UK-REACH
Understanding COVID-19 outcomes for ethnic minority healthcare workers

THE UNIVERSITY of EDINBURGH
Edinburgh Law School

HEALTHCARE WORKERS' DATA AND COVID-19 RESEARCH

UK-REACH Work Package 3 report on the legal and ethical implications of using data concerning healthcare workers and ethnicity in research during the COVID-19 pandemic

Abstract

This report analyses the legal and ethical issues arising from the collection and use of data about broad aspects of healthcare workers' lives, and provides recommendations to promote use of data that is legally and ethically acceptable. It considers the relevant legal framework and how it applies to UK-REACH. It discusses the implications of carrying out large-scale data linkage and analysis in a trusted research environment; to what extent such data, once de-identified, can be considered anonymous; and how UK-REACH should respond from a legal perspective. It also highlights the particular ethical issues arising in the context of research on healthcare workers' ethnicity and the COVID-19 pandemic. It identifies what values are at stake, and how, in consideration of these values, UK-REACH can adopt a defensible ethical approach to its research. Ultimately, this report provides a framework for using and linking sensitive data in this project in a way that is ethically, legally, and socially acceptable.

Ms Ruby Reed-Berendt, Research Associate
Dr Edward Dove, Lecturer in Health Law and Regulation
University of Edinburgh Law School

Acknowledgements: The authors would like to thank Prof Graeme Laurie for his comments on a previous draft.

BREATHE Health Data Research Hub

HDRUK Health Data Research UK

UKRI UK Research and Innovation

Contents

1. Introduction

It is now well established that healthcare workers (HCWs) from Black, Asian and Minority Ethnic (ethnic minority) groups are disproportionately affected by COVID-19. Collaborating with the General Medical Council (GMC), Nursing and Midwifery Council (NMC) and other professional bodies/associations, over a period of 12 months, the UK-REACH project is undertaking a mixed-methods study with the aim of investigating if, how, and why, ethnicity affects COVID-19 clinical outcomes in HCWs. The project is providing evidence through five interlinked work packages (WPs):

WP1 is undertaking an expedited linkage and analysis of anonymised professional registration (e.g. GMC, NMC), employment and NHS datasets, within a Trusted Research Environment (SAIL databank), to calculate the incidence of, and outcomes from, COVID-19 amongst HCWs.

WP2 is undertaking a longitudinal cohort study of ethnic minority HCWs, nested within the GMC/NMC datasets. At baseline, survey information will be collected about demographics, HCW role, interaction with COVID-19 patients, social circumstances and physical/mental well-being, Follow-up surveys will be collected over approximately twelve months to capture changes in survey responses to the ongoing and potentially further COVID-19 pandemic waves. This may include biological samples will be collected for genomic and serological profiling.

WP3 is undertaking research to understand and address legal, ethical and acceptability issues around data protection, privacy and information governance associated with the linkage of professionals' registration data and healthcare data. (This report is the principal output from this work package.)

WP4 is undertaking qualitative interviews and focus groups of ethnic minority HCWs to understand risk perceptions, support and coping mechanisms in relation to COVID-19.

Finally, WP5 comprises a multi-professional, national stakeholder group including the GMC/NMC, Royal Colleges, ethnic minority Professionals' Associations, and ethnic minority HCWs, and is facilitating rapid dissemination and translation of the research findings for HCWs, employers, and policymakers.

The UK-REACH project involves the use of different datasets and their linkage to healthcare data. This activity may be considered sensitive and raises a diverse set of ethico-legal issues. In what follows, we identify those issues and propose ways in which this work can be done in a way that is ethically, legally, and socially acceptable. We begin with an analysis of the relevant legal framework, before turning to analysis of relevant ethical considerations. We conclude with specific recommendations for UK-REACH.

2. The legal framework

UK-REACH must ensure that it is meeting its legal obligations with respect to the data it will collect and utilise relating to individuals participating in the project, be it directly through questionnaires, interviews, and/or focus groups, or indirectly through the processing of data concerning them. Three key legal frameworks (which we address in turn), with application across the UK, must be considered as part of this compliance obligation:

- The UK common law duty of confidentiality as applied to information about UK-REACH's participants and data subjects.
- Data protection law in respect of the processing of data gathered as part of the project, namely the UK General Data Protection Regulation (GDPR) and UK's Data Protection Act 2018.
- Relevant human rights of the project's data subjects and participants, specifically those rights under the UK's Human Rights Act 1998.

2.1. Duty of confidentiality

2.1.1. The common law

The common law duty of confidentiality holds that where an individual has a reasonable expectation of privacy with respect to information given in confidence, and where a person receiving that information knows or ought to know that the other can reasonably expect their privacy to be protected, such information should be kept confidential and not disclosed except under certain conditions.[1] The duty may arise through professional or contractual relationships, as well as where information is imparted in circumstances importing an obligation of confidence (e.g. employer-employee relationships, research involving human participants where information of a confidential nature is disclosed to researchers). The obligation is not to pass on the information to any third parties without justification and not to make the information (publicly) available in identifiable form.

In healthcare, the common law duty of confidentiality (owed by health professionals towards their patients) is not only well-established in case law,[2] but has been codified by professional regulators, including the GMC[3] and NMC.[4] This requires practitioners to keep patients' information confidential and only to disclose information in specific circumstances (for example, with the patient's consent or with public interest justifications). Confidential data within the health service should be managed in accordance with the eight Caldicott Principles.[5]

Confidential information also may be disclosed lawfully where it is permitted or has been approved under a statutory process that sets aside the common law duty of confidentiality. One such statutory process for England and Wales is set out in the Health Service (Control

---

[1] *Campbell v MGN Limited* [2004] UKHL 22. See also *AG v Guardian Newspapers (No.2)* [1990] 1 AC 109 at 281.
[2] *X v Y* [1988] 2 All ER 648; *Hunter v Mann* [1974] QB 767; *Ashworth Hospital Authority vs MGN* [2002] UKHL 29.
[3] General Medical Council, *Confidentiality: Good Practice in Handling Patient Information* (GMC, 2017), available at <https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/confidentiality> (accessed 15 April 2021).
[4] Nursing and Midwifery Council, *The Code: Professional Standards of Practice and Behaviour for Nurses, Midwives and Nursing Associates* (NMC, 2018).
[5] The UK Caldicott Guardian Council, *The Eight Caldicott Principles,* available at <https://www.ukcgc.uk/manual/principles> (accessed 15 April 2021).

of Patient Information) Regulations 2002 (COPI).[6] In March 2020, the Secretary of State for Health and Social Care issued NHS Digital with a Notice under Regulation 3(4) of COPI to require NHS Digital to share confidential patient information with organisations entitled to process this under COPI for COVID-19 purposes. Another COPI Notice was issued in July 2020, extended until September 2021, which requires NHS Digital to disseminate confidential patient information in respect of which it is a controller. This COPI Notice means that researchers using confidential patient information without consent in COVID-19 research do not need to apply to the Health Research Authority (HRA) Confidentiality Advisory Group (CAG) for support, although they do still need review by an HRA/NHS Research Ethics Committee (REC).

