House of Commons Science and Technology Committee Inquiry – The right to privacy: Digital data

Cancer Research UK response – January 2022

Summary

1. Cancer Research UK (CRUK) welcomes the opportunity to respond to the Science and Technology Committee's inquiry on digital data.
2. In 2020/21, CRUK invested £421 million on new and ongoing research projects into the causes and treatments of cancer, and our long-term investment in state-of-the-art facilities has helped to create a thriving network of research at 90 laboratories across the UK, supporting the work of over 4,000 scientists, doctors and nurses. Our work is shaped by the unique knowledge and experience provided by our network of 1,800 people affected by cancer and our five cancer patient panels.
3. As a research funder and advocate for cancer patients, CRUK has a strong interest in ensuring data can be effectively shared and used to improve cancer outcomes. Data and data-driven technologies can drive research into the causes of disease, improve the effectiveness of diagnosis and treatments, and optimise NHS services.
4. The UK has the potential to realise the Government's ambition of being a science superpower in this field, building on its expertise in data science and benefitting from the enviable position of having healthcare systems that generate and collect a wealth of information.
5. However, much of the potential of data is untapped. There are many reasons for this, including problems with the infrastructure for managing and linking data; uneven data quality; poor integration and interoperability; a fragmented health system; complex processes governing how data is accessed; and delays in accessing data.
6. **It is essential that improvements are made to the quality, completeness, timeliness and interoperability of data and datasets in order to unlock their full potential, along with a commitment to increase resources for training the workforce in data skills.**
7. **There must be sufficient and ongoing guidance and support for individuals who are responsible for making access decisions. In addition, consideration should be given to how to simplify and speed up access for researchers who meet certain conditions.**
8. It is positive that the Government has recognised the need to address barriers to effective data use. However, we have concerns about some proposals in recent strategies and consultations:
    - We do not believe that many of the legislative changes set out in the consultation, Data: A new direction, are necessary. **CRUK believes better guidance and training would address many of the problems identified while guarding against risks. We also strongly urge the Government to avoid changes that would put at risk UK's data adequacy agreement with the European Union, which is vital for scientific research.**
    - The proposals in NHSX's Health and Social Care Data Strategy are wide-ranging and ambitious. However, **we would like to see more clarity on the proposed sequencing and prioritisation of the commitments and budget available. In addition, more detailed consideration is needed of public engagement and inequalities.**
9. The importance of public trust cannot be overstated – nor can it be taken for granted. Consultation with people affected by cancer highlights support for data to be used to improve research, patient outcomes and patient experience, as long as certain conditions are met. These include clear communication and engagement, strong safeguards and sanctions for misuse. **The Government must ensure meaningful, transparent and ongoing public engagement related to**

data use and sharing, and this work must reach an audience that accurately reflects the demographics of the population.

10. It is essential that patients and the public have access to information on both the benefits and risks of data sharing, and that there is the chance for open conversation on matters of significant concern, such as commercial access to data. **CRUK recommends the Government and health service develop a comprehensive, coherent plan to communicate the benefits of data sharing. Organisations – particularly Government and the NHS – must also provide a more comprehensive and understandable explanation of the purposes and benefits of allowing commercial access health data.**
11. Trusted Research Environments (TREs) are a popular concept in the debate about security, as they provide a secure space for researchers to access and analyse data. We strongly support the use of TREs, and are developing our own to support work to beat cancer. **CRUK recommends close attention be paid to ensuring that different TREs can work together. Standards should be developed collaboratively across the sector, with an independent body to certify them.**
12. Finally, CRUK notes that the move of controls over access to certain data from Public Health England to NHS Digital, and soon again from NHS Digital to NHS England, has the potential to cause further delays and confusion. **Every effort must be taken to identify and protect against unintended consequences. We would urge the Committee to consider this issue in its inquiry.**

1. The potential benefits, including to research, to effectively use and share data between and across Government, other public bodies, research institutions and commercial organisations, and the existing barriers to such data sharing

