PEDIGREE

Pathology, Epidemiology, DNA, Informatics & Genetics
A Research Enabling Enterprise

Policy & Procedures
Access to Data and Biospecimens

1 Contact details

Postal address: PEDIGREE Coordinating Centre
615 St Kilda Road
Melbourne 3004

Main contact: PEDIGREE Executive Officer
Ashley Fletcher
Phone: (03) 9514 6272
Email: pedigree@cancervic.org.au

2 Table of Contents

3 Background to PEDIGREE

The Pathology, Epidemiology & DNA Information: a Genetic Research Enabling Enterprise (PEDIGREE) is a consortium formed to provide research resources and technical platforms for studying genes and environment in the causation of cancer. It comprises the following:

3.1 The Melbourne Collaborative Cohort Study (MCCS)

The MCCS was set up in the early 1990s to investigate prospectively the role of diet and other lifestyle factors in causing common chronic diseases – especially cancers of the prostate, breast and bowel – and to investigate possible interactions between these exposures and common genetic variants. Cohort recruitment was funded by *VicHealth* and *Cancer Council of Victoria* (formerly known as the Anti-Cancer Council of Victoria). Continuing data management and follow-up are funded by the Cancer Council of Victoria.

The MCCS (also known as *Health 2020*) was designed to be the largest prospective cohort study conducted in Australia. Between 1990 and 1994, 41,500 people, (24,500 women and 17,000 men) aged 40-69 were recruited to the study. The unique features of this study are: *a*) 30% of participants are southern European migrants, giving the study its wide range of lifestyle exposures; *b*) collection of blood samples and physical measurements from all subjects; and *c*) face-to-face follow-up, enabling collection of further blood samples, repeated measures of exposures and assessment of a number of non-fatal, non-cancer endpoints. The repeated measures of key lifestyle exposures and selected molecules in plasma reduce measurement error and further increase the statistical power and temporal relevance.

Extensive information was collected at baseline in face-to-face interviews that included questionnaires (diet, physical activity etc.) and physical measurements, including lean and fat mass by bioelectric impedance, and blood pressure. A food frequency questionnaire was developed specifically to measure dietary intake in the cohort, with the food list based on weighed food records in a group of around 800 men and women reflecting the main country of birth groups within the MCCS. Blood samples were drawn and whole blood and plasma stored for analysis of DNA and other molecules of interest (e.g., sex hormones and growth factors, carotenoids and fatty acids involved in disease pathways).

Cases of cancer and deaths are identified by regular matching of the MCCS to cancer registries and death indices.

Follow up 2 occurred between 2003 and 2007 with the aim of collecting similar type of data to baseline, but using up-dated survey instruments, especially for diet and physical activity. Overall 62% of men and 68% of women attended. It was also used as an opportunity to collect more of the non-fatal, non-cancer outcomes associated with ageing. This included using the Kessler 10 questionnaire as a measure of psychological distress or anxiety, questions on ability to perform activities of daily living and instrumental activities of daily living, and details about a range of health conditions and the date of diagnosis.

3.2 Australian Breast Cancer Family Registry (ABCFR)

The ABCFR is a resource of families, data, bio-specimens, researchers and community representatives established for the conduct of collaborative research on breast cancer. It is part of an international registry set up in the 1990s by the National Cancer Institute (NCI USA). Researchers from the USA, Canada and Australia have recruited volunteer families into six registries using common questionnaires and protocols. Data are collated at a centralised Informatics Support Centre. ABCFR has collected epidemiological risk factor and family history data for 8,700 participants from 2,200 families and 4,700 blood samples. Tumour material has been collected and the Breast CFR Pathology review has been conducted for 1,000 cases. Almost 50,000 DNA samples have been shipped to researchers.

3.3 Australasian Colorectal Cancer Family Registry (ACCFR)

The ACCFR is Australia's foremost resource for research into the genetic and environmental causes of bowel cancer. It has recruited colorectal cancer families and collected blood samples, epidemiological questionnaires, tumour samples, medical records, dietary questionnaires and family cancer histories from more than 28,000 individuals from over 7,400 families. This study is funded by the National Institutes of Health (USA) as part of an international consortium, the Colon Cancer Family Registry. Other collaborating institutions include the Mayo Clinic, the Fred Hutchinson Cancer Research Center, the University of Hawaii, Cancer Care Ontario, and the University of Southern California. More than 13,000 families have been recruited. Major funded projects stemming from the Colon CFR include: studies to identify new genes for bowel cancer; studies on bowel cancer aetiology including molecular pathways; studies on the risk of cancer in carriers of mutations in known high risk genes; and studies on cancer prevention including screening – as well as over 100 other studies related to bowel cancer.

