Data Management Plan

**Please complete all sections. Enter N/A if a section is not applicable.**

|  |  |
| --- | --- |
| project details | |
| PI name |  |
| Project Title |  |
| Author(s) of DMP |  |
| Version | 1.0 |
| Date |  |
| 1. What research data will be collected? | |
| Describe the types of research data that will be collected, providing information about media/formats data will be collected in, and the anticipated scale or quantity of each type of data. | |
| **Examples**  *We will collect self-reported data from 100 participants, on age, sex, ethnicity, weight and height, habitual diet, disease history and current health status, as well as the research data generated from our testing procedures, such as the task battery and measures of inflammation and metabolic activity determined by serum blood sampling. All digital data will be pseudonymised using a participant ID number. All data will be in digital format, except the consent forms and self-report questionnaires that will be in paper format.*  *Data will be collected via interviews, which will be recorded and subsequently transcribed. We anticipate recruiting between 3 to 5 families to participate in the study, up to a maximum of 25 participants.*  *Qualitative data on participant opinion will be collected via a paper questionnaire. We anticipate circa 80 questionnaires to be completed.*  *Follow-up semi structured interviews will take place and collect further qualitative data, both via note taking and where the participant agrees, audio recording. Expect circa 5 interviews.*  *This study aims to obtain 20 interview transcripts from 20 individual participants as a means of collecting primary data. These transcripts shall be obtained through conducting video, telephone or in-person semi-structured interviews with the participants. These interviews shall be recorded using the Microsoft Teams video recording function, or in the case of a telephone or in-person interview a sound recording device, such as a Dictaphone. These interviews shall be transcribed in a verbatim style, providing the opportunity for participants to review their own transcripts, and redact any information that they feel that they do not want to be included within the analysis. After which, all completed and reviewed transcripts shall be converted to PDF format.*  *Data will be collected via Online Survey/MS forms.* | |
| 2. What personal data AND confidential information will be processed? | |
| 2.1 Specify the identifying information (personal data) that will be collected (tick all that apply) | |
| **Name**  **Data of Birth/Age**  **Postal Address(es)** (to include postcodes)  **Contact telephone(s)**  **Email address(es)**  **Unique Identifiers** (to include: Student ID numbers, Staff ID numbers, Passport numbers, NHS numbers, National Insurance numbers, ORCID’s, unique research participant ID numbers, Unique applicant ID numbers, vehicle reg, driving licence numbers)  **Images of individuals, including CCTV, photos**  **Location Data** (to include any GPS location data)  **Online Identifiers** (to include IP address data)  **Economic/financial data** (relating to an identifiable individual)  **Educational records** including but not limited to records held by the University and other education providers  **Counselling records**  **Pastoral records, including Extenuating Circumstances Forms**  **Disciplinary records**  **Training records**  **Employment records to include CV’s, references**  **Nationality/Domicile**  **Dietary requirements or preferences**  **Other – Please specify below** | |
| 2.2 Specify any special category or sensitive data that will be collected (tick all that apply) | |
| **Ethnicity**  **Mental Health** (status, medical records or conditions, to include disability)  **Physical Health** (status, medical records or conditions, to include disability)  **Sexual Orientation/Sexual life**  **Genetic Data** (to include DNA data)  **Biometric data** (such as facial scan, iris scan or fingerprint data used for the purposes of identifying a participant)  **Political opinions**  **Trade Union membership**  **Religious or philosophical beliefs**  **Criminal Convictions and offences** (to include alleged offences and convictions)  **Other – Please specify below** | |
| 2.3 Specify any confidential information not specified above that will be collected, e.g. non-public information relating to a business or other organisation. | |
| **Examples**  *Aggregated and anonymous results of internal staff satisfaction surveys of the companies and organisations taking part in the study will be collected.*  *Copies of organisational policies and procedures of the companies and organisations taking part in the study will be collected.* | |
| 3. How will data be stored and transferred during the project? | |
| 3.1 Identify all locations where data will be stored, indicating for each location whether it will be used to store identifying information or de-identified research data, and providing details of access controls that will be applied. | |
