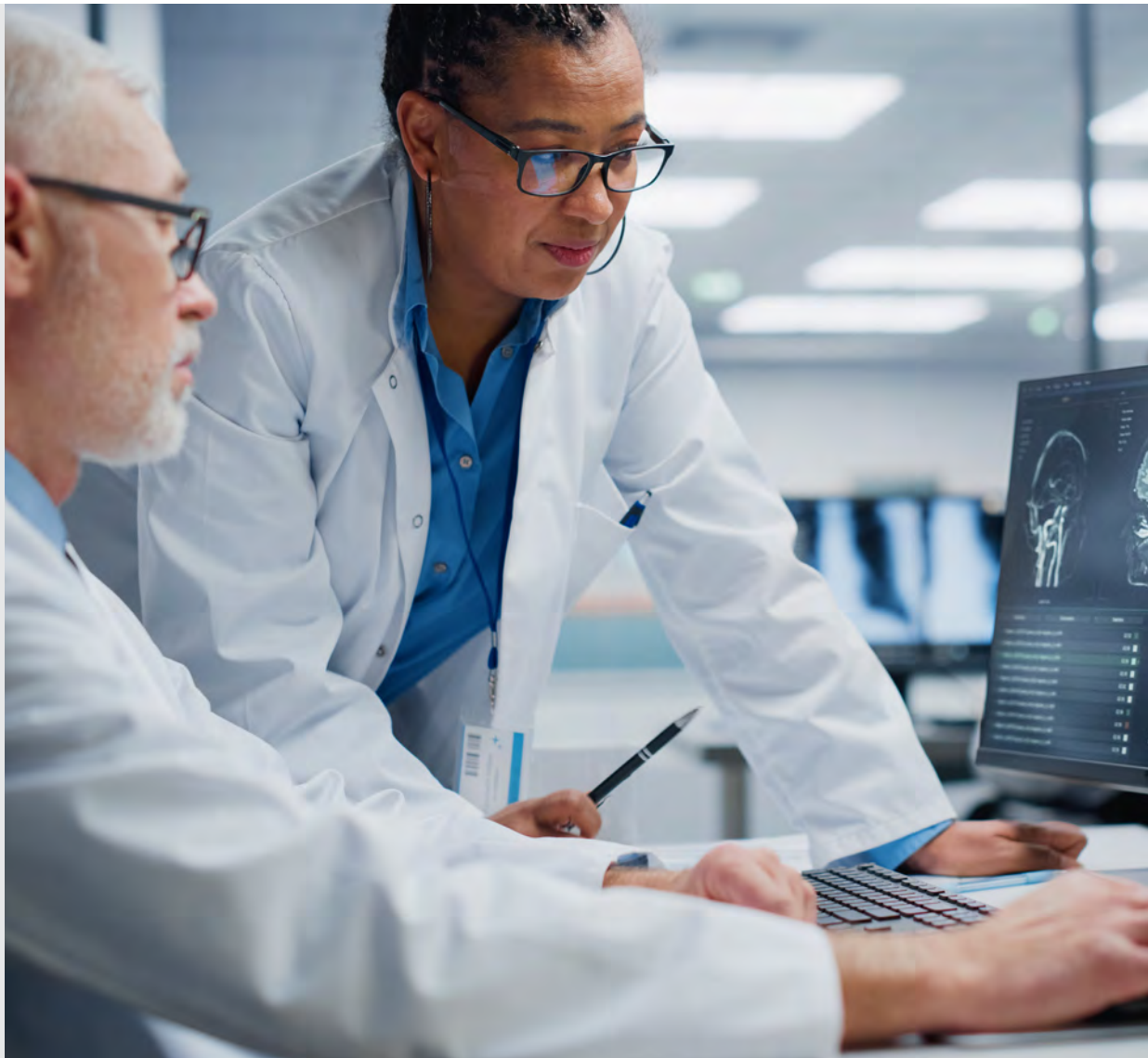
UNLEASHING THE POWER OF DATA TO BEAT CANCER

OUR RESEARCH DATA STRATEGY

Together we will beat cancer

CANCER RESEARCH UK

CONTENTS



EXECUTIVE SUMMARY

Improving how we use research data is key to answering important questions in cancer research. This will help us achieve the objectives set out in our new Research Strategy (2022) around discovery, prevention, detection and treatment of cancer.