Benefits to effective use and sharing of data

13. One in two people born since 1960 will be diagnosed with cancer in their lifetime. Better data collection, linkage and use are essential if we are to achieve our goal of beating cancer. The UK has the potential to become a world leader in the field, setting the country apart as a destination for life sciences research and stimulating international collaboration. The UK Government has recognised this potential in its Life Sciences Vision, noting that “it is a precondition to the success of this Vision that the UK seizes the opportunity provided by health data”.[i]

14. The UK has healthcare systems that generate and collect a wealth of information, creating rich, population-level datasets on a large, diverse population.[ii] Cancer data is a particularly rich dataset; England’s National Cancer Registration and Analytics Service collects data on all cases of cancer that occur in people living in England, receiving about 25 million records each year.[iii] In addition, there is data collected by national agencies and projects such as the UK Biobank, alongside information held in wearables and other citizen-generated data.

15. As the Committee notes, data sharing takes place between various actors and organisations. These flows can be complex, and the benefits will differ depending on the situation. The ability to share data across the health service can provide patients with more effective, joined-up treatment as well as creating efficiencies in the system. Data shared with non-profit organisations, like CRUK, allow us to produce statistics that help improve public understanding and advocate for evidence-based interventions.[iv] For-profit organisations help to produce diagnostics, treatments and drugs for the benefit of patients, which would not be possible without secure and responsible access to data. Inherent in the examples below is the benefit of data sharing between organisations.

16. Data can improve understanding of individual risk and the ability to detect and treat disease earlier. This is hugely important: 92% of patients with bowel cancer diagnosed at stage 1 survive their disease for at least five years, compared with 10% diagnosed at stage 4.[v] The potential to deliver an individual, dynamic cancer risk scale, based on algorithmic analyses of primary and secondary electronic health records, genetic profiling, family history, data on individual lifestyle, diet, behaviour and occupational health, as well as environmental information, has been noted in analyses of the early detection and diagnosis landscape.[vi]

17. Artificial intelligence systems can support and speed up diagnoses, allowing clinicians to spend more time with patients. CRUK jointly funded OPTIMAM, a database that contains more than 3.3 million images from mammography scans.[vii] Breast cancer detection is a good target for AI because automation can save costs and processing times in what is a highly complex process. CRUK has enabled 14 commercial organisations to access to this data, in a controlled and legally regulated way, leading to development of products that are being used in the clinic to reduce radiologist workloads and, potentially, to increase accuracy of breast screening.[viii][ix]

18. Non-clinical data can also reveal important insights, particularly when linked with clinical outcomes data. For example, a CRUK-funded study at Imperial College London aims to reduce delays in ovarian cancer diagnosis by analysing high street retailers’ loyalty card data that participants have shared with researchers. The project is assessing whether there were significant changes in purchases, such as over-the-counter medication that treats symptoms, among women who have been diagnosed with ovarian cancer, before they were diagnosed.[x]

19. Consultation with people affected by cancer shows demand for more communication of the tangible benefits of data sharing. Suggestions included instances where data highlighted

problems and how they were solved, how it could be used within the NHS to tackle inequalities, or when data collection and linkage has saved a life.

20. **CRUK recommends the Government and health service develop a comprehensive, coherent plan to communicate the benefits of data sharing that makes clear the role of all Government and health service bodies in this work.**

Barriers to effective use and sharing of data

21. Barriers to effective use and sharing of data span the whole data lifecycle, from collection to access. The data itself may be of uneven quality or contain gaps or biases. Technologies and policy interventions developed using datasets that lack information about a certain group will, at best, not serve whole populations and, at worst, exacerbate inequalities.[xi] Standardisation of data collection and access is also needed across the UK to provide more consistent data availability for digital patient records and enable cross-border benchmarking and comparisons.

