

Mexico City Prospective Study Data and Sample Sharing Policy (version 2.1)



1. Overview

The Mexico City Prospective Study (MCPS) (originally called '*Proyecto Coyoacan*') is a blood-based prospective study of 160,000 Mexican adults. Its main aims are to study the relevance of lifestyle, environmental, biochemical and genetic factors for major chronic diseases (e.g. stroke, heart disease, cancer, diabetes) in Mexican adults, to improve the prevention and treatment of these diseases.

Participants were recruited from 1998-2004. They were interviewed, had a range of physical measurements taken, gave 10 ml of blood (for long-term storage in Oxford of plasma and DNA-containing buffy coat), and gave permission for their subsequent health status to be tracked. Long-term follow-up for mortality is via electronic linkage to Mexico's national death register. This large collaborative research project was designed and is being run jointly by the Nuffield Department of Population Health (NDPH) in the University of Oxford and the National Autonomous University of Mexico (UNAM) in Mexico City.

Access to biological samples is limited by the small volume available. Consideration is given to collaborations that involve an extensive range of quality-controlled assays of all – or large numbers of – samples using high-throughput and cost-effective assay methods. The MCPS study group is actively seeking funding for assay strategies that will transform the available samples into accessible data for use by researchers in Mexico, UK and elsewhere.

At recruitment, as was consistent with standard practice at the time, the participants were not asked specifically for consent to data sharing with outside bodies. However, we understand that the consent was sufficient to permit the supply of pseudonymised data to *bona-fide* researchers of high scientific probity who have agreed to abide by the requirements described in this document and by any contractual arrangements with funders and external suppliers of the data relevant to the datasets.

Within the above constraints, the MCPS study group welcomes proposals for access to data from researchers in Mexico and around the world. This document describes the access policy and procedures, and is summarized in **Appendix 1**. It has been developed in concordance with the general principles of data sharing promoted by various research organisations worldwide including the UK Research and Innovation's common principles on data policy (www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy) and the [Data Access and Sharing Policy of the Nuffield Department of Population Health](#), University of Oxford.

2. Terminology

Data	Any MCPS study dataset, including summary datasets, baseline survey data, re-survey data, follow-up information, blood assay results, genotyping data.
Data Access Agreement	Agreement covering the terms of data access to a Requestor of Open Access Data.
Collaboration Agreement	Agreement covering the terms of data access to a Requestor working with a member of the MCPS study team in either Mexico or the UK.
Open Access Data	Data being made available to external <i>bona fide</i> researchers through the Data Access policy.
Restricted Data	Data stored in the MCPS data repository which has limitations placed on its use or wider distribution.
Requestor	An individual or group of researchers seeking access to data and or samples from the MCPS study.
Data User	An individual or group of researchers that has been granted access to data and or samples from the MCPS.

3. Principles of data sharing

As the MCPS has information on many different exposures and health outcomes over a period of years, a wide range of investigators should be involved in determining which questions to address and how best to address them. As data custodian, the MCPS study group must maintain the integrity of the database for future use and regulate data access. Data can be released outside the MCPS research group only with appropriate security safeguards and approvals. This policy on data access is based on the need to:

- Protect participants and act within the scope of their signed consent;
- Ensure compliance with UK and Mexico legal and regulatory requirements (eg, the UK Human Tissue Act, 2004, the UK Data Protection Act, 2018, the EU General Data Protection Regulation, 2018 and the Mexican General Law for Personal Data Protection, 2017).
- Ensure high quality research is fostered that will advance knowledge. Applications that include Mexican collaborators are particularly welcome since they would help to develop and strengthen the research capacity of local investigators.
- Ensure that the data security and participant confidentiality are maintained.
- Support local capacity building. There is a desire that the data from the MCPS are not only used to generate important research findings but also to help to build research capacity locally.
- Provide academic return and training for the investigators developing the study, in particular for doctoral students and early career researchers who are developing their scientific skills while working on the cohort.