2.1.2. Relevance to UK-REACH

The duty of confidentiality is directly relevant to UK-REACH as WP2-4 involve gathering of data from research participants. The duty will be engaged in respect of participant information that is disclosed to, or are otherwise obtained by, the project's research staff. As we note above, even outside the doctor-patient relationship, a duty of confidentiality may be established where the information has the necessary quality of confidence about it and where an individual has a reasonable expectation of privacy with respect to information given in confidence. For UK-REACH, a researcher-research participant relationship would likely trigger a duty of confidentiality owed by the researcher to the research participant.

For WP1 specifically, the duty of confidentiality is relevant when considering the sources of data used for analysis. The duty specific to the healthcare context is engaged through use of both primary and secondary care data. The duty of confidentiality is also engaged in the use of employment data, because as part of the contractual employment relationship, an employer (e.g. the NHS) has a duty to keep an employee's data confidential. As such, the data subjects within WP1 have a reasonable expectation that their information, be it health or employment-related, will remain confidential and not be used in ways that deviate from a reasonable expectation of privacy, unless there is a specific legal basis that lifts the duty of confidentiality (e.g. a public interest basis). This also includes an expectation that their data is handled in line with the Caldicott Principles. We note that although these principles apply primarily in respect to information collected for the provision of health and social care services (i.e. to healthcare date used in WP1), they also apply in some instances to the processing of healthcare staff information (i.e. the workforce information used in WP1).[7]

The processing of some of the information (specifically health data) in WP1 may also be possible under the COPI Notice(s) issued by the Secretary of State for Health and Social Care in March and July 2020, covering data such as NHS Digital primary care data. As noted above, the July 2020 notice has now been extended and is due to expire on 30 September 2021,[8] and as such should cover the processing of information by UK-REACH. However, should further patient data be disclosed to UK-REACH beyond that date, consideration will need to be given by UK-REACH, and specifically WP1, about how data sharing could continue if and when the COPI notices are no longer in place.

---

[6] The jurisdictional scope of COPI is limited to England and Wales, and NHS Digital is the national provider of information, data, and IT systems for commissioners, analysts, and clinicians in health and social care in England, particularly those involved with the NHS of England.
[7] The UK Caldicott Guardian Council, n 5.
[8] Health Research Authority, "Control of patient information (COPI) notice", available at <https://digital.nhs.uk/coronavirus/coronavirus-covid-19-response-information-governance-hub/control-of-patient-information-copi-notice> (accessed 15 April 2021).

2.1.3. Issues arising: maintaining the confidentiality of research participants

For WP2-4 **it is important that the confidentiality of the participants is assured [Recommendation 1].** This will be especially important for WP2 should biological data be collected in the future. To achieve this, appropriate measures have been put in place, such as anonymisation of data when they are reported, as well as ensuring that the data are stored securely and access to the data is limited. However, researchers across the WPs should be mindful of their ongoing duties towards the participants in this project and reflect on any potential risks to their confidentiality, particularly in the dissemination of research findings (which should reduce the risks as far as reasonably possible that no individual can be identified).

Returning to WP1, although this WP will be dealing with data that have been anonymised, research staff undertaking data analysis should be aware of the importance of confidentiality in the context of the datasets and ensure that use of the data is consistent with the participants' reasonable expectation of privacy. This includes consideration as to compliance with the Caldicott Principles. The de-identification process is an important (but not necessarily required) component in meeting the duty of confidentiality.[9] Rather, the key message to impart is that in the absence of patient/participant consent or a public interest ground to disclose confidential information, researcher staff in UK-REACH *must* reduce the risks as far as reasonably possible that no individual is identified or considered identifiable in the dissemination of research findings, and that information relating to them that is of a confidential nature remains undisclosed (we elaborate on this recommendation below).

Considering employment data more closely, we recognise that using staff information for research purposes may be considered novel, and therefore heightened concern may arise as to how to ensure these data are used in line with the duty of confidentiality (which as discussed clearly applies to employment data). We note that this duty does not prevent the disclosure of data where appropriate justification is in place, and measures are taken to meet participants' reasonable expectations of privacy. This requires similar mechanisms to reduce any risk of individuals being identified, and transparency for data subjects as to how their data are used and why (we expand on these measures below).

If the COPI Notices referenced above are not renewed by the Secretary of State for Health and Social Care beyond 30 September 2021, a different statutory or otherwise legal basis will be required to process patient information in WP1 insofar as it concerns England and Wales. While Regulation 3 of COPI provides specific support for identifiable patient information to be disclosed to projects such as UK-REACH, *in the absence of a valid notice, this Regulation 3 cannot be utilised*. In such a case, and assuming the COPI Notice for NHS Digital is not renewed beyond 30 September 2021, we suggest that **WP1 might consider availing themselves of Regulation 5 of COPI [Recommendation 2]**. Regulation 5 can be used to permit processing of confidential information for a range of medical purposes, broadly defined to include 'preventative medicine, medical diagnosis, medical research, the provision of care and treatment and the management of health and social care services'. Any person wishing to obtain support under Regulation 5 must submit an application to the HRA's CAG.[10] The CAG will then give advice to the relevant decision maker, which is currently the HRA for research applications and the Secretary of State for Health and Social

---

[9] *R v Department of Health ex pate Source Informatics Ltd* [2000] 1 All ER 786.
[10] See Health Research Authority, 'Confidentiality Advisory Group', available at <https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/confidentiality-advisory-group/> (accessed 15 April 2021).

Care for non-research applications. The CAG will not usually authorise disclosures under Regulation 5 to which a patient has objected. The HRA may not give an approval unless a REC has approved the medical research concerned. We note this would only be required for disclosure of further information to UK-REACH after this date (and not in respect of information which has already been provided).

2.2. Data protection law

2.2.1. The GDPR and the Data Protection Act 2018

In addition to duties under the common law, the UK also has a statutory framework for processing personal data, provided by the EU's GDPR (which has now become the UK GDPR, following the UK's exit from to EU) and the UK's Data Protection Act 2018 (DPA 2018), the latter of which is a national transposition of and supplement to the GDPR. "Processing" under the GDPR includes (among other aspects) collection, recording, organisation, storage, disclosure by transmission, dissemination and erasure or destruction.[11]

"Personal data" is information which relates to an identifiable natural person – someone who can be identified either directly or indirectly in reference to a form of identifier, such as a name, or factors specific to the person, including their physical, genetic, mental, economic, cultural or social identity.[12]

Article 4 GDPR sets out the core principles of data processing, which must be adhered to at all times and include:

- **Lawfulness, fairness and transparency**: data in relation to the data subject must be processed lawfully, fairly, and in a transparent manner;[13]
- **Purpose limitation**: the purpose of collecting the data must be specified, explicit, and legitimate, and the data must not be processed further in a manner that is incompatible with those purposes;[14]
- **Data minimisation**: the data processed must be limited to what is necessary in relation to those purposes;[15] and
- **Integrity and confidentiality**: when processed, the data must be appropriately secured and protected against unauthorised or unlawful processing, accidental loss damage or destruction.[16]

