Plan Overview

A Data Management Plan created using DMPonline

Title: Multimorbidity, Physical Activity and Sleep in Rural and Urban Populations in Malawi

Creator: Benedict Warner

Principal Investigator: Benedict Warner

Contributor: Mia Crampin, Stuart Gray, Owen Nkoka, Alison Price

Affiliation: University of Glasgow

Funder: Wellcome Trust

Template: Wellcome Trust Template

ORCID iD: 0009-0008-7654-6073

Project abstract:

Malawi has a high, and rising, prevalence of multimorbidity - the co-occurrence of two or more long-term conditions - as it undergoes a demographic and epidemiological transition from a predominantly young, rural population to an increasingly urban and older population. There is therefore an urgent need to identify opportunities for intervention that may delay or reduce to development of multimorbidity in Malawi.

Many common long-term conditions share proximate causes, including physical activity, sedentary behaviour, and sleep, which together can be described as 24-hour free living movement behaviour. While these are well described in studies from high-income settings, much less is known about the relationship between 24-hour free living movement behaviour and multimorbidity in a low-income setting such as Malawi, where physical activity is more likely to be occupational or travel-related rather than a leisure-time activity.

This study will therefore examine the relationship between 24-hour free living movement behaviours, comprising physical activity, sedentary behaviour, and sleep, and multimorbidity, through the collection of accelerometer data from 4000 participants evenly divided across an urban (Area 25, Lilongwe) and a rural (Chilumba, Karonga) site in Malawi. This data will be linked to demographic, health, and anthropometric data from MEIRU's Long Term Conditions survey, in which this study will be nested, in order to assess the association between these behaviours and multimorbidity. It will also provide baseline data for future, longitudinal analysis as part of subsequent cohort studies of the same participants.

ID: 132787

Start date: 02-08-2023

End date: 04-08-2026

Last modified: 28-06-2024

Grant number / URL: 223499/Z/21/Z

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

Multimorbidity, Physical Activity and Sleep in Rural and Urban Populations in Malawi

Data and software outputs

The data and software outputs your research will generate and/or re-use

The primary data generated by this project will be accelerometer data, taken from devices worn by participants on their wrist for 7 days at a time. This will measure movement (in miligravities) 100 times a second, 24 hours a day, for 7 days each; this allows analysis of a participants' 24-hour movement behaviour profile (including physical activity, sedentary behaviour, and sleep) but does not collect location information or heart rate. The devices collecting this data are Axivity AX3 devices. The raw data produced by the devices is in the .CWA (Continuous Wave Accelerometer) data file format, an open source format developed by Axivity.

The raw data will be extracted from the devices using Axivity's open source program, Open Movement OMGUI. It will then be processed using the open source 'R' package, GGIR, which will produce an output in .CSV (comma separated values) format, which will then be used for the statistical analysis of the data in R.

The data will be linked via unique study identifiers to the existing data gathered by MEIRU's 'Long Term Conditions Survey' (LTC survey) as part of the larger 'Healthy Lives Malawi' project. The LTC survey already gathers detailed information on participants including their demographic details, anthropometric data, self-reported behaviour data, and self-reported data on a range of health conditions. This data is already covered by a separate Data Management Plan.

The secondary data of this project will be the collection of spirometry (lung function) testing for participants in the LTC survey. Alongside the existing anthropometry data collected, participants will in addition be invited to have their lung function assessed using portable spirometers (Vitalograph COPD-6). This will be recorded as their Forced Expiratory Volume in 1 second (FEV1), Forced Expiratory Volume in 6 seconds (FEV6) and FEV1/FEV6 ratio. This information will be entered into tablets via the open-source ODK Collect application by fieldworkers, at the same time as completion of the LTC survey. The survey responses will be uploaded to a locally-hosted ODK Central server, and subsequently loaded into a Microsoft Access database on MEIRU's servers.