3.1. Key components of this data access policy:

- **Collaborations:** The MCPS research group is actively seeking and responding to requests for scientific collaborations on specific projects, especially when framed in ways that help strengthen Mexican research capacity. This model of facilitated collaboration with external researchers is adopted where it can increase the value and quality of the data. Collaborations are governed by a separate Collaboration Agreement, which: (i) identifies a dedicated project lead from within the MCPS study group in either Mexico and/or the UK; (ii) details arrangements for co-authorship or papers; (iii) covers intellectual property issues; and (iv) details financial commitments where appropriate.
- **Open Access Data Availability:** Before data is approved for any analysis, relevant members of the MCPS team responsible for generating the data must first undertake required cleaning, processing, quality control and integration. As soon as the data are clean, it will be made immediately available for open access sharing with Mexican researchers applying from within a Mexico institution, as long as the proposed project does not overlap significantly with projects already being conducted by the Mexico or Oxford-based investigators (see **Appendix**). This period of exclusive access for Mexican researchers will be for 2 years, after which the datasets will be made available for open access sharing with any *bona fide* researcher worldwide. Details of the currently available data, and a timeline for future dataset releases, are provided in the **Appendix** of this document.
- **Independent Oversight of Access:** Initial decisions on data requests are taken by the MCPS study team. The Nuffield Department of Population Health Data Access Oversight Committee provides further scrutiny and governance advice where necessary. A Requestor can appeal to this committee if they disagree with a study decision on access.
- **Protecting the Identity of Participants:** Safeguards are maintained to ensure the anonymity and confidentiality of participants' data. Researchers will need to enter a legal agreement not to make any attempt to identify participants, and the data provided to researchers will not contain any personally identifiable variables (i.e. every data set provided will be "pseudonymised" with uniquely encrypted participant identifiers [PIDs]).
- **Data Security:** All MCPS data is held on secure servers in a central data repository that is compliant with internationally recognised information governance standards. A data management team acts as gatekeepers and ensure that any shared data is delivered through a secure data delivery system and in an appropriate format.
- **Sample Preservation and Access:** Only 10 ml of blood was taken at baseline from each participant, which was divided into 1 buffy coat sample and 2-3 plasma samples that are stored in Oxford. Given the very limited amount of this depletable resource, access to these samples needs

to be carefully controlled. Consequently, rather than assay samples on a nested case-control basis (which is cost-effective for studying a particular condition but not for a resource that is to be used to study many different conditions by different researchers), assays of the samples from all participants is far preferable. Such a strategy maximises the information available to researchers while minimizing sample depletion and facilitating different comparisons across the cohort since the assay methodology and quality control would be consistent. Suggestions for particular assays to be included are welcomed, and all assay values will become part of the available dataset. In general (as is the case in UK Biobank), it is not expected that requests for direct access to samples will be agreed to by the Access Committee.

- **Fees for data access:** Open Access Data is available free of charge to applicants from Mexico. Researchers based elsewhere in the world will incur an Access Charge for each approved data request (currently £250 GBP for initial review of an application and, once approved, a further £2250 GBP prior to Data release). This contributes to the administrative costs incurred in managing and reviewing the application, and in preparing the individual datasets. Collaborating researchers may also be required to cover the costs of administering the data sharing (including legal fees if applicable), retrieving, processing and sending the data or samples. Estimated costs for a particular request will be provided during the development of the project proposal.

4. Data access process

Potential Collaborators and Data Requestors should first contact MCPS Investigators or review the MCPS study website (including previous MCPS publications) to gain an understanding of the available study data and of projects that have previously been completed and are currently being undertaken.

4.1 Collaboration requests

Researchers who are interested in collaborating on projects with MCPS researchers in Mexico or the UK are encouraged to approach the MCPS study group informally in the first instance by email to mcps-access@ndph.ox.ac.uk or to contact relevant MCPS investigators to discuss research ideas and feasibility. Enquiries should include a project title and brief outline of the research project and the relevant data of interest. Each project will require a co-investigator from within the MCPS study group who has a common interest in the project and relevant or complementary research expertise. Once identified, the collaborator and the co-investigator will co-develop a research proposal which will then be reviewed by the MCPS Steering Committee.

4.2 Open Access Data requests

Registration / eligibility

All Open Access Data Requestors are first required to complete the **MCPS Data Access Registration form** (available in [English](#) and [Spanish](#)) and send it to mcps-access@ndph.ox.ac.uk. Requestors should be employees of a recognized academic institution, health service organization or charitable research organization with experience in medical research. They should be able to clearly demonstrate, through their peer reviewed publications in the area of interest, their ability to conduct independent research.