| **Examples**  *We will use the Reading University network or University OneDrive account as the primary storage location during the active phase of the project. Both local network storage and OneDrive will provide data security, replication in separate data centres, automated backup, and 3-month file recovery. Raw data and master versions of files will be stored here as read-only files. Confidential data will also be stored in these locations per University of Reading guidance. Data collected in the field will be stored securely, backed up using local devices in the absence of an internet connection, and transferred at the earliest opportunity to the primary storage location. Data will then be deleted from local devices after transfer to primary storage. Use of personally owned devices will be only where University managed devices are not available, and only where use of the device meets the requirements of the University.*  *All identifiable data will be held on REDCap which is securely stored on a University of Reading server and managed in accordance with University Information Compliance Policies.*  *Completed screening questionnaires, food diaries, study questionnaires and signed consent forms (can be either hard or soft copies) will be stored in locked filing cabinets in a locked office at the Department.*  *For collecting and managing consent and questionnaire data, researchers will use UoR REDCap:* [*https://www.reading.ac.uk/research-services/research-data-management/managing-your-data/uor-redcap*](https://www.reading.ac.uk/research-services/research-data-management/managing-your-data/uor-redcap)*. This is a secure online data collection platform maintained on the University local network. It will be used to create data collection instruments as surveys for use by participants and data entry forms for use by project team members. It has an e-consent framework for digital collection of consent. User rights will be controlled so that participants’ details are accessible by authorised users and for other users individual records can only be viewed and exported in de-identified form.*  *The focus groups will be recorded using audio and video recording equipment which will be stored in a locked room in the department when not in use. As soon as possible following the focus groups, the recordings will be saved on OneDrive then permanently deleted from the recording devices. These recordings may be shared with a reputable professional transcription service approved by procurement to assist with transcribing the data.*  *Wherever possible, identifiable data will not be stored on portable devices. If not possible to store using on site data storage facilities, data will be stored on an encrypted USB device. Once a participant has completed their participation, this data will be transferred onto the University of Reading OneDrive, labelled only with the participant identification number, and removed from the encrypted USB drive.*  *Consent forms and devices that contain audio and video recordings from the focus groups will be stored in a filing cabinet within a locked cupboard in the department, which is only accessible to study researchers and locked when unattended. At the end of the project, signed consent forms will be stored securely in the department with access only available to authorised researchers.*  *Data will be collected via Online Survey/Microsoft Forms. The results will be stored on University managed drives in the form of a password protected spreadsheet. Only those authorised to view the personal data will have access.* | |
| 3.2 Describe any administrative measures that you will take to control the risks of inappropriate disclosure, e.g. pseudonymisation, and procedures for secure transfer between locations, e.g. using file encryption and encrypted channels. | |
| **Examples**  *Participants will be de-identified and allocated a coded participant ID # for the duration of the study, which will be used when completing all forms and experiments.*  *All participants will be assigned a six-digit random alpha-numeric study ID. The link table allowing the identification of participants will be stored securely and will be accessible by the PI only.*  *Confidential information will be managed in accordance with the University's Data Protection, Encryption, Bring your Own Device and Remote Working policies. We will password-protect or encrypt individual files. No directly identifiable data will be used in the field to safeguard the data of research participants and a pseudonym will be used for each participant. Data collected in the field will be only accessed on the researcher’s personal electronic devices through password and or finger/facial identification. Interviews will be recorded on an encrypted audio recording device. A member of the team will create pseudonymised transcripts of the audio recordings, saving materials to the secure Reading network. All recordings will be deleted from audio devices after transcription of the recording and secure transfer to the Reading network.*  *The data from this study will be pseudonymised as quickly as possible after data collection. Any direct or indirect identifiers in the data will be removed and replaced with a unique study code recorded in a restricted folder against participant personal information. Information linking participant name with the code will be known only to members of the research team. De-identified research data will be password-protected and stored on the University of Reading's OneDrive account. Researchers that work on the project will only transfer de-identified research data to their work laptops/computers using secured University network to conduct necessary analyses. Data derived from the study will be analysed in a fully pseudonymised format and the link between participant and data will be destroyed at the end of the study.*  *The file containing the participant’s personal information and the link between the participant ID code and email address will be further password protected once saved onto OneDrive with restricted access.*  *A University approved survey tool (such as Online Survey/MS Forms) will be used to collect responses. All survey responses are collected over encrypted connections ensuring that sensitive information can be transmitted securely. Data will be permanently deleted from Online Surveys at the end of the study after transferring to the shared OneDrive.*  *An approved transcription service will be used for transcribing the focus group discussions and data transfer (including audio uploads and transcript downloads) will be secure and password protected. After the transcripts have been created and checked, the electronic audio/video recordings will be deleted from OneDrive. If any of the participants’ names are used within the transcripts, these will be replaced with a participant ID code (different to their main study ID codes).* | |
