UK Statistics Authority

16th Meeting of the National Statistician's Data Ethics Advisory Committee

---

Minute, Agenda and Papers

Tuesday 7 May 2019

10:30 – 14:30

Board Room, UK Statistics Authority
London

UK STATISTICS AUTHORITY

NATIONAL STATISTICIAN'S DATA ETHICS ADVISORY COMMITTEE

Agenda

Tuesday 7 May 2019
Drummond Gate London
10:30am – 2:30pm

Chair: Dame Moira Gibb

(10:30am to 12:30pm)

1 10:30am	Minute and matters arising from the previous meeting	Presentation Simon Whitworth
2 10:50am	Update on the Ethics self-assessment	NSDEC(19)05 Emily Mason-Apps
3 11:10am	Investigating the health outcomes of hospital patients from rural and urban areas	NSDEC(19)06 Danielle Cornish
4 11:30am	Environmental and socio-economic impact assessment for siting geological disposal facilities	NSDEC(19)07
5 11:50am	Urban redevelopment baselining for urban redevelopment projects	NSDEC(19)08 Bill South
6 12:10pm	Data Protection and Ethics Compliance	NSDEC(19)09 Rhys Nadin

Lunch (12:30pm to 1:00pm)

(1:00pm to 2:30pm)

7 1:00pm	Overview from Economic Statistics	Presentation Phil Wales
8 1:20pm	Social Survey Transformation	Presentation Ian O'Sullivan
9 1:40pm	Overview from the Data Science Campus	Presentation Peter Fullerton
10 2:00pm	Overview from Public Policy Analysis	Presentation Hugh Stickland
11 2:20pm	Any other business	

National Statistician's Data Ethics Advisory Committee
Minute

Tuesday, 7 May 2019
Boardroom, Drummond Gate, London

Present
Members
Dame Moira Gibb (Chair)
Mr Stephen Balchin
Mr Rob Bumpstead
Mr Keith Dugmore
Mr Colin Godbold
Ms Annie Hitchman (on phone)
Ms Isabel Nisbet

UK Statistics Authority
Dr Simon Whitworth
Dr Emily Mason-Apps
Mr Rhys Nadin

Office for National Statistics
Ms Danielle Cornish (for item 3)
Mr Theodore Joloza (for item 3)
Mr Bill South (for item 5)
Dr Phil Wales (for item 7)
Mr Ian O'Sullivan (for item 8)
Mr Peter Fullerton (for item 9)
Mr Hugh Stickland (for item 10)

Apologies
Ms Vanessa Cuthill
Ms Marion Oswald
Dr Emma Uprichard

1. Minutes and matters arising from the previous meeting

1.1. The Chair welcomed members to the sixteenth meeting of the National Statistician's Data Ethics Advisory Committee (NSDEC).

1.2. Members approved the minutes from the previous meeting.

1.3. Dr Whitworth updated the meeting with progress on actions from previous meetings. Most actions were complete or in progress and would soon be complete.

2. Update on the Ethics self-assessment [NSDEC(19)05]

2.1 Dr Emily Mason-Apps, from the Data Governance, Legislation and Policy team at the UK Statistics Authority, briefed NSDEC on the development of the ethics self-assessment which provides researchers with an easy-to-use framework to consistently review the ethics of their projects in light of the NSDEC's ethical principles.

2.2 NSDEC approved the changes and suggested that the guidance could be further improved by defining what is meant by vulnerable adults, and providing examples of what is meant by public good.

3. Investigating the health outcomes of hospital patients from rural and urban areas [NSDEC(19)06]

3.1. Ms Danielle Cornish, from the ONS Health and Life Events team, presented a project to link Hospital Episodes, Patient Register, mortality registrations and census to investigate differences in health outcomes between those living in rural areas and those in urban areas.

3.2 This project was approved subject to Ms Cornish doing the following:

i. developing a stronger explanation in the application of the public good of this research to ensure that this research is answering questions that stakeholders want answered;
ii. making it clear in the application why the variables requested are needed to answer the research questions that are set out in the application; and
iii. making it clear in the application that this research is limited to the four research questions set out in the application and any research using the data outside of the four research questions included in the application would require future ethical approval.

4. Environmental and socio-economic impact assessment for siting geological disposal facilities [NSDEC(19)07]

4.1. The committee considered a research project from Simetrica, sponsored by Radioactive Waste Management, that wants to access deidentified data from the Annual Population Life Survey, the Community Life Survey and the National Survey for Wales to inform the likely socio-economic benefits of siting geological disposal facilities in Local Authorities that have expressed an interest in hosting these facilities. This will help inform a decision about where these facilities will be housed.

4.2 NSDEC asked for Simetrica to make the following major revisions to this research proposal:

i. making clear in the application that this research is not being done to influence Local Authorities to express an interest in hosting a geological disposal facility and that the analysis will only be conducted for those Local Authorities who have already expressed an interest;

ii. to work with any Local Authorities who have expressed an interest in hosting a geological disposal facility to get their views on what socio-economic indicators they would like the research to include;
iii. to commit to publishing more than just a summary of the research outcomes; and
iv. making clear in the application that this research should only be used to help inform the Local Authorities and therefore the local communities in their decision-making process rather than being used to directly inform the siting of the geological disposal facility.

4.3 NSDEC stated that they wanted to discuss the revised application at a future meeting before access to any data was granted and Simetrica were asked to bring this back to a future meeting.

5. Urban redevelopment baselining for urban redevelopment projects [NSDEC(19)08]

5.1 Mr Bill South presented a research project from Simetrica, sponsored by Lendlease, that wants to access Annual Population Survey data to conduct baseline socioeconomic analysis for Lendlease’s Euston train station and Silvertown Quays urban redevelopment projects in London. This will provide baseline values that allow Lendlease to monitor the creation of social value over time.

5.2 This project was approved subject to Simetrica doing the following:

i. being clear in the application about the timescale the work covers;
ii. committing to openly publishing the full results of the analysis;
iii. exposing the proposed method to independent scrutiny; and
iv. making it clear in the application whether the research will just be looking at the impacts of the developments on the residents in the area, or will they also be looking at the impacts of non-residents who interact with the area such as commuters and those working in the area.

6. Data Protection and Ethics Compliance [NSDEC(19)09]

6.1 Mr Rhys Nadin, from the Data Governance, Legislation and Policy team at the UK Statistics Authority, reported that the UK Statistics Authority would be conducting regular audits to provide assurance to the National Statistician that NSDEC’s ethical advice is being adhered to in the development and application of project proposals that have been reviewed by the committee. This will be reported in a quarterly report to the National Statistician which will be shared with NSDEC. NSDEC welcomed this work and stated their support for the work.

7. Overview from Economic Micro-data Research

7.1 Dr Phil Wales, Head of ONS Economic Micro-data Research, presented plans to build improved aggregate statistics based on the administrative data provided through Pay As You Earn and Self Assessment data. This work depends on successful future data acquisition of these data sets, and the linking of these data sets with existing data from auxiliary sources such as Census and the Business Register. The work aims to provide a new suite of statistics which will characterise labour market dynamics in a clearer, more timely and more accurate fashion than is currently possible. This work will provide research outputs and insights that will inform policy and will enable the UK to keep pace with other economies who already link data in this way.

7.2 The committee recognised the potential benefits and public good that this work will enable and it was suggested that a full application should be submitted to NSDEC. To help inform this application it was suggested that:

i. this work may benefit from more specific public engagement, given the raised sensitivities around the type of data to be linked;
ii. it would be useful to explore what the other countries who already link data in this way have done in terms of public good;
iii. this sort of data has the potential to become more sensitive over time, and
iv. serious consideration needs to be given to how to manage the data of special interest groups.

7.3 Dr Wales was asked to produce a full application for review at a future meeting.

8. Overview from Social Survey Transformation

8.1 Mr Ian O'Sullivan from ONS Social Survey Transformation presented the ongoing work in the area to move towards conducting surveys on-line. It was reported that this has been trialled with the ONS Integrated Household Survey and work has been conducted to assess the effectiveness of different invitation to participate strategies and incentives to participate. Future work will explore the feasibility of moving other surveys online.

8.2 It was suggested that it would be important to carefully review and consider the ethics and practicalities of moving some of the surveys online, along with the associated recruitment and incentives strategies. It was agreed that the proposal for the Crime Survey for England and Wales would be presented to NSDEC for independent review at a future meeting.

8.3 Mr O'Sullivan was asked to produce a full application on the transformation of the Crime Survey for England and Wales for review at a future meeting, and to work with the secretariat to ensure there is appropriate ethical review of future survey transformation and incentives strategies.

9. Overview from the Data Science Campus

9.1 Mr Peter Fullerton presented the objectives and goals of the ONS Data Science Campus (DSC), along with a summary of projects currently in the discovery or planning stage and a more detailed overview of a project on flows of tenants within social residential accommodation. It was agreed DSC would bring this to NSDEC for ethical consideration at a future meeting.

9.2 Mr Fullerton reassured the committee that all projects are subject to an ethical-self assessment, and that due to the time-frames of acquiring data, there is often flexibility for further consideration to take place when appropriate.

9.3 Mr Fullerton was asked to produce a full application on the project on Flows of tenants within social residential accommodation for review at a future meeting.

