International Journal of Population Data Science

Journal Website: www.ijpds.org

UK Longitudinal Linkage Collaboration (UK LLC): The National Trusted Research Environment for Longitudinal Research

Andy Boyd[1,*], Katharine M. Evans[1], Emma L. Turner[1], Robin Flaig[2,3], Jacqui Oakley[1], Kirsteen C. Campbell[3], Richard Thomas[1], Stela McLachlan[2,3], Matthew Crane[1], Rebecca Whitehorn[2,3], Rachel Calkin[1], Abigail Hill[1], Samantha Berman[1], David Ford[4], Martin Tobin[5,6], David Porteous[2,3], Danielle F. Gomes[7], Maria-Paz Garcia[8], Andrew Wong[9], Aida Sanchez[7], Chris Orton[4], Simon Thompson[4], John Gulliver[10], Kathryn Adams[11], Ellena Badrick[12], Chiara Batini[5], Michaela Benzeval[13], Susie Boatman[14], Gerome Breen[15,16], Shannon Bristow[15,16], Abigail Britten[14], Luke Bryant[5], Adam Butterworth[17], Archie Campbell[18], Sarah Chave[19], John Danesh[20,21,22,23,24,25], Jayati Das-Munshi[26], Karen Dennison[7], Emanuele Di Angelantonio[20,21,22,23,24,27], Thalia C. Eley[15,16,], Helen Fisher[26], Emla Fitzsimons[28], Alissa Goodman[7], Michael Gregg[19], Anna L Guyatt[5], Anna Hansell[11], Rebecca Harmston[19], Andy Heard[29], Morag Henderson[7], Rosie Hill[19], Szu-Chia Huang[19], Catherine John[5,6], Frank Kee[30], Nathalie Kingston[31,32], Jack Kneeshaw[13], Rashmi Kumar[19], Genevieve Lachance[8], Celestine Lockhart[15], Hazel Lockhart-Jones[4], Sarah Markham[19], Dan Mason[12], Bernadette McGuinness[30], Maisie McKenzie[19], Amy McMahon[20,21,23], Chelsea Mika Malouf[15,16], Mark Mumme[33], Charlotte Neville[30], Kate Northstone[33], Zoe Oldfield[34], Dara O'Neill[35], Manish Pareek[36], John Pickavance[12], Yasmin Rahman[19], Holly Reilly[36], Angela Scott[30], Deb Smith[19], Andrew Steptoe[37], Claire Steves[8], Cathie Sudlow[18], Gerald Sze[5], Nicholas L. Timpson[33], Tapiwa Tungamirai[24], Laura Venn[5], Matthew Walker[20,21,23], Neil Walker[31], Nicolas Wareham[13], Aidan Watmuff[12], Tony Webb[14], Karen Williams[19], John Wright[12], Darioush Yarand[8], George B Ploubidis[7], John Macleod[1], Jonathan AC. Sterne[1], and Nishi Chaturvedi[9]

Submission History

Submitted: 26/07/2024
Accepted: 09/12/2024
Published: 17/02/2025

[1] UK Longitudinal Linkage Collaboration, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, BS8 2PS, UK
[2] Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
[3] UK Longitudinal Linkage Collaboration, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH16 4UX, UK
[4] Population Data Science, Swansea University, Swansea, SA2 8PP, UK
[5] Population Health Sciences, University of Leicester, Leicester, LE1 7RH, UK
[6] University Hospitals of Leicester NHS Trust, Leicester, LE1 5WW, UK
[7] Centre for Longitudinal Studies, University College London, London, WC1H 0AL, UK
[8] Department of Twin Research & Genetic Epidemiology, King's College London, London, SE1 7EH, UK
[9] MRC Unit of Lifelong Health & Ageing, University College London, London, WC1E 7HB, UK
[10] Population Health Research Institute, City St George's, University of London, Cranmer Terrace, London, SW17 0RE, UK
[11] Centre for Environment, Sustainability and Health, University of Leicester, Leicester, LE1 7RH, UK
[12] Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, BD9 6RJ, UK
[13] Understanding Society, University of Essex, Colchester, CO4 3SQ, UK
[14] MRC Epidemiology Unit, University of Cambridge, Cambridge, CB2 0QQ, UK
[15] Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, SE5 8AF, UK
[16] UK National Institute for Health and Social Care Research (NIHR) Biomedical Research Centre for Mental Health, South London and Maudsley Hospital, London, SE5 8AF, UK
[17] Division of Cardiovascular Medicine, University of Cambridge, Cambridge, CB2 0QQ, UK
[18] Generation Scotland, University of Edinburgh, Edinburgh, EH4 2XU, UK
[19] UK LLC Public Involvement Programme, UK Longitudinal Linkage Collaboration, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH16 4UX, UK
[20] British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB2 0BB, UK
[21] Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, CB2 0BB, UK
[22] British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, CB2 0BB, UK
[23] National Institute for Health and Care Research Blood and Transplant Research Unit in Donor Health and Behaviour, University of Cambridge, Cambridge, CB2 0QQ, UK
[24] Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, CB10 1SA, UK
[25] Department of Human Genetics, Wellcome Sanger Institute, Hinxton, CB10 1SA, UK
[26] Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, SE5 8AF, UK
[27] Health Data Science Research Centre, Human Technopole, Milan, 20157, Italy
[28] Millennium Cohort Study, University College London, London, WC1H 0AL, UK
[29] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, Norfolk Place, London, SW7 2AZ, UK
[30] Centre for Public Health, Queens University Belfast, Belfast, BT12 6BA, UK
[31] NIHR BioResource, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK
[32] Department of Haematology, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK
[33] Avon Longitudinal Study of Parents and Children, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK
[34] Institute for Fiscal Studies, University College London, London, WC1E 7AE, UK
[35] Social Research Institute, University College London, London, WC1H 0AA, UK
[36] Department of Respiratory Sciences, University of Leicester, Leicester, LE1 9HN, UK
[37] Department of Behavioural Science and Health, University College London, London, WC1E 6BT, UK

Abstract

Introduction

The UK Longitudinal Linkage Collaboration (UK LLC) is the national Trusted Research Environment (TRE) for the UK's longitudinal research community, supporting the UK's unparalleled collection of Longitudinal Population Studies (LPS). Initially set up as a COVID-19 research resource, UK LLC is now a generic database for any research for the public good.

Objectives

UK LLC supports longitudinal research by providing record linkage and TRE services.

Methods

The UK LLC partnership provides a secure analytics environment, a trusted third-party linkage processor and a comprehensive governance framework to minimise risks to participant confidentiality. UK LLC is ISO 27001 certified and accredited by the UK Statistics Authority as a processor under the Digital Economy Act. The active involvement by members of UK LLC's public involvement programme ensures UK LLC is acceptable to LPS participants and the wider public. All UK LPS are eligible for inclusion. Researchers can apply to access the TRE via an approach that fulfils the needs of the LPS, the linked data owners and includes a review by public contributors.

Results

Twenty-two LPS have so far joined UK LLC. Where permissions allow, participants are linked to their National Health Service (NHS) England, NHS Wales and place-based records, with work ongoing to link to NHS Scotland and non-health administrative records, including Department for Work and Pensions and His Majesty's (HM) Revenue and Customs. UK LLC Explore allows potential researchers to discover the breadth of data available in the TRE. All applications are listed on UK LLC's publicly accessible Data Access Register.

Conclusions

UK LLC enables researchers to interrogate pooled LPS participant data that are systematically linked to diverse records. UK LLC remains open to additional LPS joining the partnership and will increase the breadth of data to support the longitudinal research community and attract increasing numbers of researchers across multiple disciplines, government departments and industry.

Keywords

longitudinal population studies; cohort studies; record linkage; FAIR; trusted research environment; UK LLC

---

*Corresponding Author:
Email Address: a.w.boyd@bristol.ac.uk (Andy Boyd)

https://doi.org/10.23889/ijpds.v10i1.2468

Introduction

The UK Longitudinal Linkage Collaboration (UK LLC) is the national Trusted Research Environment (TRE) for the UK's longitudinal research community. It is designed to support the UK's unparalleled collection of Longitudinal Population Studies (LPS) by providing record linkage and secure analysis and data curation services. LPS follow the lives of participant volunteers over time, including over whole lifetimes, generations of families and households. Data collected include in-depth measures of physical and mental health, lifestyle, environmental and socio-economic measures, alongside biological samples. LPS therefore provide a depth of data that gives unique insights into population wellbeing, behaviours and development. The primary funders for UK LPS share a strategic objective for maximising the value of LPS by ensuring their data are FAIR (findable, accessible, interoperable and reusable) [1] and enhanced through linking participants to their routinely collected health and administrative records and geo-spatial 'place-based' data about the environment, properties and neighbourhoods in which they live.

UK LLC is a growing partnership of many of the UK's most established LPS and, with a strong public contribution, it is led by the Universities of Bristol and Edinburgh, in collaboration with Swansea University; the University of Leicester; City St George's, University of London; and UCL. UK LLC has four main objectives:

1) To provide record linkage services to LPS. This includes bringing together de-identified data from the LPS and to systematically link these data to the participants' health, administrative, and environment and neighbourhood data in the UK LLC TRE in a manner that is compliant with all relevant legal, information security and regulatory frameworks, is publicly acceptable and is efficient to all parties.

2) To provide a secure generic research database – the UK LLC TRE – that supports efficient access from large numbers of UK-based approved researchers for approved projects, so they can conduct research for the public good using diverse data from one or more LPS linked to diverse participant records.

3) To drive improvements in research equity by co-locating many LPS' populations into one location to create a highly heterogeneous UK-wide sample, with increased statistical power to study 'rare' exposures/outcomes and to consider differential outcomes across diverse and under-served population sub-groups.

4) To offer the UK LLC TRE as part of an efficient, responsive and secure UK data science capability to support investigation of emerging policy questions and to meet future crises such as pandemics, the impacts of climate change or economic shocks.

UK LLC was established in 2020 as part of the COVID-19 Longitudinal Health and Wellbeing National Core Study to fulfil these objectives and to support the longitudinal research community to transition to a TRE way of working. Initially supported by HM Treasury to underpin high priority COVID-19 research questions, UK LLC is now funded by UK Research and Innovation, Economic and Social Research Council and Medical Research Council.

Currently, 22 LPS have joined UK LLC. This group is drawn from a collection of around 100 UK LPS with a combined estimated population size of >3 million individuals. UK LLC remains open to additional LPS joining the partnership through a formalised onboarding process. The UK LLC TRE is likely to host data about >500,000 LPS participants – similar to extensively used flagship databases such as UK Biobank – with a highly heterogenous sample and data often collected from pregnancy or soon after birth and across multiple generations. This full life-course data has significant advantages in informing research on critical health or social science effects. For example, the impacts of in utero or early life exposures on later outcomes, or generational transmission of inequalities, or identifying factors impacting on key transition periods, such as from education into employment. In addition, data hosted in the UK LLC TRE could provide significant benefits when used in combination with other databases. For example, UK LLC could provide an effective means to replicate findings initially made in UK Biobank or other large volunteer database (or vice versa), or to conduct deep phenotyping for analysis protocols designed to leverage the strengths of whole population databases (i.e. large power, high representation, but limited depth of data), through access to granular LPS data, such as health or social behaviours, that are not present in routine records.

The UK LLC protocol has been iteratively developed with key stakeholders, including public and participant contributors and national data owners. This paper summarises version 2.0 of UK LLC's protocol [2] and describes a generic research databank for longitudinal research for the public good. It replaces version 1.0 of the protocol which established UK LLC as a COVID-19 research resource [3]. This protocol paper describes UK LLC in its widest terms, including partner organisations; legal and regulatory basis; underlying technical infrastructure; role of public/participant contributors; ethico-legal basis; and approach to reproducible research. The paper then summarises the current set of contributing LPS; the types of data available; the mechanisms for making the data discoverable and available to approved researchers; and the breadth of the anticipated scientific programme.

Methods

UK LLC trusted research environment

The Trusted Research Environment (TRE) concept is established across academic [4, 5], National Health Service (NHS) [6] and wider government stakeholders [7], and organisations devoted to promoting privacy and trustworthiness of data use [8, 9], as an appropriate mechanism for population health science for public good, whilst controlling for risks to confidentiality and data misuse. Most UK TREs, including UK LLC, have adopted the 'Five Safes' governance framework [10] for the design and management of a TRE. The UK LLC TRE comprises a secure analytics environment, a trusted third party linkage processor and a comprehensive governance framework implementing technological, data, security and ethico-legal governance controls to minimise risks to participant confidentiality.