| 3.3 Specify who will be able to access the identifying information and how you will ensure they process the information securely, e.g. through training, supervision and adherence to secure data handling procedures. | |
| **Examples**  *Prof [name] and PhD student [name] will be able to access the identifying information and Prof [name] is responsible for setting up collaborative drives and controlling access. Both researchers are trained in data security and data handling and will follow best practice throughout. [Student] has also received training in data management.*  *Access will only be granted to research staff directly involved with the project. They will be trained in data collection protocols and how to securely store volunteer information.*  *Study researchers have completed [insert any relevant training]. They also have extensive experience of conducting human studies. Any student researchers involved with the study will be supervised and will receive training on data protection and confidentiality.*  *The only individuals who will be able to access the data are the named researchers on the ethics applications. The lead researcher will be the only member of the study team who will work with directly identifiable data, and all other researchers will only access data that has been allocated to a pseudonymised participant code.* | |
| 4. How will research data be preserved and shared on completion of the project? (undergraduate or pGT projects should enter N/A in this section unless results will be published) | |
| 4.1 Identify the research data that will be preserved and shared at the end of the project by deposit in a public data repository or other archiving solution. | |
| **Examples**  *Only anonymised data from the study will be shared. Shared data may include dietary intake data (e.g., quantities of foods eaten, nutritional intakes and diet quality scores), biomarkers measured from blood samples, qualitative data in the form of anonymised quotes relating to the participants’ experience of CR and the healthy eating advice received during the study, sociodemographic information (including ethnicity, if employed/retired, level of education), age (not date of birth), country-specific location (i.e. all UK), body composition data (e.g. height, weight, and body mass index), and basic physical activity data (e.g. hours per week).*  *All relevant study outcomes described in the Ethics Review Application Form will be preserved in an anonymised form. These will include demographic and anthropometric data (i.e. age, sex, BMI, blood pressure), performance scores of the battery tests assessing cognitive function and mood, and the results of the blood tests measuring immunological and metabolic function.* | |
| 4.2 Describe the measures that will be taken to ensure data are suitable for sharing, e.g. securing consent, anonymising data prior to deposit/sharing, sharing confidential or high-risk information under a controlled access policy. | |
| **Examples**  *All the data uploaded will be anonymised, without any identifiers, e.g., their unique participant ID code, name, and other identifiable data. The information in the consent form relating to this, is as follows: ‘I understand that the data collected from me in this study may be preserved and made available in fully anonymised form (completely unidentifiable and without my participant identification code) so they can be consulted and re-used by others outside of the research team in the future.’ The key ‘file’ linking IDs with participant names will be destroyed before publication of the data or sharing with other researchers ensuring the data becomes anonymous.*  *Our Participant Information Sheet (Appendix 3) and Consent Form (Appendix 4) explicitly mention that participants’ data may be shared via a public data repository only in anonymised form. No confidential or personally identifiable information will be shared with third parties or deposited publicly.*  *Participants will receive information relating to the sharing of their data in the information sheet and can tick a box on the consent form to indicate they understand that their anonymised data will be preserved and made publicly available for others to consult and use. Only the anonymised, cleaned/processed data will be shared, therefore confidential information will not be available to others in the process of sharing the data.*  *The interview transcripts will be anonymised, but given the highly sensitive subject matter, we do not consider open sharing of data is the best option in this case. In order to provide reassurance to the research participants and to ensure risks are robustly managed, we propose to deposit the anonymised transcripts as ‘safeguarded data’ with the UK Data Service ReShare repository. These data would only be available in confidence to researchers registered with the UK Data Service, under the terms of an end-user licence. The Participant Information Sheet will advise participants of our intention and the consent form will include the statement: ‘I understand that the data collected from me in this study will be preserved, and subject to safeguards will be made available to other authenticated researchers’.