

# Making Data Count





# Contents

## Our funders



|                 |          |
|-----------------|----------|
| <b>Foreword</b> | <b>4</b> |
|-----------------|----------|

|                          |          |
|--------------------------|----------|
| <b>Impact in numbers</b> | <b>5</b> |
|--------------------------|----------|

|                          |          |
|--------------------------|----------|
| <b>Making data count</b> | <b>7</b> |
|--------------------------|----------|

|  |          |
|--|----------|
| <b>Improving people's lives by unlocking the power of data</b> | <b>8</b> |
|--|----------|

|  |   |
|--|---|
| First whole-population health data study delivers huge public benefits | 9 |
|--|---|

|  |    |
|--|----|
| Community and data scientists team up to boost vaccination rates in Brazil | 10 |
|--|----|

|  |    |
|--|----|
| Global collaboration helps repair health systems in COVID-19 aftermath | 11 |
|--|----|

|  |           |
|--|-----------|
| <b>Accelerating trustworthy data use by sorting the data</b> | <b>12</b> |
|--|-----------|

|   |    |
|---|----|
| Pioneering innovations for secure data environments | 12 |
|---|----|

|  |    |
|--|----|
| Empowering the sensitive data research community | 13 |
|--|----|

|   |    |
|---|----|
| Groundbreaking data collaboration improves children's lives and futures | 14 |
|---|----|

|  |    |
|--|----|
| Tackling missing ethnicity data and biases in research | 16 |
|--|----|

|  |    |
|--|----|
| New legal protocol safeguards data security and speeds up access | 17 |
|--|----|

|   |           |
|---|-----------|
| <b>Shaping the future of health data research</b> | <b>18</b> |
|---|-----------|

|  |    |
|--|----|
| Major update to the Health Data Research Gateway | 18 |
|--|----|

|               |    |
|---------------|----|
| Sudlow Review | 19 |
|---------------|----|

|   |    |
|---|----|
| The next generation of health data scientists | 20 |
|---|----|

|  |    |
|--|----|
| Putting patients and the public at the heart of what we do | 21 |
|--|----|

|                 |           |
|-----------------|-----------|
| <b>Appendix</b> | <b>22</b> |
|-----------------|-----------|

|                         |           |
|-------------------------|-----------|
| <b>Acknowledgements</b> | <b>23</b> |
|-------------------------|-----------|

# Foreword

Our mission at Health Data Research UK (HDR UK) is to accelerate trustworthy use of health data to enable discoveries that improve people's lives. We work in partnership with the NHS, industry, charities and universities in all four nations of the UK to improve the UK's data infrastructure and deliver cutting-edge research, unlocking the power of large-scale data to improve health and care.

This year has been a period of significant change. 2024 has seen the launch of the *Sudlow Review*, examining the barriers and inefficiencies that currently delay the safe and secure use of health data. A 10-year health plan for the NHS is in the works, following political change. Artificial intelligence (AI) continues to evolve at a rapid pace. Amid this, HDR UK remains an independent, authoritative centre for excellence and through our programmes, we have equipped teams with the tools, infrastructure, and expertise needed to realise the full potential of health data.

In our first Impact Report, we share the many ways HDR UK is advancing health data research:

**1 Improving people's lives by unlocking the power of data.**

**2 Accelerating trustworthy data use by sorting the data.**

**3 Shaping the future of health data research.**

This report reveals how HDR UK is making a difference to health data science and improving lives, while celebrating the stories and successes of the HDR UK community.

These stories highlight the **all-round impacts** of our work, which extend far beyond "improved health" – to include enabling critical infrastructure, shaping policies, and advancing cutting-edge analysis methods.

At the heart of these impacts lies **team science** – bringing together diverse people with a wide range of skills, across disciplines, sectors, and career stages. Each impact showcases the power of collaboration in driving meaningful and tangible outcomes.

Our impacts focus on **public trust** and provide great examples of **open science** and **knowledge exchange**. This commitment to openness, transparency and collaboration ensures that the benefits of our research are accessible to all, with the shared goal of improving health data research.

We would like to extend our heartfelt thanks to the panel of expert reviewers who helped select the case studies featured in this report. Their valuable insights guided the selection of these standout examples, which showcase the transformative power of health data science.

As you explore these pages, we invite you to celebrate the progress that the HDR UK community is making for the benefit of patients and the public – and join us in looking ahead to a future where health data transforms lives across the UK and beyond.