UK LLC TRE partnership

UK LLC contracts Swansea University to provide a Secure eResearch Platform (SeRP UK) infrastructure for the UK LLC TRE and their partner, NHS Digital Health and Care Wales (DHCW), to conduct trusted third party linkage services. The Universities of Bristol and Edinburgh manage the governance, data curation, data application and public involvement aspects of UK LLC (and lead the programme as a whole), with the University of Bristol acting as Data Controller. The University of Leicester and City St George's, University of London provide environmental exposure modelling and geo-coding expertise; Swansea University provides additional data curation expertise; and UCL provides expert interdisciplinary guidance in longitudinal research. The active involvement in the design and operation of UK LLC by contributing LPS and participants/public via UK LLC's Public and Participant Involvement and Engagement (PPIE) programme, ensures the system is designed to meet the collective and interdisciplinary needs of LPS in a manner acceptable to participants and the wider public.

The UK LLC TRE partnership works alongside and often in collaboration with wider UK Data Science investments, including other TREs (OpenSAFELY, Office for National Statistics, Secure Anonymised Infrastructure for Linkage (SAIL) Databank), data owners (NHS, Department of Education, Department for Work and Pensions and HM Revenue and Customs), LPS infrastructures (CLOSER, UK Data Service, Dementias Platform UK), and funders and data science enabling programmes (Health Data Research UK and Administrative Data Research UK) (Figure 1).

Legal and regulatory basis

Both UK LLC and collaborating LPS maintain a UK General Data Protection Regulation (GDPR) basis for holding and processing data and for the use of special category information. For UK LLC this is Article 6(1)(e) and Article 9(2)(j), respectively. The use of participant personal data for record linkage purposes carries a Duty of Confidentiality in UK Common Law, which LPS meet by either gaining explicit participant consent or through using opt-out objection mechanisms coupled with regulatory approvals and enabling legislation, which differ across the UK. For NHS England and Wales records, UK LLC fully respects any participant objections recorded in the respective NHS national data opt out schemes, unless overridden by explicit consent. NHS Wales flows de-identified data into the SAIL Databank (based on SeRP UK infrastructure) in such a way that they are not personal data whilst in the protective controls of the SAIL Databank and SeRP infrastructure (which includes the UK LLC TRE). The flow of NHS data in Scotland into the UK LLC TRE will be assessed via a 'public interest' test conducted by NHS Scotland Public Benefit and Privacy Panel for Health and Social Care. The legal basis for the linkage and use of non-health data will be section 64 of the Digital Economy Act 2017.

Participant and public involvement and engagement (PPIE)

UK LLC has embedded public contributors, including LPS participants and members of the public, across its work. Many of UK LLC's safeguards and system design decisions were informed by extensive PPIE conducted by individual LPS with their participants over many years. UK LLC has recruited a diverse group of public contributors who input to: (i) UK LLC Strategic Advisory Committee (strategy and future-facing guidance); (ii) UK LLC Data Access Public Review Panel (review of applications to access the UK LLC TRE); (iii) UK LLC Public Advisory Group (insights on new initiatives and review of UK LLC communication materials); and (iv) UK LLC Public Involvement Network (wider public perceptions on UK LLC activity). By embedding PPIE in its operations, UK LLC aims to include the public voice in design and decision making, thereby listening to concerns and demonstrating that it is trustworthy.

Working in partnership with LPS and their participants

The trust relationship between an LPS and its participants is crucial to maintaining ongoing participation in an LPS. LPS need to demonstrate to their participants how they are maintaining their custodian role within the UK LLC TRE. The framing and 'rules' relating to this are specific to any given LPS, because all differ in key ways: purpose; start date; longevity; sample composition; and prior assurances to participants. These assurances are drawn from all information given to participants over time, including consent materials and some LPS may have provided more formalised 'social contracts' (a set of rules regarding data use, e.g. ALSPAC [11] and TEDS [12]).

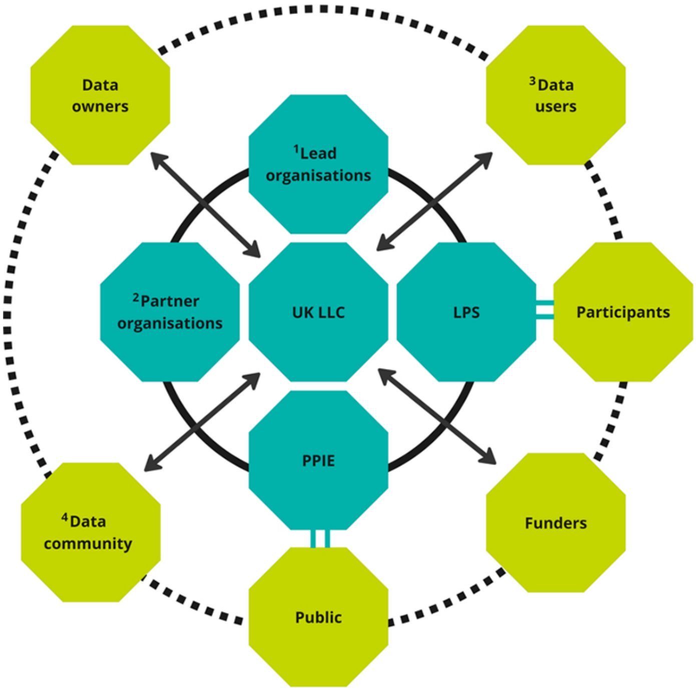
Through PPIE, UK LLC co-develops generic fair processing materials, which detail: (i) the parties and data involved; (ii) the way in which the data flow and are combined; (iii) the intended purpose for the data; (iv) how UK LLC minimises risks; (v) how UK LLC respects existing objections; and (vi) how a participant can register an objection. These materials are shared with LPS for them to tailor and implement before any data flow into the UK LLC TRE. LPS remain as Data Controllers of their data, which means at an LPS level they can decide which datasets to link to and which projects are permitted access to their data.

Split-file processing and record linkage

The SeRP UK-DHCW partnership offers 'split file' processing, which enables the UK LLC TRE to be a fully de-identified environment where hosted data are effectively anonymous to all approved researchers and the system administrators. DHCW, as UK LLC's linkage trusted third party, only handles LPS participants' identifiers and does not have access to any further individual attribute data, and SeRP UK and UK LLC only have access to de-identified LPS data and linked records (Figure 2).

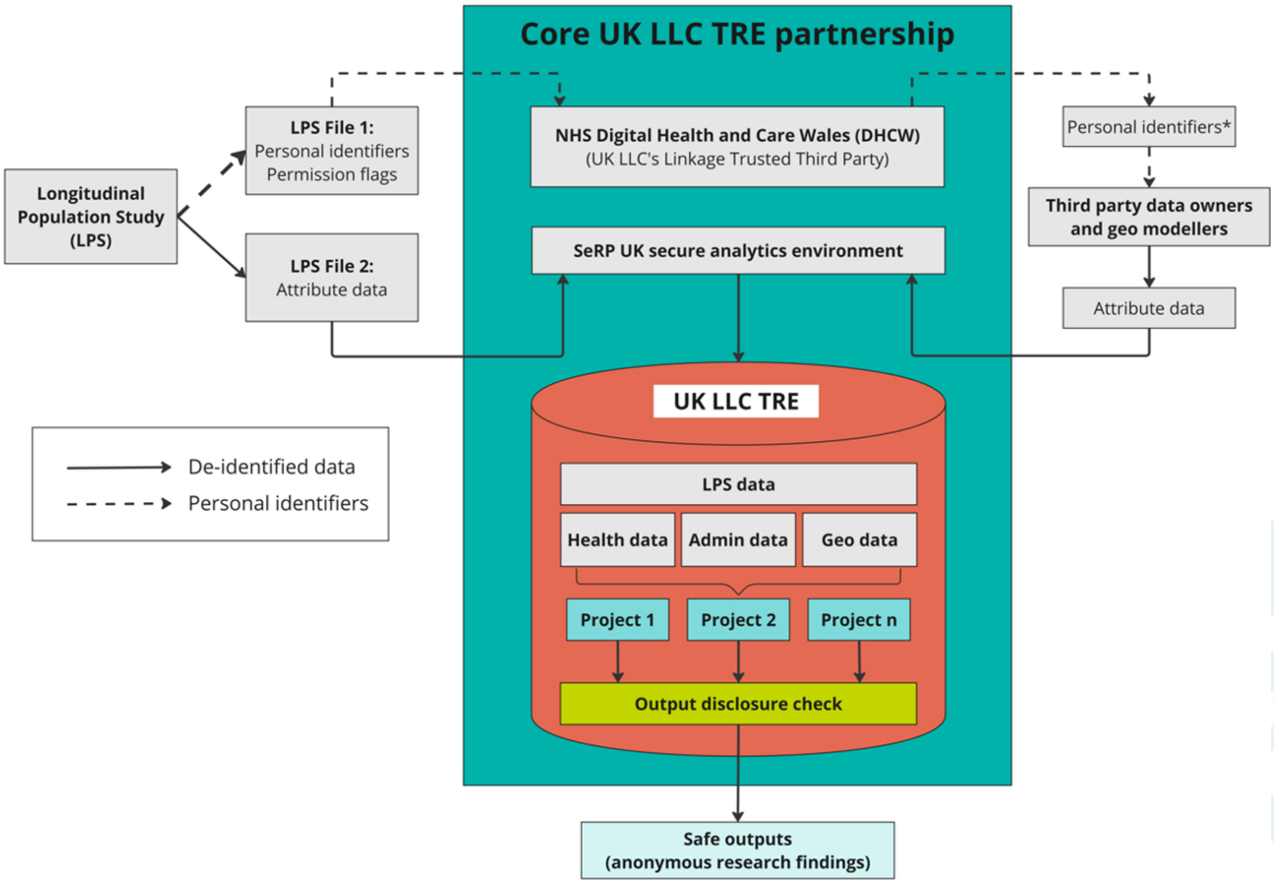
Using this methodology, the LPS data are split by LPS data managers into a file of personal identifiers ('File 1') and a file of de-identified attribute data ('File 2'). LPS send the File 1 to DHCW – this file includes current and historic names, date of birth, NHS ID, current and historic address data and

Figure 1: The wider UK LLC TRE partnership



[1]Lead organisations: University of Bristol and University of Edinburgh.
[2]Partner organisations: City St George's, University of London; Digital Health and Care Wales (DHCW); Secure eResearch Platform (SeRP UK); Swansea University; UCL; University of Leicester.
[3]Data users: Researchers interested in accessing data in the UK LLC TRE.
[4]Data community: Other data science organisations and infrastructures.
LPS: Longitudinal Population Studies; PPIE: Public and Participant Involvement and Engagement.

Figure 2: UK LLC architecture for 'split file' management of participant identifiers and de-identified LPS and linked records



permission status flags. Distinct permission flags are provided for each linked data source and for the use of address data for geo-modelling. Permission flags can be set at the LPS or participant level to indicate permissions/objections. Refreshed File 1s are sent to DHCW each quarter, enabling UK LLC to enact withdrawal/dissents, to provide updated personal identifiers to reflect name and address changes, and to add new participants for LPS with active recruitment.

DHCW then act as the linkage 'broker', facilitating linkages by sending reformatted and permission-filtered files to linked data owners and UK LLC's geo-modeller for processing. These organisations create and send File 2s containing de-identified linked attribute data to SeRP UK for import into the UK LLC TRE. The File 2s sent to SeRP UK by LPS are processed in the same way.

Secure analytics environment

The UK LLC TRE comprises a virtual desktop infrastructure running a Microsoft Windows operating system and containing standard analytics software (including R and STATA statistics software, Python programming tools, GitLab and Jupyter Notebooks documentation tools, and Microsoft Office software). Data are hosted in a Microsoft SQL Server relational database, accessible through a database management software such as SQL Server Management Studio, Eclipse and through ODBC connections to statistical software. This Windows environment is linked to a high-volume storage facility (2PB capacity) and to a High-Performance Computing cluster ($\sim$1800 processing cores and $\sim$14TB of memory) to enable computationally intensive analysis. All computing, processing and data storage is conducted on Swansea University servers meaning the data always remain in the UK.

The environment functions as a 'reading library' where the system is configured to allow approved researchers to remotely access the UK LLC TRE to conduct analyses, but blocks any mechanism for taking data out. Researchers can request the export of aggregated analytical outputs from the TRE following statistical disclosure control assessment by a team at SAIL Databank. SeRP UK, SAIL Databank, DHCW and UK LLC all have information security management systems that are certified to ISO 27001 international standard for information security. In addition, UK LLC, SAIL Databank and DHCW are accredited by the UK Statistics Authority as processing environments under the Digital Economy Act 2017.

Reproducible and reusable research

UK LLC requires all approved researchers to make analytical syntax, code lists, derived variables and project documentation available to future users in the UK LLC GitHub and GitLab, as appropriate. These research tools are catalogued and made discoverable and accessible to future UK LLC users. UK LLC has developed 'helper' syntax and software to enable commonly conducted tasks, e.g. providing the mechanism to extract data from the UK LLC database into a researcher's preferred statistical software package and to annotate these data with metadata labels. UK LLC also derives a file of core socio-economic and demographic indicators, which are harmonised in definition and encoding across all contributing LPS to expedite cross-LPS research, using both self-reported and linked information. UK LLC's Team Data Science approach is informed by the UK Reproducibility Network [13] and the FAIR Guiding Principles for scientific data management and stewardship [1].