Submission of a data request

Once approved, Data Requestors will be sent an **MCPS Project Description and Data Request form** and will need to return the completed form to mcps-access@ndph.ox.ac.uk. This form requires the Requestor to provide: a project title and abstract; scientific rationale / methodology; anticipated outputs and project timeline. Additional questions cover ethical issues, collaborators / research team, funding support and data security, and the data variables they would like to receive.

Review of a data request

Open Access Data requests will initially be assessed by the MCPS study investigators. Each application will be considered on its individual merit. If necessary, independent peer review will be sought. Approved projects will: (i) have clearly defined objectives; include a sound methodology that is likely to generate meaningful results; (ii) be based on an appropriate and available selection of data; (iii) have clearly defined timelines and outputs (e.g. 1-2 papers in peer-reviewed journals). Projects that overlap significantly with approved and/or completed projects may be rejected (see **Appendix**).

The MCPS team aim to review and respond to Data Requests within 2-4 weeks. A Requestor can appeal to the Nuffield Department of Population Health's Independent Data Access Oversight Committee if their request

is denied by the project team.

5. Terms of data access

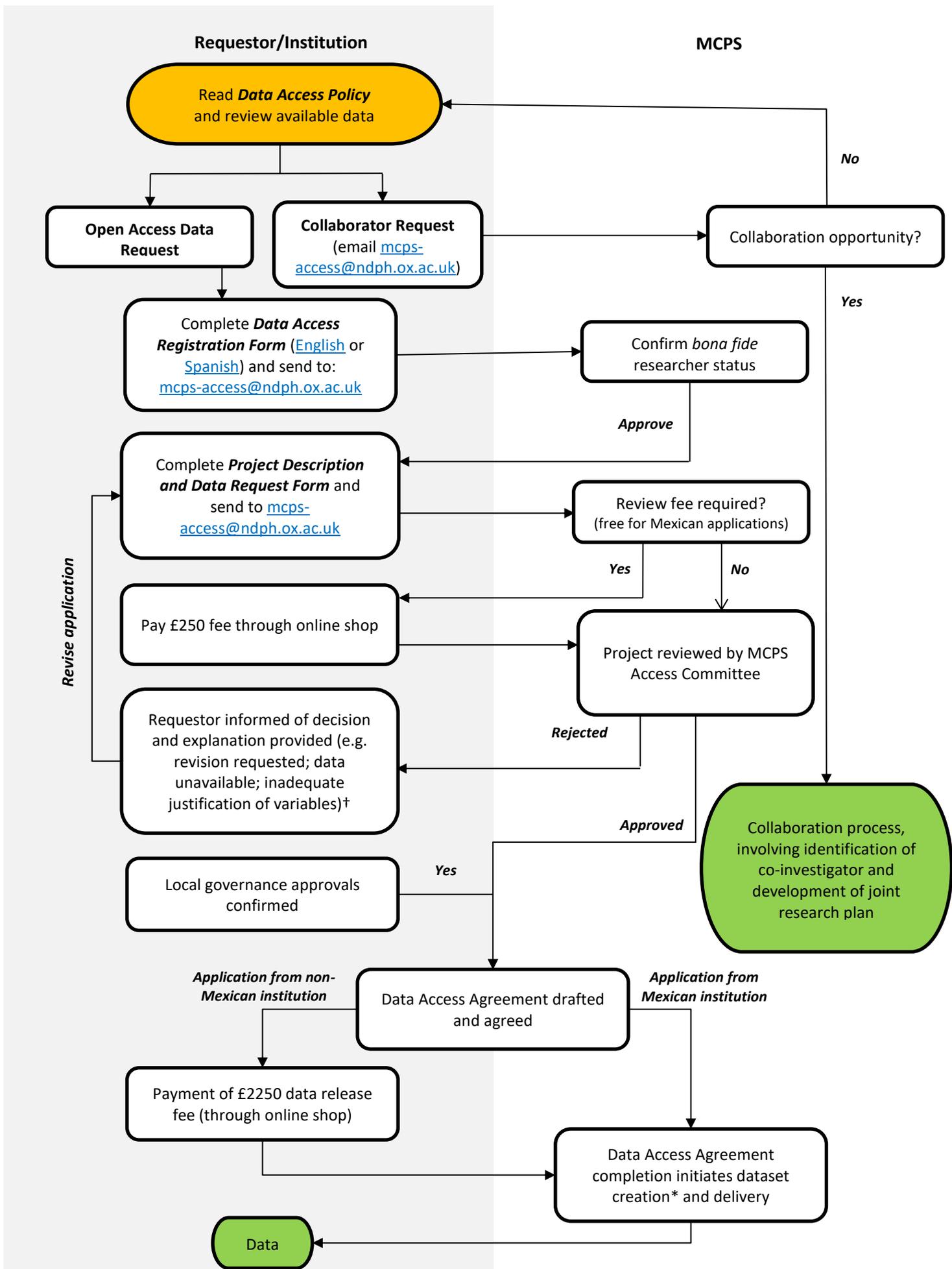
Once proposals are approved the following conditions and undertakings are required as conditions of access:

- **Data Access Agreement.** Before any data are transferred a signed transfer agreement must be in place between the Requestor's institution and the University of Oxford. This transfer agreement will be sent to the Requestor once their application is approved, and will include a copy of the approved project proposal as a Schedule.
- **Signing Authority.** Requestors should be acting as members of a recognised academic institution, research organisation or health organisation. Their request should come from a recognised email domain. Their organisation should have formal policies and procedures (i.e. IG Toolkit, ISO 27001 certification or System Level Security Policy assessment) to comply with any legal, ethical or data protection constraints and to ensure that the dataset is stored securely and used responsibly.
- **Ethics and Research Governance Approval.** Where applicable Ethics Committee approval for the research is the responsibility of the Requestor. The Requestor, in conjunction with study investigators, may also need to obtain approval from the Research Ethics Committees responsible for the MCPS study. Local Research Governance approval and R&D approvals, if required, are the responsibility of the Requestor. Approvals will need to be in place before any data are transferred.
- **Limitations on Use.** The data will be used for the purposes of medical research only and within the constraints of the consent under which the data were originally gathered, and of any contractual agreements between the MCPS study and its funders or external data sources. Access will be permitted only for research that is consistent with the originally-submitted project description, has been ethically and scientifically approved by appropriate independent reviewers and where the use of the data will be for the demonstrable benefit of health and/or social care. Data supplied may only be transferred to Requestors named at the time of the original application or in subsequent applications and specified in the Access Agreement or later amendments. Data from the collection cannot be transferred to individuals outside the Requestor's research group without formal approval by the MCPS Principal Investigators.
- **Identifying Data.** The data provided to researchers will not contain any personally identifiable variables. Data sets will be "pseudonymised" with uniquely encrypted participant identifiers (PIDs). The Access Agreement will contain confidentiality undertakings to further safeguard participants' privacy. Recipients must agree not to link the pseudonymised data provided with any other data set without permission. Recipients must not attempt to identify any individual from the data provided. Should recipients believe that they have inadvertently identified any individual, they must not record this, share the identification with any other person or attempt to contact the individual.
- **Intellectual Property.** All Intellectual Property Rights in the Data are and shall remain at all times the property of UNAM and the University of Oxford. All Arising Intellectual Property shall vest in and be owned by the Requestors. The Requestors will cover any cost for the protection of arising Intellectual Property. The Requestors shall promptly disclose any such Arising IP in writing to the Principal Investigators. UNAM and the University of Oxford will be granted rights to use all Arising Intellectual Property for academic and research purposes, including research involving projects funded by third parties provided that those parties gain or claim no rights to such Arising Intellectual Property.
- **Payment of Access Charges.** Data Requestors from institutions outside of Mexico are expected to pay Access Charges to contribute to the administrative cost to the study of reviewing the application and preparing data for sharing, etc. Where these are applied, no Data will be provided to the Data Requestor until or unless the Access Charges are received in full.
- **Data Release and Delivery.** Once the proposal is approved and the Access Agreement signed, the data and its documentation will be generated in CSV (or any other pre-specified) format, encrypted and released in a secure manner.
- **Publicity and Dissemination.** The MCPS study team reserves the right to publish the title, the names(s) and affiliations(s) of the Chief Investigator(s), a lay summary and a scientific abstract of each piece of collaborative research for which access to the resource has been granted, before identification or publication of results. Requestors who do not wish details of their study to be openly

available need to state this in their data request and give the reason. The Requestor shall not use the name or any trademark or logo of UNAM or the University of Oxford in any press release or product advertising, or for any other commercial purpose, without prior written consent.