**Dame Julie Moore**  
Chair  
Health Data Research UK



**Professor Andrew Morris**  
**CBE MD FRCP FRSE PMedSci**  
Director  
Health Data Research UK

# Impact in numbers



## 200,000

HDR UK outputs have been cited over 200,000 times and are 6x more cited than those in similar fields



## 96%

96% of HDR UK publications are open access and 82% of software, tools, and databases are open source



## 3,932

The HDR UK research community have published 3,932 journal articles and preprints



## 193

HDR UK research has influenced policy 193 times, both nationally and internationally, including membership of key advisory committees



## 110

The UK Health Data Research Alliance has convened 110 members, with 18 new members joining in 2024 alone.

### Source:

Alliance data from HDR UK. All other figures from ResearchFish Report received May 2024. Includes outputs by a subset of HDR UK members published since April 2018 (HDR UK establishment) and attributed by submitters as HDR UK Institute outputs. Snapshot of number of citations taken July 2024.



# Making data count

## A dual role for large-scale, representative datasets

Examples of HDR UK's impact lie in both leveraging and enhancing uniquely large-scale, whole-population health data. Research offering insights that can improve health and wellbeing draw on these essential building blocks, such as linked national health records and educational data. Simultaneously, work enhancing the usability of these datasets for research enables faster and safer data access, widely adopted and fit-for-purpose standards, specifically curated data, software tools, and findings that better reflect the diversity of affected populations.

See: [Tackling missing ethnicity data and biases in research](#)

16

## A COVID-19 focus

Some of the greatest impacts across HDR UK's programmes stem from research examining COVID-19's wide-ranging impacts on healthcare access, hospital admissions, educational outcomes, and mental health. This work has shed light on immediate and long-term effects of the pandemic, informing local and national policy, and providing a blueprint for using large-scale data to tackle other health conditions, such as infectious diseases, cancer, heart disease, stroke, diabetes and dementia.

See: [First whole-population health data study delivers huge public benefits](#)

## Data linkage across sectors

Linking health data with educational and social care records has become a standard approach, providing a multi-dimensional view of health outcomes and revealing intersections with social determinants, such as housing and educational outcomes. This integration strengthens our understanding of health within a broader societal framework.

See: [Groundbreaking data collaboration improves children's lives and futures](#)

14

## Tackling health inequity

Many of HDR UK's impacts prioritise underserved and high-risk groups, such as children, adolescents, individuals with chronic conditions, and those in socioeconomically disadvantaged situations – prioritising and promoting health equity and improving access to care for marginalised populations.

See: [Community and data scientists team up to boost vaccination rates in Brazil](#)

10

[Global collaboration helps repair health systems in COVID-19 aftermath](#)

11

## A community-building ethos

Convening and fostering a collaborative health data science community plays a vital role across HDR UK's impacts – bringing together experts from academia, industry, the NHS, and the public. This creates a shared commitment and aligns efforts across diverse stakeholders, promoting interdisciplinary collaboration and shared learning.

See: [Pioneering innovations for secure data environments](#)

12

[Empowering the sensitive data research community](#)

13

# Improving people's lives by unlocking the power of data

The UK is in a unique position with almost the entire population – 67 million people – receiving care from the NHS throughout their lifetime. Safe and secure use of this abundance of health data could hold the key to improving the diagnosis and treatment of so many debilitating illnesses affecting millions of patients, from cancer to cardiovascular disease; as well as reducing suffering for many people in the future.





The availability of population-wide data has allowed us to study different combinations of COVID-19 vaccines and to consider rare cardiovascular complications. This would not have been possible without the very large data that we are privileged to access and our close, cross-institution collaborations.

**Venexia Walker**

Research Fellow at the University of Bristol, and member of the BHF Data Science Centre CVD-COVID-UK/COVID-IMPACT Consortium



The CVD-COVID-UK/COVID-IMPACT Consortium Team and colleagues

## First whole-population health data study delivers huge public benefits

**Research led by CVD-COVID-UK/COVID-IMPACT Consortium has had a direct impact on thousands of lives by addressing public concerns about vaccine safety for specific demographic groups. Their studies have highlighted the risks of under-vaccination against COVID-19, providing insights to support increased vaccination rates, and evidence of the risks associated with COVID-19 vaccinations.**

With the support of the BHF Data Science Centre (BHF DSC) at HDR UK, a multidisciplinary team of researchers used electronic health records for up to 67 million people – including the first study of the whole populations of England, Scotland, Wales, and Northern Ireland – to understand characteristics of people who were under-vaccinated against COVID-19 and the associated risks of under-vaccination. The team also conducted separate studies to assess the impact of COVID-19 vaccinations on cardiovascular events like strokes, heart attacks, and blood clots.