Contributing LPS and participant sample

LPS are eligible for inclusion in UK LLC if they have/had direct follow-up with participants and linkage permissions. There are no eligibility criteria regarding the number of participants in an LPS, because it is recognised that some LPS with small sample sizes have either unique and important data or may feature under-served groups. At the time of writing, 22 LPS participate in UK LLC (Table 1), including pan-UK studies, studies based in England, Scotland or Northern Ireland alone and studies with biomedical and/or social science backgrounds. Detailed descriptions of LPS cohorts are provided in the Supplementary Materials. All 22 LPS that participated in the initial COVID-19 programme chose to continue involvement in the new phase of UK LLC (the 2023–2028 funding period). Substantial numbers of additional LPS are now seeking to join the UK LLC partnership and the first will onboard during 2025.

Managing the complexity of the participant sample

Dynamic denominator

The UK LLC denominator is the sum product of all participants whose data are provided by the contributing LPS. The denominator is highly complex, dynamic and can be assessed at multiple levels, where: (i) the sample provided by each LPS will change as new participants join or some withdraw/revoke consent; (ii) new LPS may join UK LLC and some LPS may withdraw; (iii) participants are known to take part in multiple LPS, and any new LPS joining UK LLC may result in changing numbers of multiple participants being identified as the same unique individual; and (iv) participants' changing interactions with NHS and government services (e.g. within-UK migration) may result in their appearance in new datasets.

The UK LLC denominator is fixed on a quarterly basis (in line with LPS refreshes of File 1s and the linkages based on these). Each quarter, UK LLC establishes a 'data freeze' of the UK LLC sample, which is critical to interpreting the resource, and provisions data to approved researchers based on this headline denominator and appropriate participant permissions [14]. The data tables relating to each quarterly fix are retained for archiving and reproducibility purposes.

Multiple membership of LPS and relationships

UK LLC will develop a de-identified participant 'register' for the denominator to enable researchers to understand which participants are active in multiple LPS and the relationship between different participants, and also between participants and households. This is crucial because the joint analysis of data across multiple LPS will be conducted on a statistical assumption that the samples are independent. The overlap of participant samples may introduce consent/permission ambiguities where permissions to link to and use routine records across different LPS are set in contrasting ways.

Table 1: Key information about the 22 LPS that contribute to the UK LLC TRE

Name[1]	**Owner**	**Coverage**	**Cohort**	**Years**
AIRWAVE [23]	Imperial College London	E, S, W	53,280 police officers and staff, $\geq$17 years, recruited 2004–2015	2004–
ALSPAC [24]	University of Bristol	E	c. 14,000 pregnant women recruited 1991–1992	1991–
BCS70 [25]	UCL	E, S, W	c. 17,000 babies born in a single week of 1970	1970–
BIB [26]	Bradford Teaching Hospitals NHS Foundation Trust	E	12,453 women (3,443 partners) with 13,776 pregnancies at Bradford Royal Infirmary 2007–2010	2007–
ELSA [27]	UCL	E	c. 18,000 adults, $\geq$50 years, recruitment ongoing	2002–
EPICN [28]	University of Cambridge	E	c. 30,000 adults, 40-79 years, recruited 1993–1998	1993–
EXCEED [29]	University of Leicester	E	c. 11,000 adults, recruitment ongoing	2013–
FENLAND [30]	University of Cambridge	E	12,435 adults born 1950–1975	2005–
GENSCOT [31]	University of Edinburgh	S	c. 24,000 people, $\geq$12 years, recruitment ongoing	2006–
GLAD [32]	King's College London	UK	c. 40,000 people, $\geq$16 years, recruitment ongoing	2018–
MCS [33]	UCL	UK	18,818 babies born in 2000–2002	2000–
NCDS58 [34]	UCL	E, S, W	17,415 babies born in a single week in 1958	1958–
NEXTSTEP [35]	UCL	E	c. 16,000 people born 1989–1990 recruited in 2004	2004–
NICOLA [36]	Queen's University Belfast	NI	c. 8,500 adults aged $\geq$50 years, recruitment ongoing	2013–
NIHRBIO_COPING [37]	University of Cambridge	UK	c. 150,000 people aged $\geq$16 years, recruitment ongoing[2]	2020–
NSHD46 [38]	UCL	E, S, W	5,362 babies born in a single week in 1946	1946–
SABRE [39]	UCL	E	4,858 adults aged 40–69 years recruited 1988–1991	1988–
TEDS [40] (includes E-Risk)	King's College London	E, W	13,759 pairs of twins born 1994–1996	1994–
TRACKC19	University of Cambridge	E	Up to 90,000 adults previously recruited into INTERVAL [41], COMPARE [42] and STRIDES [43]	2020–
TWINSUK [44]	King's College London	UK	c. 15,000 adults who are identical or non-identical twins, recruitment ongoing	1992–
UKHLS [45]	University of Essex	UK	c. 40,000 households recruited in 2009	2009–
UK-REACH [46]	University of Leicester	UK	c. 18,000 HCWs recruited 2021–2022	2020–2045

[1] See the Supplementary Materials for additional key references.

[2] Recruited from the general NIHR BioResource (including c. 14,000 participants from the COPING study).

AIRWAVE: The Airwave Health Monitoring Study; ALSPAC: Avon Longitudinal Study of Parents and Children; BCS70: 1970 British Cohort Study; BIB: Born in Bradford; E: England; ELSA: The English Longitudinal Study of Ageing; EPICN: The European Prospective Investigation into Cancer Norfolk Study; E-Risk: Environmental Risk Longitudinal Twin Study; EXCEED: Extended Cohort for E-health, Environment and DNA; FENLAND: The Fenland Study; GENSCOT: Generation Scotland; GLAD: Genetic Links to Anxiety and Depression Study; MCS: The Millennium Cohort Study; NCDS58: 1958 National Child Development Study; NEXTSTEP: The Next Steps Study; NI: Northern Ireland; NICOLA: Northern Ireland Cohort for the Longitudinal Study of Ageing; NIHRBIO_COPING: NIHR BioResource COVID-19 Psychiatry and Neurological Genetics Study; NSHD46: MRC National Survey of Health and Development Cohort; S: Scotland; SABRE: Southall and Brent Revisited; TEDS: The Twins Early Development Study; TRACKC19: TRACK-COVID Study; UKHLS: Understanding Society – the UK Household Longitudinal Study; UK-REACH: UK Research study into Ethnicity And COVID-19 outcomes in Healthcare workers; W: Wales

UK LLC will work with LPS to establish the most effective methodology to overcome these challenges and will provide guidance to researchers.

Data types

LPS data

LPS data cover most conceivable topics and include detailed information about subjects that are discussed openly in society through to those that are highly sensitive. These are collected directly by LPS from participants via face-to-face or remote interviews/assessments or via linkage to routine records or novel data sources (such as social media posts, personal sensors, images). LPS ensure the anonymity of all data provided to UK LLC, for example, by providing metrics that characterise the mood of a social media post (rather than providing the actual wording) and by putting time stamps into bands. LPS data in the UK LLC TRE include:

- Broad ranging quantitative data on participants' demographic, socio-economic and health status (physical and mental); family and life-course indicators; and information on diverse behaviours, aspirations and outcomes
- Existing assayed biological information (e.g. blood group type, biomarkers such as serology)
- Derived genetic, metabolomic, proteomic and epigenetic information
- Specific COVID-19 data collections (questionnaires and assayed biological data)
- Participant consent/opt-out status and history.

LPS data could extend to other classes of de-identified data, such as derived information from qualitative studies or image data (e.g. brain or organ MRIs, DEXA bone density scans, retina images) or wider smart data (e.g. sensor data from wearable devices).

Linked health data

It is optimal to: (i) extract individuals' full life-course records to support longitudinal assessments and to build health care use pathways; (ii) for the extract to be refreshed on a timely basis to ensure an ongoing assessment of healthcare use and outcomes, and for new LPS survey data to be linked to data from an equivalent time period; and (iii) to have as complete LPS population coverage as possible. However, coverage is limited due to restrictions in data availability (e.g. limited temporal coverage) and we recognise that some vulnerable and marginalised groups are systematically under-represented [15] because of limitations in linked records, e.g. specific inclusion/exclusion criteria and due to governance reasons.

All data extracted under version 1.0 of the UK LLC protocol for COVID research purposes [3] (the 'historical data') will be retained because they are directly relevant for the new research purpose and will support ongoing COVID research.

The principles for the linkage mechanism to NHS records are the same for all four nation health authorities (see figure 2), with DHCW providing lists of participant identifiers to NHS England (for English NHS records), Public Health Scotland (for Scottish NHS records) and the SAIL Databank (for Welsh NHS records), respectively. NHS authorities apply opt-out, as appropriate. Updates are provided on a quarterly basis. It is UK LLC's intention to develop governance approvals to allow linkage to Northern Irish NHS records as mechanisms to enable this are established by the devolved authorities. NHS data in the UK LLC TRE include:

- Demographic data such as date of birth, sex and entry and exit to NHS services
- Mortality and cancer registry records, and audits such as stroke, cardiac and intensive care
- Primary care data
- Secondary care data including hospital inpatients (includes intensive care and maternity), outpatients and accident and emergency records
- Mental health data about people in contact with community mental health care services
- Community services data including breastfeeding, nutrition, care event and screening
- Medicines dispensed in the community
- COVID-19 specific datasets, including testing, vaccination and outcomes.

Linked administrative data

UK LLC is an interdisciplinary database that supports research into individuals' socio-economic outcomes and understanding of health/socio-economic interactions. UK LLC will link administrative data via the Office for National Statistics and in agreement with the source data owners. Updates will be provided on an annual basis. Data from the Department for Work and Pensions, HM Revenue and Customs, and Department for Education are anticipated to flow into the UK LLC TRE, including:

- HM Revenue and Customs: Data items relating to employment payments; workplace pensions; employment cessation payments; nature and source of income; pensions; and share schemes
- Department for Work and Pensions: Customer Information System; Benefits and Income Data; National Benefits Database; and Child Benefit Extract
- Department for Education: National Pupil Database; and Pupil Level Annual School Census ages 4-18 years.

Linked place-based data

UK LLC has commissioned the University of Leicester and City St George's, University of London to model a number of environmental exposure estimates and place-based data such as information about houses and neighbourhoods. These are based on taking existing environmental sensor readings and with other inputs (e.g. traffic count data, weather pattern data) used to model pollution and other environmental exposures and to map these to participants' addresses. The file sent to UK LLC's geo modeller contains only pooled address data at a 1:3 ratio of LPS participants' addresses to masking addresses. Although still defined as personal data (all address data are considered to be personal data, whether in the public domain or not), this use does not breach participants' confidentiality because nothing sensitive or confidential can be inferred from the list.

UK LLC also collates and processes place-based public domain data for inclusion in the UK LLC TRE. All these data are processed to ensure they are de-identified prior to ingest, with no geographies other than English region and devolved nation being available within the TRE. Environmental and place-based data will include:

- Geospatial modelling and public domain datasets: these are assigned to a property, postcode or higher-level geography (e.g. a lower super output area or region)
- Modelled environmental exposure estimates of pollution, climate data (e.g. temperature, rainfall, pollen) and noise
- Modelled access to green and blue space and measures estimating the 'walkability' around a property or area
- Information about the neighbourhood (e.g. building density, land use characteristics, deprivation indices, crime rates, provision of services and availability of 'hazards' such as fast food outlets or gambling shops)
- Information about the property (e.g. building age, type, building floor, sale dates, value and energy performance records).

Discovery and access

Discovery

UK LLC is designed to be a discoverable and accessible resource for public good research. UK LLC is promoted through the contributing LPS to their research users, through data science networks, longitudinal study resources and funders. UK LLC Explore [16] is UK LLC's discoverability and data selection portal that is populated with metadata using existing automated metadata extraction software and enriched through an API to the Catalogue of Mental Health Measures [17]. UK LLC has also developed a data documentation resource, the UK LLC Guidebook [18], which is enriched through an API to the HDR UK Innovation Gateway [19] and will be expanded and maintained in collaboration with data owners and researchers. There are plans to also pull information from CLOSER Discovery [20] and UK Data Archive [21]. These approaches avoid duplication, minimise LPS burden and promote interoperability and federation.

Applying to access the UK LLC TRE

Applications to access data in the UK LLC TRE are processed using a novel delegated and distributed approach that satisfies the needs of the contributing LPS, the third-party data owners and includes a review by members of the public. Researchers submit an expression of interest, where a proposed project is assessed for feasibility, and then work with UK LLC to develop their full application, which includes a list of datasets they would like to access. The application and data request are reviewed by each contributing LPS's data access committee and the UK LLC Data Access Committee. If an application is approved, researchers will only have access to the datasets approved by the data access committees.