To process personal data lawfully, one *must* have a legal basis as stipulated in Article 6(1) GDPR. This Article 6(1) legal basis should be read in conjunction with Article 9(2), which requires 'special category' personal data (including data concerning ethnicity, genetic data and data concerning health) to meet one of ten lawful exceptions to the prohibition against processing such data. Often, the legal basis under Article 6(1) for processing data is the data subject's consent. However, other legal bases may apply. For example, where data are processed for health and social care research by universities, NHS organisations, Research Council institutes or other public authorities, the legal basis may be a 'task in the public interest'.[17] For commercial companies and charitable research organisations, national data

[11] Article 4(2) GDPR.
[12] Article 4(1) GDPR .
[13] Article 5(1)(a) GDPR.
[14] Article 5(1)(b) GDPR.
[15] Article 5(1)(c) GDPR.
[16] Article 5(1)(f) GDPR.
[17] Article 6(1)(e) GDPR.

protection authorities expect that the processing of personal data for research will be undertaken within 'legitimate interests'.[18]

For the Article 9(2) exception to special category personal data processing (e.g. health data), in the health research context, the lawful exception is often the scientific research exemption under Article 9(2)(j) GDPR, which is to be read in accordance with Article 89. We note that the DPA 2018 expressly permits special categories of personal data, such as health data, to be processed for scientific research purposes provided that the research is in the public interest and is in accordance with Article 89(1) of the GDPR as supplemented by section 19 of the DPA 2018.

The entity that holds primary responsibility for data protection law obligations is the "data controller",[19] which is defined at Article 4 to be "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data". Data controllers are those who, as a matter of fact and law, are in control of processing personal data. For the purpose of the UK-REACH project, the primary data controller is the University of Leicester, with the other entities co-involved in certain WPs who process personal data acting as joint controllers (e.g. Swansea University for WP1, University of Edinburgh for WP3, and University of Nottingham for WP4).

Alongside the data controller, the University of Swansea will also act as "data processor" for the purposes of WP2. This means they will be responsible for the processing of data in SAIL on behalf of the University of Leicester, as facilitated via written agreement between the two parties. The data controller retains overall accountability for meeting the obligations of data protection law; however, the data processor is equally responsible for implementing measures to meet the requirements of law.[20]

Under Article 13 GDPR, certain information must be provided to data subjects (i.e. in a 'privacy notice') where personal data are collected from them, and at the time when personal data are obtained. This information must include the purposes of the processing for which the personal data are intended as well as the legal basis for the processing. This means that there are certain transparency obligations imposed on data controllers when they intend to process personal data.

The GDPR is clear that the principles of data protection apply to *all* information which concerns identified or identifiable persons. This includes information which has undergone pseudonymisation but can still be identified through the use of additional information (e.g. encoding of a dataset that can be connected to a specific individual with a key code).[21] However, the GDPR will not apply to the processing of data that are anonymised, i.e. information which does not relate to an identified or identifiable person or personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable (e.g. rendering data down to an aggregated level or converted into statistics so that individuals can no longer be identified from them).[22] Whether an individual data item can be considered anonymous or not requires case-by-case evaluation. Collected material can contain detailed information on individuals (e.g. rare diseases, postcode and occupation, or a sufficient amount of different types of data) which makes them indirectly identifiable.

[18] Article 6(1)(f) GDPR.
[19] Article 24 GDPR.
[20] Article 28 GDPR.
[21] Recital 26 GDPR.
[22] Ibid.

2.2.2. Relevance to UK-REACH

The GDPR applies to the data collected and used in WP1-4, as these WPs entail the collection and processing of data relating to identified or identifiable persons (participants in WP2-4 and data subjects in WP1). The data used in all the WPs also will include special categories of data within the meaning of Article 9 GDPR, namely data relating to ethnicity and data concerning health. It is therefore necessary for each of the UK-REACH WP Co-Investigators, in addition to the PI and primary data controller (University of Leicester) to aware of, and where appropriate, comply with, the GDPR and DPA 2018.

2.2.3. Issues arising: purpose limitation, lawfulness of processing, and the limits of anonymisation

Purpose limitation

As noted above, according to the 'purpose limitation' principle, the GDPR stipulates that the purpose of collecting personal data must be specified, explicit, and legitimate (the 'purpose specification' dimension), and the data must not be processed further in a manner that is incompatible with those purposes (the 'compatible use' dimension). This means that the purposes for processing personal data should be determined from the very beginning, at the time of the collection of the personal data. This principle is achieved in UK-REACH in the data processing activities being undertaken in WPs 1-4 that involve new data collection. However, there may be consideration as to how the principle would apply in WP1 in the specific context of employment data, namely, questions as to how such data may be lawfully used for research purposes given it was gathered for a different purpose (viz. employment-related purposes). Here, beyond the necessity of having a legal basis for processing the data (see below), we note that Article 5(1)(b) GDPR states that processing personal data for a purpose other than that for which they have been collected is allowed in certain circumstances, even if this new purpose is not compatible with the first one. Phrased another way, certain reuses of data are a priori considered as compatible provided certain conditions are met. Of benefit to UK-REACH is that Article 5(1)(b) permits further processing of personal data for scientific research purposes and such purpose will be deemed compatible with the initial purpose(s) of data collection, provided the data are processed in accordance with Article 89(1) GDPR, namely with appropriate safeguards in place for the rights and freedoms of data subjects, and to ensure that technical and organisational measures are implemented to adhere to the principle of data minimisation.

The legal basis for processing personal data

Although the GDPR is no longer applicable once data are adequately anonymised (as per the standard stated in Recital 26 where a person is no longer able to be identified through "all the means reasonably likely to be used" to identify the person directly or indirectly), it *is* applicable to the processing of those data for the purposes of achieving anonymisation. In other words, when data are initially gathered and before they have been anonymised, they are subject to the full requirements of data protection law. There is, therefore, a need to stipulate the Article 6(1) legal basis for processing the data for each of the WPs, including where the data are being processed for anonymisation as part of WP1.

According to the consortium agreement, UK-REACH intends to rely on Article 6(1)(e) as the legal basis for processing personal data, and Article 9(2)(i) for the processing of special category data across WPs. These provisions stipulate, respectively, that the processing of data is necessary for the performance of a task carried out in the public interest, and for special category data, that it is necessary for reasons of public interest in the area of public health. To this end, **we recommend that UK-REACH, primarily through the University of**

Leicester as primary data controller, should clearly stipulate why the processing is necessary to perform the research and how it purports to be in the public interest in the area of public health [Recommendation 3]. Such justification must go beyond a statement that this is by definition the case. There is a clear need to explore both how the project as a whole, and each of the constituent WPs, are in the public interest, and also to ensure that only data necessary to service this purpose are processed. Equally, **we recommend that the use of data be proportionate to the aims of the various WPs, and that they provide measures to protect the fundamental rights and interests of data subjects [Recommendation 4].** This obligation falls on all data controllers equally, as well as Swansea University as data processor for the purposes of WP2.