10. Overview from Public Policy Analysis

10.1 Mr Hugh Stickland presented an overview of the strategy, previous successes, and current and future work in ONS Public Policy Analysis. This included an outline of upcoming projects for 2019/2020.

10.2 Mr Stickland was asked to work with the Secretariat to ensure that there was appropriate ethical review of the future projects including presenting full applications to NSDEC when appropriate.

11. Any other business

11.1 There was no other business.

NSDEC(19)05

UK STATISTICS AUTHORITY

NATIONAL STATISTICIAN'S DATA ETHICS COMMITTEE

Update on Ethics Self-Assessment

Purpose

1. The purpose of this paper is to present the updated versions of the ethical self-assessment form and guidance for feedback and comment from the committee, and to present our plans to roll this out across ONS, the Government Statistical Service and, where appropriate, the wider research community.

Recommendations

2. Member of NSDEC are invited to:

    i. provide comments on the revised self-assessment guidance at **Annex A**;
    ii. provide comments on the revised self-assessment form at **Annex B**; and
    iii. provide comments on the plans set out in this paper to increase engagement with this tool across ONS, the GSS and the wider research community.

Background

3. As previously discussed with NSDEC, the self-assessment was introduced with the aim of providing researchers with an easy-to-use framework to consistently review the ethics of their projects in light of the NSDEC's ethical principles. This would enable researchers to consider the ethics of their research at the research design phase, thereby promoting a culture of ethics by design. This would also enable the NSDEC secretariat to identify those projects that are most ethically contentious and could best use the independent advice provided by NSDEC.

4. The self-assessment has been trialled across ONS with variable uptake. Based on this trial, along with feedback from users, we have made several changes to both the self-assessment form and the accompanying guidance. We will outline these changes in this paper and justify our reasons for making these changes. We will also present our plans to increase engagement with the self-assessment tool across ONS, the GSS, and the wider research community.

Discussion

5. The main text of the guidance at **Annex A** has been revised and updated with the following aims:

    i. to make the guidance more relevant and accessible for the wide range of researchers that are likely to use the self-assessment, as feedback from users highlighted that the previous version was quite ONS focused;
    ii. to make it clear that researchers should not complete this self-assessment in isolation, and should discuss project proposal and ethical assessments with senior directors/managers, relevant Information Asset Owners, Legal Services, and Communications and Media relations; and

   iii. to highlight that all self-assessment forms should be shared with the NSDEC Secretariat so that we can provide NSDEC with the necessary assurance that researchers are using, scoring and interpreting the forms correctly and appropriately.

6. As explained in **Annex A**, to avoid responses that might indicate ethical issues being averaged out of the overall score of the self-assessment, the scoring scales for each item include a tolerance setting. When these tolerances are exceeded, the form (**Annex B**) will alert researchers and recommended that they consider appropriate actions to mitigate the ethical risk, and if mitigations are not possible it is recommended that the project is sent to NSDEC for a full independent ethical review. Following discussions with ONS Legal Services and the ONS Data Protection Officer about the scoring scales and tolerance settings, we have made the following changes based on their recommendations:
   i. we have changed the scales used to score items 5, 6, 8, 9, 16, 17, and 20 so that they are now scored on a 3-point scale (with response options of 1, 3, and 5) rather than a 5-point scale (with response options of 1, 2, 3, 4, or 5). These changes have been made because it was suggested that these items require a more “binary” response, rather than the more scaled or granulated response offered by the 5-point scale. Essentially, scoring these items a “1” is the only acceptable option for all but one of these items, as both other options 3 (which is crudely the ‘don’t know’ option) or a 5 will trigger a tolerance being exceeded which consequently alerts the researcher that there are potential risks that they need to discuss with the NSDEC secretariat and mitigate; and
   ii. we have lowered the tolerance triggers for items 7, 11, 13, 18, and 19, to ensure that the form will be as sensitive as possible to alerting the researcher to any potentially unethical uses or data, and also highlights when their project is likely to benefit from the full independent review offered by NSDEC.

7. In order to ensure that the updated self-assessment is used consistently and appropriately, we have conducted or planned the following activities:
   i. We are organising a series of talks across all three ONS sites to deliver open presentations to staff.
   ii. Taking over the GSR ethics module in the introductory training given to new GSRs.
   iii. Working with processors and researchers who plan to use the Research Strand of the Digital Economy Act to make them aware that this as one potential tool that researchers can use to assess the ethics of their research.
   iv. We have made some changes to the website, and will be updating the training to make sure that all the information available to all researchers is clear and consistent with the changes made to the self-assessment form and guidance.

Emily Mason-Apps and Simon Whitworth, Data Governance, Legislation and Policy, 29 April 2019

List of Annexes

Annex A Guidelines on using the ethics self-assessment process
Annex B The ethics self-assessment form

Annex A Guidelines on using the ethics self-assessment process

The guidelines on using the ethics self-assessment process can be downloaded here.

URL: https://www.statisticsauthority.gov.uk/about-the-authority/committees/nsdec/data-ethics/self-assessment-2/

Annex B The ethics self-assessment form

The ethics self-assessment form can be found downloaded here.

URL: https://www.statisticsauthority.gov.uk/about-the-authority/committees/nsdec/data-ethics/self-assessment-2/

NSDEC(19)06

National Statistician's Data Ethics Advisory Committee

Application for Ethical Review

Please consult the guidance document before filling in the application form

Project Title

Please provide a title indicative of the project

Investigating the health outcomes of hospital patients from rural and urban areas

Start Date:	May 2019	**End Date:**	May 2020

Project Sponsor(s)
Please list the project sponsor(s)

Office for National Statistics (ONS)

Project Summary

Please provide a brief high level summary of the research giving necessary background (max 250 words)

While health outcomes are generally more favourable in rural than urban areas, more sparsely populated areas with an ageing population present challenges for delivery of health and social care. The Local Government Association note that "only 55% of rural households compared to 97% of urban households are within 8km of a hospital". Living further away from a hospital can result in 'distance decay' - where service use decreases with increasing distance. Rural transport is one reason for distance decay; Age UK found that some patients needed to make 2 to 3 bus journeys to attend hospital appointments. In addition to accessibility issues, researchers have found differences in attitudes to seeking healthcare by rurality. Research commissioned by the Department for Environment, Food, and Rural Affairs (DEFRA) identified a tendency for older people living in rural areas to present to health services in moments of crisis, therefore requiring “more intensive, immediate, invasive and complex responses”. Unplanned, emergency care can be more expensive and carries greater risk of harm to the patient. Therefore, this project will explore outcomes of patients accessing hospital services from rural versus urban areas, to investigate whether outcomes are less favourable for residents of rural communities.

DEFRA has also noted that “there is no one ‘type’ of older service user – needs are variable and shift over time”. So, a second aim of this research will be to identify the characteristics of rural patients who are at a raised risk of poor health outcomes (if any differences are found between rural and urban areas). For example, diagnosis information will allow us to see whether certain health conditions are associated with worse outcomes. While much of the research focuses on older adults, this project will analyse the outcomes of patients of all ages. If certain groups, such as older adults or children, are found to have a different level of risk then breakdowns by age will be explored (while keeping disclosure and confidentiality in mind).

The analysis will aim to answer the following research questions:

1) Do the outcomes of patients differ between those who live in rural versus urban areas? Outcomes include: whether the patient was admitted or discharged, whether the patient died within 30 days of attendance, length of admission, outpatient appointments missed.
2) If any differences are found: do these vary by demographic factors such as: age, sex, ethnic group, availability of a carer, deprivation, region, occupation / socioeconomic status, long-term health problem or disability, marital status?
3) If any differences are found: do these vary by contextual factors such as: arrival mode to Accident and Emergency (A&E), source of admission, source of referral, diagnosis, underlying cause of death, place of death?
4) Does the likelihood of changing health status (perceiving oneself as having good versus poor general health) since the 2011 Census differ by rurality? This would use accessing hospital services as a proxy for poor health in the years following the Census.

The core research, analysis and outputs will be restricted to these four research questions. No outputs outside the scope of the 4 questions as specified in this application will be produced.

This analysis will benefit the public by providing bodies such as Public Health England (PHE), DEFRA, and the NHS with better information to support the provision of healthcare (see section B1). Providing data on patient outcomes by rurality and demographic characteristics will allow healthcare providers and policymakers to target at-risk groups, ultimately saving lives. Colleagues in PHE have stated that their intelligence on some inequalities (such as deprivation, sex, and some geographies) is good, but rural versus urban inequalities could be improved. New data on this will help government to identify the scale and nature of the problem, therefore allowing local authorities, NHS and other health and care providers to design and adapt services to meet rural needs and reduce rural inequalities where they exist.

Hospital Episode Statistics (HES) will be the main dataset used for analysis. HES contains details of all admissions, outpatient appointments and A&E attendances at NHS hospitals in England. A full list of HES variables required for this research, and justifications for these being sought, are available in Annex 1.

ONS are currently linking HES to mortality data and the 2011 Census, for other projects previously presented to NSDEC. Following analysis of HES alone, this linked dataset will be used to investigate rural-urban differences in those who have died following hospital attendance (using the HES-mortality linked data) and their health characteristics in 2011 (using the Census-HES linked data).

