



## DATA LINKAGE AT TWINSUK

HELP US SHAPE A HEALTHY FUTURE



Do you need this booklet in another format like large print?  
If you do, please let us know.

## WHAT DO YOU NEED TO DO?

This booklet contains important information about the TwinsUK data linkage programme. Please read through the booklet and get in touch with us if you have any questions, using the contact details at the end of the booklet.

We will send you a Data Linkage Decision Form, either online or in the post. Please tell us if you don't want us to link to any or all of your health, education and environmental records by completing and returning the Data Linkage Decision Form.

If you choose not to return the Data Linkage Decision Form, we have approval (see page 18) to proceed to link your records. The next time we see you for a clinic visit we will ask you again directly face-to-face.

**We will always respect your decision.**

If you change your mind at any time please contact us to tell us so. We can remove linked data at any point.

## WHAT IS DATA LINKAGE?

TwinsUK is looking to the future and finding new ways to use data for breakthrough scientific discoveries in health and disease. Good research needs to use the highest quality data and so we are always looking for ways to improve data collection. Data linkage, which accurately and confidentially collects information from your official records, is one way of improving the quality of the data we use in our research.

This booklet contains more detail about data linkage, including the types of records we are linking to, the benefits to you, the benefits to the research we do, and how your data will be protected. You have the chance to decide which, if any, of your data you are happy for us to link to.

## WHY IS TWINSUK LINKING DATA FROM OFFICIAL RECORDS?

Records linkage is a key strategy for current health research and is part of a recent initiative by the Wellcome Trust which funds TwinsUK Core Resource.

It's a non-intrusive way for researchers to follow up a cohort of individuals confidentially over time.

Over the years, some individuals may find it difficult to continue actively taking part in projects at TwinsUK. This may lead to changes in the cohort, and those who do continue to actively take part may not be representative of the group overall.

Linkage allows us to take account of these changes in the cohort over time. Without it we can't be sure whether our research findings apply to everyone.

Linking to data held in official records is also important because it allows us to check whether the results of large-scale records-based research, like that carried out by TwinsUK, make sense and can be applied to people more broadly. This means that the public can benefit from health research like that being carried out by TwinsUK.

## WHAT ARE THE BENEFITS?

1. Some may find it difficult to take part in projects at TwinsUK. Data linkage is an easier, effortless way to be involved in vital research.
2. Even if you have been very active in TwinsUK and attended many visits, there may still be some 'gaps' of information in what we know about you. Data linkage can help fill these and create a bigger picture so that we can accurately carry out our research.
3. Some of the most important information about us is not something we necessarily know or remember. Data linkage helps to collect this information. For example, the precise type and dose of antibiotics you may have been given in the past.
4. When small amounts of data from individuals are linked up and pooled together, researchers can look for patterns in the data. This helps them to develop new ways of predicting or diagnosing illness and identifying ways to improve care. The information can be used to help understand more about disease risks and causes, improve diagnoses, develop new treatments and prevent disease.
5. Data linkage enables you to play an important part in TwinsUK and the future of health research.

The next pages feature case studies of how data linkage has been used by others to carry out health research and examples of how TwinsUK would benefit from data linkage.

## CASE STUDY



### How data linkage has been used by others

Clopidogrel is a drug that is recommended for people who have recently had a heart attack. However, it is difficult to know whether people continue to take this for the recommended length of time after they are discharged from hospital and, if they choose to stop taking it, what happens to them.

Researchers linked together three sets of data to try to understand more about what happens to people once they are discharged from hospital, after receiving treatment for a heart attack. They linked records of someone's treatment and prescriptions in hospital with GP data showing what happened to them after discharge. They then linked these to national records of individual deaths.

The research demonstrated that stopping taking clopidogrel is common, and that this is associated with an increased number of heart attacks and death. However, while the research demonstrated an association, it could not prove a causal link. Future research will continue to investigate this.