Applications are screened using the 'Five Safes' framework [10]. All researchers must hold valid Office for National Statistics Accredited Researcher status, be based in the UK and be employed by an organisation with sufficient capacity to support good governance in research (although future enhancements will be made to extend access internationally where LPS and other data owners permit this). All projects must be for public good, ethically sound and not for profit-making purposes. All applications are listed on the publicly accessible UK LLC Data Use Register [22]. Prior to access to the UK LLC TRE, all researchers must sign a UK LLC Data User Responsibilities Agreement and their research is controlled using a Data Access Agreement (a legal contract) between each researcher's organisation and the University of Bristol.

Scientific programme

UK LLC is a generic research database with an explicit remit to support longitudinal public good research across the breadth of the sector, including supporting a broad range of both biomedical and social science research. As such, UK LLC is a FAIR resource that is open to applications from any legitimate UK-based researcher. Currently identified themes of research include investigating health and social inequalities; occupational health outcomes and assessment of outcomes in under-served groups; an ongoing COVID-19 programme; and methodological scientific programme, e.g. understanding LPS population coverage, bias and data quality, and relating to TRE ways of working such as data integration and harmonisation. These themes are non-exclusive and indicative of the wider potential of the resource.

Conclusion

UK LLC forms a novel national TRE for the longitudinal research community of globally unique depth and breadth that will enable UK-wide world-class research. For the first time, it enables researchers to interrogate pooled LPS participant data from many contributing LPS that are systematically linked to diverse routine records, enabling the study of rare exposures/outcomes, including in under-served groups. UK LLC's bespoke governance framework enables equivalency of permissions across all contributing LPS, whilst maintaining the core aspects of the participant-LPS trust relationship.

UK LLC will expand LPS membership and increase the breadth of data, sample size and diversity to support the wider LPS community and attract increasing numbers of researchers across disciplines, UK government departments and industry. UK LLC will build new functionality to enable multi-omics analyses and replicate key derived data and data formatting from other LPS and resources (e.g. UK Biobank or whole-population databases) to enable efficient cross-TRE replication analyses. As researchers use the UK LLC TRE, they will help to create an ever-increasing knowledge base that will enable transparent reproducible research and iteratively improve ease of use and functionality.

UK LLC will continue to work with the public and LPS participants to ensure the design and operation of UK LLC is developed with public input and that all activities of UK LLC are transparent to the public.

Acknowledgements

We wish to recognise and thank all LPS participants and the LPS staff that are part of the UK LLC partnership. A full list of acknowledgements, including support for each LPS, is provided in the Supplementary Materials. We thank the National Health Service (NHS) and particularly NHS England for their work in curating LPS participants' health records and for making these available for public good research designed to improve health services. In particular, we thank Garry Coleman for his input into the design of the UK LLC protocol and Mujiba Ejaz for her invaluable support and hard work on UK LLC governance. We also thank Helen Buckles, Oliver Smith, John Wigglesworth, Louise Dunn and Abigail Lucas for all their contributions. We thank the Ordnance Survey for providing AddressBase® Plus. We thank the Administrative Data Research UK (ADR UK) and Office for National Statistics (ONS) teams for their contribution to developing non-health administrative linkages, in particular Emma Gordon (ADR UK), Emily Oliver (ADR UK), Rachel Huck (ONS), Roya Shahrokni (ONS), Jen Donald (ONS), Leah Quinn (ONS); Graham Knox and Mike Daly at the Department for Work and Pensions; Mark Barry, Angela Martindale, Nike Ogunlade, Tracy Holland, Richard Millington at HM Revenue and Customs; and David Burnett at the Department for Education. We thank Carol Morris (Public Health Scotland) and the much-missed Dermot O'Reilly (ADR Northern Ireland) for their help in understanding linkages in Scotland and Northern Ireland, respectively. We thank Vicki Bowles and Claire Ainley (VWV Ltd); Clare Smith, Henry Stuart and Adam Taylor (University of Bristol); and Cynthia McNerney, Rob Garlick and Sharon Heyes (Swansea University) for their contributions to developing the UK LLC governance framework. We thank the funders of UK longitudinal research for their guidance and contributions, particularly Mary De Silva (then Wellcome Trust), Joe McNamara and Catherine Moody (Medical Research Council) and Bridget Taylor and Rebecca Perring (Economic and Social Research Council). We thank all current and past members of the UK LLC Participant and Public Involvement and Engagement (PPIE) Programme who have played invaluable roles in shaping the design of our processes and with operational decision-making. We remember and thank Dolapo (Della) Ogunleye whose constructive challenges to improve diversity in UK LLC and longitudinal research remain with us and are an important influence on our future work. Finally, we thank Sir Patrick Vallance and all of those who worked to establish the National Core Studies (NCS) programme.

Funding

UK LLC is a UK Research and Innovation (UKRI) funded infrastructure with co-funding from the Medical Research Council (MR/X021556/1) and Economic and Social Research Council (ES/X000567/1). The initial funding which established UK LLC was provided by the UKRI-funded Longitudinal Health and Wellbeing National Core Study led by UCL and University of Bristol (MC_PC_20030 and MC_PC_20059). The protocol development was informed by ABoyd's secondment to the Economic and Social Research Council (ES/S016732/1) to scope options for improving the inclusiveness of UK longitudinal research through increased use of population data; funding from the Wellcome Trust (221574/Z/20/Z) for a secretariat to co-ordinate the response to the COVID-19 pandemic; and funding for the ALSPAC birth cohort study which is core funded by the UK Medical Research Council, the Wellcome Trust and the University of Bristol (217065/Z/19/Z). The onward design, implementation and operations of UK LLC have been supported by staff from all the contributing LPS. Acknowledgements and funding for each contributing LPS are provided in the Supplementary Materials. UK LLC is a member of the DATAMIND consortium which is addressing infrastructure needs for mental health research (MRC: MR/W014386/1) and is also supported by National Institute for Health and Care Research (NIHR) funding to establish greater functionality to support occupational health research (NIHR: NIHR20671). UK LLC is also a member of the HDR UK Social and Environmental Determinants of Health driver programme, which is supporting enhanced geo-spatial linkage capabilities (HDRUK2023.0029). The views expressed are those of the authors and not necessarily those of any of the funders or the organisations contributing data into the UK LLC.

Conflicts of interest

J Danesh serves on scientific advisory boards for AstraZeneca, Novartis and UK Biobank, and has received multiple grants from academic, charitable and industry sources outside the submitted work. M Tobin, A Guyatt and C John have a funded research collaboration with Orion for collaborative research projects outside the submitted work. No other conflicts of interest were disclosed.

Ethics and informed consent

UK LLC has ethical approval from the Health Research Authority (HRA) Research Ethics Committee (Haydock Committee; ref: 20/NW/0446). Each contributing LPS has its own independent ethical basis and has established a legal and ethical basis for participant involvement in UK LLC. Participant objections and withdrawal are always upheld. UK LLC has approval from the HRA Confidentiality Advisory Group (ref: 21/CAG/0044) to support the flow of identifiers from LPS to NHS Digital Health and Care Wales where an LPS' legal basis relies on section 251 support.

Author contributions

Authors ABoyd and RF led the development of the UK LLC concept, and the initial protocol was co-developed with AC, DF, DFG, JG, AH, JM, M-PG, CO, DP, ASanchez, CSteves, MT, NJT, ST, AWong. The concept was fully developed with substantial input from data managers and principal investigators across the Vanguard group of LPS including ABritten, AButterworth, CB, EB, NC, GB, LB, MB, SBoatman, SBristow, AC, JD, JD-M, KD, EDA, TCE, EF, AG, AHeard, DFG, M-PG, MH, CJ, FK, NK, JK, CL, GL, AM, BM, CMM, DM, MMumme, CN, KN, ZO, DP, GBP, JP, ASanchez, AScott, ASteptoe, CSteves, CSudlow, GS, MT, NJT, TT, LV, AWatmuff, AWong, MW, NWalker, NWareham, JW, TW, DY. These authors also implemented the LPS-side functionality of UK LLC. The UK LLC infrastructure was developed by the UK LLC team: KME, ELT, JO, KCC, RT, SM, MC, SBerman, RW and infrastructure collaborators: DF, CO, ST, KA, ALG, JG, AHansell, HL-G. The overarching design of the Longitudinal Health and Wellbeing National Core Study – which informed the shape of this protocol – was led by NC and JS with contributions from: ABoyd, GBP, CSteves, NJT. The UK LLC public contributors were involved in the design of key UK LLC functionality described in this protocol: SC, MG, RHarmston, RHill, SC-H, RK, MMcKenzie, SM, YR, DS, KW. ABoyd drafted this protocol, with authors KME and ELT conducting significant editing. All other authors contributed to the manuscript and agreed its final form. NC secured the initial National Core Studies funding which supported the establishment and ABoyd has led subsequent funding rounds. ABoyd is responsible for UK LLC and is the guarantor for this manuscript.

Data availability statement

Data in the UK LLC Trusted Research Environment (TRE) cannot be used or shared outside this environment. UK-based researchers who hold valid Office for National Statistics (ONS) Accredited Researcher status and are employed by an organisation that can support good governance in research can apply to access the UK LLC TRE (see the process outlined in the UK LLC Data Access and Acceptable Use Policy: https://ukllc.ac.uk/governance). Researchers can explore the data available in the UK LLC TRE using UK LLC's discoverability and data selection portal (https://explore.ukllc.ac.uk/).

References

1. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3(1):160018. https://doi.org/10.1038/sdata.2016.18

2. University of Bristol & University of Edinburgh. UK Longitudinal Linkage Collaboration: The National Trusted Research Environment for Longitudinal Research (Research Protocol V2.0). Zenodo; 2023. https://doi.org/10.5281/zenodo.10868470

3. University of Bristol & University of Edinburgh. The Longitudinal Linkage Collaboration: a platform for longitudinal COVID-19 research (Research Protocol V1.0). Zenodo; 2020. https://doi.org/10.5281/zenodo.10868639

4. Burton PR, Murtagh MJ, Boyd A, Williams JB, Dove ES, Wallace SE, et al. Data Safe Havens in health research and healthcare. Bioinformatics. 2015;31(20):3241-8. https://doi.org/10.1093/bioinformatics/btv279

5. UK Health Data Research Alliance and NHSX. Building Trusted Research Environments - Principles and Best Practices; towards TRE ecosystems. 2021. https://doi.org/10.5281/zenodo.5767586

6. Department of Health & Social Care. Data saves lives: reshaping health and social care with data. Department of Health & Social Care; 2022. https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data.

7. Office for National Statistics. About the Secure Research Service: ONS; 2024. https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/secureresearchservice/aboutthesecureresearchservice.

8. Harrison T. Putting the Trust in Trusted Research Environments London: Understanding Patient Data; 2020. https://web.archive.org/web/20220810164752/https:/understandingpatientdata.org.uk/news/putting-trust-trusted-research-environments.

9. medConfidential. Analysis and Inputs Reporting: medConfidential; 2020. https://medconfidential.org/2020/.

10. Desai T, Ritchie F, Welpton R. Five Safes: designing data access for research. Bristol: University of the West of England; 2016. https://uwe-repository.worktribe.com/output/914745

11. ALSPAC. Our commitment to you 2024. https://www.bristol.ac.uk/alspac/participants/our-commitment-to-you/.

12. TEDS. The TEDS Promise 2024. https://www.teds.ac.uk/participants/tedspromise.

13. UK Reproducibility Network. The UK Reproducibility Network (UKRN) 2024. https://www.ukrn.org/.

14. Berman S, Evans K, Thomas R, Crane M, McLachlan S, Whitehorn R, et al. Summary Profile of the UK LLC Resource: Data Freeze 1. Zenodo; 2022. https://doi.org/10.5281/zenodo.10890836

15. Boyd A. Understanding Population Data for Inclusive Longitudinal Research. Bristol, UK: University of Bristol; 2021. https://www.ukri.org/publications/understanding-population-data-for-inclusive-longitudinal-research/

16. UKLLC. UK LLC Explore: UK LLC's discoverability and data selection portal 2024. https://explore.ukllc.ac.uk/.

17. KCL. Catalogue of Mental Health Measures 2024. https://www.cataloguementalhealth.ac.uk/.

18. UK Longitudinal Linkage Collaboration. UK LLC's linked data documentation and user guide 2024. https://guidebook.ukllc.ac.uk/.