Section A
Project Details

A1 Legal gateways

Please provide the assessment of the legal gateways of the project as provided by Legal Services

NHS-Digital controlled HES data

Section 45A of the Statistics and Registration Services Act 2007 (as inserted by the Digital Economy Act 2017) provides a legal gateway for all public authorities, including NHS Digital (NHS D), to share the information they hold with ONS. This legal gateway over-rides any restriction on disclosure, including patient confidentiality, though data protection legislation must still be complied with. ONS can use the data received under this gateway in pursuit of any of its statutory functions.

ONS Mortality data

Section 42 of the Statistics and Registration Services Act (SRSA) 2007 enables ONS to receive mortality data from the Registrar General and to use that data to support any of its statutory functions.

Census Data

This is collected by ONS under their powers in the Census Act 1920 (and associated subordinate legislation) and can be used by ONS for any of its functions.

A2 Ethical approval

Has the project being reviewed or is it expected to be reviewed by another ethics committee?	○ **Yes** ○ **No**
If Yes please provide the name of the committee, the outcome and the date approved	

A3 **Proposed site of research** select all that apply

☑ ONS

☐ SRS (formerly VML)

☐ HMRC Data Lab

☐ Other

please specify

☐ ADRC - England

☐ ADRC - Scotland

☐ ADRC - Northern Ireland

☐ ADRC - Wales

A4 **Data subjects to be studied**

Does the study include all subsections of the population
(i.e. all ages, sex, ethnic groups etc?)

○ **Yes** ○ **No**

If no please detail which subsections with justification(s) below:

Subsections of the population (including vulnerable groups) the project focuses on:

n/a

Justification for focusing on these subsections or groups:

n/a

A5 Please provide details of the research protocol or methodology (e.g. data linkage, web scraping etc) (max 500 words)

The methodology for creating the linked HES-Mortality-Census dataset has previously received ethical approval from NSDEC for a different research use. Therefore only a brief outline will be given here; full details are available here (paper NSDEC(18)22).

Patient register (PR) data contains all the demographic variables commonly used for linkage, as well as NHS number. PR will be used as a bridge between the Census (which has demographics, but not NHS number), and HES (which has NHS number and some demographics, but no full name). A similar process is already underway to link Census and mortality data, via PR. This work is being conducted by a data integration team, separate to the analysis team, who have been trained in data protection and are security cleared. All linkage work will be completed in the ONS secure Data Access Platform (DAP; see annex 2). The result will be a linked person level file between HES, Census and Mortality.

Once linked, identifiers will be removed from the dataset and this de-identified subset will be stored on the ONS secure DAP. Here, two or three ONS researchers will have access to undertake the analysis. **All staff who have access to the data will be substantive ONS employees, SC cleared, and trained to handle data correctly (see section B3).** The research questions to be answered by the research team are outlined in the Project Summary section. These questions will be primarily investigated by calculating age-standardised rates (e.g. of admission from A&E, of death following discharge) by rural classification. If differences between rural and urban areas are found, multivariate regression will be used to determine which patient characteristics are linked to poor outcomes.

All the variables requested (annex 1) are necessary for the analysis to take place. The variables in the list fall into 3 categories;

I. **Research level variables**
    a. Required to answer the 4 research questions
II. **Administrative variables**
    a. To enable linkage
    b. To enable quality assurance
III. **Year and rurality variables**
    a. To identify year of record to enable comparison of different periods
    b. To enable urban-rural comparison

Products of the analysis will be published on the ONS website and the aggregated tabulations and statistics will be assessed for disclosure control against the current Mortality Disclosure Control protocol. Staff will also seek advice from analysts at NHS Digital on the presentation and disclosure control of HES data specifically.

A6 Data use

Please specify the data used **by the research team** including any timeframes *e.g. LFS data 2014-15*

Type of data	**Data Level** *Please specify the name of the data set*			
	Aggregate Data	**Identifiable Data**	**De-identified personal data**	**Anonymised/ pseudo anonymised**
Administrative data *(please specify, e.g. Patient Register 2011, School Census 2012 etc, in the relevant options adjacent)*		HES 2010/11 to 2017/18 Mortality registrations 2011 to 2018 Patient Register 2011 to 2013 Only available to those conducting the linkage	The analytical dataset will contain deidentified, linked person level data. The analytical dataset, detailed below, will only be available to a maximum 3 researchers working on this project.	
Big Data *(please specify e.g. Twitter data, smart meters and mobile phones, in the relevant options adjacent)*				
Survey Data *(please specify e.g.LFS, BRES, etc in the relevant options adjacent)*				
Census Data *(please specify year, e.g. Census 2011 in the relevant options adjacent)*		2011 Census Only available to those conducting the linkage	The analytical dataset will contain deidentified, linked person level data.	

			The analytical dataset, detailed below, will only be available to a maximum of 3 researchers working on this project.	
Other *(please specify e.g. Ordinance Survey Address register in the relevant options adjacent)*				

Section B
Assessment against NSDEC ethical principles

B1 Principle 1: The use of data has clear benefits for users and serves the public good.

Please outline the proposed benefits of the project (max 500 words)

It has been identified that the health outcomes of rural and urban populations can differ. For instance, the health and wellbeing in rural areas report published by Local Government Association and Public Heath England identified that the rural populations can suffer 'distance decay'. This shows that poor access to healthcare can result in decreased use of services. This project will identify whether those people living in rural areas are arriving at hospital in a more critical condition. This would highlight whether health issues are going unnoticed in rural areas, possibly due to a lack of access until the condition requires urgent intervention. If this is found to be the case, efforts could be made to assist earlier intervention, therefore improving the health of the rural population.

The NHS Long Term Plan identifies living in remote, rural and island locations as a health inequality in the equality and health inequalities impact assessment. Analysing the demographic and contextual factors outlined in the project summary will help inform what type of care is required in these areas, which may differ from urban areas. Meeting the health needs of the rural population is important if the identified inequalities are going to be addressed. Stakeholders in Public Health England (PHE) have been exploring rural inequalities as part of their priority to reduce inequalities. Communications with PHE have revealed that they currently have relatively little on urban versus rural inequalities. Specifically, they said "Following a House of Lords Report, and as part of the ongoing All-Party Parliamentary Group to which PHE are inputting, we are looking to gather evidence as to scale and nature of the inequality. It would be used to allow local authorities, NHS and other health and care providers to design and adapt services to meet rural needs and reduce rural inequalities where they exist".

Much of the evidence available regarding rural health is focused on the cost of delivering healthcare to the rural population and the difficulty of providing services, including the rural health care research report published by Nuffield Trust. Research into the health outcomes experienced by people living in rural areas gives the opportunity for healthcare allocation to be better informed, rather than just based on the number of people in an area and the age of the population, although these remain important factors.

It was discussed at the All-Party Parliamentary Group on Rural Services that 'Local Area Performance Metrics' showed delayed discharges from hospital were more prevalent in rural areas. By analysing the diagnoses of patients, this project will indicate the amount of ongoing care that could be required in rural areas due to certain conditions. As the rural population ages the need for more homebased care could be in high demand. This analysis will provide policymakers with intelligence on the needs of rural patients, therefore allowing health and social care services to improve provision.

To summarise, understanding the health of rural populations will help shape the future healthcare for these areas to best suit the need and demand. It will also provide information to citizens who are hoping to move to rural areas to make an informed decision about the future of their health. The scoping work that has been carried out for this project has received support from Public Health England and DEFRA. Public Health England have particularly shown an interest in overlaying factors to create the full picture of rural health. Working with these organisations will assist getting the key messages of the research out to policy makers in relevant government departments and professional bodies, to gain the most public benefit.

It is therefore important to conduct this research as there is ongoing concern that rural areas are receiving inadequate health and social care services compared to urban areas. While a lot of research has been done to explore these perceived inequalities, none has provided the greater insight that this project has the potential to offer due to the linkage of HES with the Census.

Particularly, PHE have been exploring rural inequalities as part of their remit letter priorities to reduce inequalities. While they have good inequalities data on indices of multiple deprivation, male, female and administrative geography they currently have relatively little on urban vs rural inequalities. Following a House of Lords Report, and as part of the ongoing APPG to which PHE are inputting, they are looking to gather evidence as to the scale and nature of the inequality.

The analysis that we propose to do in this application will allow local authorities, NHS and other health and care providers to design and adapt services to meet rural needs and reduce rural inequalities where they exist.

At a strategic level this piece of work is also symbolic of how far producers of health and care statistics, in this case PHE, NHS and ONS have come in working collaboratively to provide better health and care solutions for the population. Previously, one of the key criticisms was that producers worked in silos, focussed on producing statistics that were key to own policy priorities while ignoring the general public good. Additionally, this contributes towards the GSS wide initiative to make health and care statistics more coherent as recommended by the Office of Statistics Regulation.

B2 Please outline any intended future use for products (such as linked data sets or tools) produced as a result of the research and how they will be accessed. (max 250 words)

The analytical data set, which will include de-identified personal information from the linked datasets, will be held securely for as long as it is needed for research and statistical purposes such as: reproducibility of the research; further analysis/research to support government policy; parliamentary questions. Access to the de-identified data will be limited to those working on this project (the research team). After the project has been completed, we will hold the data in DAP. Researchers and data managers with access to the data will have had comprehensive training on access and outputting these types of sensitive data and will have SC clearance. The results of the initial research (and any further developments) will be published on the ONS website as an article with aggregated tables that meet the ONS disclosure control policy.

