinical systems, nor linked to the national cancer registry in Scotland.

Priority should be placed in actively including more primary care data into the cancer landscape and healthcare overall. Moreover, providing breakdowns and ensure granularity of data via demographics/region.

10B. What is needed to develop an end-to-end system for providing business intelligence for health and social care organisations in Scotland?

Along with the strategy, the Scottish Government should develop an overview of the current state of the data landscape with more detailed recommendations. This overview should highlight strands of activity, who is accountable for delivery and timelines/milestones. Recommendations should highlight standards for best practices within data to support health and social care organisations to utilise health and social care data effectively and safely.

Part 3: Empowering Industry, Innovators and Researchers

Ensure use of high-quality data to drive the development of new and improved, treatments, technologies, and ways of working for public benefit.

12. When considering the ethics of accessing health and social care data for commercial, development and research purposes:

12A. How do you think health and social care data should be used by industry and innovators to improve health and social care outcomes?

Data collected by the H&SC system is fundamental for progressing medical research. It is essential that innovators have access to H&SC data to drive insights into fundamental biology and natural history of disease; identify of risk factors associated with disease; uncover potential opportunities and targets for intervention; and ensure that services and interventions are as efficient and effective as possible.

The system should prioritise data access for innovators that is overwhelmingly in the public interest – i.e., research that clearly supports improving care and outcomes. H&SC Data should be used to develop, validate and enhance services and interventions as well as understand specific gaps and challenges that need addressing.

Industry access to H&SC data is essential to unleash the power of data for improving health but must be established in a safe, secure and transparent manner. At CRUK, we have developed, in close consultation with patients, a set of guiding principles for commercial data partnerships that protect patients, our partners and our data[xiv]. Patient involvement formed the bedrock from which our guiding principles and these principles will be reviewed and updated in consultation with patients at least annually to ensure that they are up to date and reflect the evolving landscape[xv].

12B. How can industry and innovators maintain the trust and confidence of the people of Scotland when using their health and social care data for research purposes?

A fundamental part of maintaining the social license to access health-related data is earning public trust and confidence. Therefore, it is essential that all organisations involved in the collection and use of health and social care data take opportunities to engage and involve data subjects in decisions about data access, and communicate the benefits, safeguards and current problems.
Industry and innovators must be transparent to data custodians and, by extension, the public through the register data access requests, about what data they need and for what purposes – this includes being clear on what data is being collected, what the intervention they are planning to develop is going to do and why they need that data. They need to be clear about how the data will be stored and accessed. As mentioned above, the development and implementation of clear and consistent standards will help to reassure data subjects that a robust system of governance is in place to protect their data and ensure that processing is proportional and compliant. Furthermore, industry and research organisations should consult directly with patients, for example through patient involvement panels, and should involve patients directly in the design and approval of projects, through inclusion on internal research ethics committees.

As discussed in question 4B, it is essential that organisations - particularly Scottish Government and NHS Scotland- provide a more comprehensive and understandable explanation of the purposes and benefits of allowing industry to access health and social care data to the public. For instance, this could include explanation of how research is translated into impacts for patients, and the role that commercial companies play in this process.

12C. What do you believe would be unacceptable usage of Scotland's health and social care data by industry, innovators, and researchers?

CRUK strongly believes that any use of data solely for commercial interests without any health or financial benefit for the data subjects or the health and social care services would be unacceptable. We also believe that the use of identifiable data without the knowledge of the data subjects and without any independent ethical oversight of the practices would be unacceptable.

Activities that unduly bar access to the data (exclusivity clauses notwithstanding), for example, researchers or research institutions retaining data on their own systems without cataloguing/making FAIR (e.g. so as to retain control over how and when secondary research is conducted and by whom) would also be unacceptable and should be avoided.
Data use needs to meet ethical and legal standards consistently and ultimately be in the interest of the data providers in the long term. We recommend that the provision of data for commercial purposes, even where that commercial purpose is the development of a health intervention, is seen as distinct to the provision of health and social care data to non-profit health research entities.

12D. How should industry, innovators and researchers be transparent about their purposes in accessing, and the benefits of using, health and social care data?

In general, as outlined in part 1 of this consultation response, we think it is important that all organisations that collect and use health and social care data make a concerted effort to improve overall public understanding and awareness about how and why their data is used.

In addition to broader communications, clear, concise, and easily understood information leaflets should be part of the consenting process or, where consent is not the basis for processing, should be provided in the form of a privacy notice. Additional materials and resources about the work or wider work should be available via websites, apps, community engagement events, and other publicity opportunities where the participants, patients and care communities, and wider public are able to understand and gain knowledge of the initiatives with ease. Details of data release/access should be readily available in the form of a public data release register which is regularly updated and signposted to data subjects (including where to find the public data registry and clarify what it is) to enable them to feel empowered over the use of their personal data.