*  *We believe these interviews will have enduring value as a historic record, and wish to preserve transcripts with the minimum possible redaction for future use by researchers. This will entail retaining personal information relating to the interviewees. The interview transcripts will be deposited with the knowledge of interviewees in the University’s Research Data Archive as a ‘restricted dataset’. This will be held securely by the University and made accessible only if authorised by a Data Access Committee including the PI or nominated data steward and subject to a data access agreement between the University and a recipient research-performing organisation with which the researcher requesting access to the data is affiliated. We will include appropriate wording in the Participant Information Sheet and consent form to advise participants of our intention to preserve and share under safeguards the identifiable data collected from them.*  *As we will be conducting interviews with employees of [employer] in the workplace and asking them questions related to their work, we will seek consent from [employer] to conduct the interviews and to preserve and share them, after appropriate redaction to mask the identities of the interviewees and [employer] and to remove any confidential information. [Employer] will be able to review prepared transcripts, and request further redaction before they are approved for archiving. We can also offer to archive transcripts as ‘safeguarded data’ with the UK Data Service to provide additional reassurance that they will be used responsibly and in confidence by other researchers.* | |
| 4.3 Identify data repositories or other solutions that will be used to preserve and share data. | |
| **These are some examples of data repositories which may be suitable, depending on the type and identification risk of the data in question. For more guidance on suitable data repositories see** [**https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/where-to-archive-data**](https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/where-to-archive-data)**.**  *University of Reading Research Data Archive:* [*https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/uor-research-data-archive*](https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/uor-research-data-archive)*. Suitable for open data and* [*restricted datasets*](https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/uor-research-data-archive#restricted-datasets) *(containing higher-risk anonymised or identifiable data).*  *UK Data Service ReShare:* [*https://reshare.ukdataservice.ac.uk/*](https://reshare.ukdataservice.ac.uk/)*. Suitable for open data and* [*safeguarded data*](https://reshare.ukdataservice.ac.uk/legal/#Safeguarded) *(anonymised but where there is considered to be a risk of disclosure resulting from linkage to other data, such as private databases).*  *Zenodo:* [*https://zenodo.org/*](https://zenodo.org/)*. Suitable for open data.*  *OpenNeuro:* [*https://openneuro.org/*](https://openneuro.org/)*. Suitable for open neuroimaging data, including MRI, PET, MEG, EEG, iEEG.*  *European Bioinformatics Institute repositories, e.g. ArrayExpress, BioStudies, European Genome-phenome Archive (EGA), European Nucleotide Archive (ENA):* [*https://www.ebi.ac.uk/submission/*](https://www.ebi.ac.uk/submission/)*. Suitable for different types of molecular biology data, mostly as open data, but EGA holds and manages controlled access to potentially identifiable genetic and phenotypic human data.* | |
| 5. How will retention and disposal of personal data and confidential information after project completion be managed? | |
| 5.1 State how long you plan to retain personal data/confidential information after the end of the project. | |
| **Examples**  *Personal data/confidential information will be held after the end of the project for 5 years. After this, the PI will schedule regular reviews of personal data holdings to determine whether they need to be retained or can be safely destroyed.*  *Personal data will be retained beyond the end of the project but for no longer than necessary for the purpose of contacting participants for possible future studies. We will schedule regular reviews of personal data that we hold to determine whether they need to be retained or can be safely destroyed. Consent forms will be retained for at least as long as personal data are stored.*  *Paper copies of any personally identifiable data will be destroyed, and digital data will be deleted from the University-managed cloud when no longer needed, but no earlier than 5 years after the study has been completed.*  *Personal data will be retained beyond the end of the project and within scope of appropriate need to contact participants if there is need for possible future studies.* | |
| 5.2 Specify under whose authority this information will be maintained and disposed of after the project. | |
| **Examples**  *The information will be maintained and disposed of by the Principal Investigator, [name]*  *Personal Information will be maintained and disposed of under the authority of the Principal Investigator. If the PI should leave the University, the Personal Information will be under the authority of the Head of School.*  *We have put administrative measures in place to ensure any personal and confidential information will be effectively managed should the project leads change roles or leave the University.* | |