Although Article 6(1)(e) public interest may be an appropriate legal basis for WP1 and WP2, it is less clear as to whether this is the case for the data gathered in WP3 and WP4. We note that both of these WPs will be seeking explicit consent from their participants through participant information sheets and consent forms – which contains information therein on how data will be processed – and as such, they would be able to rely on this consent as the lawful basis for data processing, within the meaning of Article 6(1)(a) and 9(2)(a). The participant information sheets and consent forms clearly emphasise the voluntary nature of participation and ability to withdraw, and do not mention the public interest as a legal basis for data processing. The focus on consent in WP3 and WP4 may lead participants to reasonably conclude that their data are being processed in line with the parameters of their consent, contrary to the actual legal basis of Article 6(1)(e) public interest. This situation carries a risk of misalignment between the expectations of data subjects as to how their data will be processed and used, and the legal basis actually relied on by UK-REACH as per the consortium agreement and potentially other agreements. Such a mismatch could be problematic because it could lead to UK-REACH processing data which does not meet a participant's reasonable expectation of privacy (as discussed above), or indeed does not protect their interests within the terms of the GDPR. For example, if a participant withdrew from WP3, it would still be lawful to use their data under a public interest justification, but this would not be keeping with their expectations of how their data would be used.[23]

It is possible to rely on a different legal basis even when seeking consent, but it is important that this is made clear to participants. This is not currently the case in the participant-facing information for WP3 and WP4. **We therefore recommend that the parties involved in WP3 and WP4 carefully consider what the legal basis for data processing should be, to ensure that participants interests are appropriately protected [Recommendation 5].**

Identifiable data and the limits of anonymisation

Looking at WP1 specifically, there also remains a need to consider the processing and use of the data once the identifiable data have been removed. Even without direct identifiers such as name, a significant amount of information about a given individual will be available to WP1 researchers, which go across aspects of data subjects' lives, including their job, their health, socioeconomic status, and nationality (see Appendix A, which lists the full range of data being collected and analysed in WP1).

This is important because, given the significant amount of information that is being made available to WP1 researchers, there is a risk that the data remain identifiable, in accordance with the Recital 26 standard. This is possible because the more detailed a dataset is about

[23] See also Edward Dove and Jiahong Chen, "Should Consent for Data Processing be Privileged in Health Research? A Comparative Legal Analysis" (2020) 10 International Data Privacy Law 117-131.

any one person (i.e. the more values there are), the more unlikely it becomes that anybody else will possess all same combinations of values in the set.[24] From this, individuals within the set may be identified even in absence of a name or other direct identifier. This even holds for less detailed information. For example, legal scholar Paul Ohm details how the combination of ZIP code, sex, and date of birth can be used to identify the majority of the population in the USA – a much larger population, of course, than the UK.[25] This demonstrates the surprising ease with which ostensibly anonymised data may in fact be re-identified. **We therefore recommend that WP1 carefully and thoroughly consider on an ongoing basis whether all datasets are 'anonymous' for the purposes of the GDPR/DPA 2018 (i.e. as a legal standard of reasonable risk management) [Recommendation 6].**

Moreover, the risk of re-identification is amplified through the linkage of previously unconnected datasets, because the linked data will be a far richer account of the individual than each individual set,[26] revealing more about who they are and their lives than each dataset in isolation could. Linkage also further decreases the likelihood of more than one individual sharing all the same characteristics. This makes it impossible to guarantee the anonymity of the dataset and means that the privacy of individuals cannot be assured merely through the removal of identifiers.[27]

On this basis, it would be inappropriate to state that because WP1 data are being de-identified (which, we note, is not necessarily the same thing as "anonymisation" under the GDPR), the principles of the GDPR do not apply. Even once identifiers are removed, respect of the interests of the data subjects and safeguarding of their fundamental rights should be the guiding principles. This does not mean that anonymisation is without utility – the removal of key identifiers remains one form of protecting the privacy of data subjects. Instead, two key modifications in our understanding of anonymisation are required.

First, **we recommend anonymisation must be seen as part of a framework for good information governance [Recommendation 7].**[28] The other elements of this framework should be respect for the core principles of the GDPR. Key in controlling the risk of re-identification are the principles of data limitation, integrity and confidentiality, and data minimisation. Respecting these principles means limiting the number of people who access the data and the amount of data available. There are several additional measures that WP1 utilises which reflect these principles. Most significant in terms of protection measures is the use of a trusted research environment such as SAIL. This gives greater security to the data by limiting access, as well as requiring training for those who will access it. It also gives heightened security to the data by keeping it out of the public domain. The process of de-identification, including the encryption of identifiers and deletion once they are no longer required, is in keeping with the principles of data minimisation and limitation. The continued use of these measures will assist in complying with the requirements of GDPR and good information governance throughout WP1. But those in control of the data in question must

---

[24] Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization" (2010) 57 UCLA Law Review 1701-1769.
[25] Ibid.
[26] Ibid.
[27] Nuffield Council on Bioethics, *The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issues* (NCB, 2015) <https://www.nuffieldbioethics.org/publications/biological-and-health-data> (accessed 15 April 2021).
[28] Miranda Mourby, "Anonymity in EU Health Law: Not An Alternative to Information Governance" (2020) Medical Law Review 478-501.

always remain vigilant to the prospect that data might at some point in the future become identifiable.

Following from the last point, **we encourage a reframing of the anonymisation process to make it more *dynamic* [Recommendation 8].** Techniques to anonymise data should not use pre-determined measures to de-identify, as these may fail to properly reflect the context of the data, including shifting contexts over time.[29] Instead, what is identifiable should be considered by reference not just to what is identifiable in most contexts, but also what is identifiable in conjunction with data which are readily available or may be available to anyone seeking to re-identify data.[30] This becomes particularly important when the findings of WP1 are reported and disseminated. It is vital to ensure that the data presented in the findings cannot be re-identified once they are publicly available.

Equally, anonymisation should not be viewed as a one-off event (so-called "release-and-forget" anonymisation).[31] A watching brief is required to ensure the data remain de-identified throughout their lifecycle in the research endeavour and that the risk of data triangulation is mitigated. This is important for the planned WP1 sub-studies, where a more granular analysis will be undertaken, as closer attention is paid to individual attributes and the risk of identification becomes higher. It is also important should the data be retained for use in further research or where further linkage is undertaken.

By understanding anonymisation in this nuanced, dynamic manner, and seeking to control risks to re-identification on an ongoing basis, UK-REACH can ensure robust compliance with the GDPR and core principles of data protection law, and more importantly, ensure it works with sensitive data in a way that accords with the reasonable expectations of health workers.