Additionally, as discussed in question 4B, consultation with patient panels shows demand for more communication of the tangible benefits of data sharing. Suggestions included examples of instances where data highlighted problems and how they were solved, how it could be used within the health and social care system in Scotland to tackle inequalities, or when data collection and linkage has saved a life.

13. We want to create an infrastructure that supports access to data for research and innovation in a safe, secure, and transparent way:

13A. How should the Scottish Government seek to store and share health and social care data for research in order that it can best facilitate easier access that is still safe and secure?

Ensuring that the data storage and processing facilities are secure, useable, sustainable and reliable is key. This will involve recruiting suppliers that work to the highest independently certified standards (e.g., ISO 27000 Series and preferably ISO 9000 series). Where possible, there should be support for healthcare providers to develop their own solutions as appropriate, but it is likely more reliable to work with industrial suppliers for infrastructure and ensure best value for money as well as high quality services. It is crucial that whatever systems are employed, they are interoperable, auditable and work transparently.

In general, CRUK is very supportive of the move towards secure data environments, and particularly, trusted research environments. Having data users come to the data rather than the data being shared to many different places, which risks creating silos, is positive for data security, privacy and transparency and should ultimately be beneficial for gaining public confidence. However, there is still a technical gap which needs to be filled in order for researchers and commercial organisations to be able to use their tools without barriers within these systems, stifling innovation/access in the short to mid-term.

13B. What do you believe are the key data needs and gaps that are faced by industry, innovators, and researchers when it comes to Scotland's health and social care data?

Cancer datasets

There are significant limitations to using some of the key datasets for cancer due to issues with completeness and timeliness. In its first 3 years, the strategy must prioritise providing timely, high-quality, transparent, and integrated data to enable a better understanding of the cancer landscape in Scotland and particularly where variation exists. This is crucial to understand where improvements need to be made, and to measure and drive forward the progress set out in the rest of the strategy.

Specifically, CRUK recommends that the strategy commits to:

- Maintain high quality reporting of cancer incidence and mortality, wherever possible broken down by tumour site and socio-economic status.
- Extend publicly available staging data to all cancers combined as well as a wider range of cancer site specific data. Data should be published routinely and regularly, with information on individual staging categories available.
- Consider additional measures which place renewed focus on diagnosis for patients alongside maintaining and further understanding variation in Cancer Waiting Times data.
- Continue to improve the completeness, accuracy and consistency of primary care data relevant to informing the cancer pathway and patient care and integrate this with other patient data to allow for analysis of the whole cancer pathway.
- Consider routinely publishing data on routes into diagnosis, especially emergency presentations by cancer tumour site and Health Board
- Consider integrating PROMs into QPIs rather than treating as separate priorities.
- Improve the quality and availability of data relating to health inequalities and utilise this to improve population health and patient outcomes. Data should be routinely collected to better understand any inequity of access to cancer care and research for those living in Scotland’s urban areas and rural communities.
- Consider collaboratively developing a set of agreed minimum indicators for understanding the entire cancer pathway in Scotland.

Informatics and analysis

Broadly speaking, the strategy for health and social care data in Scotland should seek to align with and support the delivery of the transformative plans captured in the Scottish Cancer Intelligence Framework (SCIF), which is being delivered primarily via Public Health Scotland. This includes delivering on the ambitions to:

- Deliver a ‘Once for Scotland’ approach, joining up intelligence across systems, and support interoperability at local and national levels.
- Place an emphasis on financial resourcing to attract and retain high quality analytical staff.
- Ensure any adjunct cancer workforce plan includes cancer analysts and data scientists as a distinct resource category.
- Ensure proportionate emphasis on effective resourcing of MDT data management and Health Board level data collection.
- Ensure GPs and other primary care health professionals and services are aware of, engaged with and able to utilise the primary care data opportunities presented by SPIRE, both across the full primary care pathway and at the interface with secondary care.

CRUK welcomes the recent developments to the Scottish Cancer Registration and Intelligence Service (SCRIS) that aim to make cancer data more readily available and make data on waiting times, screening, diagnosis, and treatment more easily linked to the Registry. CRUK recommends that the strategy builds on this momentum to:

- Harness the increased data and intelligence available through SCRIS and continue to develop SCRIS ambitiously and incrementally to meet patient and clinical needs, and support service improvement.
- Continue to progress the ambition to harness primary, secondary and social care data.
- Ensure sufficient capacity and capabilities to generate intelligence from data.
- Realise the benefits of having a national view of SACT data in Scotland.

