# Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** An investigation to understand the effects depth of context has upon the perception of quality and comprehension of harmonised data: Protocol for a mixed methods study

**Creator:**Esmond Urwin

**Principal Investigator:** Esmond Urwin, Phil Quinlan

**Affiliation:** University of Nottingham

**Funder:** National Institute for Health and Care Research (NIHR)

**Template:** University of Nottingham Staff/PGR Data Management Plan

**ORCID iD:** 0000-0003-4626-2172

**ORCID iD:** 0000-0002-3012-6646

**Project abstract:**

**C4HD (Context for Harmonised Data):**

Utilising a mixed methods approach, this study will conduct a number of case studies utilising questionnaires and interviews. The initial part of the study will focus upon a questionnaire for study participants. Results from the questionnaire will undergo inductive qualitative analysis to assess whether the proposed Data Context Model (see next section) can capture and convey healthcare data contextual information that influences perceptions of quality in transformed data. The second part of the study will conduct a number of interviews to elicit further information and insight from study participants. Inductive qualitative analysis will again be used to assess the results and formulate outcomes. The input to the study will be the an OMOP Data Quality Dashboard Report and a Data Context Report formed from the proposed Data Context Model (detailed in the next section), a study questionnaire and an interview questionnaire. It will employ a number of independent case studies (each case study equates to one participant), each of which is situated within a different context (that being the respective participant's context). All case study participants will be asked to complete a questionnaire, the results of a which will be inductively analysed to produce set of preliminary results. After which a number of participants will be invited to attend an interview. Outputs from the case studies and interviews will be collated and analysed (again utilising inductive analysis techniques) to produce a set of results. These results will then be assessed by the authors of this paper and the case study participants, from which the study outcomes will emerge.

The study questionnaire will be used to elicit and capture the responses, thoughts and perspectives of the participating personnel for each case study and differing contextual setting. Questions will be asked to elicit subjective individual assessments of the study, the Data Context Model and the objectives of the study. The results of those tasks will be captured digitally using Microsoft Forms (MSForms).

Interviews will seek to capture richer and more nuanced responses from the study participants than can be captured using questionnaires. The aim is to delve into participants' experiences and viewpoints around context and what it means from their perspective. A semi-structured interview will be used, composed of closed and open questions, enabling free flowing interviews with scope to diverge and explore insights. Interviews will follow Blandford, Furniss and Makri's 5 step approach to semi-structured interviews. Questions will focus upon whether detailed contextual information can change participants' perspectives and understanding of harmonised data and what potential differences it might produce. All interviews will be recorded using Microsoft Teams and results will be captured digitally using Microsoft Forms (MSForms). Participants will be interviewed once and receive a copy of the results of the interview questions that they took part in, to review the responses recorded. Accordingly, participants are free to withdraw from interviews at any point if they wish to, moreover, they are free not to respond to interview questions.

**ID:** 203380

**Start date:** 01-02-2027

**End date:** 16-06-2028

**Last modified:** 31-10-2026

**Grant number / URL:** NIHR203310

**Copyright information:**

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# An investigation to understand the effects depth of context has upon the perception of quality and comprehension of harmonised data: Protocol for a mixed methods study

## Data description

### What data will you create?

Data Collection:

- Questionnaire [closed & open ended questions] using Microsoft Forms.
- Interviews [closed & open ended questions] using Microsoft Forms.
- The survey will be distributed study participants through a UoN email account and collected through the Microsoft Forms database. This will be downloaded onto an Excel spreadsheet (.xsls) and saved on Microsoft Teams.
- Personal data collected will be name, email address, organisation and job title. Email addresses will be used for results distribution.

Data Types:

- Qualitative data - free text answers for both closed & open ended questions
- Quantitative data - likert scale questions & answers

Data Volume:

- ~25 x questionnaires (1MB per questionnaire).
- ~10 x Interviews (~500MB per interview).
- ~10 x Interview transcripts (~1MB per interview transcript .pdf).
- Total Data Size (before analysis) = ~5.5GB (Raw Data - before deletion of MSTeams interview recordings).
- Total Data Size (after analysis) = ~100MB (Synthesised Data - after deletion of MSTeams interview recordings).

Data Scope: Healthcare Data Transformation - OMOP Common Data model

Software: Microsoft Forms, Microsoft Excel, Microsoft Word, Adobe Acrobat

File formats:  .csv, .xlsx, .docx, .pdf

Data Validity: chosen file formats ensure long-term validity and interoperability.