Their research has directly informed policy-makers, including the Medicines and Healthcare products Regulatory Agency (MHRA), the Joint Committee on Vaccination and Immunisation (JCVI), the UK Government’s Chief Medical Officer (CMO), and the Chief Scientific Adviser (CSA); shaping critical decisions on responses to COVID-19, and improving vaccination coverage, reducing under-vaccination, and mitigating COVID-19-related hospitalisations and deaths.

The BHF DSC streamlined data curation and analyses, facilitating access to secure data environments (SDEs), and equipping researchers with resources to accelerate analyses. Researchers responded by openly sharing analysis scripts, phenotyping algorithms and outputs, via platforms such as GitHub and the HDR UK Phenotype Library – putting open science at the heart of this project, and amplifying accessibility and advancing the field.

By harmonising and mobilising data at a national level, the Consortium has laid the groundwork for future UK-wide, whole population, health data research – far beyond COVID-19.



**What the panel of expert reviewers said:**

**We recognize this project’s remarkable contributions to data science and public health. We were particularly impressed by the project’s large-scale team science approach, mobilizing data at a national level to directly influence COVID-19 vaccination policy and public discourse. The project had a clear and tangible impact on thousands of lives.**

## Community and data scientists team up to boost vaccination rates in Brazil



One Brazil-based study led to a major increase in vaccination uptake for some of the poorest and most vulnerable people in Rio de Janeiro. This impressive outcome was achieved through a highly grassroots approach, working closely with the local community and NGOs. The team produced further vital insights into COVID-19 vaccination approaches, which significantly improved local and regional vaccination programmes in the country. They also devised strategies to reach the wider community with health and COVID-19 messaging, at a time when public communication was limited – reaching more than two million people on social media, and gaining wide dissemination of public health messaging in local and international media.

The study used resources from the International COVID-19 Data Alliance (ICODA), a global partnership of research organisations convened by HDR UK, that worked to harness the power of health data in response to the COVID-19 pandemic.

The researchers were able to leverage ICODEA-enabled data infrastructure and resources, including secure data access and governance approaches, to compare the severity of respiratory disease in COVID-19 patients across 43 hospitals. They also analysed vaccination rates in adults in 5,500 Brazilian municipalities – determining that

socio-economic disparities impacted on first dose vaccination and disease protection. They also found that primary care providers were critical in providing equal access to vaccines for vulnerable people in deprived areas.

The team embedded their research activities in socially vulnerable communities, alongside NGOs working on the ground, in Rio de Janeiro’s favelas. The team worked hard to build public trust and make their findings relevant for the communities most affected by their research. This also allowed them to ensure the communities’ priorities were addressed in their analysis – and that community members understood how the results could improve disease outcomes, both for the COVID-19 pandemic and other public health issues.

A further legacy of the team’s work is a website that shares a toolkit and approaches for communities and researchers to work together on prioritising and co-developing health research studies. It provides a framework for communities and health practitioners to work together to address a wide range of health challenges – such as cancer, heart disease, and stroke – as well as engage in future research in a trustworthy, systematic and productive way.



Sharing the leadership of the project, with its mistakes, successes, challenges, and virtues, on an equal footing between residents and researchers, and recognizing existing knowledge and leadership in the territory, proved to be a great innovation.

The DP-EFFECT Team



What the panel of expert reviewers said:

**This research directly informed policy both nationally and more locally. We were particularly impressed by the team’s decision to “invert the research triangle”, which resulted in meaningful impacts in public trust and health.**

## Global collaboration helps repair health systems in COVID-19 aftermath

An international research collaboration, supported by HDR UK's International COVID-19 Data Alliance (ICODA) programme, drove decision-making on pandemic policy across ten countries in Africa, Latin America and Asia. The team guided efforts to monitor disruptions in essential healthcare services during the pandemic – as well as provide evidence and recommendations to strengthen resilience against future emergencies.

The international research team assessed the impact of COVID-19 on healthcare services and institutional mortality across 31 health services in ten countries, including low-, middle-, and high-income nations.

The team used shared approaches (such as common code books, shared tools for data harmonisation, and a version control system for statistical code); standardised data; and innovative statistical techniques to evaluate disruptions in health service delivery, as well as the effects of pandemic containment policies.



It is heartening to know that this project helped support policy-making, provided insights on health system resilience, and improved our understanding of how to work with health data during a crisis.

**Catherine Arsenault**

Assistant Professor | Global Health,  
Milken Institute School of Public Health,  
The George Washington University

The study revealed significant disruptions in essential services, such as cancer screenings, tuberculosis (TB) testing, and maternal health care – but with no correlation between the level of disruption and a country's income or pandemic intensity. These insights directly influenced policy (such as Mexico's National Strategy for Health Services Recovery) and informed targeted interventions to strengthen health system resilience.