Acting on the opportunities outlined here in the short to medium term would provide an evidence base to measure and monitor progress in the longer-term timeframe of the strategy. Longer term, the strategy must ensure there are continual improvements in cancer data collection and use. This should include working towards real-time or near real-time datasets as well as improving the ability to collect, store and integrate diverse patient datasets such as imaging data, genomic data, GP records and electronic health records.

Data is fundamental to research into cancer risk, biology, the development of effective ways to diagnose and treat cancer, and optimising cancer care. Researcher access to cancer data is integral to reducing cancer incidence and improving patient outcomes. Therefore, the strategy for health and social care in Scotland must include an overarching commitment to improve access to data for research in a timely manner. This should include commitments to:

- The development of service level standards for data access for research
- Collaboratively developing a data sharing and access framework, bringing together the health system, data custodians, academic and industrial researchers, and patients.
- Working with Scottish data controllers and authorities on ensuring that data can be accessed within TREs and secure environments. It is essential that these are interoperable with other TREs, including CRUK’s TRE that is UK wide in scope.

Fundamental to the above, the Scottish government must ensure Information Governance arrangements maximise data utilisation for the benefit of the public. We also recommend that NHS Scotland explore a streamlined, rapid process for appropriately accredited stakeholders including researchers, third sector and NHS. It will be essential to embed meaningful public involvement and engagement in such exploration from the outset.

We propose that NHS Scotland considers introducing a new cancer waiting times standard for 95% of people who receive an urgent suspected cancer referral or have been picked up through cancer screening, to have cancer ruled out or receive a diagnosis within 28-days. Given ongoing difficulties in cancer services, exacerbated by the pandemic, it would be appropriate to phase the introduction of this target – for example initially aiming to meet an 80% threshold, while setting out a clear plan and timeline for meeting the higher target.

In introducing the new target, the Scottish Government should evaluate the implementation of the Single Cancer Pathway in Wales and the Faster Diagnosis Standard in England, for example looking at how to secure alignment between different standards and cancer pathways and considering which patients may not be captured.

Measuring the time taken for someone to receive a definitive diagnosis or ruling out of cancer could offer significant benefits. This would be a meaningful target – both clinically and to patients – that could help catalyse and monitor diagnostic performance, providing further incentives to accelerate a part of cancer pathway that is meaningful to patients and encouraging the faster diagnosis of more cancers. It would enable evaluation of how the full diagnostic pathway is performing. Using one clear headline performance figure would also be helpful for clear communication with patients helping improve patient experience by setting clear expectations and helping ensure they receive a diagnosis as soon as possible during what can be an incredibly anxious time.

14. Used appropriately and well, technologies such as Artificial Intelligence can help to improve decision making, empower health workers and delivery higher quality health and social care services to citizens, improving how you receive health and social care services:

14A. What are your views on the benefits of using AI to improve the delivery of health and social care services?

AI in health and social care has significant, positive potential to improve and optimise the delivery of health and social care services. The benefits of AI includes speed and accuracy in clinical decision making but also the ability to derive novel insights not possible with humans very important for the management of long term or complex diseases where there is significant data from a range of different sources to be crunched.

For instance, AI can be particularly useful in recognising anomalies in images that are hard or impossible for the human eye and this can flag to a human radiologist for example cases that should be followed up during screening services. Given the massive learning capacity of deep learning algorithms, it qualifies them to handle such variance and detect characteristics well beyond those considered by humans[xvi]. Moreover, the use of AI in digital pathology setting to make predictions regarding treatment response can enable the selection of more effective treatments for patients[xvii].

However, to optimise the benefits, the development of AI needs to be explainable so that it can be trusted by clinicians and effectively used. Priority should be placed in clearly communicating that AI will be a tool for clinicians and not replacement of human experts, including care teams, diagnostics and public service provision managers.

Additionally, care should be taken when using AI in health and social care services. For example, if AI is used for scheduling appointments, it may unfairly bias against certain people, for example, if they have had a history of failing to attend appointments (e.g., an AI algorithm wouldn't necessarily be trained to detect when someone had missed an appointment because their child was ill or they had had an accident).

AI supported interventions should prioritise where machine learning algorithms are better adept at analysing data but with the caveat that its outputs need to be managed and interpreted by the professionals that are using it, and this includes cases where patients are self-managing and being guided by any such intervention. Moreover, it needs to ensure that datasets used in AI training are diverse so that the application is generalisable to different demographics and equipment types.

Overall, AI when used appropriately can support clinicians to effectively deliver health and social care services. AI is to be used to work in a more supportive way to process data from different systems within the hospital and ensure that the clinician is seeing the most important aspects to direct clinical decision making.

14B. What safeguards do you think need to be applied when using AI?