We're already seeing excellent research in the field of data science across our funding portfolio, but we want to do more to scale up our existing activity.

This Research Data Strategy sets out three stages for how we will capitalise on our position as a leading funder of cancer research to maximise the reuse of research data and unleash its potential to beat cancer.

Stages

1. **Lay the foundations:**
   We will establish the foundations of the strategy and enable cancer researchers in the UK to better generate, curate and use data.
2. **Enable our vision:**
   We will secure effective partnerships and embed clear communication channels to make our vision a reality.
3. **Exemplify our ambition:**
   We will develop exemplar research programmes that demonstrate the value of taking a data-driven approach to cancer research and embody the foundations of the strategy to leave a legacy that improves the data ecosystem.

INTRODUCTION

INTRODUCTION

Data science is already accelerating the generation of insights into the biology, prevention, detection and treatment of cancer and laying the foundations for improvements in patient outcomes.

Our refreshed Research Strategy (2022) identifies data science as a new opportunity to help us achieve our four objectives: discover, detect, prevent and treat. So we have developed our first ever Research Data Strategy, which sets out how we intend to unleash the enormous potential of research data and data science.

As the leading funder of cancer research in the UK, we are in a unique position to influence the whole data lifecycle. We will champion data standards to facilitate the integration of different modes of data and translate the insights from big data into better outcomes for people affected by cancer.

An important part of this will be to help patients and the public understand the value of their data in the fight against cancer.

Patient privacy and data protection

Patient data is an incredible resource for cancer research to drive improvements in prevention, detection and treatment.

Access to patient data for research purposes is strictly governed so that it is used safely in cancer research.

We always respect patient privacy in the use of patient data for research – one of the ways in which we can do this is by making it anonymous.

We champion fairness and transparency in data-sharing partnerships, and place patient involvement at the heart of these discussions.

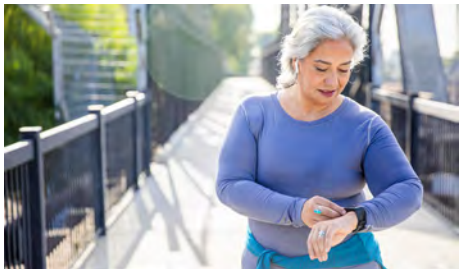
WHAT IS RESEARCH DATA?

Data generated by cancer research can include non-human data, such as data collected about mouse models, as well as human data from:

blood, cells or tissue samples

imaging (eg x-ray, MRI and CT scans)

clinical trials



health and fitness devices

Other data types used for cancer research can include:

electronic health records – for example, data collected as part of NHS health and social care, such as effectiveness of treatments and how early cancers are diagnosed

administrative and environmental data – for example, analysing outdoor air pollution levels and lung cancer incidence estimates that this kind of pollution causes roughly 1 in 10 cases of lung cancer

OUR VISION

Our vision is to maximise the reuse of research data and unleash the potential of using big data to beat cancer.

We want to get the most value out of every piece of research data to discover further insights about cancer and amplify patient and public benefit.

And we know that leveraging data science and advanced analytics (including artificial intelligence) to glean insights from large-scale health-relevant data will help us make discoveries leading to more and better ways to prevent, detect and treat cancer.

There are great examples where this is already happening. Our strategy sets out how we aim to scale up this activity and lead the way in using data science in cancer research.

> “What’s important is that we continue to develop research and continue to develop treatments for patients and the only way to do that is through research and through gaining access to appropriate data that is actually going to change policies, and change operational systems and change treatment for patients”
>
> Patient and member of patient advocacy organisation 'use MY data'

THE CHALLENGE

THE CHALLENGE

Emerging technologies hold great promise for new discoveries in cancer research.

But there are a number of challenges associated with using data science in cancer research.

So we have developed our Research Data Strategy with a vision to overcome these challenges and realise the full power of data science to beat cancer.