B3 Principle 2: The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.

Please outline how data security, confidentiality and informed consent is safeguarded in this project(max 500 words)

All research outcomes will be subject to statistical disclosure controls to mitigate the potential risk of disclosures of small populations.

The linked data is stored securely on the ONS Data Access Platform (DAP). Annex 2 provides detail on the security of the DAP solution. The research will take place in DAP. Record level data will not leave this secure environment.

ONS researchers in the Health Analysis and Life Events Team have signed confidentiality agreements and received a high level of training in handling person level data, this includes:

- additional security vetting to the level of Security Cleared. This has been provided due to the sensitive nature of the data they have access to
- trained in data protection through office-wide courses and have a thorough understanding of the mortality disclosure control policy
- the research and outputs from the analysis will be overseen by the Assistant Deputy Director of the division, an expert in Health-Related Disclosure Control
- all staff are highly aware of the sensitive nature of the data we hold

Aggregated data will be published using the Rural-Urban Classification (2011) of Output Areas in England and Wales, initially breaking the data down by a rural versus urban split. Then researchers will assess to see if tables can be produced including all 10 areas of classification: rural hamlets and isolated dwellings, rural hamlets and isolated dwellings in a sparse setting, rural town and fringe, rural town and fringe in a sparse setting, rural village, rural village in a sparse setting, urban city and town, urban city and town in a sparse setting, urban major conurbation, urban minor conurbation. Tables using regional breakdowns will also be analysed, and the possibility of publishing data disaggregated by both region and rural-urban classification will be assessed, considering disclosure of individuals. All of this will be carried out following the published Mortality Disclosure Control policy. This applies threshold rules to table cells to ensure that individuals are not able to be identified in the aggregated tabulations.

Results from the research will only be released by ONS staff once they have confirmed that they contain no risk of identifying an individual. This is assured by the statistical disclosure controls the ONS currently have in place.

Consent is not being sought for this study. Consent is not required by law (see processing conditions set out in section B5), nor would it be practical or proportionate to obtain it for ethical reasons. The data being used has already been collected, so seeking consent would require trying to retrospectively contact all individuals. Further, the study is not concerned with individuals; identifying information is being used for linking only and then being removed before analysis takes place, so there will be no impact on individuals.

B4

Principle 3: The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.

Please describe how the any risks from new technologies are been mitigated as well as any quality assurance activities in the project (max 500 words)

No new technologies are to be used. We are confident that the matching should have a high success rate – previous ONS research has shown that a good match rate can be achieved between the Census and the 2011 patient register, and NHS number quality (completeness, accuracy) on HES and death registrations since 2011 is good. Work so far to link Census to deaths has achieved a link on 90% of death registrations. However, until we start the linkage work with HES it is not possible to be absolutely certain how good this link will be. All processes will have adequate human control to ensure appropriate quality assurance arrangements are in place.

B5

Principle 4: Data used and methods employed are consistent with legal requirements such as the Data Protection Act, the Human Rights Act, the Statistics and Registration Service Act and the common law duty of confidence

Please describe the legal frameworks pertinent to this project (max 500 words)

The legal basis for acquiring the HES data was Section 45A of the SRSA 2007 as amended by the Digital Economy Act 2017. This removes any other duty of confidence on the supplier, including common law duty of confidence. However, the data share must be needed for ONS functions, and not contravene data protection legislation.

To assess the former, there is a code of practice underpinning this power that the data share is assessed against. Acquiring HES data is in line with this code of practice, and internal scrutiny was provided by the ONS Data Governance Committee which approved ONS acquiring HES data for a wide range of purposes including this one. In terms of the data protection legislation the necessary conditions of fair processing have been met as follows:
Main condition:
(a) Public task: the processing is necessary for you to perform a task in the public interest or for your official functions, and the task or function has a clear basis in law.
Special category condition:
(b) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

To further comply with data protection legislation:
1) ONS will only obtain and use the minimum data needed for the success of project outcomes. One specific example of this is that only the minimum diagnosis information required to enable

robust feasibility testing of the methods being developed by this (and other planned ONS health projects) will be acquired from NHS Digital at this stage. Only once it is clear that robust statistics that will enable better public health decision making for the public good can be produced successfully, will additional diagnosis information be sought.
2) Researchers will only handle anonymised data
3) Strict, published disclosures control protocols will be adhered to with any outputs produced
4) There is a need for data controllers to inform individuals how their data are being used. For ONS we consider it would be disproportionate effort to notify each person individually, since there will be no impact on them. Instead details of the data shares will be published on the ONS website along with other privacy information on how we use data and how we keep it secure. In addition, NHS D will publish the Data Access Request Service application form that ONS will complete as part of the process to access the HES data.

With regards to the Human Rights Act there is no interference with the right to family life and privacy (HRA Article 8). No information will be collected about family members and they will not be contacted or identified in any circumstances. The data will be fully anonymised and subjected to strict Statistical Disclosure Controls so there is minimal risk of causing any harm or distress by a breach of confidentiality. The research methods used are sanctioned by law (see above) and proportionate to the public interest in the protection of health.

B6 Collaboration and Sponsors

Please describe the project sponsors and the **legal gateways** to acquire, process use and share their data

List of Collaborators/Sponsors	**Details and relevant documentation relating to collaboration (you may attach copies of relevant documentation)**

B7 **Principle 5: The views of the public are considered in light of the data used and the perceived benefits of the research**

Please list any public engagement activities (max 250 words)

While we have not sought the views of the public for this specific project, the research is of clear public benefit as per section B1.

ONS owns and regularly analyses Census and mortality data. Therefore, this section will focus in on the public acceptability of sharing and statistical research involving health data such as HES:

The 2017 National Data Guardian Review of Data Security, consent and opt outs, and the subsequent Department of Health and Social Care (DHSC) public consultation, included public acceptability research on the sharing of health data. The NDG review stated:

> ‘people hold mixed views about their information being used for purposes beyond direct care. Some are concerned primarily with privacy and are suspicious that information might be used by commercial companies for marketing or insurance.
>
> Others prioritise the sharing of information to improve health and social care, and for research into new treatments. There is broad support for data being used in running the health and social care system when the benefits of doing so are clearly explained, but people think that anonymised information should be used wherever possible.
>
> The Review also heard very strong views from providers, commissioners, researchers and public bodies that high quality person-level data is needed to run the health and social care system, and to support research…
>
> Because of the importance of earning public trust, the Review concluded that people should be able to opt out of their personal confidential data being used for purposes beyond their direct care unless there is a mandatory legal requirement or an overriding public interest.’

The NDG review also recommended that data flowing to ONS for the purposes of official statistics be exempt from any opt out, and DHSC have since adopted this position in policy – meaning the data will still flow even if someone has opted out of their data being shared beyond their direct care.

This, alongside the passing of the statistics section of the Digital Economy Act 2017(DEA) itself, shows the will of parliament is for the national statistics institute to be able to acquire and link data for better statistics to support better decisions in the public interest (and that there is trust in ONS to keep the data secure and be independent). Health data are in scope for the statistics strand of the DEA powers, and there was extensive and transparent public consultation during the parliamentary passage of these powers and the codes of practice which underpin them.

For ONS’s part, research was conducted into the public acceptability of ONS acquiring and linking admin data for the Admin Data Census programme. It found that:

- Around three quarters of people do not object to data held by other government departments being shared with ONS
- The public are supportive of data sharing when personal or public benefits can be demonstrated and are communicated effectively

- Data linking and storage are more acceptable if personal data are anonymised (i.e. name, address and other personal identifiers are removed)
- Public confidence in ONS is high, with 78 per cent stating that they think the organisation adequately protects the confidentiality of the personal information it collects
- When provided with reassurance about security and privacy, the public broadly support ONS re-using administrative data to produce statistics

With the introduction of the Digital Economy Act 2017 amendments to the SRSA 2007, which broadens ONS's powers to acquire data, further public acceptability work is planned.

In summary, all the above is of course a fine balancing act between privacy and research for the public good. The inclusion of health data in the statistics strand of the new DEA powers was important to enable such research, and in at least some cases we should consider that it may be unethical to not complete work – for example where it could save lives. On the other hand, we must be transparent about such uses, ensure the data are kept secure, and welcome external challenge. This ethics committee has an important role to play here.

B8 Principle 6: The access, use and sharing of data is transparent, and is communicated clearly and accessibly to the public

How will the findings of the research be disseminated? (max 500 words)

The HES data will be acquired under section 45A of the SRSA 2007, as amended by the Digital Economy Act 2017. All datasets acquired under section 45 of the SRSA will be explained and justified on the ONS website. In addition, NHS D will publish the Data Access Request Service application form that ONS will complete as part of the process to access the HES data. These actions are in line with the transparency principle within the code of practice that underpins these powers.

The non-disclosive aggregated tabulations of the results of this project will be initially shared with analytical colleagues in NHS D for quality assurance and we will consult them on disseminating the findings moving forward. No microdata will be shared outside of the small group of researchers in the Health Analysis and Life Events division in ONS who will access the de-identified linked data.

Outcomes of this work will be published on the ONS website, badged as experimental statistics in an article. We will work with the other organisations who publish analyses of HES to ensure an overall coherent picture is presented – for example, providing links to and a short explanation of the ONS research alongside their analysis.