AI can have significant benefits in optimising the delivery of health and social care services. However, there is a need to ensure and implement appropriate safeguards to ensure appropriate and safe use.

Firstly, current statistical methods struggle with incomplete, inconsistent, heterogeneous and unstructured data. Utilising AI within this setting may help, but there is a need to develop new approaches. Thus, the limitations of AI in the context of health and social care data should be identified and communicated to clinicians, and health and social care staff in a clear, concise and simplified manner.

Moreover, there is a need for an adaptable regulatory system that is able to (1) Validate performance before integration into care pathways (2) Quickly adapt as the system continues to learn and improve.

Example of key challenges for AI include:

- Population bias in the source data, where certain demographics will be less represented than others and the outputs may not be attuned to their needs or situation;
- Lack of explainability for AI driven interventions, which may mean it is impossible to assure that they are working correctly or that the results they are giving are safe, accurate and representative; this causes issues for transparency with data subjects as well.
- Lack of autonomy – especially where data about people is being used to train algorithms without their knowledge and potentially to their detriment;
- Safety and accuracy – especially where black box training is concerned (per above).

---

[i] Edelstein, Michael et al. “Strengthening Global Public Health Surveillance through Data and Benefit Sharing.” Emerging Infectious Diseases vol. 24,7 (2018): 1324–1330. doi:10.3201/eid2407.151830 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6038731/

ii Scottish Government. “Survey shows high levels of trust in Scottish Government” (2020) https://www.gov.scot/news/survey-shows-high-levels-of-trust-in-scottish-government/

iii Gidman W, Ward P, McGregor L. “Understanding public trust in services provided by community pharmacists relative to those provided by general practitioners: a qualitative study” BMJ Open 2012;2:e000939. https://doi.org/10.1136/bmjopen-2012-000939

iv Linda A. Jones et al, (2021): “Public opinion on sharing data from UK health services for clinical and research purposes without explicit consent”. 21260635; https://doi.org/10.1101/2021.07.19.21260635

v NHS Health Research Authorities (HRA), (2019). “Sharing anonymised patient-level data where there is a mixed public and private benefit - a new report” https://www.hra.nhs.uk/about-us/news-updates/sharing-anonymised-patient-level-data-where-there-mixed-public-and-private-benefit-new-report/

vi Centre for Data Ethics and Innovation (CDEI) “Public attitudes to data and AI” (PADAI) – Tracker survey (2021)
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1092140/Public_Attitudes_to_Data_and_AI_-_Tracker_Survey.pdf

vii Understanding Patient Data “What does 'anonymised' mean?”
https://understandingpatientdata.org.uk/what-does-anonymised-mean

[viii] Katikirreddi et al (2018) “Assessment of health care, hospital admissions, and mortality by ethnicity: population-based cohort study of health-system performance in Scotland”
https://www.sciencedirect.com/science/article/pii/S2468266718300689

[ix] Jones et al (2021). “Public opinion on sharing data from health services for clinical and research purposes without explicit consent: an anonymous online survey in the UK”
https://bmjopen.bmj.com/content/12/4/e057579.abstract

[x] Use My Data https://usemydata.org/about.php

[xi] Understanding Patient Data (2020) “What do people think about third parties using NHS data?”
https://understandingpatientdata.org.uk/what-do-people-think-about-third-parties-using-nhs-data

[xii] HRA & University of Sheffield (2019). “Sharing anonymised patient-level data where there is a mixed public and private benefit, HRA and University of Sheffield” https://www.hra.nhs.uk/about-us/news-updates/sharing-anonymised-patient-level-data-where-there-mixed-public-and-private-benefit-new-report/

[xiii] HRA & University of Sheffield (2019). “Sharing anonymised patient-level data where there is a mixed public and private benefit” https://www.hra.nhs.uk/about-us/news-updates/sharing-anonymised-patient-level-data-where-there-mixed-public-and-private-benefit-new-report/

[xiv] Cancer Research UK (CRUK). “Our guiding principles for data partnerships”
https://www.cancerresearchhorizons.com/our-guiding-principles-data-partnerships

[xv] Cancer Research UK (CRUK). “Our guiding principles for data partnerships”
https://www.cancerresearchhorizons.com/our-guiding-principles-data-partnerships

[xvi] Hosny, Ahmed et al. “Artificial intelligence in radiology.” *Nature reviews. Cancer* vol. 18,8 (2018): 500-510.
https://doi.org/10.1038/s41568-018-0016-5

[xvii] Hosny, Ahmed et al. “Artificial intelligence in radiology.” *Nature reviews. Cancer* vol. 18,8 (2018): 500-510.
https://doi.org/10.1038/s41568-018-0016-5