19. Health Data Research UK. HDR UK Innovation Gateway 2024. https://healthdatagateway.org/en.

20. UCL. CLOSER Discovery 2024. https://discovery.closer.ac.uk/.

21. UKDA. UK Data Archive 2024. https://www.data-archive.ac.uk/.

22. UK Longitudinal Linkage Collaboration. UK Longitudinal Linkage Collaboration (UK LLC) 2024. https://ukllc.ac.uk/.

23. Elliott P, Vergnaud AC, Singh D, Neasham D, Spear J, Heard A. The Airwave Health Monitoring Study of police officers and staff in Great Britain: rationale, design and methods. Environ Res. 2014;134:280-5. https://doi.org/10.1016/j.envres.2014.07.025

24. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort Profile: the 'children of the 90s'–the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol. 2013;42(1):111-27. https://doi.org/10.1093/ije/dys064

25. Sullivan A, Brown M, Hamer M, Ploubidis GB. Cohort Profile Update: The 1970 British Cohort Study (BCS70). Int J Epidemiol. 2023;52(3):e179-e86. https://doi.org/10.1093/ije/dyac148

26. Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, et al. Cohort Profile: the Born in Bradford multi-ethnic family cohort study. Int J Epidemiol. 2013;42(4):978-91. https://doi.org/10.1093/ije/dys112

27. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort profile: the English longitudinal study of ageing. Int J Epidemiol. 2013;42(6):1640-8. https://doi.org/10.1093/ije/dys168

28. Hayat SA, Luben R, Keevil VL, Moore S, Dalzell N, Bhaniani A, et al. Cohort Profile: A prospective cohort study of objective physical and cognitive capability and visual health in an ageing population of men and women in Norfolk (EPIC-Norfolk 3). Int J Epidemiol. 2013;43(4):1063-72. https://doi.org/10.1093/ije/dyt086

29. John C, Reeve NF, Free RC, Williams AT, Ntalla I, Farmaki AE, et al. Cohort Profile: Extended Cohort for E-health, Environment and DNA (EXCEED). Int J Epidemiol. 2019;48(3):678-9j. https://doi.org/10.1093/ije/dyz073

30. Lindsay T, Westgate K, Wijndaele K, Hollidge S, Kerrison N, Forouhi N, et al. Descriptive epidemiology of physical activity energy expenditure in UK adults (The Fenland study). IJBNPA. 2019;16:126. https://doi.org/10.1186/s12966-019-0882-6

31. Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, et al. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. Int J Epidemiol. 2013;42(3):689-700. https://doi.org/10.1093/ije/dys084

32. Davies MR, Kalsi G, Armour C, Jones IR, McIntosh AM, Smith DJ, et al. The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. Behav Res Ther. 2019;123:103503. https://doi.org/10.1016/j.brat.2019.103503

33. Connelly R, Platt L. Cohort Profile: UK Millennium Cohort Study (MCS). Int J Epidemiol. 2014;43(6):1719-25. https://doi.org/10.1093/ije/dyu001

34. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). Int J Epidemiol. 2005;35(1):34-41. https://doi.org/10.1093/ije/dyi183

35. Next Steps. A guide to the re-deposit of sweeps 1 to 7 datasets. Institute of Education, UCL; 2020. https://doc.ukdataservice.ac.uk/doc/5545/mrdoc/pdf/next_steps_userguide_to_the_redeposit_of_sweeps_1to7_may2020.pdf

36. Neville C, Burns F, Cruise S, Scott A, O'Reilly D, Kee F, Young I. Cohort Profile: The Northern Ireland Cohort for the Longitudinal Study of Ageing (NICOLA). Int J Epidemiol. 2023;52(4):e211-e21. https://doi.org/10.1093/ije/dyad026

37. Young KA-O, Purves KL, Hübel C, Davies MR, Thompson KN, Bristow S, et al. Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK. Psychol Med. 2023;53(12):5428-41. https://doi.org/10.1017/S0033291722002501

38. Kuh D, Pierce M, Adams J, Deanfield J, Ekelund U, Friberg P, et al. Cohort profile: updating the cohort profile for the MRC National Survey of Health and Development: a new clinic-based data collection for ageing research. Int J Epidemiol. 2011;40(1):e1-9. https://doi.org/10.1093/ije/dyq231

39. Jones S, Tillin T, Park C, Williams S, Rapala A, Al Saikhan L, et al. Cohort Profile Update: Southall and Brent Revisited (SABRE) study: a UK population-based comparison of cardiovascular disease and diabetes in people of European, South Asian and African Caribbean heritage. Int J Epidemiol. 2020;49(5):1441-2e. https://doi.org/10.1093/ije/dyaa135

40. Lockhart C, Bright J, Ahmadzadeh Y, Breen G, Bristow S, Boyd A, et al. Twins Early Development Study (TEDS): A genetically sensitive investigation of mental health outcomes in the mid-twenties. JCPP Adv. 2023;3(2):e12154. https://doi.org/10.1002/jcv2.12154

41. Di Angelantonio E, Thompson SG, Kaptoge S, Moore C, Walker M, Armitage J, et al. Efficiency and safety

of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. Lancet. 2017;390(10110):2360-71. https://doi.org/10.1016/s0140-6736(17)31928-1

42. Bell SA-O, Sweeting MA-O, Ramond AA-O, Chung R, Kaptoge S, Walker M, et al. Comparison of four methods to measure haemoglobin concentrations in whole blood donors (COMPARE): A diagnostic accuracy study. Transfus Med. 2021;31:94-103. https://doi.org/10.1111/tme.12750
43. McMahon A, Kaptoge S, Walker M, Mehenny S, Gilchrist PT, Sambrook J, et al. Evaluation of interventions to prevent vasovagal reactions among whole blood donors: rationale and design of a large cluster randomised trial. Trials. 2023;24(1):512. https://doi.org.10.1186/s13063-023-07473-z
44. Verdi S, Abbasian G, Bowyer RCE, Lachance G, Yarand D, Christofidou P, et al. TwinsUK: The UK Adult Twin Registry Update. Twin Res Hum Genet. 2019;22(6):523-9. https://doi.org/10.1017/thg.2019.65
45. Buck N, McFall S. Understanding Society: design overview. Longitudinal and Life Course Studies. 2012;3(1):5-17. https://doi.org/10.14301/llcs.v3i1.159
46. Bryant L, Free RC, Woolf K, Melbourne C, Guyatt AL, John C, et al. Cohort Profile: The United Kingdom Research study into Ethnicity and COVID-19 outcomes in Healthcare workers (UK-REACH). Int J Epidemiol. 2022;52(1):e38-e45. https://doi.org/10.1093/ije/dyac171

Abbreviations

DHCW:	NHS Digital Health and Care Wales
FAIR:	Findable, Accessible, Interoperable, Reproducible
GDPR:	General Data Protection Regulation
HM:	His Majesty's
LPS:	Longitudinal Population Study
NHS:	National Health Service
PPIE:	Public and Participant Involvement and Engagement
SAIL:	Secure Anonymised Information Linkage
SeRP:	Secure eResearch Platform
TRE:	Trusted Research Environment
UK LLC:	UK Longitudinal Linkage Collaboration

Supplementary Materials – LPS descriptions and acknowledgements

AIRWAVE: Airwave Health Monitoring Study

Description of Study Population (including citations and references if required)	The Airwave Health Monitoring Study was established in 2004 in response to the Stewart report; it is a long-term occupational cohort study following up the health of police officers and staff across Great Britain in relation to TETRA use and other exposures (Elliott et al., 2014). The cohort was described in detail here: https://pubmed.ncbi.nlm.nih.gov/25194498/
Acknowledgements	The Study is funded by the Medical Research Council (MRC), (MR/R023484/1), the National Institute for Health Care Research (NIHR) Health Protection Research Unit in Chemical and Radiation Threats and Hazards (NIHR-200922), the Imperial College Biomedical Research Centre (BRC) 2017-22, Imperial College London, and the Imperial College Healthcare NHS Trust. The initial phase of the study, including participant recruitment, was funded by the Home Office (780-TETRA; 2003-18). Views expressed are those of the authors and not necessarily those of the study sponsors. We thank all study participants for their involvement.
Ethics	REC reference is 19/NW/0054
Further information	https://police-health.org.uk/applying-access-resource

ALSPAC: Avon Longitudinal Study of Parents and Children

Description of Study Population (including citations and references if required)	Pregnant women resident in a defined area of the former county of Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992 were invited to take part in the study [1,2]. The initial number of pregnancies enrolled is 14,541 (14,676 foetuses), resulting in 14,062 live births and 13,988 children who were alive at 1 year of age. Further recruitment took place after the age of 7 years, the total sample size for analyses using any data collected after the age of 7 is therefore 15,454 pregnancies, resulting in 15,589 foetuses. Of these 14,901 were alive at 1 year of age. [1] Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, Molloy L, Ness A, Ring S, Davey Smith G. Cohort Profile: The 'Children of the 90s'; the index offspring of The Avon Longitudinal Study of Parents and Children (ALSPAC). International Journal of Epidemiology 2013; 42: 111-127. [2] Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, Henderson J, Macleod J, Molloy L, Ness A, Ring S, Nelson SM, Lawlor DA. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. International Journal of Epidemiology 2013; 42:97-110.
Acknowledgements	We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.
Ethics	Ethical approval for the study was obtained from the ALSPAC Law and Ethics committee and local research ethics committees (NHS Haydock REC: 10/H1010/70).
Further information	http://www.bristol.ac.uk/alspac/researchers/access/

BCS70: 1970 British Cohort Study	
Description of Study Population (including citations and references if required)	The 1970 British Cohort Study (BCS70) follows the lives of more than 17,000 people born in England, Scotland and Wales in a single week of 1970. Over the course of cohort members' lives, BCS70 has collected information on health, physical, educational and social development, and economic circumstances, among other factors. Since the birth survey in 1970, there have been nine 'sweeps' of all cohort members at ages 5, 10, 16, 26, 30, 34, 38, 42 and most recently at 46 (a biomedical data collection). The Age 51 Sweep is currently in the field (2022). Data have been collected from a number of different sources, including the midwife present at birth, parents of the cohort members, head and class teachers, school health service personnel and the cohort members themselves. The data have been collected in a variety of ways, including via paper and electronic questionnaires, clinical records, medical examinations, biological samples, physical measurements, tests of ability, educational assessments and diaries. The study is conducted by the Centre for Longitudinal Studies.
Acknowledgements	BCS70 is core-funded by the ESRC.
Ethics	Ethics approval has been obtained for each follow-up from an NHS Research Ethics Committee (REC) since 2000. In addition, separate REC approval is in place to cover the ongoing activities of the study in between major sweeps of data collection (i.e. Keeping in touch with and tracing cohort members; cleaning, documenting and providing access to the data for research; and linking data from administrative sources to survey data to increase the utility of the data for research).
Further information	https://cls.ucl.ac.uk/cls-studies/bcs70/

BIB: Born in Bradford

Description of Study Population (including citations and references if required)	Born in Bradford (BiB) is a prospective pregnancy and birth cohort study of the children, mothers and fathers from 13,776 pregnancies from between 2007 and 2011, based in Bradford, UK. The study was established to examine how genetic, nutritional, environmental, behavioural and social factors affect health and development during childhood, and subsequently, adult life in a deprived multi-ethnic population[1]. From 2017 to 2021, a full follow-up of the cohort was conducted[2] to investigate the determinants of children's pre-pubertal health and development, including through understanding parents' health and wellbeing, and to obtain data on exposures in childhood that might influence future health. In 2022, the Age of Wonder study was launched to complete another full follow-up through adolescence into adulthood, focusing on priority areas of physical and mental health, growth, identity, cognition, socioeconomic status and environmental exposures[3]. [1] John Wright, Neil Small, Pauline Raynor, Derek Tuffnell, Raj Bhopal, Noel Cameron, Lesley Fairley, Debbie A Lawlor, Roger Parslow, Emily S Petherick, Kate E Pickett, Dagmar Waiblinger, Jane West, on behalf of the Born in Bradford Scientific Collaborators Group, Cohort Profile: The Born in Bradford multi-ethnic family cohort study, International Journal of Epidemiology, Volume 42, Issue 4, August 2013, Pages 978–991, https://doi.org/10.1093/ije/dys112 [2] Rosemary R C McEachan, Gillian Santorelli, Aidan Watmuff, Dan Mason, Sally E Barber, Daniel D Bingham, Philippa K Bird, Laura Lennon, Dan Lewer, Mark Mon-Williams, Katy A Shire, Dagmar Waiblinger, Jane West, Tiffany C Yang, Deborah A Lawlor, Kate E Pickett, John Wright, Cohort Profile Update: Born in Bradford, International Journal of Epidemiology, Volume 53, Issue 2, April 2024, dyae037, https://doi.org/10.1093/ije/dyae037 [3] Shire KA, Newsham A, Rahman A et al. Born in Bradford's Age of Wonder cohort: protocol for adolescent data collection [version 1; peer review: 2 approved]. Wellcome Open Res 2024, 9:32 (https://doi.org/10.12688/wellcomeopenres.20785.1)
Acknowledgements	Born in Bradford has received funding from the Wellcome Trust [101597/Z/13/Z and 223601/Z/21/Z]; a joint grant from the UK Medical Research Council (MRC) and UK Economic and Social Science Research Council (ESRC) [MR/N024391/1; a British Heart Foundation Clinical Study grant [CS/16/4/32482]; the National Institute for Health Research under its Applied Research Collaboration Yorkshire and Humber [NIHR200166]; and a Health Foundation COVID-19 Award [2301201]. Born in Bradford is only possible because of the enthusiasm and commitment of the Children and Parents in BiB. We are grateful to all the participants, health professionals, schools and researchers who have made Born in Bradford happen. **NMR metabolomics data – additional acknowledgement** Researchers who use the NMR metabolomics data should cite the following data note: https://wellcomeopenresearch.org/articles/5-264 The funding acknowledgement should be: Funding for the metabolomics analyses in BiB has been provided by the US National Institutes of Health [R01 DK10324]; the European Research Council (ERC) under the European Union's Seventh Framework Programme [FP7/2007-2013] / ERC grant agreement no 669545; and the UK Medical Research Council [MC_UU_00011/6].
Ethics	Study: Born in Bradford Age of Wonder: a co-produced mixed methods longitudinal exploration of health and wellbeing trajectories through adolescence and young adulthood in the multi-cultural city of Bradford, UK. REC: 21/YH/0261 Study: Born in Bradford's Growing up Family Study REC: 16/YH/0320 Study: Born in Bradford: A longitudinal cohort study of babies born in Bradford and their mothers and fathers REC: 07/H1302/112
Further information	https://www.borninbradford.nhs.uk