Research outputs will be published by ONS as aggregated data tables and statistics with the relevant statistical disclosure controls applied.

Section C
Responsible owner and applicant details

C1 Responsible Owner

Full Name: Theodore Joloza	**Position:** Head of GSS Leadership and Policy Response Team, Health Analysis, ONS
Address:	**Email:**
	Telephone:
	Organisation: ONS

Declaration to be signed by the responsible owner

I have met with and advised the applicant on the ethical aspects of this project design (applicable only if the responsible owner is not the Applicant).

I understand that it is a requirement for all researchers accessing the data to have undergone relevant training and to have either relevant security clearances or approved researcher status in order to access the data.

I am satisfied that the research complies with current professional, departmental and other relevant guidelines.

I will ensure that changes in approved research protocols are reported promptly and are not initiated without approval by the National Statistician's Data Ethics Advisory Committee.

I will provide notification when the study is complete if it or fails to start or is abandoned.

I will ensure that all adverse or unforeseen problems arising from the research are reported in a timely fashion to the National Statistician's Data Ethics Advisory Committee.

I will consider all advice received from the National Statistician's Data Ethics Advisory Committee and should I be unable to implement any of the recommendations made, I will provide reasoning in writing to the Committee.

Signature: .... **Date:** ....

C2 Applicant Details (if applicant is not the responsible owner)

Full Name: Danielle Cornish	**Position:** Senior Research Officer
Address:	**Email**:
	Telephone:
	Organisation: ONS

Annex 1. Variables to be held on the analysis file

Question	How measured	Variable	Dataset
Do the outcomes of patients differ between those who live in rural versus urban areas?			
	Patients attending Accident & Emergency:		
	Admitted to hospital / Discharged / Referred on / Died / Left without treatment	AEATTENDDISP - Attendance disposal	HES
	Death within 30 days of attendance	ARRIVALDATE - Arrival date	HES
		DoD - date of death	Mortality
	Admitted patients:		
	Method of admission – emergency versus elective	ADMIMETH - Method of admission	HES
	Episode/spell duration – number of days	EPIDUR - episode duration	HES
		SPELDUR - duration of spell	HES
		EPIS_CFL - Episode start date check flag	HES
		EPIE_CFL - Episode end date check flag	HES
		POSTDUR - Postnatal stay	HES
	Bed days	BEDYEAR - Bed days within the year	HES
	Whether had operation(s)	OPDATE_NN - Date of operation	HES
	Death in hospital	DISMETH - Method of discharge	HES
	Death within 30 days of discharge	ADMIDATE - date of admission	HES
		ADM_CFL -Admission date check flag	HES
		DoD - date of death	Mortality
		Doddy - Day of death	Mortality
		Dodmt - Month of death	Mortality
	Delayed discharge	DISREADYDATE - discharge ready date	HES
		DISDATE - date of discharge	HES
		DIS_CFL - Discharge date check flag	HES
	Number of times admitted	EPIORDER - Episode order	HES
		ADMIFLAG - Admission episode flag	HES
	Maternity outcomes	BIRSTAT_N - Birth status	HES
		BIRWEIT_N - Birth weight	HES
		GESTAT_N - Length of gestation	HES
		HNEOIND - Healthly Neonate Indicator	HES
		MATERNITY_EPISODE_TYPE - Episode Type - Maternity	HES
	Outpatients:		
	Number of outpatient appointments	ATTENDKEY - Record identifier	HES
		ATTENDKEY_FLAG - Attendence Key Flag	HES
	Attended appointment or not	ATTENDED - attended or did not attend	HES
		ATENTYPE - attendance type	HES
	Level of priority given	PRIORITY - priority type	HES
If any differences are found: do these vary by demographic factors?			
	Age	ACTIVAGE - age at activity date	HES
		ADMIAGE - Age on admission	HES

	ARRIVALAGE - Age on arrival	HES
	STARTAGE - Age at start of episode	HES
	ENDAGE - Age at end of episode	HES
	MATAGE - Mother's age at delivery	HES
	Agegroup 2 - age	Mortality
	Age in years - age	Mortality
	AGE - Age	Census
Sex	SEX - Sex of patient	HES
	SEXBABY_N - Sex of baby	HES
	SEX - Coded sex of deceased	Mortality
	SEX - sex	Census
Ethnic group	ETHNOS - Ethnic category	HES
	ETHRAW - Ethnic character (audit version)	HES
	ETHRAWL - Ethnic category (audit version)	HES
	ETHNICID - Ethnic group	Census
	ETHPUK11 - Ethnicity	Census
Religion	RELIGION - Religion	Census
	RELPUK11 - Religion (Grouped)	Census
Whether patient has carer support	CARERSI - Carer support indicator	HES
Living in care/nursing home	POSITION - Position in a communal establishment	Census
	CECTMCEWS11 - Communal establishment management and type	Census
Tenure and accommodation type	TENHUK11 - Tenure of household	Census
	TYPACCOM - Accommodation type	Census
Deprivation	IMD04 - IMD Index of Multiple Deprivation	HES
	IMD04_DECILE - IMD Decile Group	HES
	IMD04C - IMD Crime Domain	HES
	IMD04ED - IMD Education Training and Skills Domain	HES
	IMD04EM - IMD Employment Deprivation Domain	HES
	IMD04HD - IMD Health and Disability Domain	HES
	IMD04HS - IMD Barriers to Housing and Service Domain	HES
	IMD04I - IMD Income Domain	HES
	IMD04IA - IMD Income affecting Adults Domain	HES
	IMD04IC - IMD Income affecting Children Domain	HES
	IMD04LE - IMD Living Environment Domain	HES

	IMD04RK - IMD Overall Rank	HES
Region of England	RESGOR_ONS - Government office region of residence (ONS)	HES
	NER_TREATMENT - NHS England region of treatment	HES
Occupation / socioeconomic status in 2011	NSSEC -National Statistics Socio-economic Classification	Census
	SCGPUK11 - Approximated social grade	Census
	OCCPUK112 - Occupation (Sub-major group)	Census
General health in 2011	HEALTH - General Health	Census
Long-term health problem or disability in 2011	DISABILITY - Long-term health problem or disability	Census
Migrant status	LRESPUK11 - Length of residence in the UK	Census
	COB - Country of Birth	Census
	Ctryir - Country indicator code for usual residence of deceased	Mortality
	YRARRPUK11 - Year of arrival in the UK	Census
Marital status	MARSTAT - Marital status of deceased	Mortality
	MARSTAT - Marital and Civil Partnership status	Mortality
(for deceased patients) occupation / socioeconomic status	Empsecdm - Employment status (ONS code) of deceased or mother of deceased juvenile for SOC2000	Mortality
	SOC2KDM - Standard occupation classification (2000) for deceased or mother of deceased child	Mortality
	SECCATDM - NS SEC operational category for deceased or mother of deceased.	Mortality
If any differences are found: do these vary by contextual factors?		
Arrival mode – brought in by ambulance/helicopter versus other mode	AEARRIVALMODE - arrival mode	HES
Source of admission – usual place of residence, care home, other hospital etc.	ADMISORC - Source of admission	HES
Source of referral – e.g. self-referral, GP, Police etc.	AEREFSOURCE - Source of referral for A&E	HES
Diagnosis. *Note: not available for all years*	DIAG_NN - A&E diagnosis	HES
	DIAG_COUNT - Count of diagnoses	HES
	DIAG_NN - All Diagnosis codes	HES
	CAUSE - cause code	HES
	AEPATGROUP - patient group	HES
Destination on discharge -e.g. care home	DISDEST - Destination on discharge	HES
Birth/delivery episodes versus other types	EPITYPE - episode type	HES

	MATERNITY_EPISODE_TYPE - Episode Type - Maternity	HES
(for deceased patients) underlying cause of death	Fic10und - Final underlying cause of death (ICD10)	Mortality
	Fimdth10 - Final manner of death (ICD10)	Mortality
	Ic10f001-15 - Final ICD10 code	Mortality
(for deceased patients) place of death (e.g. communal establishments)	PCDPOD - Postcode of place of death	Mortality
	NHSIND - NHS establishment indicator	Mortality
	PCDR - place of usual residence of deceased	Mortality
	Esttyped - Communal Establishment type where death occurred	Mortality
	Cestrss - Communal establishment code	Mortality
	Cestrssrs - Communal establishment code	Mortality
	Ceststay - Duration of stay in communal establishment	Mortality
Does the likelihood of changing health status (perceiving oneself as having good versus poor general health) since the 2011 Census differ by rurality?		
General health in 2011	HEALTH - General Health	Census
Long-term health problem or disability in 2011	DISABILITY - Long-term health problem or disability	Census
Admitted to hospital / Discharged / Referred on / Died / Left without treatment	AEATTENDDISP - Attendance disposal	HES
Bed days	BEDYEAR - Bed days within the year	HES
Episode/spell duration – number of days	EPIDUR - episode duration	HES
	SPELDUR - duration of spell	HES
	EPIS_CFL - Episode start date check flag	HES
	EPIE_CFL - Episode end date check flag	HES
	POSTDUR - Postnatal stay	HES
Year and rurality breakdowns		
Geographies *(Note: only output area is needed for assigning rural-urban classification, but all other geography variables included are aggregations of OA so are no more sensitive. Having access to these will mean geography aggregations don't need to be re-calculated based on OA)*	OACODE11 - Census Output Area, 2011	HES
	CURRWARD_ONS - Current electoral ward (ONS)	HES
	LSOA11 - Lower Super Output Area (LSOA11)	HES
	MSOA11 - Middle Super Output Area, 2011	HES
	RESCTY_ONS - County of residence (ONS)	HES
	RESGOR_ONS - Government office region of residence (ONS)	HES
	RESLADST_ONS - Local authority district (ONS)	HES