- **Authorship and Approvals.** Collaboration Agreements and Access Agreements will specify expectations regarding authorship and acknowledgements on research outputs. Collaborations require at least one co-author from the MCPS study group. For Open Access Agreements no authorship from the MCPS team is required, but Requestors are nevertheless asked to submit proposed publications to the MCPS team for review not less than 30 days in advance of the submission for publication.
- **Publications and Open Access.** All publications of the Results in a peer-reviewed journal, or as a scholarly monograph or book chapter, must be made available from PubMed Central and Europe PubMed Central as soon as possible and no later than six months from the date of final publication. Journal requirements for data release and deposition that may be requested following publication of an article must be discussed with and approved by the Principal Investigators prior to submission of a manuscript.
- **Integration of the Data.** After completion of work using released MCPS data, the original dataset as well as any derived dataset and/or variables generated during the research must be returned to the MCPS central data repository for archiving and/or merging with the main database for future use. If considered appropriate, the MCPS staff may carry out independent checks and/or validation of the data and results to ensure the continued data integrity and reliability of the study findings.
- **Monitoring and Accountability.** The Data Requestor shall be required to submit annual reports and any other information reasonably requested to evidence the work undertaken in connection with the proposed project. If there is substantial deviation or change in the planned use of the data, further approval will be needed. If there is substantial delay or difficulty in completing the planned research, the MCPS study team will have the right, after consultation with the Nuffield Department of Population Health Data Access Oversight Committee, to terminate the work if in its view there is little chance that the problem will be rectified. Under such circumstances, all Data that have been provided must be deleted and a deletion certificate provided.

Appendix 1: Flow diagram of data request process



† The requestor may appeal the decision as described in the NDPH Data Access Policy.

* All data sets pseudonymised with unique participant identifiers, download uses secure encryption.

Appendix 2: Details of currently available data (for researchers worldwide)

Baseline characteristics	
Month and year of recruitment	Reproductive history (women)
Socio-demographic characteristics	<ul style="list-style-type: none">• Menopausal status• Hysterectomy• Oophorectomy• Hormone replacement therapy• Contraceptive use• Age at first sexual relationship• Age at first pregnancy• Number of pregnancies
<ul style="list-style-type: none">• Age• Sex• Area of residence• Marital status• Educational achievement• Occupation• Income• Health service provider	Physical measurements
Lifestyle characteristics	<ul style="list-style-type: none">• Height• Weight• Waist circumference• Hip circumference• Systolic blood pressure• Diastolic blood pressure
<ul style="list-style-type: none">• Smoking• Passive smoking• Alcohol consumption• Physical activity• Sleep duration• Fruit/vegetable intake• Fried food intake• Type of cooking oil used	Blood sampling data
Prior diseases (self-reported)	<ul style="list-style-type: none">• Time of blood sampling• Time since last meal
List of medications taken (self-reported)	Glycosylated haemoglobin
Mortality data (up to 1st Jan 2016)	
<ul style="list-style-type: none">• Time from recruitment to death / end of follow-up• ICD-10 underlying cause• ICD-10 contributory causes• Timing/duration of diseases• Location of death• Seen by doctor before death	

Appendix 3: Anticipated timeline for future data releases

- **June 2020 (Mexican researchers). Resurvey data on 10,000 survivors recruited between 2015 and 2019.** Same questions/assessments as at baseline plus: additional questions related to diabetes consequences and control; MMSE questionnaire; questions related to fractures and falls; additional dietary questions; bioimpedance measures; spirometry.

Appendix 4: Protected research projects (using currently available data [appendix 2])

- General and central adiposity and cause-specific mortality.
- Smoking, alcohol consumption and cause-specific mortality.
- Blood pressure and risk of death from a vascular or renal cause.
- Physical activity, diet and cause-specific mortality.
- Cardiovascular mortality risk prediction equations.
- Factors associated with the control and treatment of diabetes.
- Obesity, diabetes and cognitive function.