ELSA: English Longitudinal Study of Ageing

Description of Study Population (including citations and references if required)	The English Longitudinal Study of Ageing (ELSA) is a unique and rich resource of information on the dynamics of health, social, wellbeing and economic circumstances in the English population aged 50 and older [1]. The original sample was drawn from households that had previously responded to the Health Survey for England (HSE) between 1998 and 2001. The main fieldwork began in March 2002. The same group of respondents have been interviewed at two-yearly interviews. [1] Banks J, Batty GD, Breedvelt JJF, Coughlin K, Crawford R, Marmot M, Nazroo J, Oldfield Z, Steel N, Steptoe A, Wood M, Zaninotto P (2021) English Longitudinal Study of Ageing: Waves 0-9, 1998-2019
Acknowledgements	The English Longitudinal Study of Ageing was developed by a team of researchers based at University College London, NatCen Social Research, the Institute for Fiscal Studies, the University of Manchester and the University of East Anglia. The data were collected by NatCen Social Research. The funding is currently provided by the National Institute on Aging (Ref: R01AG017644) and by a consortium of UK government departments: Department for Health and Social Care; Department for Transport; Department for Work and Pensions, which is coordinated by the National Institute for Health Research (NIHR, Ref: 198-1074). Funding has also been provided by the Economic and Social Research Council (ESRC).
Ethics	https://www.elsa-project.ac.uk/ethical-approval
Further information	https://www.elsa-project.ac.uk/data-and-documentation

EPIC-Norfolk: European Prospective Investigation into Cancer – Norfolk study

Description of Study Population (including citations and references if required)	EPIC-Norfolk is a prospective population-based cohort study that was established as part of the European Prospective Investigation of Cancer (EPIC) a large multi-centre cohort across 23 centres in 10 different European countries. EPIC-Norfolk was primarily established to examine the links between diet, lifestyle factors and cancer risk, but the study was broadened from the outset to include other chronic diseases and health conditions. At baseline from 1993-97, 25,000 people aged 40-79 years were recruited into EPIC-Norfolk and provided a baseline blood sample. Participants were recruited from the sampling frame of 42 GP practices through Norfolk. Participants have been followed up in a total of five health checks and have additionally provided information through questionnaires over 25 years. Extensive record linkage has enabled collection of health endpoint data on a continuing basis.
Acknowledgements	The EPIC-Norfolk study (https://doi.org/10.22025/2019.10.105.00004) has received funding from the Medical Research Council (MR/N003284/1 MC-UU_12015/1 and MC_UU_00006/1) and Cancer Research UK (C864/A14136). The genetics work in the EPIC-Norfolk study was funded by the Medical Research Council (MC_PC_13048). We are grateful to all the participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research.
Ethics	The EPIC-Norfolk study was approved by the Norwich Local Ethics Committee (previously known as Norwich District Ethics Committee) (REC Ref: 98CN01); all participants gave their informed written consent before entering the study.
Further information	https://www.epic-norfolk.org.uk/for-researchers/data-sharing/data-requests/

EXCEED: Extended Cohort for E-Health, Environment and DNA	
Description of Study Population (including citations and references if required)	EXCEED is a longitudinal population-based cohort which facilitates investigation of genetic, environmental and lifestyle-related determinants of a broad range of diseases and of multiple morbidity through data collected at baseline and via electronic healthcare record linkage. Recruitment has taken place in Leicester, Leicestershire and Rutland since 2013 and is ongoing, with 11,000 participants. Participants provided a DNA sample, have consented to follow-up for up to 25 years through electronic health records and additional bespoke data collection is planned. Data available includes baseline demographics, anthropometry, spirometry, lifestyle factors (smoking and alcohol use), multi-omics data, and longitudinal health information from primary and secondary care records. Patients have consented to be contacted for recall-by-genotype and recall-by-phenotype sub-studies. Further details about the study can be accessed in the Cohort Profile Paper[1], with additional information about our COVID-19 Focus available as a Data Note Paper[2]. [1] Catherine John, Nicola F Reeve, Robert C Free, [. . .] Edward J Hollox, Louise V Wain, Martin D Tobin. Cohort Profile: Extended Cohort for E-health, Environment and DNA (EXCEED). International Journal of Epidemiology, Volume 48, Issue 3, June 2019, Pages 678–679j, https://doi.org/10.1093/ije/dyz073 [2] Lee PH, Guyatt AL, John C et al. Extended Cohort for E-health, Environment and DNA (EXCEED) COVID-19 focus [version 1; peer review: awaiting peer review]. Wellcome Open Res 2021, 6:349, https://doi.org/10.12688/wellcomeopenres.17437.1
Acknowledgements	EXCEED is funded by the University of Leicester, the NIHR Leicester Respiratory Biomedical Research Centre, the NIHR Clinical Research Network East Midlands, the Medical Research Council (grant G0902313) the Wellcome Trust (grant 202849) and HDR UK BREATHE- Health Data Research Hub for Respiratory Health (grant MC-PC_19004). EXCEED gratefully acknowledges the support of all participants and staff who have contributed to the study.
Ethics	The study is led by the University of Leicester, in partnership with University Hospitals of Leicester NHS Trust and in collaboration with Leicestershire Partnership NHS Trust, local general practices and smoking cessation services. Ethical approval for the study was obtained from the Leicester Central Research Ethics Committee (13/EM/0226).
Further information	www.exceed.org.uk/research

Fenland Study	
Description of Study Population (including citations and references if required)	The Fenland Study is a population-based detailed quantitative metabolic trait cohort study designed to investigate the interplay between environmental, behavioural and genetic factors in determining the risk of obesity, type 2 diabetes, and related metabolic disorders. At baseline from 2004-15, 12,435 participants were recruited from an age-sex sampling frame from GP practices in the Fenland area and completed extensive metabolic phenotyping. From 2015 to 2020, study participants were invited for a follow up study and nearly 8,000 repeated the same measurements. In 2020/21, 4,000 participants in the Fenland cohort were recruited to a COVID-19 remote monitoring study in which they provided blood samples at home at 3-monthly intervals that were used to quantify COVID-19 antibodies. Information about symptoms and signs, collected using objective methods, were ascertained via an App or through completion of online questionnaires.
Acknowledgements	The study is supported by the Medical Research Council (grant MC_UU_00006/1). We are grateful to the Fenland Study participants for their willingness and time to take part. We thank all members of the following teams responsible for practical aspects of the study; Study Coordination, Field Epidemiology, Anthropometry Team, Physical Activity Technical Team, IT and Data Management.
Ethics	Cambridge East (ref 04/Q0108/19).
Further information	https://www.mrc-epid.cam.ac.uk/research/studies/fenland/information-for-researchers/

Generation Scotland	
Description of Study Population (including citations and references if required)	The Generation Scotland Scottish Family Health Study has 24,000 adult volunteers recruited in Scotland 2006-2011, with consent for linkage to medical records and recontact for further studies[1]. Smith, B. H., Campbell, A., Linksted, P., Fitzpatrick, B., Jackson, C., Kerr, S. M., … Morris, A. D. (2013). Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. International Journal of Epidemiology, 42(3), 689-700. http://doi.org/10.1093/ije/dys084
Acknowledgements	Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006] and is currently supported by the Wellcome Trust [216767/Z/19/Z].
Ethics	Generation Scotland obtained Research Tissue Bank approval from the East of Scotland Research Ethics Service (on behalf of NHS Scotland). Reference number 20/ES/0021.
Further information	www.generationscotland.org

GLAD: Genetic Links to Anxiety and Depression Study dataset including the Eating Disorders Genetics Initiative (EDGI) and COVID-19 Psychiatry and Neurological Genetics Study (COPING) study	
Description of Study Population (including citations and references if required)	The Genetic Links to Anxiety and Depression (GLAD) Study is an NIHR BioResource funded project assessing the genetic and environmental links to anxiety and depression (N∼35,000; Davies et al. 2019). The Covid-19 Psychiatry and Neurological Genetics Study (COPING) was set up as a COVID-19 specific study investigating the mental health impact of individuals living in the UK (N=30,450). COPING participants were recruited from the existing GLAD, EDGI and NIHR BioResource studies to provide pandemic relevant data (Young et al. 2021). The Eating Disorders Genetics Initiative (EDGI) study is an NIHR BioResource funded project assessing the genetic and environmental links to all eating disorders (N∼5,000; Monssen et al. 2023). **References** Davies, M. R., Kalsi, G., Armour, C., Jones, I. R., McIntosh, A. M., Smith, D. J., & Breen, G. (2019). The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. *Behaviour Research and Therapy*, *123*, 103503. Young, K. S., Purves, K. L., Hübel, C., Davies, M. R., Thompson, K. N., Bristow, S., & Breen, G. (2021). Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK. Monssen, D., Davies, H. L., Kakar, S., Bristow, S., Curzons, S. C., Davies, M. R., & Breen, G. (2023). The United Kingdom Eating Disorders Genetics Initiative. *International Journal of Eating Disorders*.
Acknowledgements	The NIHR BioResource Centre Maudsley and the NIHR BioResource should be acknowledged in all publications resulting from any study that we have supported. Please use the following wording: We thank the GLAD/EDGI/COPING and NIHR BioResource volunteers for their participation, and gratefully acknowledge NIHR BioResource centres, NHS Trusts and staff for their contribution. We thank the National Institute for Health Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme. This work was supported by the National Institute of Health Research (NIHR) BioResource Centre Maudsley and is part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. Patient and public involvement groups and services were provided by the NIHR KCL-Maudsley Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.
Ethics	GLAD – 18/LO/1218 – London Fulham REC EDGI – 19/LO/1254 – London Fulham REC COPING – 20/SW/0078 – South West Central Bristol REC
Further information	https://gladstudy.org.uk (changing Autumn 2022) https://edgiuk.org

MCS: Millennium Cohort Study	
Description of Study Population (including citations and references if required)	The Millennium Cohort Study (MCS) is following the lives of young people born across England, Scotland, Wales and Northern Ireland in 2000-02. The study began with an original sample of 18,818 cohort members. The study is designed and led by the Centre for Longitudinal Studies (CLS) at University College London. The broad aim of the study is to examine the impact that circumstances and experiences at one stage of life have on outcomes and achievements in later life. Since the baseline survey at age 9 months, there have been six major ‘sweeps’ at ages 3, 5, 7, 11, 14 and 17. The next sweep, at age 22, is currently under development. Data have been collected from a number of different sources, including the cohort members and their parents and teachers. The data have been collected in a variety of ways, including via paper and electronic questionnaires, biological samples, physical measurements, tests of ability, and linked educational attainment and health records. The information collected forms a high quality data resource for scientific investigations across a full range of domains of individuals’ lives and across different points in time in them. The study has been designed to ensure comparability with other major cohort studies both in the UK and internationally and to permit the examination of links between social change and the changing experiences of different cohorts. https://www.llcsjournal.org/index.php/llcs/article/view/410/0 https://academic.oup.com/ije/article/43/6/1719/703283
Acknowledgements	MCS is core-funded by the ESRC and co-funded by a consortium of government departments.
Ethics	Ethics approval has been obtained for each follow-up from an NHS Research Ethics Committee (REC). In addition, separate REC approval is in place to cover the ongoing activities of the study in between major sweeps of data collection (i.e. Keeping in touch with and tracing cohort Members; cleaning, documenting and providing access to the data for research and linking data from administrative sources to survey data to increase the utility of the data for research.)
Further information	https://cls.ucl.ac.uk/cls-studies/mcs/

