



Administrative Data  
Research Network

An ESRC Data  
Investment

# Ethics and administrative data

*Guidance for research ethics committees*

**Guidance for researchers**

## Contents

Introduction .....	3
Foreword.....	3
<b>1. An introduction to the Network</b> .....	<b>4</b>
What does the Network aim to do? .....	4
Why use administrative data for research? .....	4
What is a Network project? .....	4
How will the Network keep data safe / protect privacy?.....	5
<b>2. Guidance for Research Ethics Committees and for researchers</b> .....	<b>6</b>
Why is ethical review needed? .....	6
Who are the research participants?.....	7
Is consent being gained from data subjects / can consent be withdrawn? .....	7
What do you mean by de-identified data?.....	8
What are the conditions of use? .....	9
How will data be kept secure and transferred? .....	9
Will data be shared outside the proposed research team?.....	9
Can I change my research project after it has been approved by the Approvals Panel?9	
How will the data be managed to ensure confidentiality, privacy and data protection?. 9	

## Introduction

This guidance has been created to help both researchers who are applying to use the Administrative Data Research Network, and the ethical committees who review their projects.

The Network aims to increase the use of administrative data in the UK, as part of the Economic and Social Research Council's wider efforts to exploit the opportunities of 'big data'. This is a relatively new area of research, so it may be necessary to think about the ethics of it in a new way. This information, for example, was not collected specifically for research, so needs to be handled differently to survey data.

Demonstrating an ethical review has been successfully completed is one of the criteria which has to be met for the ADRN Approval Panels to Approve a project. It is important that ethical committees are given an accurate explanation of the ADRN processes when appropriate within ethical review applications. This guidance is intended both to prompt researchers to think carefully about the ethical dimensions of the work they are planning, and to prompt research ethics committees to ask the right questions of projects they are asked to review.

The guidance is split into two sections:

1. Background information about the Administrative Data Research Network. This is equally useful to ethics committees and to researchers applying to the Network. It explains the Network's purpose and how an ethical review process fits within the project application and approval process.
2. Specific information that relates to some of the areas commonly covered by ethics committees. This is intended to help with difficulties that might occur when a traditional approach to ethical review is applied to a project that uses and links administrative data. It is also intended to guide researchers in preparing their submission to an ethics committee.

## Foreword

The Administrative Data Research Network is the first phase of the ESRC's £64 million investment in big data. Four new innovative Administrative Data Research Centres and an Administrative Data Service have been created to strengthen the UK's use of big data, with the aim of creating real benefits for the public.

The Network will make routinely collected administrative data accessible for research – which we believe will provide a sound evidence base for developing, implementing and evaluating policy. It will also make sure the data is handled in a way that prevents people being identified. This requires not just a safe, secure and efficient system for preparing, managing and analysing administrative data, founded on secure technology, but also trust between data providers, researchers and other interested parties, including the public.

I am proud to serve on the independent Approvals Panel which assesses applications to the Network, and to play a part in increasing the UK's use of big data. There will be benefits for researchers, government, local communities and the public. Indeed, there is the potential for a revolution in our ability to answer a host of questions that were previously intractable.

Professor Sir Ian Diamond  
Principal and Vice Chancellor, University of Aberdeen  
Chair, ADRN Approvals Panel

## 1. An introduction to the Network

The Administrative Data Research Network includes universities, national statistics authorities, the Economic and Social Research Council, government departments and agencies, and researchers.

We enable researchers to carry out social and economic research using administrative data (routine information government departments and agencies collect when they register people and carry out transactions). Summaries of the projects currently being undertaken within the ADRN are available via the ADRN's website under [Approved Projects](#)<sup>1</sup>

The Network is coordinated by the Administrative Data Service, led by the University of Essex. We have four Administrative Data Research Centres – one in each country of the UK:

- ▶ England, led by University of Southampton
- ▶ Northern Ireland, led by Queen's University Belfast and the University of Ulster
- ▶ Scotland, led by University of Edinburgh
- ▶ Wales, led by Swansea University

### What does the Network aim to do?

We want to see informed research which improves our understanding of society – and which can inform governments, councils and charities so they can shape policies and target their work in ways which will benefit society.