The metadata and documentation that will accompany the outputs

An 'Accelerometer Field Log' will be completed by field workers at the time of distribution of the accelerometers. This will include a unique participant ID, serial number of the device being delivered, date and time of the device being delivered to and collected from the participants. The unique participant ID (the 'Accelerometer Link Number') will at the same time be entered into the LTC survey on the ODK Collect application on the tablet, to allow linkage to the pseudonymised data from the LTC survey, for subsequent analysis alongside LTC survey data such as data on long term conditions.

Spirometry data will be entered directly into the ODK Collect application on the tablet, to allow linkage with the LTC survey that includes participant age, sex, height and weight to allow further analysis of lung function against reference ranges.

The files that will be downloaded from the accelerometers will be downloaded to a secure folder on MEIRU's servers, and will be named in the format 'xxxxx_00000xxxxx.cwa' where the first 5 digits are the serial number of the device, and the last 5 digits are the participant's 'Accelerometer Link Number'.

The data captured via the ODK Collect application will be uploaded to MEIRU's ODK Central server, and then loaded into a Microsoft Access Database which stores the LTC survey data. Every participant recruited to any MEIRU study is assigned a unique identifier (their 'ident') which persists across studies, and in addition is assigned a study identifier ('STID') for the purpose of the study.

A data dictionary has been made that documents all the variables collected within the LTC survey. This has been updated to include names and descriptions of variables added by this study, specifically, those variables extracted from the accelerometer data, and the values generated by the spirometry (FEV1, FEV6, and FEV1/FEV6 ratio). Study-level documentation for the LTC has also been produced, in the form of a Data Documentation Initiative (DDI) codebook and MEIRU's Data Management SOP.

Data tables will be documented using Nesstar publisher and this documentation will be displayed on the MEIRU internal data catalogue, along with the latest version of the data table from the database in Stata (and potentially other) format. The data will be replaced on the catalogue on a regular/ad hoc basis. The data documentation without the datasets will also be available on the external version of the catalogue where the study team can direct external interested parties for more information and chance to request data.

Code in R used for statistical analysis, and anonymised processed data for analysis, will be deposited in the University of Glasgow's 'Enlighten: Research Data' repository.

Files deposited in Enlighten: Research Data will be given a Digital Object Identifier (DOI) and the associated metadata will be listed in the University of Glasgow Research Data Registry and the DataCite metadata store. Metadata about datasets held in the University Registry will be publicly searchable and discoverable and will indicate how and on what terms the dataset can be accessed. Information about datasets from the Registry will be displayed on researcher profile pages on the University of Glasgow webpages which will also increase the visibility of the datasets. Only anonymised data for analysis will be transferred to University of Glasgow servers; all pseudonmyized data will remain on MEIRU's secure database.

When you intend to share your data and software

Raw data will be deposited within MEIRU's databases, and subsequent anonymised processed data will be deposited within the University of Glasgow's Enlighten: Research Data repository at the time of submission of the PhD thesis based on this research project and at the time of publication of any work based on this data. A 'Transfer of Data' agreement between the University of Glasgow and MEIRU was signed on 30th April 2020.

Where your data and software will be made available

Raw, pseudonymised data will be deposited in MEIRU's database, which has a detailed data access policy document that describes the general processes and procedures involved in accessing the MEIRU data, available on request from info@meiru.mw.

Anonymised, aggregated processed data will be deposited in the University of Glasgow's 'Enlighten' data repository and will be available at the time of publication of any output based upon this research.

All publications based on this data will be done so via open access arrangements to ensure they are openly available.

How your data and software will be accessible to others

Preliminary data exports will be carried out on a regular/ad hoc basis by the data scientist who will combine the data to run preliminary analytical checks which may feed into questionnaire update/ field protocol refining/re-training etc. Data will be communicated securely from data scientist to scientific team members (i.e., by University of Glasgow FTP Transfer Service); participant identifiers such as name will not be included on exports.