2.3. Human rights

2.3.1. European Convention on Human Rights

In addition to the law on confidentiality and data protection, due consideration must be given to relevant human rights protected under the European Convention on Human Rights (ECHR). These rights form part of UK Law under the Human Rights Act 1998. Rights to highlight are as follows:

Article 8 protects the right to respect for private and family life. This right is interpreted broadly to include aspects of an individual's identity, both physical and social.[32] It includes privacy of information, and will be engaged through the storing of data about individuals,[33] as well as how their data is collected and used by third parties.[34] Any interference with Article 8 rights must be justified as necessary and proportionate to the legitimate interest pursued.[35]

---

[29] Nuffield Council on Bioethics, n 27.
[30] Ibid.
[31] Ohm, n 24.
[32] European Court of Human Rights "Guide on Article 8 of the European Convention on Human Rights" (Council of Europe, 2020) <https://www.echr.coe.int/documents/guide_art_8_eng.pdf> (accessed 15 April 2021).
[33] *S and Marper v United Kingdom* (30562/04) [2008] WLUK 117.
[34] *LH v Latvia* (2015) 6 EHRR 17.
[35] Article 8(2) ECHR. "There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others."

Article 9 protects the rights freedom of thought, conscience, and religion. The disclosure of personal information about religious convictions will engage both Article 8 and Article 9,[36] and under Article 9, individuals have a right to object to the inclusion of their data on moral or religious grounds. Again, interference with Article 9 must be justified in the same manner as Article 8.

Article 14 holds that all the rights within the Convention must be guaranteed without discrimination. Article 14 is not a free-standing right, but is engaged where the enjoyment of other rights is impacted by discrimination (i.e. differential treatment with respect to an identifiable characteristic).[37] However, there is no need to demonstrate a violation of the relevant substantive right in order to prove a violation of Article 14.[38] The prohibition of discrimination applies to both direct and indirect discrimination – any seemingly neutral course of action which has a disproportionate impact on a protected group will still violate Article 14.[39] If there is differential treatment on the basic of an identified characteristic, this must be objectively and reasonably justified.

2.3.2. Relevance to UK-REACH

Article 8 is engaged through all the WPs because they involve the collection, storage, and use of data relating to the private lives of individuals.[40] The relevance of Article 8 is greatest for WP1 and WP2 because both involve the collection and use of far more detailed information about individuals and their lives. For WP1, it is the magnitude of data (see Appendix A) about private life, as well as the reliance on a legal basis other than consent, which makes it relevant. For WP2, it is the longitudinal nature of the study, which means potential interference over a greater length of time, as well as the inclusion of highly detailed information, including genetic data.

An awareness of Article 9 is important for WP2-5, as it is possible for an individual to object to the inclusion of data about them in the study on moral grounds, though this issue is mitigated by the advance notice (e.g. in a participant information sheet) of the purposes for which the data are being processed, and providing participants an opportunity both to not participate and to withdraw at any point. However, we note that even once identifiers are removed, an individual may still exert their rights over their data and request its exclusion from the dataset.

Article 14 is particularly important given UK-REACH's focus on race and ethnicity, two of Article 14's identified characteristics. The analytical focus on ethnicity creates a potential risk of differential treatment or disproportionate negative impact on a particular group. Because of the health and employment context, this impact could affect individual's livelihoods and/or health status – both of which engage Article 8. Even if UK-REACH does not purport to treat participants differently based on characteristics, as WP5 aims to inform policy development, it is vital to consider any indirect discrimination arising from the findings of UK-REACH. Examples of indirect discrimination might include a policy for COVID-19 physical distancing restrictions, or work restrictions, which at first glance apply to everyone in the same way, but in fact have a worse effect on some people than others, such as members of ethnic minority communities.

---

[36] *Folgerø and others v Norway* (15472/02) (2008) 46 EHRR 47.
[37] *Carson and others v United Kingdom* (42184/05) (2010) 51 HER 13.
[38] Ibid.
[39] *Biao v Denmark* (38590/10) [2016].
[40] *S and Marper v United Kingdom* (30562/04) [2008] WLUK 117.

2.3.3. Issues arising: assuring respect for human rights

Given the potential for interference with Article 8, it is necessary, as per Recommendation 3 above, that the use of data by UK-REACH is proportionate (i.e. adequate, relevant and limited to what is necessary in relation to the purposes of the project) and that the interests pursued are legitimate. This is particularly required in the context where the consent of participants will not be sought to use their data (WP1), either because an alternative legal basis will be sought or because certain datasets will be anonymised.

Considering proportionality, the safeguards highlighted above for WP1 – that is, the use of dynamic anonymisation, and respect for the principles of GDPR (data minimisation, integrity and confidentiality) – will be key in limiting any potential for interference with Article 8 (for example, a breach of privacy). However, **all WPs should be mindful of the data subjects' and participants' Article 8 ECHR rights [Recommendation 11].** This in turn will ensure compliance with the GDPR to allow the use of data to remain lawful and to respect the fundamental rights of participants.

For Article 9 ECHR, given the right to object to the inclusion of personal data on moral grounds**, it is important for the purposes of WP2-5 to be mindful of the potential for – and right of – participants to object to the inclusion of data about them [Recommendation 12]**. This right is necessarily qualified by what is possible and practical in the circumstances: once the data are anonymised and integrated, it may be very difficult, if not impossible, to identify which data are a particular data subject's/participant's and remove them from the dataset. Participants must therefore be advised at the outset of their opportunities to opt out (and the fact that opt out might not be an option in such circumstances) and must be informed what will happen to their data (as is the case for WP3 and 4). This request would, however, need to be weighed against the feasibility and proportionality of finding the relevant data within the set and removing and destroying them, without damaging the dataset as a whole. This again speaks to the need to be transparent with participants up front about what is involved and how, when, and under what conditions they can maintain some say over what happens to their data.

Regarding Article 14 ECHR, all WPs across UK-REACH must ensure that in the collection, storage and dissemination of research findings, there is no discriminatory impact on the participants or communities involved in the project [Recommendation 13]. The potential for indirect discrimination is particularly important to assess in the context of possible policy implications arising from WP5. Even if the policies arising purport to be neutral, it is key to consider possible disproportionate and prejudicial impacts arising from the WP5 recommendations, such as group harms caused by particular findings on the impact of COVID-19 on ethnic minority communities and any ensuing policy recommendations (this will be commented on further below).

Having summarised what we consider to be the relevant legal framework and its three constituent elements, we now turn to analyse the relevant ethical considerations in the project.

3. Ethical considerations

3.1. The requirement for an ethical approach

Aside from the legal issues highlighted above, due regard must be given to any ethical concerns arising from the work of UK-REACH and assurance given that an ethical approach is taken to the project. We highlight two key reasons why an ethical approach is important.

First, considering in general the use of significant quantities of data in health research, although there is an opportunity to benefit from using and linking data as proposed by UK-REACH, it is not without risk. It is entirely possible that decision-making and actions in terms of data use, and subsequent policy informed by this use, can both benefit and harm individuals or groups.[41] And, though there may be full legal compliance with confidentiality, data protection, and human rights, there may still be unethical practice. Phrased another way, just because the legal obligations are met does not mean that the project owes no obligations to its participants and subjects. It remains necessary to fully consider and weigh up the risks and benefits of the project and use of data to ensure ethical acceptability.

In the context of UK-REACH specifically (i.e. research conducted in the midst of an unprecedented pandemic), an ethical approach is key to supplement the legal framework and meet any challenges the project encounters that cannot be resolved through legal or regulatory means. The pandemic present numerous unique challenges to not just UK-REACH but also to its data subjects and participants. For co-investigators in WP3 and 4, the pandemic has created new challenges to undertaking research. Interviews must be held remotely rather than face-to-face, creating new issues around access to and reliance on technology, and how to ensure confidentiality when a private space may not be guaranteed, due to home working.