NCDS: National Child Development Study	
Description of Study Population (including citations and references if required)	The National Child Development Study (NCDS) is a continuing longitudinal study that seeks to follow the lives of all those living in Great Britain who were born in one particular week in 1958. Conducted by the Centre for Longitudinal Studies, the aim of the study is to improve understanding of the factors affecting human development over the whole lifespan. It collects information on physical and educational development, economic circumstances, employment, family life, health behaviour, wellbeing, social participation and attitudes. The broad aim of the study is to examine the impact that circumstances and experiences at one stage of life have on outcomes and achievements in later life. Since the birth survey in 1958, there have been ten 'sweeps' of all cohort members at ages 7, 11, 16, 23, 33, 42, 44/5 (a biomedical collection) 46, 50 and most recently at 55. The Age 62 Sweep is currently in the field (2022). Data have been collected from a number of different sources, including the midwife present at birth, parents of the cohort members, teachers, doctors and the cohort members themselves. The data have been collected in a variety of ways, including via paper and electronic questionnaires, clinical records, medical examinations, biological samples, physical measurements, tests of ability and educational assessments. The information collected forms a high quality data resource for scientific investigations across a full range of domains of individuals' lives and across different points in time in them. The study has been designed to ensure comparability with other major cohort studies and to permit the examination of links between social change and the changing experiences of different cohorts. https://cls.ucl.ac.uk/cls-studies/1958-national-child-development-study/
Acknowledgements	NCDS is core-funded by the ESRC.
Ethics	Ethics approval has been obtained for each follow-up from an NHS Research Ethics Committee (REC) since 2000. In addition, separate REC approval is in place to cover the ongoing activities of the study in between major sweeps of data collection (i.e. Keeping in touch with and tracing cohort members; cleaning, documenting and providing access to the data for research; and linking data from administrative sources to survey data to increase the utility of the data for research).
Further information	https://cls.ucl.ac.uk/cls-studies/ncds/

Next Steps	
Description of Study Population (including citations and references if required)	Next Steps (previously known as the Longitudinal Study of Young People in England (LSYPE1)) is a major longitudinal study that follows the lives of around 16,000 people born in 1989-90. The first seven sweeps of the study (2004-2010) were funded and managed by the Department for Education and mainly focused on the educational and early labour market experiences of young people. The study began in 2004 and included young people in Year 9 who attended state and independent schools in England. Following the initial survey at age 13-14, the cohort members were interviewed every year until 2010. In 2013, the management of Next Steps was transferred to the Centre for Longitudinal Studies (CLS) at the IOE, UCL's Faculty of Education and Society. The first sweep conducted by CLS aimed to find out how the lives of the cohort members had turned out at age 25. It maintained the strong focus on education, but the content was broadened to become a more multi-disciplinary research resource. The Age 32 Sweep is currently in the field (2022). https://doc.ukdataservice.ac.uk/doc/5545/mrdoc/pdf/next_steps_userguide_to_the_redeposit_of_sweeps_1to7_may2020.pdf https://doc.ukdataservice.ac.uk/doc/5545/mrdoc/pdf/nextsteps_age25_survey_user_guide_v3.pdf https://cls.ucl.ac.uk/cls-studies/next-steps/
Acknowledgements	Next Steps now is core-funded by the ESRC.
Ethics	Ethics approval is obtained for each follow-up from an NHS Research Ethics Committee (REC). In addition, separate REC approval is in place to cover the ongoing activities of the study in between major sweeps of data collection (i.e. keeping in touch with and tracing cohort members; cleaning, documenting and providing access to the data for research; and linking data from administrative sources to survey data to increase the utility of the data for research).
Further information	https://cls.ucl.ac.uk/cls-studies/next-steps/

NICOLA: Northern Ireland Cohort for the Longitudinal Study of Ageing	
Description of Study Population (including citations and references if required)	NICOLA is a large-scale longitudinal study in Northern Ireland designed to investigate ageing. Set up in 2013, the study visited households where at least one member was $\geq$50 years old and in its first wave recruited a representative sample of 8,478 men and women living in private residential accommodation in Northern Ireland. Wave 1 data collection involved four components: a computer assisted face-to-face home interview, a self-completion questionnaire, a health assessment (during which blood and urine samples were collected) and subsequently a dietary questionnaire. The participants who consented to follow up for Wave 2 (2017-2019) were again invited to take part in a face-to-face home interview and complete a self-completion questionnaire. From these participants, 5,925 did not opt-out of additional linkage and were mailed an additional self-completion Covid-19 questionnaire (3,149 completed). Neville C, Burns F, Cruise S, Scott A, O'Reilly D, Kee F, Young I. Cohort Profile: The Northern Ireland Cohort for the Longitudinal Study of Ageing (NICOLA). Int J Epidemiol. 2023 Aug 2;52(4):e211-e221. http://doi.org/10.1093/ije/dyad026. PMID: 37011634; PMCID: PMC10396407. Neville CE, Young IS, Kee F, Hogg RE, Scott A, Burns F, Woodside JV, McGuinness B. Northern Ireland Cohort for the Longitudinal Study of Ageing (NICOLA): health assessment protocol, participant profile and patterns of participation. BMC Public Health. 2023 Mar 10;23(1):466. http://doi.org/10.1186/s12889-023-15355-x. PMID: 36899371; PMCID: PMC9999338.
Acknowledgements	We are grateful to all the participants of the NICOLA Study, and the whole NICOLA team. The study has received funding from The Atlantic Philanthropies, the Economic and Social Research Council, the UKCRC Centre of Excellence for Public Health Northern Ireland, the Centre for Ageing Research and Development in Ireland, the Office of the First Minister and Deputy First Minister, HSC Research and Development Division of the Public Health Agency, the Wellcome Trust/Wolfson Foundation and Queen's University Belfast which provide core financial support for NICOLA. The authors alone are responsible for the interpretation of the data and any views or opinions presented are solely those of the authors and do not necessarily represent those of the NICOLA Study team.
Ethics	Ethical approval for NICOLA was obtained from the School of Medicine, Dentistry and Biomedical Sciences Ethics Committee, Queen's University Belfast.
Further information	https://www.qub.ac.uk/sites/NICOLA/InformationforResearchers/

NIHR BioResource: National Institute of Health Research BioResource COVID-19 Psychiatry and Neurological Genetics (COPING) Study

Description of Study Population (including citations and references if required)	The NIHR BioResource is a recallable resource of over 200,000 volunteers from the general population, and patients with rare and common diseases. Participants provide information about their health and lifestyle, together with biological samples, including DNA, and consent for access to their health records and for re-contact. The BioResource is one of four key infrastructures supporting population level genomic projects in the UK Life Science Industrial Strategy. Key unique features of the NIHR BioResource are its focus on recall of participants for experimental medicine studies by genotype and/or phenotype, and the inclusion of both healthy volunteers and patients with common and rare diseases. Participants represented in UK LLC are respondents to an online recall study, the Covid-19 Psychiatry and Neurological Genetics (COPING) study[1]. [1] https://www.maudsleybrc.nihr.ac.uk/posts/2020/may/covid-19-psychiatry-and-neurological-genetics-coping-study/
Acknowledgements	We thank NIHR BioResource volunteers for their participation, and gratefully acknowledge NIHR BioResource centres, NHS Trusts and staff for their contribution. We thank the National Institute for Health and Care Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. BioResource is funded by the NIHR (NIHR203312 and RG94028).
Ethics	Research Tissue Bank (REC REF: 17/EE/0025).
Further information	https://bioresource.nihr.ac.uk/using-our-bioresource/academic-and-clinical-researchers/apply-for-bioresource-data/

NSHD: Medical Research Council National Survey of Health and Development

Description of Study Population (including citations and references if required)	The MRC National Survey of Health and Development (NSHD) is a socially stratified birth cohort of 2,547 women and 2,815 men. It is a sample of all births in England, Scotland, and Wales that occurred in one week in 1946, and consists of all singleton births to married women with a husband in non-manual and agricultural employment and 1 in 4 of all comparable births to women with a husband in manual employment.[1] The study members have been followed up in over 25 data collections. Regular interviews with the mothers were conducted by health visitors, with additional assessments by school doctors and teachers. In adult life, research nurses conducted home visits at ages 26, 36, 43, 53 and 69, a detailed clinic visit took place between ages 60-64, as well as clinical sub studies focusing on the heart (Myofit46) and brain (Insight46). [1] Kuh et al. Cohort profile: updating the cohort profile for the MRC National Survey of Health and Development: a new clinic-based data collection for ageing research. Int J Epidemiol. 2011 Feb;40(1):e1-9. http://doi.org/10.1093/ije/dyq231.
Acknowledgements	The UK Medical Research Council provides core funding for the MRC National Survey of Health and Development (MC_UU_00019/1; MR/Y014022/1). We are extremely grateful to the NSHD study members for their lifelong participation and continuing support; and to past and present members of the study teams, who helped to collect and process the data.
Ethics	Ethical approval for the study was obtained from the UK Research Ethics Committee (REC).
Further information	https://skylark.ucl.ac.uk/

SABRE: Southall and Brent Revisited

Description of Study Population (including citations and references if required)	SABRE (Southall and Brent Revisited) is a population-based cohort study consisting of White British people and first-generation migrants of South Asian or African Caribbean heritage recruited from West London[1] The focus of the study is on health in ageing, including health of the heart and cardiovascular system, physical and brain function, metabolic health (for example diabetes). The study examines how mid-life health, lifestyle and social factors affect health in older age and aims to improve understanding of the reasons underlying ethnic group differences in health. Participants were aged 40-69 years at baseline (1988-1991) and comprised 2,346 people of European, 1,710 people of South Asian and 801 people of African Caribbean ethnicity respectively. A total of 3,571 participants were followed up in the second wave (2008-2011) and 1,412 participants (including partners of index cases and a booster sample of people of African Caribbean ethnicity) were seen in clinic during the third wave (2014-2018). [1] Jones S et al. Cohort Profile Update: Southall and Brent Revisited (SABRE) study: a UK population-based comparison of cardiovascular disease and diabetes in people of European, South Asian and African Caribbean heritage. Int J Epidemiol. 2020 Oct 1;49(5):1441-1442e. http://doi.org/10.1093/ije/dyaa135.
Acknowledgements	The study was funded at baseline by the Medical Research Council, Diabetes UK, and the British Heart Foundation, and at follow-up by the Wellcome Trust, the British Heart Foundation and NIHR. We are extremely grateful to all the people who took part in the study, and past and present members of the SABRE team who helped to collect and analyse the data.
Ethics	Ethical approval for the study was obtained from the UK Research Ethics Committee (REC).
Further information	Data sharing applications are welcome. Please contact mrclha.swiftinfo@ucl.ac.uk with an outline of proposed analyses or query.

TEDS: Twins Early Development Study	
Description of Study Population (including citations and references if required)	**TEDS** The Twins Early Development Study (TEDS) is a population-based sample of twins born in England and Wales between January 1994 and December 1996. Families were identified using electronic birth records and invited to take part though the Office of National Statistics (ONS). Of the families contacted, 13,759 parents registered their twins in the study. TEDS parents, twins and teachers have provided reports on a range of cognitive, behavioural and emotional measures at regular intervals throughout the study (A full list can be found in the TEDS data dictionary: https://www.teds.ac.uk/datadictionary/home.htm). Lockhart, C., Ahmadzadeh, Y., Breen, G., Bright, J., Bristow, S., Boyd, A., Downs, J., Hotopf, M., Palaiologou, E., Rimfeld, K., Maxwell, J., Malanchini, M., McAdams, T. A., Plomin, R., & **Eley, T. C.** (2023). Twins Early Development Study (TEDS): A genetically sensitive investigation of mental health outcomes in the mid-twenties. *JCPP Advances, 3*. http://dx.doi.org/10.1002/jcv2.12154; PMC10519737 **E-Risk** The Environmental Risk (E-Risk) Longitudinal Twin Study is a sub-study of TEDS involving 1,116 of the TEDS families with same-sex twins selected based on residential location throughout England and Wales and mother's age. These 2,232 twins were assessed via home visits initially at age 5 and then ages 7, 10, 12 and 18 (participation rates were 98%, 96%, 96% and 93%, respectively). These comprehensive assessments covered behaviour, cognition, mental and physical health, victimisation, family and neighbourhood environment, plus collection of biological samples. The E-Risk twins are currently being assessed in online interviews at age 30. Moffitt, T.E., & E-Risk Study Team (2002). Teen-aged mothers in contemporary Britain. *Journal of Child Psychology & Psychiatry 43*(6), 727-742.
Acknowledgements	**TEDS** We are very grateful to the TEDS twins and their families for all of their time and effort given to our research over the years. TEDS has been funded by the UK Medical Research Council (MRC) with 6 consecutive programme grants, the most recent being awarded to T.C. Eley (current grant ref: MR/V012878/1; previously: MR/M021475/1, G0901245, G0500079, G9424799, and G9424799). **E-Risk** We are extremely grateful to the E-Risk study mothers and fathers, the twins, and the twins' teachers and neighbours for their participation. Our thanks to the E-Risk team for their dedication, hard work, and insights. The E-Risk Study is funded by the Medical Research Council [G1002190; MR/X010791/1]. Additional support was provided by the National Institute of Child Health and Human Development [HD077482], the Jacobs Foundation, and the Economic and Social Research Council Centre for Society and Mental Health at King's College London [ES/S012567/1].
Ethics	**TEDS** TEDS has been granted ethical approval by the King's College London Ethics Committee (Ref: PNM/09/10–104; Ref: HR/DP-20/21-22060). **E-Risk** The Joint South London and Maudsley and the Institute of Psychiatry Research Ethics Committee approved phases 5-18 of the E-Risk study (1997/122) and King's College London Ethics Committee approved phase 30 (HR/DP-23/24-39753).
Further information	**TEDS** https://www.teds.ac.uk/researchers/teds-data-access-policy **E-Risk** https://eriskstudy.com/data-access/