You can find out more by visiting [understandingpatientdata.org.uk](http://understandingpatientdata.org.uk)

## CASE STUDY



### How data linkage could improve TwinsUK research methods

Mrs Jones is a twin in our study. One morning she received an email asking her to complete an online questionnaire from TwinsUK about antibiotic use in her life so far. The study was looking at the effect of antibiotic use over our lives on our gut microbes and on health conditions such as obesity and type 2 diabetes. Like most people, Mrs Jones had not kept notes on her antibiotic use throughout her life.

With data linkage, we would be able to see when exactly Mrs Jones has been prescribed antibiotics, which type, and for how long, much more easily and accurately. This would give us better quality data for our research study, and without needing to disturb Mrs Jones.

# CASE STUDY

## How data linkage could benefit TwinsUK research

Twins have joined the TwinsUK registry at different points since it was set up in 1992. Even if twins attended all of their twin visits, we would only have a few weight and blood pressure measurements for them since they joined the TwinsUK registry.

By linking to other health records however like those held by a twin's GP, we would be able to see more weight and blood pressure measurements taken throughout their lifetime. This would help us fill in any gaps in our data from twins who joined TwinsUK later on, or those who are no longer able to attend visits in person. We would also be able to look further back to weight and blood pressure measurements from before the twins joined TwinsUK. This would give us a lot of valuable information for our research.

TwinsUK Data Alone										
Do we have weight and blood pressure information?										
	1970-74	1975-79	1980-84	1985-89	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
Participant 1					✓	✓	✓	✓	✓	✓
Participant 2					✓			✓		
Participant 3						✓	✓			✓
Participant 4							✓			
Participant 5								✓	✓	✓
Participant 6									✓	✓

TwinsUK Data With Linked Data										
Do we have weight and blood pressure information?										
	1970-74	1975-79	1980-84	1985-89	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
Participant 1	✓			✓	✓	✓	✓	✓	✓	✓
Participant 2		✓	✓	✓	✓		✓	✓	✓	✓
Participant 3			✓	✓		✓	✓	✓	✓	✓
Participant 4	✓	✓	✓		✓	✓	✓	✓		✓
Participant 5				✓	✓	✓		✓	✓	✓
Participant 6			✓		✓		✓	✓	✓	✓

## YOUR HEALTH RECORDS

Unless you have opted out online or through your GP, your health records will already be being used to plan healthcare nationally and for healthcare research. Details on this national data use can be found on page 16 of this booklet. Health Records Linkage is all about existing studies, such as TwinsUK, being able to link into this data.

TwinsUK aims to discover the causes of important health problems and the best way to treat and prevent them. Much of the information that is most useful to TwinsUK researchers is held in your official health records. Health records include those held by your general practitioner (GP) and hospital records, mental health data, cancer registrations and mortality data held by NHS Digital. We may also request hospital records held by NHS Wales Informatics Service.

We are bound by the same rules of confidentiality as your GP – your name is never linked to the information taken from your records when used by researchers. Such records may contain information some people find particularly sensitive to disclose openly, for example about mental health or sexual health issues.

Your patient health information includes:

- Details of visits to your doctor, diagnoses made and any treatment you were given.
- If you've ever been to hospital; why you were there and what happened whilst you were there. This includes visits to accident & emergency, if you were admitted to hospital for care or if you received treatment as an outpatient (e.g. to see a specialist consultant or nurse).
- Details of community care you may receive, whether it is for a particular illness or for reasons to do with your mental health.
- Specific information about you, such as your ethnic group.



## YOUR EDUCATION RECORDS

Socioeconomic factors, including education and care you may have received, can tell us how your environment may affect your future. There is a strong link between educational attainment and health. Our questionnaire data is often incomplete and lacks detail. Linkage to your education records held by the Department for Education can improve the data we have on your educational background.

Your education records include:

- Your exam results.
- Information about your attendance.
- Information about support and care you received e.g. if you have a special educational need.
- Information about whether you received free school meals.
- Specific information about you such as your ethnic group or the languages spoken in your home.