For data subjects and participants, it has led to drastic legal and policy changes which have imposed numerous restrictions on what they are able to do in their day-to-day lives. These changes in turn may affect the level of trust placed in government policy, law, and regulation.[42] The pandemic also has the potential to change individual anxiety and perception of risk, which in turn may influence individuals' concerns about the use of their data and law's ability to guarantee this. A fast-paced regulatory environment is also likely to bring novel challenges around data use. For example, the advent of Track and Trace mobile applications raised new challenges concerning data protection and required a response from the Information Commissioner's Office.[43]

This fluid environment raises the possibility that UK-REACH may encounter challenges regarding data use which have not previously been identified and considered. Such challenges may not be easily considered within a legal framework but do raise concerns of ethical importance. Defining an ethical approach for UK-REACH should allow for novel challenges to be considered and evaluated appropriately.

We highlight one overarching ethical consideration arising from this complex landscape which is vital to consider for UK-REACH: how the activities of the project could impact on

---

[41] Nuffield Council on Bioethics, n 27.
[42] Daniel Devine, Jennifer Gaskell, Will Jennings and Gerry Stoker "Trust and the Coronavirus Pandemic: What are the Consequences of and for Trust? An Early Review of the Literature" (2020) Political Studies Review (forthcoming), doi: https://doi.org/10.1177%2F1478929920948684.
[43] Elizabeth Denham "Blog: Data protection considerations and the NHS COVID-19 app" (Information Commissioner's Office, 2020) <https://ico.org.uk/global/data-protection-and-coronavirus-information-hub/blog-data-protection-considerations-and-the-nhs-covid-19-app/> (accessed 15 April 2021)

vulnerabilities of particular groups, particularly ethnic minority health workers. In line with Xafis and colleagues, we understand vulnerability to mean having "elevated susceptibility to systematic disadvantage."[44] Such vulnerabilities are of particular ethical concern where they are exacerbated by socially determined factors, or pre-existent structural inequalities.[45]

Similar to what we have highlighted above, concerns about vulnerability arise both in the context of data usage and the COVID-19 pandemic. First, in terms of the use of data, issues of vulnerability may arise from the unequal distribution of power between those who are the subject of the data and those who control the data (the "big data divide"),[46] as well as the possibility of harm to a group occurring from the use or misuse of the data. Second, further layers of vulnerability arise out of the pandemic which link closely to ethnicity, socioeconomic status, and job role (i.e. healthcare workers), which include, among other things, an ostensibly heightened risk of contracting and suffering the ill-effects of the virus.[47] Although UK-REACH aims to address vulnerabilities arising from COVID-19 for ethnic minority healthcare workers, it is also important to consider where these vulnerabilities could in fact be exacerbated. An ethical approach will allow UK-REACH to do so.

As such, UK-REACH and all the constituent WPs must work towards full legal *and* ethical compliance in all its research activities and across the research lifecycle, from recruitment to data collection and analysis to research findings, communications, and implementable policy proposals. Therefore, there remains a need to consider what morally relevant interests are engaged, where ethical concerns arise, and how these ought to be addressed.[48]

3.2. An ethical framework

To identify and consider the relevant interests that are engaged in this project, we adopt the "Ethics Framework for Big Data in Health and Research" proposed by Xafis and colleagues.[49] This framework allows us to identify the ethical values which are central to UK-REACH and its participants and data subjects, and consider how to weigh these on an ongoing basis and throughout all data management, use, access, and data-related decision making. UK-REACH as a project falls into the scope this framework because it is research in the health context (considering both physical and mental health in relation to the COVID-19 pandemic) and involves the use of 'Big Data'[50] as follows:

- Volume: WP1 aims to gather highly detailed data for healthcare workers across the UK (in excess of 170,000 people). WP2 will gather even more detailed data, (possibly including genetic data in the future) from around 30,000 people.
- Variety: The data forms gathered across the WPs are diverse, from interviews and focus groups to health databases to potentially genetic data. In WP1 and WP2, the sources of data are diverse and range from health to employment to registration data.

---

[44] Nuffield Council on Bioethics, n 27.
[45] Wendy Rogers, Catriona Mackenzie, and Susan Dodds, "Why Bioethics Needs a Concept of Vulnerability" (2012) 5 International Journal of Feminist Approaches to Bioethics 11-38.
[46] Vicki Xafis and others, "An Ethics Framework for Big Data in Health and Research" (2019) 11 Asian Bioethics Review 227-254.
[47] Agomoni Ganguli-Mitra, "The Need to Unpack Vulnerability in a Pandemic" (2020) BMJ opinion, <https://blogs.bmj.com/bmj/2020/07/03/agomoni-ganguli-mitra-the-need-to-unpack-vulnerability-in-a-pandemic/> (accessed 15 April 2021)
[48] Nuffield Council on Bioethics, n 27.
[49] Xafis and others, n 46.
[50] Ibid.

- Velocity: the aim of WP1 is to carry out expedited linkage and analysis of multiple, existent datasets, requiring sharing with multiple organisations in varying timescales.

3.2.1. Substantive values

Although all the substantive values highlighted by Xafis and colleagues are relevant to UK-REACH, we highlight below the values we consider to be of greatest relevance to the project:

Justice: the value of justice ties in strongly with the concerns highlighted above about vulnerability. It is key to address any issues of inequity or group harms which may arise from the research. The focus of the research in UK-REACH is on ethnic minority groups who already experience structural inequality, marginalisation, and discrimination. From a data perspective, ethnic minority groups have traditionally (especially in the health context) had unequal access to Big Data research projects, and have been largely excluded from their benefits.[51] Arising from this may be issues of trust and heightened concern about potential misuse of data, which would be important to consider and mitigate. This is particularly crucial in the context of a pandemic which has seen a disproportionate impact on ethnic minority groups, where they may be heightened concern about data misuse or research which inappropriately focuses on ethnic minorities.

Justice in the context of UK-REACH requires, at a minimum and in the negative sense, that data are not used in a way that exacerbates discrimination against, and power asymmetries between, different groups of health professionals, and different genders and ethnicities in UK society. In the positive sense, data should be used to provide greater distributive justice through making hitherto neglected or invisible communities more visible. More generally, due regard should be given to fair access, participation, and representation in the project so that those most vulnerable to discrimination and invisibility have an opportunity to contribute meaningfully to the project's design, delivery, and dissemination.

Specifically, due consideration should be given to the possibility that research outputs could result in group harms such as heightened social stigma arising from the dissemination of findings[52] on ethnic minority healthcare workers, or resulting in unfair differential treatment in the workplace, even if this is unintended. For example, should particular healthcare roles, such as working in an ITU, be considered "too risky" for ethnic minorities because of COVID-19, this could impact on choices available to ethnic minority individuals in terms of career pathways, or indeed limit opportunities for progression for trainee doctors, adversely affecting them and raising concerns of injustice. Linked to this is the value of ***harm minimisation*** – any harms arising should be identified and appropriately mitigated.