22. There are well-documented problems that persist with both Government and health service technology systems, which are often outdated and inefficient.[xii] [xiii] Further to this, there is a lack of integration between different types of electronic health records (e.g. GP and hospital records), and fragmentation and poor interoperability of data systems between NHS practices, hospitals and trusts. Alongside this there is a lack of streamlined, centralised points of data access for real-time clinical data.[xiv]

23. Some of the above barriers are exacerbated by a lack of knowledge of data handling processes or awareness of the importance of data, and the best collection practices, within the health service.[xv]

24. **Government and the health service must deliver improvements in the quality, completeness, timeliness, linkage and interoperability of health-related data, and commit to increase resources for training the workforce in data skills.**

25. A significant barrier to effective data use, and particularly research, is access, with governance processes being inconsistent and expensive, and application processes complex.[xvi] In 2019, CRUK submitted data access applications to a national data holder for two research projects with almost identical design, but one was approved after two months, the other after eight. The reason for the difference is unclear. Delays to legitimate researcher access to data can be significant and can be a blocker to discovery and validation. We are also concerned about the effects that this could have on the development of research talent.

26. **We recommend that the Committee considers how best to simplify and speed up access for researchers who meet certain conditions, for instance by exploring the possibility of a streamlined process for accredited researchers.**

27. We have also experienced delays in our efforts to progress the COVID-RT study to understand why changes in radiotherapy schedules were implemented during the pandemic and explore the impact of these changes on patient outcomes.[xvii] This study seeks to bring together data from radiotherapy centres across the UK and requires data from all four of the UK cancer data custodians. Our discussions to gain access to the data specifically collected for this study by the centres have been ongoing with some data custodians for 15 months and, so far, we have only received data from the one centre in Northern Ireland. Having to navigate four different processes to access these data is a challenge, but it has been further exacerbated in England by the delays caused by the transfer of Public Health England into NHS Digital and the lack of a data access process for two months, which brought delays to the processing of all applications for far longer. Given the ongoing importance of research into Covid-19 and the nature of a fast moving and evolving pandemic, delays of this kind act as a serious barrier to research and, ultimately, threaten the viability and usefulness of studies which are time sensitive.

28. **We recommend that the Committee considers how to ensure timely access to data that is appropriate for the organisation requesting it and the purpose of the study, and how the different processes across the four UK nations affect access to data – and therefore delay research outputs which could improve patient outcomes.**

29. In our experience, a lack of expertise or confidence interpreting complex data protection legislation among data custodians can lead to an overly risk-averse approach in making decisions to grant access to data. In particular, we believe a lack of guidance and support for Caldicott Guardians has led to a risk averse approach at Trust level.

30. **CRUK strongly recommends sufficient and ongoing guidance and support for individuals who are responsible for making access decisions, to avoid what we currently see as an overly rules-based approach.**

31. We also note that, to facilitate the transfer of datasets from Public Health England to NHS Digital, access requests were put on pause for two months. Although requests are now being processed, there has been little information on the size of the backlog or how requests will be prioritised. And there will be another upheaval this year, as NHS Digital is merged into NHS England.[xviii] This has the potential to cause further delays and confusion, and there is not yet much public information on the move. However, it may also offer an opportunity to streamline and improve data access, which we urge the Government and NHSE to take.

32. **Every effort must be taken to identify and protect against unintended consequences of the merger of NHS Digital into NHS England. We would urge the Committee to prioritise this issue in its inquiry.**

33. It is important to stress that good information governance is not in itself a barrier to data use. Rather, it is essential to ensure that patient data is used in a secure and privacy protecting manner. This is fundamental to demonstrate trustworthiness and therefore build and maintain public trust – damage to that trust could become a barrier to effective data sharing. We discuss this in more detail in question 3.

2. The extent to which data issues are appropriately addressed by the Government's National Data Strategy, its draft strategy, data saves lives: reshaping health and social care with data, and its consultation Data: a new direction

34. It is positive that the Government has recognised the barriers to effective data use and is seeking to address them. However, CRUK has concerns about some of the proposals recently set out by Government. Below is a summary of the key points we made in our full responses to the Health and Social Care Data Strategy[xix] and Data: A New Direction[xx].

Health and Social Care Data Strategy

35. The strategy is wide-ranging and ambitious, which we welcome. **However, we want to see more clarity on the proposed sequencing, prioritisation, funding and resourcing commitments available to ensure effective delivery**.

36. We support commitments to improve researchers' access to data; to improve the completeness, quality and interoperability of datasets and governance around collection and management; and actions to upskill healthcare professionals and data custodians. Other areas seek to support clinical research, service planning, innovation and improvements to technical infrastructure appear to be going in the right direction but lacked detail.