TRACK-COVID Study

Description of Study Population (including citations and references if required)	The study employed a two-stage design. During the first stage of this effort, ~90,000 individuals previously recruited into the INTERVAL[1], COMPARE[2] and STRIDES[3] studies (i.e., National Blood Donor Studies[4]) were invited via email to participate and to provide Covid-19-related information using an online questionnaire. During the second stage, participants were asked to provide self-collected finger-prick capillary blood sample every 6 weeks over a period of 18 months. [1] Di Angelantonio E, Thompson SG, Kaptoge SK, Moore C, Walker M, Armitage J, Ouwehand WH, Roberts DJ, Danesh J, INTERVAL Trial Group. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. Lancet. 2017 Nov 25;390(10110):2360-2371. [2] Bell S, Sweeting M, Ramond A, Chung R, Kaptoge S, Walker M, Bolton T, Sambrook J, Moore C, McMahon A, Fahle S, Cullen D, Mehenny S, Wood AM, Armitage J, Ouwehand WO, Miflin G, Roberts DJ, Danesh J, Di Angelantonio E, COMPARE Study Group. Comparison of four methods to measure haemoglobin concentrations in whole blood donors (COMPARE): A diagnostic accuracy study. Transfus Med. 2020 Dec 20. [3] McMahon A, Kaptoge S, Walker M, Mehenny S, Gilchrist PT, Sambrook J, Akhtar N, Sweeting M, Wood AM, Stirrups K, Chung R, Fahle S, Johnson E, Cullen D, Godfrey R, Duthie S, Allen L, Harvey P, Berkson M, Allen E, Watkins NA, Bradley JR, Kingston N, Miflin G, Armitage J, Roberts DJ, Danesh J, Di Angelantonio E. Evaluation of interventions to prevent vasovagal reactions among whole blood donors: rationale and design of a large cluster randomised trial. Trials. 2023 Aug 10;24(1):512. http://doi.org/10.1186/s13063-023-07473-z. [4] http://www.donorhealth-btru.nihr.ac.uk/
Acknowledgements	The TRACK-COVID study recruited participants from the STRIDES Bioresource, INTERVAL and COMPARE trials and the academic coordinating centre would like to thank blood donor centre staff and blood donors for their participation. The academic coordinating centre at the Department of Public Health and Primary Care for INTERVAL, COMPARE, STRIDES BioResource and TRACK-COVID was supported by core funding from the: NIHR BTRU in Donor Health and Genomics (NIHR BTRU-2014-10024), NIHR BTRU in Donor Health and Behaviour (NIHR203337), UK Medical Research Council (MR/L003120/1), British Heart Foundation (SP/09/002; RG/13/13/30194; RG/18/13/33946) and NIHR Cambridge BRC (BRC-1215-20014; NIHR203312) [*]. We thank NIHR BioResource volunteers for their participation, and gratefully acknowledge NIHR BioResource centres, NHS Trusts and staff for their contribution. We thank the NIHR, NHS Blood and Transplant (NHSBT) and Health Data Research (HDR) UK as part of the Digital Innovation Hub Programme. We thank Leeds Teaching Hospitals NHS Foundation Trust for their contribution to the SARS-COV-2 antibody analysis. Participants in the COMPARE study were recruited with the active collaboration of NHSBT (www.nhsbt.nhs.uk). Funding was provided by NHSBT and the NIHR BTRU in Donor Health and Genomics (NIHR BTRU-2014-10024). DNA extraction and genotyping were co-funded by the NIHR BTRU and the NIHR BioResource (http://bioresource.nihr.ac.uk). A complete list of the investigators and contributors to the COMPARE study is provided in reference [**]. Participants in the STRIDES BioResource study were recruited with the active collaboration of NHSBT (www.nhsbt.nhs.uk). Funding was provided by NHSBT (17-01-GEN) and the NIHR BTRU in Donor Health and Genomics (NIHR BTRU-2014-10024). A complete list of the investigators and contributors to the STRIDES BioResource trial is provided in reference [***]. Participants in the INTERVAL randomised controlled trial were recruited with the active collaboration of NHSBT, which has supported fieldwork and other elements of the trial. DNA extraction and genotyping were co-funded by the NIHR, the NIHR BioResource (http://bioresource.nihr.ac.uk) and the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014) [*]. A complete list of the investigators and contributors to the INTERVAL trial is provided in reference [****]. Persons from the University of Cambridge academic coordinating centre were funded by the NIHR Blood and Transplant Research Unit (BTRU) in Donor Health and Genomics (NIHR BTRU-2014–10024) and are now funded by the NIHR BTRU in Donor Health and Behaviour (NIHR203337) and the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014; NIHR203312) [*].

	John Danesh holds a British Heart Foundation Professorship and a NIHR Senior Investigator Award This work was supported by core funding from the: British Heart Foundation (RG/18/13/33946: RG/F/23/110103), NIHR Cambridge Biomedical Research Centre (NIHR203312), Cambridge BHF Centre of Research Excellence (RE/18/1/34212) and BHF Chair Award (CH/12/2/29428), and by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome. This work was supported by Health Data Research UK (Molecules to Health Records programme), which is funded by the Medical Research Council (UKRI), the National Institute for Health Research, the British Heart Foundation, Cancer Research UK, the Economic and Social Research Council (UKRI), the Engineering and Physical Sciences Research Council (UKRI), Health and Care Research Wales, Chief Scientist Office of the Scottish Government Health and Social Care Directorates, and Health and Social Care Research and Development Division (Public Health Agency, Northern Ireland). *The views expressed are those of the authors and not necessarily those of the NIHR, NHSBT or the Department of Health and Social Care. **Bell S, Sweeting M, Ramond A, Chung R, Kaptoge S, Walker M, Bolton T, Sambrook J, Moore C, McMahon A, Fahle S, Cullen D, Mehenny S, Wood AM, Armitage J, Ouwehand WO, Miflin G, Roberts DJ, Danesh J, Di Angelantonio E, COMPARE Study Group. Comparison of four methods to measure haemoglobin concentrations in whole blood donors (COMPARE): A diagnostic accuracy study. Transfus Med. 2021 Apr;31(2):94-103. *** McMahon A, Kaptoge S, Walker M, Mehenny S, Gilchrist PT, Sambrook J, Akhtar N, Sweeting M, Wood AM, Stirrups K, Chung R, Fahle S, Johnson E, Cullen D, Godfrey R, Duthie S, Allen L, Harvey P, Berkson M, Allen E, Watkins NA, Bradley JR, Kingston N, Miflin G, Armitage J, Roberts DJ, Danesh J, Di Angelantonio E. Evaluation of interventions to prevent vasovagal reactions among whole blood donors: rationale and design of a large cluster randomised trial. Trials. 2023 Aug 10;24(1):512. ****Di Angelantonio E, Thompson SG, Kaptoge SK, Moore C, Walker M, Armitage J, Ouwehand WH, Roberts DJ, Danesh J, INTERVAL Trial Group. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. Lancet. 2017 Nov 25;390(10110):2360-2371.
Ethics	20/EM/0121 - East Midlands - Nottingham 2 Research Ethics Committee
Further information	http://www.donorhealth-btru.nihr.ac.uk/

TwinsUK	
Description of Study Population (including citations and references if required)	TwinsUK is the largest adult twin registry in the UK and the most clinically detailed in the world. The national, population-based study was founded in 1992 and aims to investigate the genetic and environmental basis of a range of complex diseases and conditions. TwinsUK currently consists of over 15,700 volunteer adult twins (both monozygotic and dizygotic) who are between 18 to 104 years of age from around the UK (mean age 59)[1]. The cohort is predominantly female, and disease prevalence is broadly reflective of the UK population. Over 750,000 biological samples and extensive phenotypes have been collected longitudinally over 30 years. [1] Verdi S, Abbasian G, Bowyer RCE, et al.: TwinsUK: The UK Adult Twin Registry Update. Twin Res Hum Genet. 2019; 22(6): 523–529.
Acknowledgements	We thank TwinsUK members for their participation and the TwinsUK operations team for coordinating and undertaking twin clinic visits and data and sample collections. TwinsUK is funded by the Medical Research Council (MRC), Wellcome LEAP, Wellcome Trust, EPSRC, BBSRC, Versus Arthritis, European Commission, Chronic Disease Research Foundation (CDRF), Zoe Ltd, the National Institute for Health and Care Research (NIHR) Clinical Research Network (CRN) and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.
Ethics	All collections of TwinsUK data have received ethical approval associated with TwinsUK Biobank (19/NW/0187), TwinsUK (EC04/015) or Healthy Ageing Twin Study (H.A.T.S) (07/H0802/84) from NHS Research Ethics Committees. Linkage to health and environmental records is also covered by approval from the Health Research Authority (19/CAG/0223).
Further information	https://twinsuk.ac.uk/resources-for-researchers/access-our-data/

UK-REACH: The United Kingdom Research study into Ethnicity And COVID-19 outcomes in Healthcare workers	
Description of Study Population (including citations and references if required)	The UK-REACH cohort was established to understand why ethnic minority healthcare workers (HCWs) are at risk of poorer outcomes from COVID-19 when compared with their White ethnic counterparts in the UK.[1] The cohort comprises 17,891 HCWs aged 16–89 years (mean age 44 years). Online questionnaires include information on demographics, COVID-19 exposures at work and home, redeployment in the workforce due to COVID-19, mental health measures, workforce attrition and opinions on COVID-19 vaccines. [1] Luke Bryant, Robert C Free, Katherine Woolf, et al. Cohort Profile: The United Kingdom Research study into Ethnicity and COVID-19 outcomes in Healthcare workers (UK-REACH). Int J Epidemiol, Volume 52, Issue 1, February 2023, Pages e38–e45. http://doi.org/10.1093/ije/dyac171
Acknowledgements	UK-REACH is supported by a grant to the University of Leicester from the MRC-UK Research and Innovation, and National Institute for Health Research (NIHR) rapid response panel to tackle COVID-19 and by core funding provided by NIHR Leicester Biomedical Research Centre – a partnership between the University of Leicester and University Hospitals of Leicester NHS Trust.
Ethics	UK-REACH has been granted ethical approval by the London – Brighton & Sussex Research Ethics Committee (Ref: 20/HRA/4718).
Further information	https://uk-reach.org/main/data_sharing

Understanding Society – the UK Household Longitudinal Study	
Description of Study Population (including citations and references if required)	Understanding Society, the UK Household Longitudinal Study, is a longitudinal survey of the members of ∼40,000 households (at Wave 1, 2009-10) in the United Kingdom. The survey sample consists of a large General Population Sample (∼26,000 households) plus three other components: the Ethnic Minority Boost Sample (∼4,000 households), the former British Household Panel Survey sample (∼8,000 households) and the Immigrant and Ethnic Minority Boost Sample (∼2,900 households, added at Wave 6). Household and individual interviews are conducted annually. The study is multi-topic and multi-purpose. From April 2020 to September 2021, participants from the main Understanding Society sample were asked to complete nine short web-surveys (with a telephone option in some months). The COVID-19 study covered the changing impact of the pandemic on the welfare of UK individuals, families and wider communities. ∼18,000 individuals provided a full or partial interview at Wave 1 (April 2020). At Wave 8 of the COVID-19 study, 8477 participants provided consent to link their survey data to administrative health records.
Acknowledgements	Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public The COVID-19 study (2020-2021) was funded by the Economic and Social Research Council and the Health Foundation. Serology testing was funded by the COVID-19 Longitudinal Health and Wealth – National Core Study. Fieldwork for the web survey was carried out by Ipsos MORI and for the telephone survey by Kantar.
Ethics	The University of Essex Ethics Committee has approved all data collection on Understanding Society main study, COVID-19 surveys and innovation panel waves, including asking consent for all data linkages except to health records. Approval for asking consent for health record linkage and for the collection of blood and subsequent serology testing in the March 2021 wave of the COVID-19 study was obtained from London – City & East Research Ethics Committee (21/HRA/0644).
Further information	https://ukllc.ac.uk https://ukdataservice.ac.uk