## YOUR ENVIRONMENTAL DATA – NEIGHBOURHOOD AND THE NATURAL ENVIRONMENT

Your surrounding environment can influence your health and wellbeing. Information about your neighbourhood and environment is derived from **your postcode only**, from organisations such as the Office of National Statistics (ONS) and the Centre for Hydrology & Ecology (CHE). TwinsUK aims to understand how your neighbourhood and your environment impacts on your health and wellbeing.

Natural environment and neighbourhood information includes:

- Information about your community, such as crime figures, house prices and employment figures.
- Information about houses in your neighbourhood, such as insulation levels, house age and house price.
- Information about pollution and the type of land around your neighbourhood and where you work.
- Information about services around you, such as distance to the nearest GP, parks, libraries and other facilities.



## HOW WILL TWINSUK LINK YOUR HEALTH, EDUCATION AND ENVIRONMENTAL DATA?

TwinsUK will follow strict guidelines when linking data, which will be in accordance with the Confidentiality Advisory Group's (CAG) guidance and approved by the Health Research Authority (HRA). The CAG is an independent body which provides expert advice on the use of confidential patient information in health research without consent. The HRA ensures that research is ethically reviewed.



To accurately identify your records, we use information we have such as NHS ID numbers, date of birth, name and postcode to make an accurate link to your health, education and environmental records.



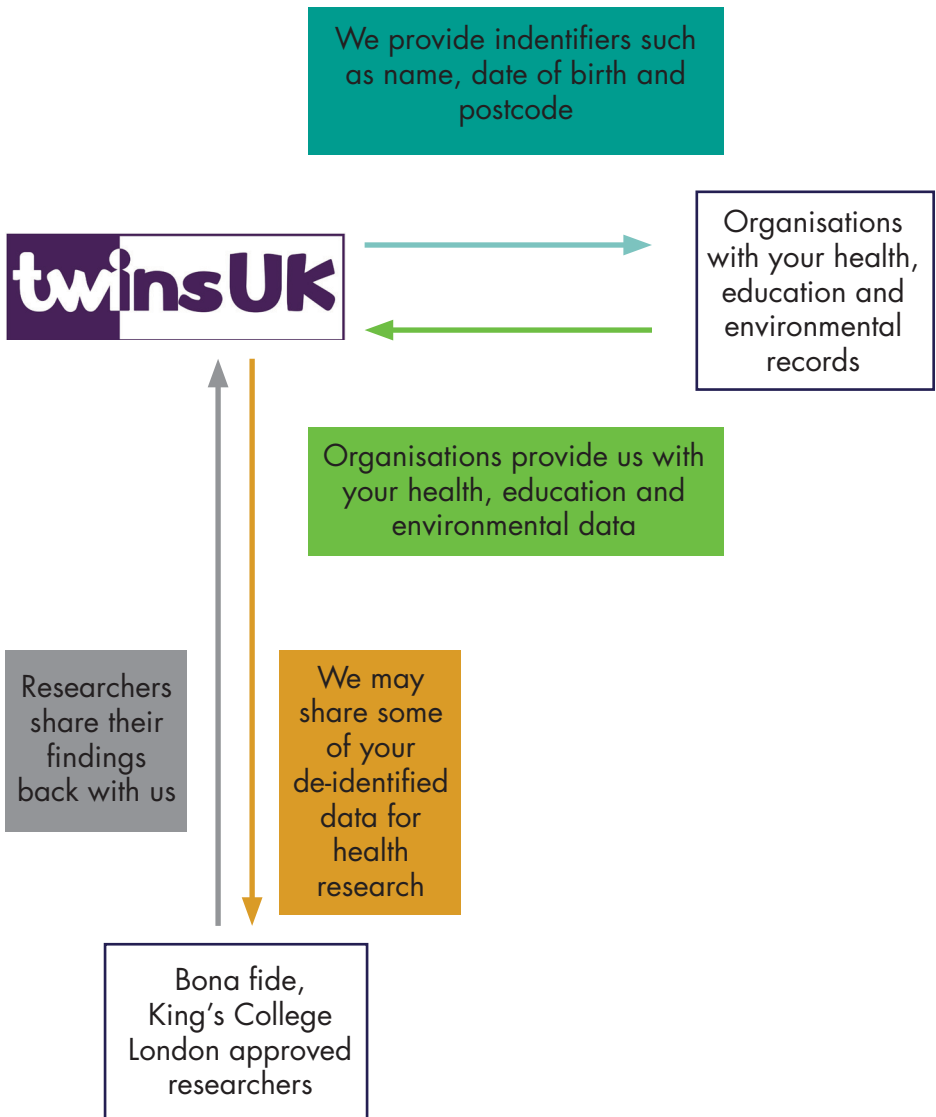
We will provide this information to organisations that hold your health, education and environmental data and they will provide us with the data they hold on your health, education and environmental records, as outlined on pages 7-9 of this booklet.