Cancer research data is not sufficiently Findable, Accessible, Interoperable (meaning the ability of computer systems to exchange and make use of information) and Reusable. These characteristics are collectively referred to as the FAIR data principles.

The cancer data science community could benefit from being better connected to a more collaborative community to be better positioned to tackle common challenges collectively, share best practice and drive data culture change.

While research using patient data is essential for better outcomes for people affected by cancer, it can be difficult to build and maintain public trust that their data will be used responsibly and securely. This is made even harder as data science is a particularly complex and technical.

There aren't many visible role models and case studies that exemplify the use of data science in cancer research.

Although the demand for data scientists and other data roles is high, we need more of these experts applying their skills to cancer research.

There is a risk of exacerbating health inequalities, such as developing data-driven clinical decision-making tools that don't benefit all parts of society equally.

The enormous digital storage and high-performance computing power required for data science can potentially have a huge environmental impact.

Overall, there is a lack of coordination, expert advice and strategy for the use of big data in cancer research.

OUR APPROACH

OUR APPROACH

1. Lay the foundations

We will establish the foundations of the strategy to enable cancer researchers in the UK to better generate, curate and use data. We will:

- earn and maintain public and patient trust, confidence and support
- maximise the intellectual value of cancer data
- strengthen the national cancer data science environment
- build a collaborative and supportive cancer data science community
- translate insights from big data into patient benefits
- attract and train cancer data experts
- ensure equality, diversity and inclusion
- improve the sustainability and efficiency of data science

2. Enable our vision

We will secure effective partnerships and embed clear communication channels to make our vision a reality.

Partnerships

Communication

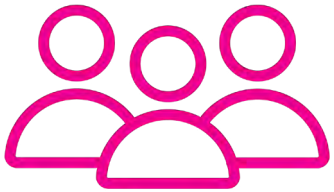
3. Exemplify our ambition

We will develop exemplar research programmes that demonstrate the value of taking a data-driven approach to cancer research and embody the foundations of this strategy to leave a legacy that improves the data ecosystem.

LAY THE FOUNDATIONS

We intend to create the appropriate research environment, culture, tools and support that will facilitate the timely, well-informed, transparent, sustainable, inclusive and community-minded generation, curation and use of cancer research data.

We have identified eight elements that will lay the foundations for us to achieve this:

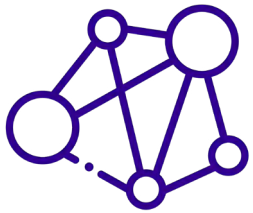


Earn and maintain public and patient trust, confidence and support

We can't achieve our research goals without public trust that we are using their health data responsibly to achieve better patient outcomes.

We've heard from the patient panels consulted on this strategy that they want to know more about the benefits, risks and mitigation measures that relate to data science in cancer research and they want to be involved in discussions and decision-making.

So we are putting people affected by cancer at the heart of this strategy, supporting their understanding of how patient data is used, and amplifying their voices and opinions. This will include a new bespoke public and patient involvement panel to specifically advise on activities under this strategy.



Maximise the intellectual value of research data

We will generate research data that:

- meets FAIR data principles: Findable, Accessible, Interoperable and Reusable
- is securely controlled where patient data is used

For example, we're building a trusted research environment (a secure computing environment that holds data and enables access to it only for strictly approved users) to support researchers across the UK who don't have a suitable safe set-up to store and analyse patient data.

We're also reviewing our data-sharing policies and considering how to make the data we generate more findable and shareable across research communities.

LAY THE FOUNDATIONS

Strengthen the national cancer data science environment

We will improve the UK's research and development environment to better enable data science by influencing policy.

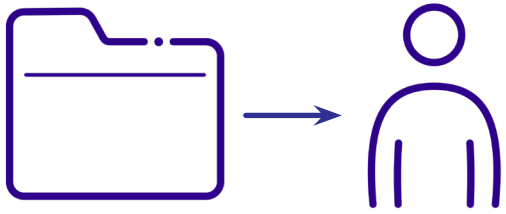
A priority here is to streamline access for cancer researchers to nationally-held cancer-relevant health data.

This aligns with our new strategic approach to science and research influencing which has identified new policy areas of focus and new approaches to take over the next five years.