## Data collection / generation

### What are your methodologies for data collection / generation? How will you ensure data quality?

Utilising a mixed methods approach, this study will conduct a number of case studies utilising questionnaires and interviews. The initial part of the study will focus upon a questionnaire for study participants. Results from the questionnaire will undergo inductive qualitative analysis  to assess whether the proposed Data Context Model (see next section) can capture and convey healthcare data contextual information that influences perceptions of quality in transformed data. The second part of the study will conduct a number of interviews to elicit further information and insight from study participants. Inductive qualitative analysis will again be used to assess the results and formulate outcomes. The input to the study will be the an OMOP Data Quality Dashboard Report and a Data Context Report formed from the proposed Data Context Model (detailed in the next section), a study questionnaire and an interview questionnaire. It will employ a number of independent case studies (each case study equates to one participant), each of which is situated within a different context (that being the respective participant’s context). All case study participants will be asked to complete a questionnaire, the results of a which will be inductively analysed to produce set of preliminary results. After which a number of participants will be invited to attend an interview. Outputs from the case studies and interviews will be collated and analysed (again utilising inductive analysis techniques) to produce a set of results. These results will then be assessed by the authors of this paper and the case study participants, from which the study outcomes will emerge.

The study questionnaire will be used to elicit and capture the responses, thoughts and perspectives of the participating personnel for each case study and differing contextual setting. Questions will be asked to elicit subjective individual assessments of the study, the Data Context Model and the objectives of the study. The results of those tasks will be captured digitally using Microsoft Forms (MSForms).

Interviews will seek to capture richer and more nuanced responses from the study participants than can be captured using questionnaires. The aim is to delve into participants’ experiences and viewpoints around context and what it means from their perspective. A semi-structured interview will be used, composed of closed and open questions, enabling free flowing interviews with scope to diverge and explore insights. Interviews will follow Blandford, Furniss and Makri’s 5 step approach to semi-structured interviews. Questions will focus upon whether detailed contextual information can change participants' perspectives and understanding of harmonised data and what potential differences it might produce. All interviews will be recorded using Microsoft

Teams and results will be captured digitally using Microsoft Forms (MSForms). Participants will be interviewed once and receive a copy of the results of the interview questions that they took part in, to review the responses recorded. Accordingly, participants are free to withdraw from interviews at any point if they wish to, moreover, they are free not to respond to interview questions.

## Personal Data

**Where any research project involves the collection, processing or storage of Personal Data, a DPIA screening must be undertaken to determine whether a Data Protection Impact Assessment (DPIA) may be required.**

A DPIA has been applied for via the UoN Information Compliance Team.

**Project Name:** A mixed methods study protocol to evidence if context influences the perception of quality and comprehension of harmonised data

**Short Title:** C4HD (Context for Harmonised Data)

## Data storage and security

**Where and how will your data be stored, backed-up, transferred, and secured during the active phase (short to medium term) of research?**

**Data Collection**: data will be collected as follows:

- Questionnaires = Microsoft Forms (MSForms)
  - Immediate questionnaire responses will be stored by the online Microsoft Forms database, which will only be accessible through Esmond Urwin's UoN login credentials.
  - Once all responses have been collected, the core dataset will be downloaded as an Excel file (.xlsx) and stored on Microsoft Teams (licensed by UoN) allowing for secure and controlled sharing of data among the research team. Microsoft Teams encrypts data both in transit and at rest and is approved
    against the University's Handling Restricted Data Policy. The service provides several layers of automatic back up and, in a disaster scenario, files can be recovered. Access to data stored in MS Teams is via secure log-in with multi-factor authentication and will only be granted to Esmond Urwin and Phillp Quinlan.
- Interviews = Microsoft Teams (MSTeams) & Microsoft Forms (MSForms):
  - Each interview will be performed using Microsoft Teams and a MSForms interview questionnaire. Each interview will be recorded in a Microsoft Teams meeting, the subsequent recording be stored in the personal secure online UoN OneDrive account for Esmond Urwin, which will only be accessible through Esmond Urwin's UoN login credentials. Each interview MSForms questionnaire will have the questionnaire responses stored by the online Microsoft Forms database, which will only be accessible through Esmond Urwin's UoN login credentials.
  - Once all responses have been collected, the core dataset will be transcribed into an Excel file (.xlsx) and stored on Microsoft Teams (licensed by UoN) allowing for secure and controlled sharing of data among the research team. Microsoft Teams encrypts data both in transit and at rest and is approved
    against the University's Handling Restricted Data Policy. The service provides several layers of automatic back up and, in a disaster scenario, files can be recovered. Access to data stored in MS Teams is via secure log-in with multi-factor authentication and will only be granted to Esmond Urwin and Phillp Quinlan.