The data is imported into our database with any information that could directly identify you removed. The data is linked with your current TwinsUK pseudonymised data using a unique, de-identified number. The data is NEVER linked with any identifiable data.

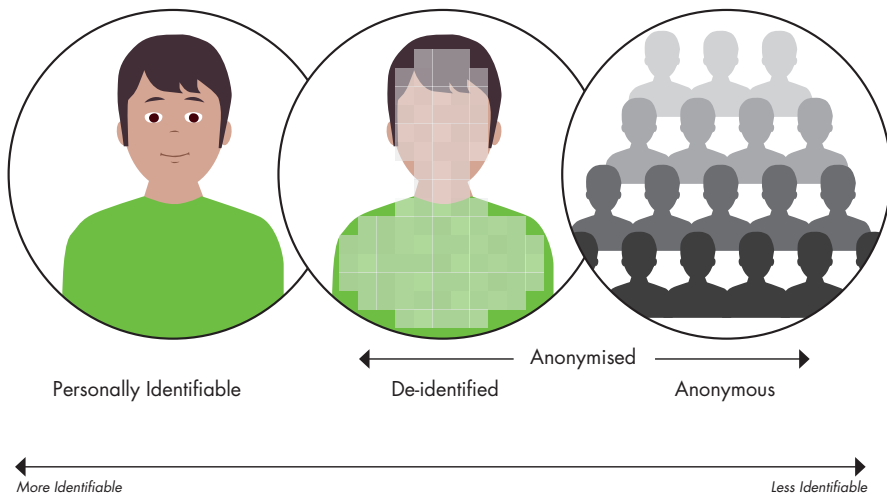


This de-identified data will be kept in a secure research environment controlled by the Department of Twin Research, King's College London. TwinsUK Research Executive Committee will review all applications from the wider research community to use this data and only bona fide researchers approved by King's College London working in the public interest will be granted access.



The figure above describes how data moves between the different organisations.

# SPECTRUM OF IDENTIFIABILITY



De-identified data has had any information that could identify the individual to which it belongs removed. This includes information such as the individual's name, date of birth and address.

Pseudonymised data is anonymous to the people who hold or receive it (e.g. a research team) but contains information or codes that would allow others (e.g. those responsible for the individual's care) to identify an individual from it.

More information about the CAG and how we link to health records can be found on the HRA website by searching "CAG": [www.hra.nhs.uk](http://www.hra.nhs.uk)

## HOW WILL TWINSUK PROTECT YOUR HEALTH, EDUCATION AND ENVIRONMENTAL DATA?

We will keep all information collected about you as part of TwinsUK strictly confidential. We will also store it according to the General Data Protection Regulation 2018, and the UK Government's data security and protection standards for health and care organisations.

We take appropriate measures to ensure that the information disclosed to us is kept secure, accurate and up to date and kept only for so long as is necessary for the purposes for which it is used.

More information about how your data is kept are available here: [www.gov.uk/government/publications/data-security-and-protection-for-health-and-care-organisations](http://www.gov.uk/government/publications/data-security-and-protection-for-health-and-care-organisations).