Build a collaborative and supportive cancer data science community

Pockets of incredible data expertise exist across the Cancer Research UK community, so we will bring them and others together in a coordinated community of practice that's integrated with relevant disciplines, such as mathematics and informatics, to:

- share best practice
- tackle common challenges
- foster the kind of data culture that best serves the research community and ultimately people affected by cancer



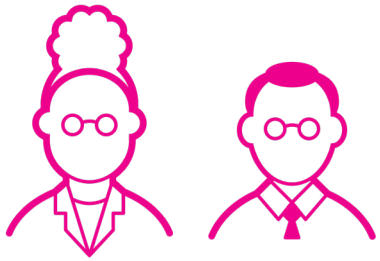
Translate insights from big data into patient benefits

We will continue to support translation of insights from data science into benefits for people with cancer and the public through the work of our centres and institutes.

We will also develop new ways to collaborate with the private sector transparently, with fair returns back into cancer research.

The challenges we will prioritise tackling in this area are public concern around commercial access to data and issues around consent and data quality.

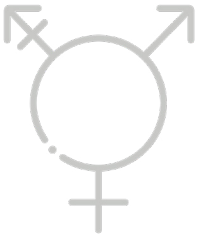
LAY THE FOUNDATIONS



Attract and train data experts

We've heard from the research community that it can be a struggle to recruit data scientists and other data roles, so we will attract and support new talent to the cancer data science workforce and upskill the research community by offering data training opportunities throughout their careers.

We will consider how to strengthen, better recognise and shine a light on the diverse career pathways of data experts in cancer research.



Ensure equality, diversity and inclusion

We will ensure that all parts of society benefit equitably from data science and proactively support the reduction of health inequalities, for example, by ensuring that data-driven clinical decision-making tools serve us all.

We're also looking at how we can improve the diversity of data and capture that diversity in metadata.

As part of this, we're exploring how we can recruit and retain a diverse cancer data science workforce.

This aligns with and complements our broader Equality, Diversity and Inclusion Strategy.

Improve the sustainability and efficiency of data science

Data science often involves running complex algorithms on very large datasets using supercomputers, which uses a lot of power and can result in a significant carbon footprint.

We will explore ways to reduce this environmental impact and encourage environmentally and financially sustainable data science at scale.

ENABLE OUR VISION

Our strategy will be enabled through:

Partnerships

We will develop partnerships with like-minded organisations to leverage our shared expertise, tools, data and resources towards common goals.

International alignment and partnerships are critical to pooling expertise and resources across borders, particularly for research into rare cancers.

Communication

We will embed communication that enables a multidisciplinary cancer data science community – including mathematicians, cancer biologists, clinicians and patients – to understand each other and work effectively together.

EXEMPLIFY OUR AMBITION

We're developing proposals for ambitious exemplar programmes that embody our eight foundations and demonstrate the potential of data science in cancer research.

These exemplar programmes will unleash the potential of big data and data science to answer important questions about cancer. They will deliver on the research objectives set out in our Research Strategy – discover, prevent, detect and treat – with a focus on approaches such as prioritising cancers that affect children and young people, tackling cancers of unmet need, driving further progress in early detection and driving therapeutic innovation.

> “This new Research Data Strategy has the potential to transform the landscape for data science in cancer research for patient benefit, informing better prevention, detection and treatment. I am privileged to be chairing the new Research Data Advisory Board who will be guiding Cancer Research UK’s journey toward this vision”

Bissan Al-Lazikani, Professor of Genomic Medicine and Director of Therapeutics Data Science at MD Anderson Cancer Center

ACKNOWLEDGEMENTS

Our thanks to members of the research community, organisations that are active in the field of data science and, in particular, people affected by cancer and members of the public, who all provided their time, insights and perspectives to help shape this strategy.

Get in touch
Dr Melissa Lewis-Brown,
Head of Research Data Strategy,
Cancer Research UK
ResearchData@cancer.org.uk

To read our research strategy, visit
cruk.org/research-strategy

Cancer Research UK
2 Redman Place
London E20 1JQ

T: +44 (0)20 7242 0200
cruk.org/science

CANCER
RESEARCH
UK

Together we will beat cancer