Third, UK-REACH will be engaging with a broad variety of HCWs from varying staff groups, each with their own individual background and social context. Concerns about trust and uneven access to data will vary across staff groups. Careful thought should be given as to how to engage various groups to ensure appropriate representation, as well as how to capture and reflect the values of groups and feeding those values into decision-making processes in UK-REACH. It is equally key, given this diversity, that groups are not inappropriately homogenised, or findings generalised in a manner which fails to recognise relevant and intersectional experiences.

---

[51] Kenneth Boyd, "Ethnicity and the Ethics of Data Linkage" (2007) 7 BMC Public Health 318.
[52] Ibid.

Considering any concerns about injustice and seeking to mitigate these will also assist in meeting the Article 14 ECHR prohibition of discrimination.

Privacy: The key privacy issues have already been highlighted above. However, it is important to recognise that privacy, as well as a legal requirement, is also of ethical importance and that participants have an interest in the assurance of their privacy, including use of data in line with their morally reasonable expectations.[53] An ethical view of privacy goes beyond legal compliance to consider the value of privacy as both a means to an end (e.g. privacy facilitates human flourishing) and an end in itself (e.g. privacy is an inherent good we are owed and duty-bound to uphold), and also the norms underlying it that ought to be respected by individuals, society, and the state. Equally, novel issues surrounding privacy due to the increased individual surveillance for public health measures (as discussed above) mean it is vital to consider privacy as more than a simple legal obligation.

Benefit for publics: UK-REACH should continue to reflect upon and clearly articulate the benefits that are likely to arise from the study and what they might be for various publics (thereby considering the importance of diversity and the contextualised nature of benefits in this area of research),[54] and how this should be weighed alongside considerations of privacy and justice. Key UK-REACH benefits for publics might include both immediate answers relating to the clinical outcomes of COVID-19 in ethnic minority HCWs, as well as a better evidence-based understanding of COVID-19 that can inform responses to current/future pandemic waves, and a framework within which researchers and policymakers are able to investigate longer-term clinical sequelae (viz. physical health and mental health). Clearly articulating this will also assist UK-REACH in justifying its use of Article 6(1)(e) and Article 9(2)(i) GDPR as the legal basis and special category exception, respectively, for data processing.

Proportionality: the complexities of UK-REACH, as well as the sensitivities engaged though focusing on a disproportionate impact on ethnic minority healthcare workers, mean that proportionality is not just a legal requirement, but also a value of ethical importance. It is vital to ensure that decision-making in each of the WPs strike a balance between the envisaged benefits and protection of the rights and interests of the participant/data subject. As such, the sharing, access, and data use through UK-REACH must be appropriate to the aims of the project, while also accounting for the competing individual interests. Part of assuring a proportionate approach will also be consideration of ***harm minimisation*** measures to limit the risks to privacy or group harms arising.[55]

3.2.2. Procedural values

The key procedural values to consider are as follows:

Engagement: Given the justice concerns highlighted above, it is vital to ensure early, sustained, and meaningful engagement of stakeholders from groups with morally relevant interests,[56] especially groups who may have uneven access to data or do not traditionally engage with research. Such engagement, which ought to commence at the earliest stages of research design and continue throughout the lifecycle of the project in order to be afforded committed support and not merely serve as a rubber-stamp exercise at the end of an

---

[53] Nuffield Council on Bioethics, n 27.
[54] Graeme Laurie, "Cross-sectoral Big Data: The Application of an Ethics Framework for Big Data in Health and Research" (2019) 11 Asian Bioethics Review 327-333.
[55] Ibid.
[56] Nuffield Council on Bioethics, n 27.

already-decided process, will allow for stronger consideration of potential justice concerns arising as well as development of harm minimisation measures.

Reasonableness: Because consent is not being relied upon in all the WPs, it is important that the use of data remains reasonable and acceptable by reference to widely recognised standards. Public engagement research about acceptability of data linkage by Xafis[57] and for the Scottish Government[58] demonstrate general support for data linkage projects and a recognition of the benefits of doing so. However, the research also shows that the public may have heighted concern about specific types of data being linked (e.g. employment data)[59] and the potential for misuse of data, or the potential for labelling and stigmatisation to arise.[60] Consideration should therefore be given in this context as to how to assure use is reasonable and accounts for such concerns.

Trustworthiness and Transparency: In the context of gathering large quantities of data about individuals, it is key that UK-REACH demonstrate through its approach to data management, project governance arrangements, and stakeholder engagement, that it is a trustworthy and dependable project. Transparency and openness to scrutiny of actions and processes (both internally and publicly) form a key aspect of this as well.

Reflexivity: Given the dynamic and evolving nature of the data and the pandemic itself, it is necessary to continue to reflect on its limitations, any uncertainties within the dataset, and the ongoing management of competing interests arising. This is especially important because of the ongoing flow of data between organisations and any changes to the aims and scope UK-REACH. There is a need to continue to reflect on why the data are gathered, how they are used, where they are held, and what the risks are. It is equally key for each of the WPs to continue to reflect on their approach to the analysis of data gathered, including any potential personal biases or interests arising. The value of reflexivity should complement the watching brief on anonymisation discussed above and afford regular review of processes and governance mechanisms.

3.3. Defining an ethical approach for UK-REACH

In light of each of these identified values, we consider that an ethical approach for UK-REACH would seek to clearly articulate and reflect upon the purported benefits of the research, and consider on an ongoing basis how this should be balanced against individual interests (e.g. privacy interests) in a proportional manner. Particular attention should be paid to issues of justice, with consideration of any potential group harms or inequity arising from the gathering, sharing, and use of participant data. The identification and minimisation of any such issues should be done through engagement with key stakeholders to ensure that those with morally relevant interests are able to participate in discussions and influence any responses. UK-REACH should seek to be open and transparent in its processes and ongoing engagement is required to allow for scrutiny of the research, and that the use of data remains reasonable and publicly acceptable throughout the lifetime of the project.

Utilising this approach, UK-REACH can ensure the ethical and legal acceptability of the project, even in the uncertainty caused by the COVID-19 pandemic. Taking an engaged,

---

[57] Vicki Xafis, "The Acceptability of Conducting Data Linkage Research Without Obtaining Consent: Lay People's Views and Justifications" (2015) 16 BMC Medical Ethics 79.
[58] Sara Davidson and others, *Public Acceptability of Cross-Sectoral Data Linkage: Deliberatative Research Findings* (Scottish Government Social Research, 2012).
[59] Xafis, n 57.
[60] Davidson and others, n 58.

transparent, and reflexive approach will also allow the project to demonstrate what ongoing measures are being taken to protect participants' data and interests, as well as enabling response to any pandemic-specific concerns that are raised. It will also allow any new challenges to be rapidly and properly considered to ensure harms are identified and adequately addressed where required.