## WHO WILL HAVE ACCESS TO YOUR HEALTH, EDUCATION AND ENVIRONMENTAL DATA?

The information from your health, education and environmental records can be used to help understand more about diseases. But TwinsUK can't do all the data analysis on its own. We work in partnership with researchers from across the health research spectrum, including academic institutions, charities and regulators, provided the research is in the public interest and is governed by data sharing agreements. The results of the research must be made public.

Researchers from outside of TwinsUK may ask us for permission to use the data we hold. TwinsUK reviews each and every one of these requests carefully, and only bona fide researchers approved by King's College London and working for the public good are granted access. We will never share your personal data such as name or date of birth.

**We will not share your linked health, education and environment data with commercial companies.**



# THE UK LONGITUDINAL LINKAGE COLLABORATION



TwinsUK is one of many population cohort studies in the UK. As a result of the COVID-19 pandemic, we have been working closely with other cohort studies in order to advance COVID-19 research. One project is the UK Longitudinal Linkage Collaboration (UK LLC). This is part of the UK government-funded National Core Studies, set up to help defeat COVID-19 and inform public policy. For the first time, data from many UK cohort studies will be brought together into one secure research server. The data will be held securely at the University of Swansea and managed by the University of Bristol.

The UK LLC will hold COVID-19 research data on cohorts' participants and link to health records.

The UK LLC will also link to environmental records, such as air pollution levels, based on your postcode only. All participants' data held in the UK LLC will be de-identified at all stages, meaning that researchers and the UK LLC staff cannot identify anyone from the data.

Researchers doing COVID-19 research can apply to use any of the data held in the UK LLC. The TwinsUK Resource Executive Committee will review every request to use TwinsUK data and only allow access to bona fide researchers carrying out COVID-19 research for the public good.

The diagram opposite shows how the UK LLC operates.



TwinsUK will provide 'identifiers' such as name, address and NHS number to the NHS Wales Informatics Service (NWIS), a trusted third party.



NWIS will create and assign an ID number to all participants. This will help to collect your health and environmental records securely, from organisations including NHS Digital, Public Health Scotland and Health & Social Care Northern Ireland Business Development Services.



Your records and relevant research data will be combined, de-identified and sent to the UK LLC. No organisation will have both your identifiers and your combined records & research data.



Your de-identified data will be held safely within the UK LLC secure research server.



The TwinsUK Resource Executive Committee will review every request for access to TwinsUK participants' data and only approve bona fide researchers carrying out projects for the public good.

*Note: The UK LLC staff have no access to the key which links the UK LLC ID to participant identifiers. Researchers and UK LLC staff cannot identify anyone from the data held at the UK LLC.*

The UK LLC has been reviewed and approved by the Health Research Authority on the advice of the Confidentiality Advisory Group (reference 21/CAG/0044).

## HOW CAN I TAKE PART IN THE UK LLC?

Your COVID-19-relevant research data, health records and environmental records will be included in the UK LLC unless you opt out of both health AND environmental records linkage. The table below shows how TwinsUK participants will be included in the UK LLC, depending on which types of records you are happy for TwinsUK to link to.

<b>I have reviewed the Data Linkage Decision Form and...</b>	<b>Will I be included in the UK LLC?</b>	<b>Will the UK LLC receive my linked health records?</b>	<b>Will the UK LLC receive my linked environmental records?</b>
<b>I am happy for TwinsUK to link to my health AND environmental records – I have not opted out of either.</b>	Yes	Yes	Yes
<b>I have opted out of health records linkage ONLY.</b>	Yes	No	Yes
<b>I have opted out of environmental records linkage ONLY.</b>	Yes	Yes	No
<b>I have opted out of BOTH health and environmental records linkage.</b>	No	No	No

The UK LLC will not receive your education records.

If you decide you do not wish for TwinsUK to link to your environmental and your health records, we will not be able to include you in the UK LLC.



## WHAT ARE THE RISKS OF DATA LINKAGE?

The way we keep your health, education and environmental data is regulated in order to make any risks as low as possible.