	Year of record	INYRFLAG - In Year flag	HES
		PARTYEAR - Year and month of data	HES
		FYEAR - Financial Year	HES
		regyr - year of registration of the death	Mortality
		DoD - date of death	Mortality
		DoR - Date of registration	Mortality
Administrative variables			
	For linkage and quality assurance	AEKEY - record identifier	HES
		AEKEY_FLAG - AEKEY Flag	HES
		ATTENDKEY - Record identifier	
		ENCRYPTED_HESID - Encrypted HES ID	HES
		EPIKEY - record identifier	HES
		HESID_ORIG - Patient ID - HES generated (original)	HES
		MATCH_RANK - MATCH_RANK	HES
		MATCHID - Patient identifier (HES generated) - basis of match	HES
		REGDETS - ONS unique identifier for death records	
	For exclusion of APC episodes where patient is still in hospital	DISFLAG - Discharge episode flag	HES
	For exclusion of APC episodes where episode is unfinished	EPISTAT - Episode status	HES
		FAE - Finished Admission Episode	HES

Annex 2: Overview of Data Access Platform (DAP)

Statistical data processing

Advances in technology and statistical methods create enormous opportunities to use new sources of data for the public good. With the Digital Economy Act, there is now a legal framework for ONS to increase its data sources from external organisations, both within Government and in the commercial sector.

To support this data transformation, ONS has developed an integrated, single environment – the Data Access Platform (DAP) – to host data and analytical applications. DAP facilitates the processing and analysis of more data in richer and more complex forms, integrating administrative and commercial data sources supported by appropriate methods and standards.

Data security

Security of data is a key priority for ONS. DAP takes a robust approach to security that is risk-based and holistic, covering people, process and technology. DAP security is based on two key security governance and management layers. The first layer is a set of security principles to inform design and operation; the second layer distils these principles into specific security controls within the platform.

All security controls have been developed following recognised security standards and guidance from within Government, including from the Cabinet Office, the National Cyber Security Centre and the Centre for the Protection of National Infrastructure. They also adhere to international standards and best practice, as set by ISO 27001 and the Information Security Forum.

The key security elements of DAP are:

- Overarching security controls, covering all aspects of physical, technical, people and process security within the ONS framework
- A system designed to securely store sensitive data in bulk;
- Data specifically assessed for sensitivity to direct protection requirements and access management;
- Cloud-based, virtualised environment segregated from the ONS organisational network;
- Virtual desktop infrastructure with dedicated separate accounts for user and IT management access;
- Comprehensive monitoring of user and system activity to detect security events and alerts;
- Dedicated ingest route supported by governance, business approval processes and a secure staging zone;
- Dedicated export route supported by governance, business approval processes and a single transfer mechanism.

NSDEC(19)07

UK Statistics Authority
National Statistician's Data Ethics Advisory Committee

Environmental and socio-economic impact assessment for siting geological disposal facilities

This project is undergoing major revisions and will be published in due course

NSDEC(19)08

National Statistician's Data Ethics Advisory Committee

Application for Ethical Review

Please consult the guidance document before filling in the application form

Project Title

Please provide a title indicative of the project

Urban Redevelopment Baselining for Urban Redevelopment Projects

Start Date:	15.06.19	**End Date:**	01.08.19

Project Sponsor(s)

Please list the project sponsor(s)

Lendlease

Project Summary

Please provide a brief high level summary of the research giving necessary background (max 250 words)

Lendlease has commissioned Simetrica to conduct baseline socioeconomic analysis for their Euston train station and Silvertown Quays urban redevelopment projects in London.

These development projects are aimed at improving the transport infrastructure and housing supply in their respective areas, with a focus on economic inclusion, community development and green spaces.

Lendlease value the impact of these urban redevelopment projects on key social outcomes including reduced noise pollution and improved journey times. They also have a mandate from the local authorities involved to assess the social context around the redevelopments.

Simetrica's work will provide baseline values that allow them to anticipate the impact of planned development. This baselining will also allow them to monitor the creation of social value over time, and future applications may be made to extend the project into longitudinal analysis when newer waves of data become available.

Simetrica will use ONS Annual Population Survey data for the analysis, which will provide household and person data, including information on subjective wellbeing. Publicly available air pollution and traffic data (from DEFRA and Highways England, respectively) will be used to map APS data to traffic and pollutant contours by postcode.

The analysis will reveal wellbeing inequalities within each project area and identify which outcomes have the largest impact on individual wellbeing. Wider analysis will also shed light on possible environmental concerns.

Section A
Project Details

A1 Legal gateways

Please provide the assessment of the legal gateways of the project as provided by Legal Services

The ONS Approved Researcher scheme is the legal gateway being used to access the ONS-owned data. This complies with the Statistics and Registration Service Act 2007.

The ONS Microdata Release Panel (MRP) have approved the proposal, as there is a legal gateway to access the data, it is deemed appropriate use of ONS data and a public benefit has been demonstrated.

A2 Ethical approval

Has the project been reviewed or is it expected to be reviewed by another ethics committee?	○ **Yes** ● **No**
If <u>Yes</u> please provide the name of the committee, the outcome and the date approved	

A3 Proposed site of research select all that apply

☐ ONS	☐ ADRC - England
☑ SRS (formerly VML)	☐ ADRC - Scotland
☐ HMRC Data Lab	☐ ADRC - Northern Ireland
☐ Other	☐ ADRC - Wales
please specify	

A4 Data subjects to be studied

Does the study include all subsections of the population?
(i.e. all ages, sex, ethnic groups etc)

● **Yes** ○ **No**

If no please detail which subsections with justification(s) below:

Subsections of the population (including vulnerable groups) the project focuses on:

N/A

Justification for focusing on these subsections or groups:

N/A

A5 Please provide details of the research protocol or methodology (e.g. data linkage, web scraping etc) (max 500 words)

Simetrica's approach will build on methodologies they have applied in previous studies such as the A303 Stonehenge Stage 1 contingent valuation and the quality of life impact study of Heathrow and Gatwick Airport expansion. These include ongoing projects previously approved by the committee regarding water flooding and roadworks (for Anglian Water*), Superfast broadband (for DCMS[1]) and odour (for Thames Water*). Each of these linked subjective wellbeing from the Annual Population Survey (APS) to other geographic data.

The Annual Population Survey is a subset of the national Labour Force Survey that contains household and person information. It includes four wellbeing-related questions[2] which ask respondents[3] to rate, on a scale of 1 to 10:

- How satisfied they are with their life nowadays
- To what extent they feel the things they do are worthwhile
- How happy they felt yesterday
- How anxious they felt yesterday

The project will also use publicly available air pollution and traffic data from Department for Environment, and Highways England respectively. This will be imported into the Secure Research Service (SRS) so that the researchers can match it to the APS data. ONS statistical disclosure controls (SDC) will be applied to the outputs to ensure that no matching of data results in disclosive outputs.

Air pollution data will be matched to APS by year and geography (matching DEFRA locational grid to household postcodes in the APS). Traffic data is defined along road links. Based on a household's postcode, Simetrica will identify the most relevant road links and match them based on distance (limited to the relevant time period).

While the infrastructure improvements will affect residents and commuters both, the focus of this use of public data is, due to limitations of the APS, on nearby residents only (for example, in Camden or Newham). In the final reports, analysis will be supplemented by primary data collection being conducted in both project areas, where residents will be distinguished from visitors, commuters or people working in the area. Using both data sources together will result in a holistic picture.

By using postcode-level data from the APS, the researchers will be able to map them within well-defined traffic and pollutant contours. They will assess wellbeing variables as well as types of paid work and socioeconomic indicators in the area. This will serve as a benchmark

---

* Both previously approved by NSDEC

[1] Department for Culture, Media and Sports

[2] https://www.ons.gov.uk/file?uri=/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/labourforcesurveyuserguidance/vol22019final.pdf page 23

[3] Over 16s only

for future analysis on the impact of Lendlease's development work in the areas and to anticipate the impacts of the planned developments.

Initially Simetrica will be using the data provided to enrich their benchmarking studies. If the data proves robust enough they hope to produce a future project application to take advantage of new waves of data for longitudinal study.