HDR UK supported the research by providing an Ethics and Governance Framework to guide ethical and trustworthy practices; and policies for output review, attribution, and publication – an approach which underpinned collaboration and knowledge exchange among researchers and policy-makers worldwide.

By developing common tools and practices, and strong engagement of policy-makers, DP-REHCORD has notably influenced policy across multiple countries – demonstrating not just national but global impact, and underscoring the significance of data stewardship for future research.



**What the panel of expert reviewers said:**

**The project's outputs have notably influenced policy across multiple countries, demonstrating "global impact". The implementation of common approaches across different health systems was particularly impressive.**



# Accelerating trustworthy data use by sorting the data

We know that the health data ecosystem is complex and fragmented, meaning access to datasets is difficult or slow and holds back progress that could improve lives. Our work is making it easier for researchers to find and access the data they need to generate new discoveries, using streamlined systems that are designed to ensure the privacy and security of people's data.

## Pioneering innovations for secure data environments

Five projects in the DARE UK (Data and Analytics Research Environments) programme have delivered much-needed practical solutions for key issues facing secure data environments (SDEs). The projects provide a blueprint for overcoming barriers to secure and rapid data sharing across multiple SDEs, which delay data access. They have also delivered (semi)-automated systems to better guarantee the security and privacy of the sensitive data. Several of these initiatives have already been implemented – in the UK and internationally.

DARE UK is a programme run by HDR UK and Administrative Data Research UK (ADR UK), working to establish safe systems and processes for research on cross-domain sensitive data. For example, health, education, housing and incomes data.

The Standardised Architecture for Trusted Research Environments (SATRE) project created – for the first time – a standardised specification for SDEs to adopt as they support sensitive data projects. Achieved through partnerships between the public and the research community, it has been adopted in all Scottish SDEs, as well as several commercial SDEs. The NHS Research Secure Data Environment Network in England has also adopted SATRE as a baseline.

The Semi-Automated Checking of Research Outputs (SACRO) project created semi-automated tools to support output-checking for SDEs, reducing reliance on manual load. SACRO's solutions are being tested and adopted across the UK, Germany and the US.

The TRE-FX and TELEPORT projects developed two different mechanisms to enable analysis of de-identified data held across multiple SDEs within a single project, thereby increasing efficiency and facilitating research which requires datasets held in different SDEs. TRE-FX solutions have been used by Dementia's Platform Australia, and the BY-COVID project.

The Semi-Automated Risk Assessment (SARA) project semi-automated the process of recording how data is processed as it moves through SDE systems, in order to provide researchers who use the data with a dynamic data trail, improving data reliability and accessibility. In addition, the SARA project also looked at ways to automate privacy risk assessment in text data, to improve secure use of free-text data, such as clinical notes.

# Empowering the sensitive data research community

The DARE UK Community Groups initiative has also helped drive improvements in SDE infrastructure development, through enabling better collaboration within the sensitive data research community. Partnership working across the community is essential, if we are to develop and enhance a compatible infrastructure for sensitive data research. Recognising that resourcing issues have hindered such collaboration, the DARE programme has funded a number of community groups supporting rapid expansion in partnership working and enabling the sharing of expertise across the sector.



The DARE UK Phase 1 Driver Projects demonstrate the power of collaboration in tackling Secure Data Environment (SDE) challenges, from data linkage to access. Their adoption across the UK and beyond signals a new era of unified, efficient sensitive data research for public benefit—and we’re proud to be at the forefront of this transformation.

The DARE UK Delivery Team

This has already ensured that standards and tools developed by the DARE UK driver projects have been co-developed with the input of the sensitive data research community, ensuring they meet the needs of the sector going forward. Funding this community collaboration has also resulted in the creation of resource hubs for the sector, to consolidate and align new developments, as well as provide routes and guidance on engaging the public.



What the panel of expert reviewers said:

**Through a set of funding opportunities, the team have... been instrumental in “bringing people together” to establish the UK TRE community as a truly open, connected, and collaborative ecosystem, breaking down silos and fostering shared practices across institutions.**

The DARE UK Delivery Team





The ECHILD Team

## Groundbreaking data collaboration improves children's lives and futures

**A unique data resource influenced early years' policies, advanced research on childhood health and education, and improved decision-making across government departments.**

The Education and Child Health Insights from Linked Data (ECHILD) Research Database is a pioneering dataset, which links health, education, and social care data for 20 million children in England. Established through collaboration across multiple partners including HDR UK, it is the largest dataset of its kind, containing comprehensive data from birth to adulthood. By enabling secure, pseudonymised data access, ECHILD has transformed how researchers address policy-relevant questions, and improve outcomes for children and families.