The possible risks and the steps to reduce them include:

Risks	How we ensure your data is looked after
Invasion of privacy, or personal information being revealed to others	<ul style="list-style-type: none"><li>• All our data is held on secure servers. Your personal data (name, date of birth and address) is kept completely separately and is only ever accessible by a very small number of individuals in the core TwinsUK team for the purpose of contacting you and recruiting you to studies.</li><li>• Research data (for example our questionnaire and visit data and linked health and other data) DOES NOT include your name and contact details. Only a series of numbers and codes taken from a record is used in research.</li><li>• Research data can only be accessed by King's College London approved researchers with official credentials. There is a clear audit trail showing who has accessed the data and when.</li><li>• Only the minimum amount of information necessary is passed on to any researcher relevant for their specific research.</li><li>• We treat it as a potential criminal offence to misuse personal data, including trying to re-identify someone without permission.</li></ul>
Loss of control if data is passed outside the TwinsUK	<ul style="list-style-type: none"><li>• Sharing of health records data must be in accordance with stipulations by NHS Digital and are subject to audit.</li><li>• TwinsUK will never share data linked with your personally identifiable data.</li><li>• Whoever the user, there are regulations to access your data. Researchers have to sign contracts setting out what they can and cannot do with the data, including limits on passing data to other third parties. Data must be stored securely, with controlled access and robust IT systems to keep data safe, and there are strong sanctions if data is misused.</li></ul>

## WHO HAS REVIEWED AND APPROVED THE RESEARCH?

You have received this information leaflet as you signed up to TwinsUK before 2021 and we have not seen you for a clinic visit in 2021.

We have been given permission to link to your health records, and to take part in the UK Longitudinal Linkage Collaboration, by the Health Research Authority on the advice of the Confidentiality Advisory group (CAG) (references 19/CAG/0223 and 21/CAG/0044 respectively) unless you tell us otherwise using the Data Linkage Decision Form.

The function of the CAG is to provide independent advice on the use of confidential patient information in health research without consent. The Health Research Authority exists to protect and promote the interests of patients and the public in health and social care research. They make sure that research is ethically reviewed and approved.

Overall, TwinsUK has been reviewed and approved by the NHS – London Bridge Research Ethics Committee (REC reference <<EC/04/015 >>), and by Guy's and St Thomas' NHS Foundation Trust Research and Development (R&D) in 2012. TwinsUK BioBank has been reviewed and approved by the North West – Liverpool East Research Ethics Committee (REC reference <<19/NW/0187>>) in 2019.

We also have a Volunteer Advisory Panel (VAP) consisting of at least 24 members of TwinsUK, who review and offer advice on our research programme, including this data linkage programme.



## NATIONAL OPT-OUT

If you have opted out of allowing your data to be used for research or planning through your GP surgery or online, then even if you decide to agree to our linkage, we will not be able to do so. You can find out more about the National Opt-out and set your permissions here:

[www.digital.nhs.uk/services/national-data-opt-out-programme](http://www.digital.nhs.uk/services/national-data-opt-out-programme)

If you have any further questions, please get in touch using the contact details below.

## WHO CAN I CONTACT FOR FURTHER INFORMATION?

If you have any questions or concerns, please contact the TwinsUK admin team on the details below.



020 7188 5555



[twinsuk@kcl.ac.uk](mailto:twinsuk@kcl.ac.uk)



[www.twinsuk.ac.uk/twinzone](http://www.twinsuk.ac.uk/twinzone)



Department of Twin Research  
St Thomas' Hospital Campus  
4th Floor South Wing D Block  
Westminster Bridge Road  
London  
SE1 7EH

## WHO CAN I CONTACT IF I WANT TO MAKE A COMPLAINT?

Please contact the TwinsUK admin team using the contact details above. If you remain unhappy and wish to complain formally, you can do this through the Guy's and St Thomas' Patients Advice and Liaison Service (PALS) on 020 7188 8801, [pals@gstt.nhs.uk](mailto:pals@gstt.nhs.uk). The PALS team are based in the main entrance on the ground floor at St Thomas' Hospital and on the ground floor at Guy's Hospital in the Tower Wing.

**Thank you for taking the time to read this information sheet**