### Why use administrative data for research?

Administrative data are collected by government departments and other organisations when they register people and carry out transactions, usually while delivering a service. As these data collections already exist, using them for research is efficient and economical – there's no extra cost for collecting the information.

These data collections are also highly inclusive. Administrative data cover the whole of the UK's population, giving us a detailed picture, leading to greater understanding of our society and better-informed government policy. They often contain information about people who are under-represented in conventional surveys, but who are equally important when making policy decisions.

However, only a few research projects in the UK have ever used linked administrative data. They have traditionally been under-used because they have been very difficult to access. The creation of the Administrative Data Research Network establishes a UK-wide network to facilitate this research.

### What is a Network project?

A Network project will be research that requires the use of very detailed administrative data, typically at individual level. In many cases projects will require administrative data from two or more sources to be combined together.

These projects will have to pass an approval process, in which an independent Approvals Panel decides whether or not a project will go through the Network, based on a number of criteria. Further information is available via the ADRN's website under [Approvals Panel](#)<sup>2</sup>.

---

<sup>1</sup> **Approved Projects:** <https://adrn.ac.uk/research-projects/approved-projects/>

<sup>2</sup> **Approvals Panel:** <https://adrn.ac.uk/getting-data/approvals-panel/>

A Network project must:

- ▶ be purely non-commercial research
- ▶ be feasible, viable, ethical and have a clear potential public benefit
- ▶ make a case for using administrative data to carry out the research
- ▶ show that the data can only be accessed through the Network, rather than through alternative arrangements (for example Farr Institute, UK Data Service Secure Lab, or longitudinal studies)
- ▶ not be research which a government department or agency would carry out as part of its normal operations
- ▶ make its results public through this website

The researchers who will undertake the work have to undergo Safe Users of Research Data Environments (SURE) training and complete an accreditation process. Further information is available via the ADRN's website under [SURE training](#)<sup>3</sup>

## How will the Network keep data safe / protect privacy?

We aim to have the highest standards of secure data sharing, which will be consistent across the Network. This means:

- ▶ **safe projects:** only projects approved by our Approvals Panel will have access to our services. The panel reviews and approves all research proposals. Among other things, they make sure each project is for non-commercial research and has a clear potential public benefit. We expect researchers to be from academia, charities, third sector, government or an independent research organisation recognised by our governing board. We will not currently consider proposals from researchers from commercial organisations, but it may be possible, for example, for a PhD student sponsored by a business to use the Network if their work has clear public benefit and the findings are going to be published. If the research is funded by a business and is not for publication, it would not be eligible.
- ▶ **safe people:** only accredited researchers will have access to the Network's services. We give researchers mandatory training in how to access and use administrative data safely, lawfully and responsibly, and provide secure facilities where they can do their research. We also work with government departments to enable researchers to get permission to link and access the data.
- ▶ **safe data:** data will be de-identified – that is, researchers will not be able to see information which directly identifies any individual. We make sure all the information that can directly identify a person, household or business (names, addresses, etc.) is removed from the data before researchers access it. We review all research outputs before they are published to make sure privacy is protected.
- ▶ **secure environments:** we provide a secure environment where researchers can access the data, and state-of-the-art secure information technology and procedures provide physical, hardware and software security across the whole Network. Researchers cannot take anything in or out of the room with them (including mobile phones, memory sticks or even pen and paper), and they will not be able to copy, download or disseminate the data in any way.

The Centres which make up the Network all follow nationally recognised security standards for handling sensitive data, and all abide by the Network's information assurance and data security policies. Data which identifies individuals directly will travel

---

<sup>3</sup> **SURE training:** <https://adrn.ac.uk/getting-data/researcher-training/sure-training/>

by an accredited, secure route from data provider to trusted third party (using a third party to link identifiers ensures that no organisation ever sees the entire linked dataset. Specific organisations are commissioned by the network and are therefore 'trusted' to undertake this function) before data are matched at the Administrative Data Research Centre facility where the researcher will work. Research teams will work in the secure facilities and will not transfer data outside of them.