A6

Data use

Please specify the data used **by the research team** including any timeframes *e.g. LFS data 2014-15*

Data Level

Please specify the name of the data set

Type of data	**Aggregate Data**	**Identifiable Data**	**De-identified personal data**	**Anonymised/ pseudo anonymised**
Administrative data *(please specify, e.g. Patient Register 2011, School Census 2012 etc, in the relevant options adjacent)*	DEFRA Air pollution data[4] Highways England traffic data[5]			
Big Data *(please specify e.g. Twitter data, smart meters and mobile phones, in the relevant options adjacent)*				
Survey Data *(please specify e.g.LFS, BRES, etc in the relevant options adjacent)*			APS: Wellbeing 2011-2015 APS: Households 2004-2016 APS: Person 2004-2017	
Census Data *(please specify year, e.g. Census 2011 in the relevant options adjacent)*				

---

[4] https://uk-air.defra.gov.uk/data/

[5] https://data.gov.uk/dataset/9562c512-4a0b-45ee-b6ad-afc0f99b841f/highways-england-network-journey-time-and-traffic-flow-data

Other *(please specify e.g. Ordinance Survey Address register in the relevant options adjacent)*				

Section B
Assessment against NSDEC ethical principles

B1 Principle 1: The use of data has clear benefits for users and serves the public good.

Please outline the proposed benefits of the project (max 500 words)

The Euston station redevelopment project aims to provide a framework for the long-term development of Euston station. The focus will be on improving rail services for the passengers travelling from London to the Midlands, the North of England and Scotland. The project also aims to provide housing for circa 2,000 people and create up to 19,000 jobs around the Euston area in the Borough of Camden. In Silvertown, Lendlease has approval for 7 million square feet of residential and commercial use, including the restoration of the historic Millenium Mills and the grade 2 listed Silo-D.

These projects are aimed at improving the transport infrastructure and housing supply in their respective areas, with a focus on economic inclusion, community development and nature and development.

Lendlease value the impact of these urban redevelopment projects on key social outcomes (economic inclusion, community development and nature and development), with specific statistical reference to reduced noise pollution and improved journey times. Simetrica will carry out a quasi-experimental estimate of the wellbeing impact of numerous factors associated with urban redevelopment projects. The initial use of the results will be to aid Lendlease's planning of the redevelopments.

Within each geographical project area, the analysis will reveal possible wellbeing inequalities and identify which outcomes have the largest impact on individual wellbeing. The wider analysis they contribute to will also shed light on environmental concerns in those areas (eg: air pollution levels).

Having specific statistical evidence will inform the wellbeing and quality of life aspects of Lendease's development strategy. If the studies indicate particular risks or opportunities, then the redevelopment work can pivot accordingly.

Project outputs will also serve as a benchmark for the areas of interest to inform future asessments of the impact of the redevelopment work. These assessments will identify any dips in social value over time for attention, as well as making Lendlease accountable to the local authorities commissioning the redvelopments.

The outcome of this project will be shared with the relevant local authorities to evidence Lendlease's obligations to address local quality of life. This will ensure transparency in the contractual relationship. Information on local wellbeing indicators will be of further interest to the local authority and can be used to inform their own local development plans.

B2 Please outline any intended future use for products (such as linked data sets or tools) produced as a result of the research and how they will be accessed. (max 250 words)

The findings will be used to produce an Integrated Social Value Assessment tool to be distributed among Simetrica and Lendlease employees. This tool will assess the impact of different factors on wellbeing, allowing Lendlease to monitor the creation of social value over time.

The data held within the tool will be the aggregated statistical outputs of the project and not disclosive of individuals.

B3 Principle 2: The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.

Please outline how data security, confidentiality and informed consent is safeguarded in this project (max 500 words)

Access to, analysis and use of the data will only take place within the ONS Secure Research Service (SRS) environment. All outputs will be checked by the ONS Statistical Support team to ensure that they do not disclose confidential or personal identifiable data and to confirm that they are proportional and necessary to the project. This process ensures that disclosure control and the confidentiality of data subjects is protected. Outputs that identify small groups or individuals will not be allowed out of the SRS environment.

The researchers are accredited as ONS Approved Researchers. ONS will have sight of the final report before it is published.

B4 Principle 3: The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.

Please describe how the any risks from new technologies are been mitigated as well as any quality assurance activities in the project (max 500 words)

No new technologies are being employed.

B5 Principle 4: Data used and methods employed are consistent with legal requirements such as the Data Protection Act, the Human Rights Act, the Statistics and Registration Service Act and the common law duty of confidence

Please describe the legal frameworks pertinent to this project (max 500 words)

Access to the potentially disclosive data will be in the Secure Research Service environment and via an approved legal gateway (Approved Researcher scheme). This complies with the Statistics and Registration Service Act 2007 and the methods used comply with the principles in the Data Protection Act.

B6 Collaboration and Sponsors

Please describe the project sponsors and the **legal gateways** to acquire, process use and share their data

List of Collaborators/Sponsors	**Details and relevant documentation relating to collaboration (you may attach copies of relevant documentation)**
Simetrica	Contracted research organisation
Lendlease	Sponsor

B7

Principle 5: The views of the public are considered in light of the data used and the perceived benefits of the research

Please list any public engagement activities (max 250 words)

Whilst the views of the public have not been sought with regards to the research, there is a clear public benefit for the analysis (as described in B1 above). A public consultation on the Approved Researcher scheme in 2015/16 recommended that commercial organisations should be allowed to access ONS research data where there is a clear public benefit.

B8

Principle 6: The access, use and sharing of data is transparent, and is communicated clearly and accessibly to the public

How will the findings of the research be disseminated? (max 500 words)

ONS has a commitment to transparency and publishes a register of all Approved Researchers on their website. This includes their organisations, current projects, sponsors and the datasets being used.

Results of the study will be published on the Simetrica or Lendlease websites in the form of short reports. Outcomes will also be published in papers authored by WPI economics. Analysis of the Euston project will be published in the Economics of Transportation journal. Analysis of the Silverton project will be published in the Journal of Urban Regeneration and Renewal. The open publication of the project's outcomes and methodologies will allow third parties to critique and, if necessary, use public data to replicate WPI's findings.

Section C
Responsible owner and applicant details

C1 Responsible Owner

Full Name: Sebastien Peytrignet	**Position:** Lead Researcher
Address:	**Email:**
	Telephone:
	Organisation: Simetrica

Declaration to be signed by the responsible owner

I have advised the applicant on the ethical aspects of this project design (applicable only if the responsible owner is not the Applicant).

I understand that it is a requirement for all researchers accessing the data to have undergone relevant training and to have either relevant security clearances or approved researcher status in order to access the data.

I am satisfied that the research complies with current professional, departmental and other relevant guidelines.

I will ensure that changes in approved research protocols are reported promptly and are not initiated without approval by the National Statistician's Data Ethics Advisory Committee.

I will provide notification when the study is complete if it or fails to start or is abandoned.

I will ensure that all adverse or unforeseen problems arising from the research are reported in a timely fashion to the National Statistician's Data Ethics Advisory Committee.

I will consider all advice received from the National Statistician's Data Ethics Advisory Committee and should I be unable to implement any of the recommendations made, I will provide reasoning in writing to the Committee.

Signature: .... **Date:** ....

C2 Applicant Details (if applicant is not the responsible owner)

Full Name:	**Position:**
Address:	**Email:**
	Telephone:
	Organisation:

NSDEC(19)09

UK STATISTICS AUTHORITY

NATIONAL STATISTICIAN'S DATA ETHICS COMMITTEE

Data Protection and Ethics Compliance

Purpose

1. The purpose of this paper is to outline the role of the Data Protection Compliance Team in providing assurance that the ONS research projects considered and approved by NSDEC continue to adhere to the ethical principles and recommendations throughout the project life span.

Recommendations

2. Members of the NSDEC are invited to:

   i. Consider the projects listed for review and approve high-level Data Ethics Compliance Review Plan for 2019/20 at **Annex A**; and
   ii. agree the process for the reporting, monitoring and escalation to the committee and the National Statistician on the level of compliance with the ethical principles.

Background

3. In order to provide the organisation with assurance that it is complying with data protection legislation, a function has been set up to support the Data Protection Officer to provide data protection auditing and compliance monitoring services. An audit/compliance reporting programme has been developed to cover all activities across the Office from a data protection perspective.

4. Good ethical practice and data protection legislation promote the appropriate use of data with appropriate safeguards. The organisation and the projects (involving the use of personal data) it sponsors, are required to adhere to data protection legislation where appropriate. However, good ethical practice ensures that the access, use and sharing of public data, for research and statistical purposes by the ONS, is ethical and for the public good.

5. Whilst this is activity is beyond the strict scope of personal data/data protection legislation, we are keen to ensure that the Office's activities in certain areas are in line with NSDEC's ethical principles as an additional layer of assurance around how we handle personal, or sensitive personal, data.

Discussion

6. We intend to provide assurance to National Statistician that the ethical principles are being adhered to in the development and application of project proposals via a quarterly report. We will therefore undertake a programme of systematic reviews of research projects which have been considered by the committee.

7. To give the committee an idea of the type of work the Data Protection Compliance Team will undertake, we have carried out a short review into the implementation of recommendations made by NSDEC in the approval of the Crime Survey for England and Wales (CSEW) in

January 2018 (**Annex B**). However, we have not carried out a complete review of ethics against the CSEW, just the NSDEC's recommendations to the project proposal itself.