**Data Storage & Backup**: data will be stored and backed using UoN provided storage, for our working data. UoN licenses Microsoft Teams, allowing for secure and controlled sharing of data among the research team. Microsoft Teams encrypts data both in transit and at rest and is approved against the University's Handling Restricted Data Policy. The service provides several layers of automatic back up and, in a disaster scenario, files can be recovered. Access to data stored in MS Teams is via secure log-in with multi-factor authentication.

Data will be stored at the following 2 locations:

- UoN OneDrive: Esmond Urwin (personal secure folder)
- UoN Teams/SharePoint: Centre for Health Informatics (strictly access controlled secure folder - access only for Esmond Urwin and Phil Quinlan)

**Devices**: data will be captured and processed using a UoN owned Macbook Laptop.

**Field Work**: Work will not take place in the field.

## Data management, documentation, and curation

**What are your principles (e.g. FAIR), systems, and major standards for data management and creation? What metadata and documentation will you keep? What file formatting and naming conventions will you adopt?**

Data & Metadata:
We will follow UK Data Service best practice for file naming - 25 characters, use underscores to separate elements and include dates in YYYY.MM.DD (ISO 8601).

An example file name would be "2026.04.10_DCM-Extract_Interview_01.xlsx" with raw dataset extracted from "2026.04.10_MSForms_DCM-Raw_Interview_01.xlsx."

Folders will be organised into a hierarchy and folder names will be formatted to section-contents-dataset in order to easily identify which datasets are kept where, and whether they are raw data or active. Each folder will have a sub-folder for completed work to avoid any over-writing or duplication.

The word 'raw' will be used to identify unedited data which is to be preserved as a back up file.

Supporting study-level documentation: Any codes or key words will be documented and saved in a folder named "Key information for this project". This document would allow anyone to pick up the data and understand it without input from myself. It will be stored on Microsoft Teams in 'Data Discovery' then the subfolder 'OMOP Work'. The document will also contain key aims, objectives and
hypotheses.

## Ethics and privacy

**Are there any ethical, commercial, or privacy issues that will affect the collection and storage of your data? E.g. For research involving humans, human tissue samples and animals.**

Ethics approval is currently being sought through the University of Nottingham's Faculty of Medicine and Health Sciences (FMHS) Research Ethics Committee for the questionnaire data capture, interview process, capture of data and subsequent analysis thereof.

Personal data will be collected during this project, and the project has considered ethical and legal implications in its data storage, as well as appropriate security of personal data. All participants will, agree to data collection, data handling, and data distribution in the form of a report, publication or presentation. Research will follow standard ethical procedures of the Faculty Medicine and Health Sciences and the University of Nottingham. Specific aspects will be considered by the Faculty ethics committee as appropriate. In particular, the creation of data from a survey will require ethical approval, including consent forms outlining the storage and use for research purposes of data, including access to those data by project researchers and other researchers, both during and after the life of the project. Participants will be informed that they can withdraw their participation at any stage during or after the interviews. As we will be working with personal data we will ensure that we comply with the Data Protection Act 2018, including GDPR (**Regulation (EU) 2016/679**) requirements. This will include providing research participants with the relevant privacy information and ensuring appropriate safeguards for the storage and handling of data are in place.

## Data preservation and archiving

**How will you ensure the long-term storage and preservation of data?**

All data will be managed according to the University of Nottingham's
data management policy to ensure that all research data be managed in
a manner that supports its authenticity, reliability, security, discoverability and, where appropriate, accessibility for re-use. In particular, we will use the Nottingham Research Data Management Repository (https://rdmc.nottingham.ac.uk/) which will provide a DOI for the published data for easy retrieval.

The University of Nottingham will retain and preserve research data in
line with the University of Nottingham requirements for a minimum of 5 years, but data will be retained for longer periods of time where it is of continual value to users.