In supporting secure data use while safeguarding privacy, the ECHILD project has benefited significantly from shared expertise, infrastructure, and frameworks. The project has also drawn on shared knowledge on best practice and public engagement – and in response has contributed to modules on HDR UK Futures, a learning and development platform for data scientists.

ECHILD has also expanded through partnerships, knowledge-sharing forums, and collaborations with driver programmes. By streamlining data access and supporting capacity-building through seminars and training, partnership working has helped ensure that the ECHILD dataset is used effectively for research that improves outcomes for children, influences national policies, and fosters a culture of collaboration and open science.

ECHILD's success demonstrates the transformative impact of linking health, education, and social care data to address complex societal challenges. It also provides a blueprint for other countries, to develop similar resources.

“

ECHILD represents an exciting step change in the availability of data for research to improve the lives of children and families. Linking together data across different services allows us to bring together different parts of the puzzle to ultimately let services make decisions in more joined up ways.

**Katie Harron**

Professor of Statistics and Health Data Science, UCL



What the panel of expert reviewers said:

**This is a resource with significant research outputs, alongside impact that continues to grow – evidenced by its visibility among significant entities such as the NIHR and the House of Lords. The development of this initiative stood out to the panel as a leading example of a data consortium operating effectively at both national and international levels, showcasing multi-stakeholder engagement.**

---



## Tackling missing ethnicity data and biases in research

The UK Health Data Research Alliance (the Alliance), a sector-wide collective of over 100 patient/public groups, data custodians, health care professionals and more, is helping policy-makers, researchers and the NHS to improve health and social care. Supported by a HDR UK secretariat, an Alliance working group has published a White Paper, detailing key steps that must be taken to improve data quality for use in research, clinical practice and intelligence by removing potential widespread biases, through better collection and use of ethnicity data. By tackling barriers to accurate and comprehensive data collection, the group's work guides policy-makers on how to make decisions based on a clearer picture of health at local and population levels – paving the way for fairer and more effective healthcare.



Everyone deserves access to high-quality healthcare regardless of their background. We need to work harder and smarter to address inequalities, associated with ethnicity, in healthcare access and health outcomes.

**Kamlesh Khunti**

Professor of Primary Care Diabetes and Vascular Medicine, University of Leicester. Co-Chair HDR UK Ethnicity Coding Working Group

The working group identified that ethnicity data is often a missing component in health data analysis, and health and social care research. Good quality and comprehensive data sources are vital if we are to have the best outcomes from research to meet the needs of the whole community and reduce disparities. The White Paper, which is supported by the National Health Service (NHS) and Office for National Statistics (ONS), gives policy-makers a multi-faceted blueprint to increase representativeness in health research, as well as transparency and communication with the public.

Drawing on widespread expertise as well as crucial public input, the working group recommendations deliver a strategy to tackle the issue across the data science landscape – from policy-making to data collection and research.

This work has the power to improve decision-making, training and skills, as well as public trust, health and wellbeing. Embedding a variety of ethnic groups at the centre of decisions about how sensitive information is shared with clinicians empowers the public – and makes clear that they are included and have a choice, as well as ensuring representation in the data.

The project is an excellent example of bringing a variety of groups and contributors together around shared challenges and opportunities, with a multidisciplinary, community-driven approach and wide engagement.



What the panel of expert reviewers said:

**The team, led by a dedicated subgroup of the UK Health Data Research Alliance, has taken on the critical challenge of tackling coding standards as a key aspect of inequality in health data and is recognised as a “truly multidisciplinary group of professionals who adopted co-creation at its heart”. Their work stands as a testament to the power of collaborative, community-driven efforts in tackling complex issues.**



“

We have never appreciated the importance of utilising data for public benefit more than throughout the COVID-19 pandemic, yet data research continues to be delayed by complex and divergent contracting arrangements. Widespread adoption of the DAA template will not only help build public trust but also improve the efficiency of research, ultimately enabling discoveries that will improve and save lives.

**Rachel Brophy**

Head of Information and Research Governance (Interim), HDR UK

## New legal protocol safeguards data security and speeds up access

A new and innovative Data Access Agreement (DAA) template, which standardises data governance for secure data environments (SDEs), has been designed to significantly improve the process for health data researchers working with SDEs. The new template streamlines and accelerates the timeline from project approval to working with data. At the same time it maintains rigorous data protection standards, while reducing cost and increasing efficiency – leading to faster research outcomes.

One of the most frustrating barriers faced by data researchers is the variation in processes to access data between data custodians. This adds delays, inconsistencies in approvals, and administrative burden – without any benefits to privacy or security. The new DAA template, an open source resource, significantly reduces the need for repeated development of agreements and the associated legal burden. It was developed with versatility in mind. With customisable annexes and alignment with the Five Safes Framework on safe data access principles, the template is flexible while maintaining comprehensive data security.