Measures already in place to meet this ethical approach are as follows:

The WP5 stakeholder groups: UK-REACH's Public Involvement Strategy through WP5 is vital in ensuring an ethical approach to the research activities and data processing in the project. This will ensure the value of *meaningful* engagement is at the centre of the progress of each of the WPs, as well as supporting reasonableness, trustworthiness, and transparency by opening the discussions of the project up to public scrutiny. As WP5 also seeks to inform policy, it will also allow UK-REACH to clearly assess the public benefit of the research.

Use of a trusted research environment: the use of SAIL in WP1 promotes the trustworthiness, proportionality, and reasonableness of the project's data processing and research activities. It gives heightened security to the data, thus more strongly protecting privacy interests. The additional protections allow the use of data to be limited to what is necessary. It will also serve to increase public acceptability and trust by assuaging concerns or doubts about misuse of data. The transparency around how SAIL holds data and the data flow processes will also support an ethical approach. As the project gets underway, UK-REACH should reflect throughout the process on what data is held in SAIL so that these values continue to be supported.

Further measures we suggest for consideration are as follows:

Considerations for each work package

To promote compliance with data protection legal and ethical obligations, we recommend that all WPs facilitate ongoing discussion and reflection regarding the benefits of their research and how they can adequately account for the relevant interests at stake. This should include consideration of what information is being gathered, where it is held, and whether it is necessary and proportional to hold it, and if so, under what conditions and for how long. Efforts should be made in particular to identify relevant vulnerabilities and continue to assess what impact the research has on these, both positive and negative.

This process should continue throughout the lifespan of the project to allow the WPs to respond to any new concerns or issues arising.

Provision of public-facing information

We encourage the UK-REACH website to be regularly updated to ensure that the aims, methods, and research findings from each of the WPs are reported rapidly and in easy-to-understand language. Beyond text-based information, disseminating findings through other forms of media (podcasts, blogs, etc.) ought to be considered. In addition, UK-REACH as a group should ensure its findings are made available to SAGE and other policymakers in a timely manner so that policy decisions can be made in near real-time.

These measures will allow UK-REACH to ensure that the research is ethically acceptable, and that any ethical concerns arising, particularly those due to the pandemic, can be appropriately identified and mitigated.

4. Recommendations

Having now considered the ethical and legal implications of UK-REACH, we set out below our key recommendations for the project, several of which we have already highlighted in the analysis above:

Recommendation 1	Confidentiality of the participants ought to be assured across all WPs.
Recommendation 2	Should the current COPI notice for NHS Digital not be renewed beyond 30 September 2021, WP1 ought to consider availing themselves of Regulation 5 of COPI to justify continued use of confidential NHS data (at least as concerns data in England and Wales).
Recommendation 3	When relying on public interest as the legal basis for processing data, each work package should clearly stipulate why the processing is necessary to perform the research and how it purports to be in the public interest in the area of public health. This justification should be made publicly available.
Recommendation 4	The use of data ought to be proportionate to the aims of each work package (e.g. WP1 should only collect and use datasets that are necessary in relation to its aims); for WPs that involve anonymisation processes, measures ought to be provided to protect the fundamental rights and interests of data subjects when processing data to remove identifiers.
Recommendation 5	WP3 and WP4 should carefully consider what the lawful basis for data processing should be to ensure that this is consistent with the reasonable expectations of participants and that their interests are appropriately protected.
Recommendation 6	WP1 should carefully and thoroughly consider on an ongoing basis whether all datasets are 'anonymous' for the purposes of the UK GDPR/DPA 2018 (i.e. as a legal standard of reasonable risk-management).
Recommendation 7	Anonymisation ought to be seen as part of a framework for good information governance.
Recommendation 8	Anonymisation should be approached on a dynamic basis rather than through a set of pre-determined measures, treating anonymisation as if it were a one-off event. Such a dynamic basis includes keeping a watching brief on what is required to ensure the data remains anonymous throughout the lifecycle of its use in the UK-REACH project.
Recommendation 9	Where data and datasets are deemed to be anonymised according to the UK GDPR/DPA 2018 and commonly accepted standards (e.g. those promulgated by the UK's Information Commissioner's Office), UK-REACH should continue to consider what legal and ethical obligations arise even after anonymisation has been achieved. This includes consideration of confidentiality of data, risk of re-identification, human rights, and other relevant interests.

Recommendation 10	Because anonymisation in the context of linking large datasets across sectors is not always sufficient to protect privacy adequately, UK-REACH should seek to establish and consider any ongoing privacy risks and continue to apply the principles of data protection legislation.
Recommendation 11	All WPs should be mindful of the data subjects' and participants' Article 8 ECHR rights.
Recommendation 12	With respect to Article 9 ECHR, given the right to object to the inclusion of personal data on moral grounds, the investigators in WP2-5 should be mindful of the potential for participants to object to the inclusion of data about them.
Recommendation 13	With respect to Article 14 ECHR, the UK-REACH project as a whole should ensure that in the collection, storage and dissemination of research findings, there is no discriminatory impact (directly or indirectly) on the participants and communities involved in the project.
Recommendation 14	UK-REACH should aim to clearly state to publics the expected benefits of each of the WPs and balance those expected benefits with any relevant interests.
Recommendation 15	UK-REACH should be mindful of relevant vulnerabilities of individual participants in the WPs, as well as groups of participants (e.g. sectors of healthcare workers, ethnic minority communities) and the risk of group harms such as stigma, mitigating these where appropriate to ensure compliance with Article 14 of the European Convention on Human Rights (prohibition of discrimination). Particular attention should be paid to vulnerabilities arising from the COVID-19 pandemic.
Recommendation 16	UK-REACH should continue to engage with and encourage participation of individuals with morally relevant interests as key stakeholders. The makeup of the stakeholder groups should be reviewed on an ongoing basis to ensure those with morally relevant interests are included at all stages and afforded a meaningful opportunity to influence decision-making and the direction of the project.
Recommendation 17	UK-REACH should aim to update its website regularly in plain language to make its aims, methods, and findings accessible to the public, and consider use of other media where appropriate.
Recommendation 18	A watching brief is required on the relevant ethical values at play to ensure that the project's activities continue to be proportional to the benefit sought, and to identify any new issues of concern that may arise. This recommendation ought to be undertaken by each of the investigators across the WPs.

Appendix A: Data used by UK-REACH WP1 after de-identification

Demographic data

- Ethnicity
- Age in years
- Sex
- Disability (where applicable)
- Religious belief
- Socioeconomic depravation IMD measure (calculated by postcode however postcode itself will not be accessed)
- Nationality
- Immigration status

Health data

- BMI
- Underlying health conditions (e.g. cardiovascular disease, diabetes etc.)
- Management of health conditions including medications
- Hospital admissions or operations during the relevant period
- Medical observations and interventions for those admitted to Intensive Care
- Cause of death for any who have died
- Covid-19 diagnosis
- Outcome of Covid-19 diagnosis (e.g. hospitalisation, death etc.)

Workplace data

- Occupational Group
- Workplace Region
- Area of work
- Role description
- Full/part time status
- Patient/non patient facing
- Time in role
- Organisation type
- Salary grade
- Length of service
- Seniority
- Absence information
- Total hours worked