37. There were also areas that we felt were not considered in sufficient detail:

- Fragmentation within the system: We would like to see proposed mechanisms to foster collaboration across departments, and national, regional and local actors, and ensure all these actors understand their different roles and responsibilities.
- Bias and inequalities: The strategy should more clearly recognise the impact that commitments may have on certain groups, and state how it will mitigate those risks. This includes digital inequalities in terms of public access to and use of technology.
- Public and patient involvement and engagement: It was disappointing not to see a more significant section on public engagement in the strategy, but we expect that this will be addressed in the final document.

Data: A new direction

38. We support the Government's ambition to improve researchers' access to data, and acknowledge that different interpretations of the UK General Data Protection Regulation (GDPR) within and between organisations can cause confusion and delays to sharing patient data.

39. CRUK is concerned that there is a misconception that regulations for protecting data and research governance are hindering innovation. **CRUK, in common with the wider medical research sector, does not believe there is a strong case for the legislative changes proposed in Chapter 1 of the consultation. Better guidance and training would address many of the problems identified while guarding against risks:**

- Improved, coordinated cross-sector guidance on issues including broad consent, identification of appropriate legal basis, the use of legitimate interests, and further processing.
- Education and upskilling of data custodians, researchers and host organisations. Indeed, where a lack of confidence or knowledge poses a barrier, changes could cause greater confusion.

40. We are also concerned about the risk posed to the UK's data adequacy agreement with the European Union. Joint UK-EU clinical trials rely on routine international data transfers – 32% of CRUK clinical trials involve an EU member state. Without an adequacy agreement, UK-based researchers would need to arrange alternative mechanisms that would increase the cost of research through added complexity and legal fees, which could deter EU research studies from involving UK-based researchers.

41. **We strongly urge the Government to avoid making unnecessary legislative changes that will affect the UK's adequacy agreement with the EU.**

3. The ethics underpinning the use and sharing of individuals' data in health and care contexts

42. A key part of any discussion of ethics in use and sharing of data is the issue of public trust. Consultation with people affected by cancer highlights strong support for data to be used to improve research, patient outcomes and patient experience. They emphasise the need for patients to know how and why data will be used; that the organisations that have access to it are ethical, responsible and accountable; and clarity on safeguards, controls and sanctions for misuse. We have found this messaging to be consistent over time: whether in our most recent e-consultation for the Health and Social Care Data Strategy or our larger Review of Informed Choice for Cancer Registration in 2016.[xxi] It is also consistent with results from broader public attitudes studies. [xxii]

43. The Government must ensure meaningful, transparent and ongoing public engagement related to data use and sharing. Strategies and proposals regarding use of personal data

should include a detailed plan for public engagement that reaches an audience that accurately reflects the demographics of the population.

44. Many public attitudes studies demonstrate concern about private companies' use of health data, and some show low awareness of how patient data is used or why companies need access to patient data.[xxiii]

45. **It is essential that organisations - particularly Government and the NHS - provide a more comprehensive and understandable explanation of the purposes and benefits of allowing commercial researchers or companies to access health data. 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.**

46. We also support the Association of Medical Research Charities recommendation that the use of health data for purposes beyond health must reflect public expectations and consider patient benefit.

47. Mistakes – real or perceived – could negatively impact research and outcomes. If large numbers of people opt out of sharing their data it risks reducing the completeness and quality of the datasets. This could exacerbate existing health inequalities if certain groups are more likely to opt out. In the first full month after concerns were raised about the GP Data for Planning and Research programme, National Data Opt Outs increased by more than a million.[xxiv]

48. **Public trust cannot be taken for granted, and Government must acknowledge and learn from its mistakes.**

4. The extent to which appropriate safeguards and privacy are applied in the usage and sharing of individuals' data

49. Appropriate safeguards and privacy measures are essential to protect everyone's personal data, and indeed many of these are already in place. But data sharing can never be risk-free and it is important that both patients and the public have access to clear and understandable information on the risks and safeguards in place, as well as on the benefits of data sharing.