HDR UK provided essential legal governance and expertise in developing the DAA template, in partnership with academic institutions, NHS bodies, and legal and industry advisors across the UK. Public involvement was central to the DAA's development, crucial for both building trust and incorporating different perspectives. The team plans to use its networks to drive adoption of the template across the UK's SDE landscape.

The team have also worked collaboratively and shared resources with networks such as the UK Health Data Research Alliance, who have established the DAA template across the sector. The initiative highlights HDR UK's mission to advance data access and governance to support impactful research, while upholding public trust and transparency.



What the panel of expert reviewers said:

**The DAA project has made notable strides in streamlining and standardizing data access agreements, a crucial step towards greater collaboration and efficiency across the ecosystem... The diversity of the team stood out to the panel, with members from a range of universities, organizations, and non-science backgrounds. This varied expertise not only enriched the project but also fostered a spirit of inclusivity and collaboration that is often difficult to achieve.**

---

# Shaping the future of health data research

As this report demonstrates, exciting times lie ahead for the expanding health data landscape. At HDR UK we've already taken major steps in infrastructure innovation for researchers, and we are at the forefront of keeping health data science high on the public policy agenda. But the true power of health data can only be achieved within a wider ecosystem that actively seeks to connect people, data and organisations. HDR UK is committed to embracing diverse perspectives and skills, which are vital for tackling the health challenges facing our 21st century communities. We never forget that behind each data point lies a person, a loved one, a family. Our strategy reflects this ethos, as we strive to make patients and the public essential partners in shaping our work and our future strategy.

## Major update to the Health Data Research Gateway

This year we introduced a major new version of the Health Data Research Gateway (the Gateway) with significant enhancements designed to empower the research community and accelerate scientific discovery for public benefit. The Gateway platform was first launched in 2020, aimed at making health data more discoverable to researchers so they can spend more time addressing important research questions. The 2024 update goes much further, advancing beyond data discovery to make the entire researcher journey – from finding datasets to analysing data and sharing outputs – smoother and more efficient. This major new release brings a suite of enhancements to improve user experience, search functionality, resource discovery, data access enquires and metadata management. The platform is also being used by NHS England's Data for Research and Development Programme to provide a 'single front door' for research access to the NHS Research Secure Data Environment Network.

The newly launched Gateway simplifies the process of finding and using health data so that researchers can spend more time answering important questions that drive improvements to people's health. Our vision is that the Gateway becomes the go-to place for discovering, exploring, and accessing health data and other shareable resources, evolving into a platform that's co-created by the community, builds upon existing technological solutions, and aligns to latest advancements in health data science. We are committed to continuously improving the Gateway, working closely with our partners to meet the evolving needs of researchers while expanding the number of resources accessible to them and the wider health data community.



## Sudlow Review

November 2024 saw the launch of Professor Cathie Sudlow's independent review, *Uniting the UK's Health Data: A Huge Opportunity for Society*. Commissioned by the Chief Medical Officer for England and UK Government's Chief Medical Adviser, Professor Sir Chris Whitty, NHS England's National Director of Transformation, Dr Timothy Ferris (now Dr Vin Diwakar, Medical Director for Transformation), and the UK's National Statistician Professor Sir Ian Diamond, with the support of the Chief Medical Officers in the devolved nations. Cathie was supported in carrying out the review by a team at HDR UK.

The Sudlow Review sets out five key recommendations for overcoming the barriers and inefficiencies that currently delay the safe and secure use of health data to improve lives, prevent suffering, and accelerate the race against disease.

HDR UK launched the Review at the Royal Society, where the Review's recommendations were enthusiastically welcomed by senior figures from across academia, industry, research, NHS, regulators and government officials. HDR UK also co-ordinated the communications strategy for the Sudlow Review, which drew significant digital engagement – with nearly 2000 views of the report on launch day alone. It captured attention across health and science media, including the BMJ, Pulse, and Research Fortnight, as well as the pharmaceutical and tech press. Mainstream coverage amplified its reach, with highlights such as Professor Sudlow's interview on BBC Radio 4's Today programme.



# Black Internship Programme

Helping transform the prospects of talented Black health data scientists in the UK



Mfoniso Udo, Black Internship Programme alumna

## The next generation of health data scientists

Public trust and excellence in research are fostered through building a secure health data infrastructure – but also through diversifying the data science workforce and ecosystem. Now in its fifth year, our Black Internship Programme (BIP), in collaboration with the 10,000 Black Interns Foundation and UK Health Data Research Alliance, has created huge opportunities for under-represented groups – providing 308 internships with partners across the health data sector. The programme has been immensely popular, with over 1600 applicants in 2024 alone, and 66 partner hosts – with more signing up each year.