8. We propose, where necessary, to provide recommendations to projects where we identify non-compliance with NSDEC's ethical principles. However, we will also seek to report on the justification for not following the ethical guidance given by the committee.
9. We will provide quarterly updates on the status of previous recommendations to NSDEC. These recommendations will be monitored by the Data Protection Compliance Team until completion, and updates requested from action owners identified in management responses.
10. Through discussion with the NSDEC secretariat, we have identified a number of projects to initially review, the first of which are included in the work programme at **Annex A.**

Conclusion

11. The Data Protection Team's ethical compliance reviews will provide NSDEC and the National Statistician with assurance that the project proposals which have been reviewed by the committee continue to adhere to the ethical principles and recommendations throughout the project life span. They also offer the opportunity to give one view on the impact of the committee. We would also welcome the committee and the National Statistician to suggest any specific reviews into ethical compliance of ONS projects.

Rhys Nadin, Data Protection Compliance Manager, April 2019

List of Annexes

Annex A	**Data Ethics Compliance Reviews 2018/19 – 2019/20**
Annex B	**Data Ethics Compliance Review of Crime Survey of England and Wales**

Annex A Data Ethics Compliance Reviews 2018/2019 – 2019/20

Proposed Review Area	Project Approved by NSDEC	2019/20
Crime Survey England and Wales: child experience of online abuse	January 2018	Q1
Small area income estimates CTP: Estimating income from administrative data	January 2017	Q2
Students' suicide statistics	January 2018	Q3
Study on the effects of migrant workers on London's economy	January 2017	Q4

Data Ethics Compliance Review Report

Review of Crime Survey for England and Wales

April 2019

Table of Contents

Executive Summary

Introduction

A data ethics compliance review of the Crime Survey for England and Wales was agreed as part of the 2018/19 Data Protection Compliance Programme and agreed with the Data Protection Officer.

The National Statistician's Data Ethics Advisory Committee (NSDEC) provides ethical advice that the access, use and sharing of public data for research and statistical purposes is ethical and for the public good. In reviewing project and policy proposals, NSDEC provides advice to the National Statistician on the basis of the following six ethical principles:

1. The use of data has clear benefits for users and serves the public good.
2. The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.
3. The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.
4. Data used, and methods employed, are consistent with legal requirements such as the Data Protection Act, the Human Rights Act, the Statistics and Registration Service Act and the common law duty of confidence.
5. The views of the public are considered in light of the data used and the perceived benefits of the research.
6. The access use and sharing of data is transparent and is communicated clearly and accessibly to the public.

In January 2018, the ONS Crime Team presented a re-submitted application to review the Crime Survey for England and Wales (CSEW), a ONS survey conducted by a third party (Kantar Public) on behalf of the ONS. This ethical review was focused on the abuse during childhood module, which is included in the CSEW for the April 2018 to March 2019 survey year. At the meeting, members agreed that significant improvements had been made in the application to support its compliance with the ethical principles outlined above. However, it was suggested that the following further improvements should be made:

I. ONS should collaborate with the National Society for the Prevention of Cruelty to Children (NSPCC) to strengthen the language in the supporting material and advance letter to make clear the voluntary nature of the survey;
II. members requested more clarity on how parental consent was obtained for survey participants between 16-17 years old;
III. the committee asked for further assurance that survey participants were only re-contacted when they had provided their permission to do so; and
IV. members recommended that it be made clear to respondents that information would be passed to other research organisations only if respondents opted to be re-contacted for further research.

NSDEC approved the project design around the abuse during childhood module for the Crime Survey for England and Wales subject to these minor revisions.

Scope

The scope of this review is to provide assurance to NSDEC that the ONS Crime Team have implemented the following actions agreed at NSDEC in January 2018:

I. ensure that informed and voluntary consent is obtained, especially when parental consent is required, and this is clearly communicated to participants in the advance letter;
II. clarify in the application how parental consent is acquired for participants aged between 16-17 years old;

III. provide assurance that the confidentiality of data subjects is protected and that participants cannot be re-contacted without their permission after completing the survey; and
IV. consult with NSPCC to make clear the voluntary nature of the survey in the advance letter.

The limitations for the scope of the review were:

- This review focussed on the actions agreed above and is not a complete review of ethics against the CSEW.

Conclusion

The Crime Survey for England and Wales has taken the advice of the committee on board, and the review has found that all the actions agreed at the NSDEC meeting in January 2018 have been satisfactorily implemented. The application for ethical review has been updated to clarify how parental consent is acquired for participants aged between 16-17 years old. Informed and voluntary consent has been clearly communicated to participants and confirmed that consultation with the NSPCC has been undertaken. However, we identified that from April 2019, the word ‘voluntary’ was removed from the advance letter, as response rates for the survey fell by approximately 2%. In addition, the crime survey advance letter was not consistent with approach taken by other ONS household surveys and the recommended wording of the ONS Materials Strategy Advisory Group (they provide oversight of the development of respondent-facing materials used across ONS). As a result, the wording was changed to the agreed harmonised approach used on all ONS household surveys. Whilst it still makes clear that the survey is voluntary it no longer uses the word ‘voluntary’ itself. The relevant wording was changed from “We rely on people’s co-operation with this voluntary survey to produce information about the extent of crime in England and Wales, which is used to help reduce crime and make communities safer. Everything you tell us is treated as confidential” to “We rely on people’s help in taking part to be able to produce information about the extent of crime in England and Wales. This information is used to help reduce crime and make communities safer. Everything you tell us is treated as confidential”.

Detailed Findings

Ref	ACTION	FINDINGS	IMPLEMENTED
i	Ensure that informed and voluntary consent is obtained, especially when parental consent is required, and this is clearly communicated to participants in the advance letter.	Through review of the advance letter, we confirmed that it includes information on what the survey is about and the voluntary nature of the survey. The letter also encourages respondents to read guidance material on the survey website around its content, which includes: • Why the respondent has been chosen; • the type of questions asked; • the voluntary nature of the survey, • the obtaining of permission from a parent or guardian before an interview is conducted with anyone aged 10-15, and; • how the results of the survey are used. An accompanying leaflet has also been included with the letter outlining the types of questions asked and the requirement to obtain consent from a parent or guardian before interviewing anyone under the age of 18. Additionally, a letter is is given to all parents and guardians of 16- and 17-year old respondents (when selected for interview) in advance of the interview taking place which outlines: • the obtaining of permission from a parent or guardian before an interview is conducted; • the type of questions asked; • Confidentiality, and; • Contact details.	Yes
ii	Clarify in the application how parental consent is acquired for participants aged between 16-17 years old.	Through review of the application, we confirmed that it had been updated to include how parental consent is acquired if a 16 and 17-year-old is selected as the main respondent for a household. The application states that consent must be obtained from a parent or guardian before the interview can take place. Permission is verbal but will be recorded on the survey instrument by the interviewer. Additionally, through our testing in ref i. we confirmed that a letter is given to all parents and guardians of selected respondents outlining the requirement to obtain consent.	Yes
iii	Provide assurance that the confidentiality of data subjects is protected and that participants cannot be re-	The CSEW reviewed the assurances given to respondents about the information they provide and in relation to being re-contacted. The CSEW repeatedly gives assurances about the data and its confidentiality, including assurances that participants cannot be re-contacted without permission after completing the survey.	Yes

	contacted without their permission after completing the survey.	In the advance letter the CSEW states: *“For more information about the types of questions included, and information about how we keep your data safe, you can also visit www.CrimeSurvey.co.uk.”* The website links to a page dedicated to data security and also includes a dedicated Crime Survey Privacy Policy page. In the Crime Survey leaflet that sits alongside the advance letter states: *“The survey is confidential - The information that is collected will be used solely for research purposes. Your name and address details are kept separate from your answers and will not be passed on to any other organisation or matched to your responses without your specific permission. No individual will be identifiable from the results. Your answers will be combined with others that take part. You will not receive any ‘junk mail’ as a result of taking part”.* In the Adult reassurance leaflet it states: *“The survey is confidential - All information is used solely for research purposes. Your name and address details are kept safely and separately from your answers and are not passed onto any other organisation. No individual person will be identifiable from the results as all the answers are combined. You will not receive any ‘junk mail’ as a result of taking part”*	
iv	Consult with NSPCC to make clear the voluntary nature of the survey in the advance letter.	The advance letter for the 2017/18 CSEW was amended to read: *“We rely on people’s co-operation with this voluntary survey to produce information about the extent of crime in England and Wales, which is used to help reduce crime and make communities safer. Everything you tell us is treated as confidential”.* However, due to this change, response rates for the survey dropped by 2%. KANTAR (the field contractors for the survey) requested that the word ‘voluntary’ be removed from the letter and the voluntary nature of the survey made clear using different language. In addition, the crime survey advance letter was not consistent with approach taken by other voluntary ONS household surveys and the recommended wording of the ONS Materials Strategy Advisory Group. The wording was therefore changed from April 2019 to the agreed harmonised approach used on all voluntary ONS household surveys. Whilst it still makes clear that the survey is voluntary it no longer uses the word itself: *“We rely on people’s help in taking part to be able to produce information about the extent of crime in England and Wales. This information is used to help reduce crime and make communities safer. Everything you tell us is treated as confidential”.*	Yes

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

Overview from Economic Statistics

Presentation

Phil Wales

UK Statistics Authority

National Statistician’s Data Ethics Advisory Committee

Social Survey Transformation

Presentation

Ian O’Sullivan

Overview from the Data Science Campus

Presentation

Peter Fullerton

UK Statistics Authority
National Statistician’s Data Ethics Advisory Committee

Overview from Public Policy Analysis
Presentation
Hugh Stickland

Any other business