50. Many of the concepts of privacy and safeguards for data use – for instance anonymisation, pseudonymisation, de- and re-identification – are complex and not well understood, even within the sector. Understanding Patient Data (UPD) is working to produce public guides that explain large-scale health data sets and pseudonymised patient data in a clear and balanced way.[xxv] This sort of work is essential to ensuring meaningful public discussions.

51. **The Government and organisations across all sectors must consider how they can carry out and contribute to such work, especially as UPD will not be funded beyond 2022.**

Trusted Research Environments

52. Anonymity is often virtually impossible to achieve for health data, especially when multiple datasets are linked together, and we should be vigilant to the risk of re-identification. Furthermore, harnessing the full potential of data – including for cancer research – often requires some degree of linkage, and so completely anonymous data would not properly benefit researchers or patients.

53. Trusted Research Environments (TREs) provide a secure space for researchers to access and analyse data without the data leaving that environment. Working within a TRE provides the necessary security required when working with data that may be identifiable, while assuring compliance with GDPR and other regulations.

54. CRUK strongly supports the development and use of TREs. We have been using a TRE internally for our analyst teams for some time and have established a TRE to support our funded researchers and other partners who want to work with us to beat cancer sooner. We are happy to provide the Committee with more information on how we are developing our TRE.

55. As numerous TREs are developed by different organisations it is vital that they are able to work together, and that both NHS and non-NHS data can be linked within them. Equally important are the creation of common standards and certification for TREs.

56. CRUK recommends close attention be paid to ensuring that different TREs can work together. Standards should be developed collaboratively across the sector, with an independent body to certify them, as well as alignment with wider international efforts in this space.

5. The effectiveness of existing governance arrangements, e.g. the Centre for Data Ethics and Innovation

57. The office of the National Data Guardian (NDG) is a significant and positive contribution to the sector. The NDG effectively carries out its function of providing advice and challenge to the health care system, as well as overseeing the important role of Caldicott guardians within NHS Trusts, and contributing essential research and public dialogues to the benefit of all those within the sector. **The National Data Guardian's work must be appropriately resourced.**

About CRUK

Cancer Research UK (CRUK) is the world's largest cancer charity dedicated to saving lives through research. We support research into over 200 types of cancer, and our vision is to bring forward the day when all cancers are cured. 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, Cancer Research UK invested £421 million on new and ongoing research projects into the causes and treatments for cancer.

CRUK is a patient-centred organisation that routinely partners with people affected by cancer to improve the relevance, quality and impact of our work. We consult, collaborate and partner with a network of 1,800 people affected by cancer who shape our work by sharing their unique knowledge and experience of cancer. Alongside our network we run five cancer patient panels, including one jointly run panel with the British Heart Foundation that focuses on the use of, and access to, patient data.

In addition to this, CRUK also uses patient data in its work, analysing national cancer datasets to improve cancer patients care and outcomes, and has been doing so for almost two decades. We apply for access to patient data from NHS Digital, Public Health England and other organisations that hold datasets on patients. Within CRUK, there are strict rules in place around who can use the data, what we can use it for and how long we can have it.

We are one of only two health charities with 'safe haven' status, which allows for enhanced access to high risk data, including our direct access to data stored on the Health and Social Care Network. Our accreditation is supported by our ongoing compliance with the NHS Data Security Protection Toolkit, the provision of a Cancer Intelligence Safe Room from which to access this data, CRUK's proven track record in storing and processing data compliantly through our internal Secure Data Environment, and our continued position at the forefront of driving robust policies and processes for the secure handling of data (such as our early adoption of the '5 safes' model for data security which is now promoted by HDR-UK as the gold standard for secure research).