The HDR UK internship... inspired me to aim higher than I ever thought possible. Thank you for all the incredible impact you've had on my career. I wouldn't have reached this point without your support.

**Emmanuel Ughoo**  
BIP alumnus

Applicants and hosts hail from all four corners of the UK, and 56% of interns did not attend a Russell Group university. Alumni have gone on to varied career paths within health data and wider data science fields, including joining PhD programmes, working as health data scientists and analysts, joining local government and civil service departments, pharmaceutical industry roles, and more.



It's been such a joy to meet talented, driven young people excited about a future in healthcare and data science. They all articulated a clear sense of mission and ambitions to put their careers to good use.

**Partner provider Phil Walmsley**  
Department of Health and Social Care

## Putting patients and the public at the heart of what we do

As host for, and active partner in, the Public Engagement in Data Research Initiative (PEDRI), our Public and Patient Involvement and Engagement (PPIE) team has helped deliver huge progress in promoting good standards in effective engagement and involvement of the public in data for research and statistics; and increasing collaboration between data scientists and the public. Both are vital for building public trust – the latter a fundamental component for conducting effective health data research, creating health policy, and implementing findings.

PEDRI unites different UK organisations working with data and statistics to improve how we all work with the public, and has grown significantly. Starting with just three founding members, it has now expanded to represent 15 key organisations, with the UK Longitudinal Linkage Collaboration (LLC) and the Office for Statistics Regulation (OSR) joining in the last few months. PEDRI's partnerships and collaborations also now include industry leaders, such as IQVIA and the Association of the British Pharmaceutical Industry (ABPI).

Nurturing collaboration with public representatives and public-led organisations, such as Use MY Data, has been key to PEDRI's success – ensuring that data research initiatives are grounded in the needs and perspectives of the public. PEDRI's Public Advisory Group ensures the public voice is weaved into all PEDRI's work.

PEDRI has recently produced an [updated version](#) of the Good Practice Standards (GPS) guide for public involvement and engagement (PIE) in data research and statistics, along with an [easy read version](#). A recent consultation with over 50 organisations informed the creation of a new [Resource Hub](#), which will serve as the primary source for public information and engagement information in data and statistics.

PEDRI recently hosted a roundtable to discuss public engagement in Secure Data Environment access – an initiative which has formed the bedrock to new work on identifying actionable insights and deciding which future initiatives are needed, to ensure more public trust and collaboration with health data research.



# Appendix

HDR UK, together with all the studies' authors, extend our heartfelt thanks to our many contributors, partners and supporters.

## Improving people's lives by unlocking the power of data

### First whole-population health data study delivers huge public benefits

#### Research team and collaborators:

- Alan Keys
- Alexia Sampri
- Amitava Banerjee
- Angela Wood
- Ashley Akbari
- Aziz Sheikh
- Cathie Sudlow
- Christopher Robertson
- Christopher Sullivan
- David W
- Declan Bradley
- Efosa Omigie
- Elsie Horne
- Emanuele Di Angelantonio
- Fatemeh Torabi
- Frank Kee
- Genevieve Cezard
- Hoda Abbaszanjani
- Jan D
- Jennifer Cooper
- Jim McMenamin
- Jonathan Sterne
- Kamlesh Khunti
- Karen Mooney
- Kirsty Morrison
- Lewis Ritchie
- Lynsey Patterson
- Rachel Denholm
- Ronan Lyons
- Sam Hollings
- Samaira Khan
- Samantha Ip
- Siobhan Murphy
- Spencer Keene
- Spiros Denaxas
- Steven Kerr
- Stuart Bedston
- Teri-Louise North
- Thomas Bolton
- Venexia Walker
- William Whiteley
- Yangfan Li

### Community and data scientists team up to boost vaccination rates in Brazil

#### Research team and collaborators:

- Beatriz Rache
- Fernanda Baião
- Fernando Bozza
- José Cerbino
- Leonardo S.L. Bastos
- Otavio Ranzani
- Paola Maçaira
- Rudi Rocha
- Silvio Hamacher
- Soraida Aguilar

### Global collaboration helps repair health systems in COVID-19 aftermath

#### Research team and collaborators:

- Catherine Arsenault
- Margaret E. Kruk
- Munir Kassa Eshetu
- Roody Thermidor
- Sebastian Bauhoff
- Shogo Kubota
- Suresh Mehata
- Svetlana Doubova
- Data custodians and collaborators at the University of KwaZulu-Natal and at the Medical Research Council of South Africa
- Data custodians from Ghana Health Services, including health information specialists