Summary data regarding numbers recruited/preliminary findings will be printed and shared with the wider LTC/MEIRU team on a regular basis. Summary data and feedback based on the data which require actions by the field, data or lab teams will also be printed and shared on an ad hoc basis.

Subsets of the full data will be available on request from MEIRU by external researchers, who can see the external version of the catalogue on MEIRU's website.

Whether limits to data and software sharing are required

Raw data deposited within MEIRU's databases will be linked to pseudonymised data from the LTC survey. Access to this will be strictly controlled and limited to MEIRU's data management team in order to maintain the privacy and confidentiality of study participants. However, it will be possible for anonymised datasets to be shared with external researchers through application to MEIRU's data access committee. Participants will be asked at the time of recruitment to consent to having their data archived by MEIRU and linked to data from other studies.

How datasets and software will be preserved

All data collected will be deposited within MEIRU's repository on MEIRU's servers. This will only be accessible by MEIRU's data team, or securely communicated to researchers in the UK via the University of Glasgow's encrypted 'Transfer' FTP service and securely stored on the University of Glasgow's OneDrive service during the project.

MEIRU's servers are backed up weekly by the University of Glasgow's FTP based backup facility (RCB FTP). MEIRU's data team will be responsible for preserving the data from the project as part of the complete LTC survey dataset for a minimum of 10 years.

Anonymised, processed data will be deposited within Enlighten: Research Data. Data in Enlighten: Research Data, the University of Glasgow's Data Repository, will be issued with a Digital Object Identifier (DOI). This can be included as part of a data citation in publications, allowing the datasets underpinning a publication to be identified and accessed. DOIs will also be linked with appropriate records in Enlighten: Publications, the University’s publication repository, to enhance visibility of datasets. The associated metadata will be listed in the University of Glasgow Research Data Registry and the DataCite metadata store. The retention schedule for data in Enlighten: Research Data will be 10 years from date of deposition in the first instance, with extensions applied to datasets which are subsequently accessed. This complies with both University of Glasgow guidance and funder policies.

Enlighten: Research Data is backed by commercial digital storage with is audited on a twice-yearly basis for compliance with the ISO27001 Information Security Management standard.

Research materials

What materials your research will produce and how these will be made available

This project will produce the following materials:

- Raw accelerometer data, in .CWA format; this will be deposited in MEIRU's repository, pseudonymised and linked to the LTC survey data
- Processed accelerometer data, in .CSV format; this will be deposited in both the University of Glasgow's Enlighten repository, and in MEIRU's repository
- Spirometry data, which will be stored in MEIRU's repository as part of the LTC survey's Microsoft Access database

In addition, any publications based on the analysis of the above data will be made accessible through Open Access publishing.

Resources required

You should consider what resources you may need to deliver your plan and outline where dedicated resources are required.

For collection of the accelerometer data, the project has budgeted for and purchased 115 Axivity AX3 devices. An equivalent of 2 months of programme/data manager time, 24 months of field and data office staff time, and 4 months of driver time has been estimated and budgeted for the project.

For the spirometer data, 14 portable COPD-6 spirometers with disposable, single-use mouth filters will be required. The personnel requirements are included in the above estimates.

Additional contributions toward the overheads for MEIRU's servers and other infrastructure are included in the budget.

Intellectual property

What IP your research will generate

It is not anticipated that this research will produce any significant IP.

How IP will be protected

N/A

How IP will be used to achieve health benefits

N/A

Provide the name and contact details for the person in your organisation (e.g. Technology Transfer Officer or Business Development executive) who can act as a point of contact for Wellcome in connection with the protection and commercialisation of this IP

The University of Glasgow's Research and Innovation Services: College of Medicine, Veterinary & Life Sciences (MVLS) IP & Innovation Managers, Rachel Colman (Rachel.Colman@glasgow.ac.uk) and Natasha Tian (Natasha.Tian@glasgow.ac.uk)