No-one is allowed to share data outside the proposed research team. Analytical outputs can only be shared if they have undergone statistical disclosure control and been cleared by the Network's expert staff.

## 2. Guidance for Research Ethics Committees and for researchers

Research Ethics Committees have an established track record of considering the ethical implications of social science projects, and will also consider the reputation of their own institution when looking at any project. Often, the main area of interest would have been in the way in which interactions between the research team and participants are established and managed. In some areas, Network projects will differ from the majority of social science projects that ethics committees will previously have seen. De-identification of data is part of the ADRN's measures to minimise the risk of statistical disclosure; however, Research Ethics Committees will also need to consider any wider ethical considerations in relation to the project.

Researchers applying to use the Network are likely to have made previous submissions to Research Ethics Committees. They will be accustomed to considering particular ethical issues in designing their work and will have experience of producing submissions to ethics committees which have been successful.

However, using administrative data for research presents a range of different ethical issues to consider and, in some cases, may mean some of the questions ethics committees ask will need to be answered differently and some of the answers provided by applicants may appear unusual.

The remainder of this document is intended to provide explanation and information which enables the right questions to be asked by research ethics committees, and accurate answers to be provided by applicants.

### Why is ethical review needed?

The Network has identified ethical review as a crucial element of the overall project approval process for the network. Our independent Approvals Panel requires researchers to provide evidence of a supportive verdict from an institutional ethical review on their research proposals.

This is because there are ethical factors to consider when using and linking administrative data and we want to be sure that due consideration has been given to the context of the research project.

There are three main factors:

1. Consent  
These data are not initially collected for research, but by government departments and agencies during the course of their normal business. Individuals contained in the data collection will not be explicitly aware of the research, because their personal identities do not directly appear in the research data. This lack of explicit research consent means we have a particular moral responsibility

to safeguard the data, even though it is de-identified and not classified as 'personal data' under the Data Protection Act.

2. Risk of identification

There is a very remote risk of 'statistical disclosure', when specific information from a de-identified data collection can be associated with a particular individual, household or business. We take this seriously, and have procedures in place to keep the risk to an absolute minimum.

3. Sensitivity

Although our researchers cannot directly identify anyone whose details are contained in a data collection, the data involved may contain detailed and potentially sensitive information. The Network has strict procedures in place to address the issue of privacy but it remains important to have an independent check on the ethical basis of the work.

We address the first two of these below, but the question of sensitivity is one on which ethics committees will make their own judgments.

## Who are the research participants?

Depending on the definition of 'participants' that an ethics committee adopts, the answer to this question, or associated questions about 'participants' may not be straightforward and may not appear to make sense.

Research using administrative data will use aggregate or de-identified individual records which relate to the general public. An individual project is likely to want to use specific cohorts in their analysis but may also look to use data about the population as a whole, either across individual countries or across the UK as a whole.

Some ethics committees consider participants to be people who have actively engaged in a data collection process for the project (e.g. they've been interviewed, surveyed, provided biological samples, worn devices which collect physiological data, etc.). When a researcher uses administrative data they are reusing data which was collected for another purpose; the data subjects can be passive in the process so there may be instances where they are not be considered a 'participant', although there will be scenarios where data subjects have been compelled to provide their data (e.g. through the Criminal Justice System). This ethical dimension may already be familiar to ethics committees through scenarios such as ethnography research involving non-participatory observation.

## Is consent being gained from data subjects / can consent be withdrawn?

Within the Network, administrative data is reused for secondary analysis without the data subjects' consent. The network only processes data where there is a legal right to access without explicit consent. Only de-identified data is made available to the ADRN, protecting privacy of individuals in line with the Anonymisation Code of Practice issued by the Information Commissioners Office in 2012.