## Accelerating trustworthy data use by sorting the data

### Pioneering innovations for secure data environments

#### Research team and collaborators (principal investigators):

- Arlene Casey
- Balint Stewart
- Carole Goble
- Christian Cole
- Fergus McDonald
- Jim Smith
- Michelle Amugi
- Rob Baxter
- Simon Ellwood-Thompson
- Westley Igbo

### Empowering the sensitive data research community

#### Research team and collaborators:

- Balint Stewart
- Emily Jefferson
- Fergus McDonald
- Michelle Amugi
- Rob Baxter
- Westley Igbo

### Groundbreaking data collaboration improves children's lives and futures

#### Research team and collaborators:

- Ania Zylbersztejn
- Bianca De Stavola
- Difei Shi
- Farzan Ramzan
- Kate Lewis
- Katie Harron
- Louise Mc Grath-Lone
- Matt Jay
- Ruth Blackburn
- Ruth Gilbert
- Tony Stone
- Vincent Nguyen

### Tackling missing ethnicity data and biases in research

#### Research team and collaborators:

- Alastair Denniston
- Angela Wood
- Ashley Akbari
- Claire Argent
- Jonathan Valabhji
- Kamlesh Khunti
- Marta Pineda Moncusi
- Paola Quattroni
- Rose Drummond

- Sara Khalid
- Uwaye Ideh
- Vahe Nafilyan
- Participants of the Alliance Ethnicity Coding Standards working group

### New legal protocol safeguards data security and speeds up access

#### Research team and collaborators:

- Adam McArthur
- Alan Harbinson
- Alison Afrifa
- Alison Knight
- Allison Noble
- Amanda Rudczenko
- Amir Mehrkar
- Amy Osborne
- Andrew Morris
- Andy Boyd
- Brian Berry
- Carole Morris
- Cassie Smith
- Claire Edgeworth
- David Evans
- Doug Akins
- Edel McNamara
- Ishbel MacPherson
- James Squires
- Janet Valentine
- Juliet Underdown
- Katharine Evans
- Kevin Willis
- Konstantinos Kaouras
- Laura Goold
- Laura Marchant
- Liam Hart
- Liz Merrifield
- Lucy Turner
- Maeve Groot Bluemink
- Marie Devlin
- Mary Fitzgerald
- Matthew Bonam
- Munisa Hashimi
- Pamela Linksted
- Paola Quattroni
- Pete Stokes
- Rachel Brophy
- Rajive Sharma
- Roger Gibb
- Ruth Boardman
- Sharon Heys
- Sophie Baines
- Stephen Burrows
- Tony Plant
- Yemi Macaulay

# Acknowledgements

With special thanks to the expert panel of reviewers who helped us assess and select our impacts: James O'Shaughnessy (Chair, HDR UK Board); Patsy Wilkinson (HDR UK Board), Arun Sujenthiran (FlatIron Health), Rebecca Cosgriff (formerly NHS Research and Development), Melissa Lewis-Brown (Cancer Research UK), David Goll (10DS, Downing St), Alison Paprica (ICES Canada), Barbara Moloney-Oates (Health Research Authority), Evelina Gabasova (Turing Institute), Ekin Bolukbasi (Wellcome, acting in a personal capacity), Elliot McClenaghan (London School of Hygiene and Tropical Medicine; Early Career Researcher), Munisa Hashimi (HDR UK Public Advisory Board), Chris Monk (Public Contributor), Aisha Kekere-Ekun (HDR UK Public Advisory Board), Anmol Arora (University of Cambridge; Early Career Researcher).

Also thanks to HDR UK's Impact Committee, comprised entirely of early career researchers, who lead on HDR UK's approach to impact: Anmol Arora (University of Cambridge), Aditya Kale (University of Birmingham), Elliot McClenaghan (London School of Hygiene and Tropical Medicine), Jon Kennedy (Swansea University), Ina Kostakis (Portsmouth Hospitals University NHS Trust), Louise Marryat (University of Dundee), Jeannie Collins (University College London), and Kieran Zucker (University of Leeds).

Printed by Kingfisher Press  
on FSC® certified paper.

Kingfisher Press is an EMAS certified company  
and its Environmental Management System  
is certified to ISO 14001.

This document is printed on Soporset Offset,  
a paper containing 100% virgin fibre sourced  
from well-managed, responsible, FSC® certified  
forests and other controlled sources.

Designed and produced by **Made Noise**.

Health Data Research UK is the national institute for health data science and registered charity (no. 1194431), funded by UK Research and Innovation, the Department of Health and Social Care in England and equivalents in Northern Ireland, Wales and Scotland, and leading medical research charities.

**[www.hdruk.ac.uk](http://www.hdruk.ac.uk)**