For more information on this response, please contact Rebecca Hill (science policy advisor) on Rebecca.Hill@cancer.org.uk

---

References

[i] Life Sciences Vision, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1013597/life-sciences-vision-2021.pdf

[ii] Life Sciences Industrial Strategy, 2017 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/650447/LifeSciencesIndustrialStrategy_acc2.pdf

[iii] National Cancer Registration and Analysis Service https://www.ndrs.nhs.uk/wp-content/uploads/2020/07/National_Cancer_Registration_Guide-1.pdf

[iv] Cancer Statistics for the UK, CRUK https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk

[v] Cancer Research UK. Bowel cancer statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer

[vi] Priestman, E. 2020. Early Diagnosis and Detection of Cancer: A health economist's perspective. https://www.medtechnavigator.co.uk/app/uploads/2020/11/White-Paper_Early-Diagnosis-Detection-of-Cancer.pdf

[vii] OMI-DB OPTIMAM Medical Image Database https://medphys.royalsurrey.nhs.uk/omidb/about-omi-db/

[viii] McKinney, S.M., Sieniek, M., Godbole, V. *et al.* International evaluation of an AI system for breast cancer screening. *Nature* **577,** 89–94 (2020). https://doi.org/10.1038/s41586-019-1799-6

[ix] Cancer Research UK, 2021. Speaking a thousand words – how a cancer image collection is set to improve AI diagnosis https://news.cancerresearchuk.org/2021/09/08/speaking-a-thousand-words-how-a-cancer-image-collection-is-set-to-improve-ai-diagnosis/

[x] Imperial College London. Cancer Loyalty Card Study. https://www.clocsproject.org.uk/

[xi] Nordling, L. 2019. A fairer way forward for AI in health care. https://www.nature.com/articles/d41586-019-02872-2

[xii] Gallagher, Y. Six reasons why digital transformation is still a problem for government https://www.nao.org.uk/naoblog/six-reasons-why-digital-transformation-is-still-a-problem-for-government/

[xiii] National Audit Office, 2020. Digital transformation in the NHS. https://www.nao.org.uk/wp-content/uploads/2019/05/Digital-transformation-in-the-NHS.pdf

[xiv] Association of the British Pharmaceutical Industry, 2020. Unlocking the promise of UK health data. https://www.abpi.org.uk/media/8497/abpi_health-data-report.pdf

[xv] Cancer Research UK, 2020. Cancer Research UK's Vision for NHS health data in England https://www.cancerresearchuk.org/sites/default/files/cruk_vision_for_data_jan_20_1.pdf

[xvi] Harwich, E and Lasko-Skinner, R. 2018. Making NHS data work for everyone. https://reform.uk/sites/default/files/2018-12/Making%20NHS%20data%20work%20for%20everyone%20WEB_1.pdf

[xvii] National Cancer Research Institute, 2020. NCRI's CTRad to lead COVID RT to study the impact of COVID-19 on radiotherapy https://www.ncri.org.uk/covid19-radiotherapy-initiative/

[xviii] Wade-Gery, L. 2021. Putting data, digital and tech at the heart of transforming the NHS. https://www.gov.uk/government/publications/putting-data-digital-and-tech-at-the-heart-of-transforming-the-nhs

[xix] Cancer Research UK, 2021. Cancer Research UK response to Health and Social Care Data Strategy. https://www.cancerresearchuk.org/sites/default/files/july_2021_-_cancer_research_uk_response_to_health_and_social_care_data_strategy.pdf

[xx] Cancer Research UK, 2021. CRUK submission to Department for Digital, Culture, Media and Sport consultation, Data: A new direction consultation https://www.cancerresearchuk.org/sites/default/files/2021_nov_-_cruk_response_to_dcms_data_reforms_consultation.pdf

[xxi] Cancer Research UK and Macmillan, 2016. Review of Informed Choice for Cancer Registration. https://bit.ly/3ryuag7

[xxii] For instance, Wellcome Trust Monitor Report, April 2016 https://wellcome.org/sites/default/files/monitor-wave3-full-wellcome-apr16.pdf

[xxiii] Understanding Patient Data. How do people feel about the use of data? https://understandingpatientdata.org.uk/how-do-people-feel-about-use-data

[xxiv] NHS Digital. National Data Opt-out, July 2021. https://digital.nhs.uk/data-and-information/publications/statistical/national-data-opt-out/july-2021

[xxv] Jesper-Mir, E. and Asher, R. 2021. Working with us to develop resources on datasets and pseudonymisation. https://understandingpatientdata.org.uk/news/work-us-develop-resources-datasets-pseudonymisation