Some research projects will use survey data which has been linked with administrative data. Survey data refers to information which have been proactively gathered from members of the public, usually through completion of a questionnaire or an interview process. In these circumstances it will be essential that specific consent was gained from survey participants for their data to be reused for statistical analysis.

If the researcher wants to add administrative data to survey data they themselves are collecting, they should seek explicit consent from their survey participants.

## What do you mean by de-identified data?

When preparing data for researchers to use, the Network takes a number of steps to remove information that can directly identify individuals, households or businesses.

The organisation providing the data splits it into two parts:

1. information (such as names, full addresses, dates of birth, etc.) which can directly identify individuals (identifier data)
2. the rest of the information (attribute data) which is of interest to the researcher (for example, information on tax paid, benefits received or educational qualification achieved) without any direct identifiers

Both parts are given a reference number.

A [trusted third party](#)<sup>4</sup> matches the information by using the unique reference numbers and the identifying information. The identifying information is then destroyed, leaving only the matched unique reference numbers – called a linkage key. The Administrative Data Research Centre then links the de-identified attribute data using the linkage key.

Researchers can then access the linked attribute data in a secure environment. The linked data file has all the information needed for research but no directly identifying information.

Because the Network only handles the de-identified data we would not be able to identify individual records in order to remove them.

## How were the data collected? How will they be analysed?

The Network helps researchers to use administrative data, which are collected during the course of normal government business or service activity. They include tax, benefits, education, medical and other records. Data from different sources are securely linked by a trusted third party, and the researcher sees only de-identified data (all direct identifiers, such as name, address, etc. removed). However, the data may be highly detailed, so there is a theoretical risk of identifying individuals from a unique combination of variables. Because of this residual risk, we have a number of extra controls in place governing access to and analysis of the data.

All data analysis will take place in a secure setting which has been independently accredited to handle data at the level of sensitivity proposed. Researchers will not be able to remove data from the setting, and we will check all research outputs before release to make sure individuals' privacy is protected. Further information is available via the ADRN's website under [ADRN Terms of Use](#)<sup>5</sup>.

---

<sup>4</sup> **Trusted Third Party:** <https://adrn.ac.uk/getting-data/de-identification/trusted-third-party/>

<sup>5</sup> **ADRN Terms of Use:** [https://adrn.ac.uk/media/1399/adrn021-termsofuse\\_v00-10\\_pub.pdf](https://adrn.ac.uk/media/1399/adrn021-termsofuse_v00-10_pub.pdf)

## What are the conditions of use?

The Network will obtain data for projects, and permission to link them, on a case by case basis on the researcher's behalf. We will agree and implement any additional conditions of use put on the data by the data providers.

## How will data be kept secure and transferred?

There are details on our website of how we keep data secure, further information is available via *Data security policies* under [Policies and Procedures](#)<sup>6</sup>. The Centres which make up the Network are all accredited to handle these data collections, and all abide by the Network's information assurance and data security policies. Data will travel via one of a variety of accredited secure methods from data provider to trusted third party and to the Administrative Data Research Centre facility where the researcher will work. Research teams will work in the secure facilities and will not transfer data outside them.

## Will data be shared outside the proposed research team?

Access to the underlying data is only granted to the researchers associated with the specific project. No-one is allowed to share data outside the proposed research team. Analytical outputs can only be shared if they have undergone statistical disclosure control and been cleared by the Network's expert staff.

## Can I change my research project after it has been approved by the Approvals Panel?

The ADRN has a specific process for considering requests to amend projects. Depending on the extent of the change proposed it may be necessary for project to be reconsidered by the Approvals Panel. It may also be deemed necessary for the project to be referred back to the ethical committee for reconsideration.

## How will the data be managed to ensure confidentiality, privacy and data protection?

All researchers using the Network are required to undertake accreditation training, which will educate them about their legal and ethical responsibilities while handling and analysing data. All research outputs will undergo formal statistical disclosure control before being released for publication. After the research is complete, the linked data collections will be securely destroyed.

---

<sup>6</sup> **Policies and Procedures:** <https://adrn.ac.uk/getting-data/resources/documentation>