## Data sharing and access (internally and externally)

**How will the data generated be shared and published securely? Have all the required agreements been considered?**

All data will be managed according to the University of Nottingham's data management policy to ensure that all research data be managed in a manner that supports its authenticity, reliability, security, discoverability and, where appropriate, accessibility for re-use. In particular, we will use the Nottingham Research Data Management Repository (https://rdmc.nottingham.ac.uk/) which will provide a DOI for the published data for easy retrieval.

Anonymised, transcribed data will be suitable for sharing on request from the relevant university bodies. This will be shared via university SharePoint accounts, to ensure secure sharing of data.

## Data deletion and destruction

**What data must you keep and what data can be deleted/destroyed?**

All interview recordings will be transcribed and then those recordings will be deleted after a period of 6 months from the interview date.
Questionnaires and transcriptions will be kept as these files are of a very small size (.docx, .xlsx, .pdf).

## Roles and responsibilities

**Who will be responsible for managing data, data security, and data quality both during the project and once the project has finished?**

Esmond Urwin will:

- be responsible for the management of data, including:
  - data quality, security and metadata creation.
- create and upload the data.
- ensure compliance with this Data Management Plan.

## Relevant policies

**What are the relevant institutional, departmental or study policies on data sharing and data security?**

The University of Nottingham (UoN) Faculty of Medicine and Health Sciences, as part of the university, adheres to strict data security and sharing policies, primarily governed by the UoN Data Protection Policy and the Code of Research Conduct and Research Ethics. These policies ensure compliance with UK GDPR (**Regulation (EU) 2016/679** ) and the Data Protection Act 2018, emphasizing secure storage, controlled access, and ethical data sharing

## Intellectual property

**Who will own the copyright and IPR of any data that you will collect or create? Will you create a licence(s) for its use and reuse? If you are planning to use existing data as part of your research, do any copyright or other restrictions determine its use?**

The following Creative Commons license will apply to the research work:
**CC BY 4.0**
**Attribution 4.0 International**
You are free to:

1. · Share — copy and redistribute the material in any medium or format for any purpose, even commercially.
2. · Adapt — remix, transform, and build upon the material for any purpose, even commercially.

3. · The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

1. · Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
2. · No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

**Notices:**

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

## Budgeting

**What are the anticipated costs for managing, storing, archiving and sharing research data?**

There will be no additional costs for the managing or storing of the research data.

## Governance

**Completed DMPs must be uploaded into RIS, document module.**

**The DMP must be considered a living document and updated with the relevant details throughout the life of the project.**

Question not answered.

## Further Help

**Tick the box below if you would like your DMP to be reviewed by specialists in Libraries. Please note this is not a necessary step for Ethics approval - Libraries can offer guidance on best practice around DMPs, but are not involved in the Ethics process.**

**Saving this plan after checking the "Yes" box will immediately notify Libraries DMP review service, please only do this when you are ready for review.**

Question not answered.

**Would you like a reminder and further guidance on depositing your data? If so, indicate when would be most useful.**

**Guidance is sent out twice a year, but you can contact library-researchsupport@nottingham.ac.uk at any time for further support.**

Question not answered.

**Tickbox Testing**

**Does this research project involve the collection of personal data?**

- Yes

Ethics approval is currently being sought through the University of Nottingham’s RFaculty of Medicine and Health Sciences (FMHS) Research Ethics Committee at the University of Nottingham for the questionnaire data, interview process, capture of data and subsequent analysis thereof.

A DPIA has been applied for via the UoN Information Compliance Team.

Personal data will be collected during this project, and the project has considered ethical and legal implications in its data storage, as well as appropriate security of personal data. All participants will, agree to data collection, data handling, and data distribution in the form of a report, publication or presentation. Research will follow standard ethical procedures of the Faculty Medicine and Health Sciences and the University of Nottingham. Specific aspects will be considered by the Faculty ethics committee as appropriate. In particular, the creation of data from a survey will require ethical approval, including consent forms outlining the storage and use for research purposes of data, including access to those data by project researchers and other researchers, both during and after the life of the project. Participants will be informed that they can withdraw their participation at any stage during or after the interviews. As we will be working with personal data we will ensure that we comply with the Data Protection Act 2018, including GDPR (**Regulation (EU) 2016/679**) requirements. This will include providing research participants with the relevant privacy information and ensuring appropriate safeguards for the storage and handling of data are in place.