3.4 Australian Prostate Cancer Family Program (ACPFP)

The APCFP was established in 1997, with epidemiological risk factor data, family history data and biospecimens from 4,000 participants from 1,500 population-based and 50 clinic-based prostate cancer families from Melbourne, Sydney and Perth. Since then it has been augmented with data and biospecimens from >1,000 MCCS participants diagnosed with prostate cancer and with a population-based series of men diagnosed at an early age and their family members. Between 1998 and 2008, all men with histopathologically confirmed carcinoma of the prostate were identified though the population-complete Victorian Cancer Registry. The Early Onset Prostate Cancer Family Study (EOPCFS) included all men younger than 55 years at diagnosis but the upper age limit varied over time according to available numbers, with quotas being filled by randomly sampling additional cases aged from 55 to 59 years. Overall, 68% of eligible men approached participated in the study.

4 Access to PEDIGREE information

4.1 Release of data and biospecimens

For the purposes of this document, "data" refers to information collected from PEDIGREE study participants, proxies, other organisations or agencies and as a result of analysis of biospecimens. "Biospecimens" refer to any biological samples collected from participants.

PEDIGREE data and biospecimens may only be released to collaborating researchers in accordance with the PEDIGREE Privacy Policy. PEDIGREE studies collect, use and disclose information concerning participants within the relevant Commonwealth and State privacy legislation and according to the requirements of the Human Research Ethics Committees (HREC) of Cancer Council Victoria or the University of Melbourne. The PEDIGREE Privacy Policy is available on the PEDIGREE website.

Information and biospecimens are released on the understanding that:

- they shall only be collected, used, disclosed and stored within the terms of the memorandum of understanding (MOU) and *Additional Terms of Agreement* and/or materials transfer agreement (MTA),
- they shall not be matched, in whole or in part, with any other information or biospecimens for the purposes of attempting to identify individuals, nor shall any other attempt to identify an individual be made, and
- any new information generated shall be returned via the PEDIGREE Coordinator for incorporation with the central database.

4.2 Access to different types of information

These definitions are from the NHMRC's National Statement on Ethical Conduct in Research Involving Humans:

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72.pdf, accessed on 17 September 2012

- **Individually identifiable data**, where the identity of a specific individual can reasonably be ascertained. Examples of identifiers include the individual's name, image, date of birth or address;
- **Re-identifiable data**, from which identifiers have been removed and replaced by a code, but it remains possible to re-identify a specific individual by, for example, using the code or linking different data sets;
- **Non-identifiable data**, which have never been labelled with individual identifiers or from which identifiers have been permanently removed, and by means of which no specific individual can be identified. A subset of non-identifiable data are those that can be linked with other data so it can be known that they are about the same data subject, although the person's identity remains unknown.

The term "de-identified" is used frequently to refer to information from which only names have been removed. Such information, depending on the size of the data set may remain "re-identifiable".

The table below delineates how to access different types of information from PEDIGREE.

Type of Information	Participant Contact	Participant Consent	HREC Approvals*	Data Supplied
Individually Identifiable	Yes	Yes	Yes	Names and addresses from relevant participants who agreed to participate
	No	Yes	Yes	PEDIGREE collects additional information on behalf of the researcher; then they release data in re-identifiable form
Re-identifiable	No	No	Yes	Released with a unique identifier to enable PEDIGREE to re-identify the participant if necessary
Non-identifiable	No	No	PEDIGREE **or** the researcher's institution	PEDIGREE erases the unique identifier used so that the data set can never be linked with the main PEDIGREE database

Data shall be released in "as close to de-identified form as possible", i.e., dates of birth and postcodes will only be released when a project cannot proceed without them. All identifiable and potentially identifiable data will carry a PEDIGREE identification number to enable linkage with the main database.

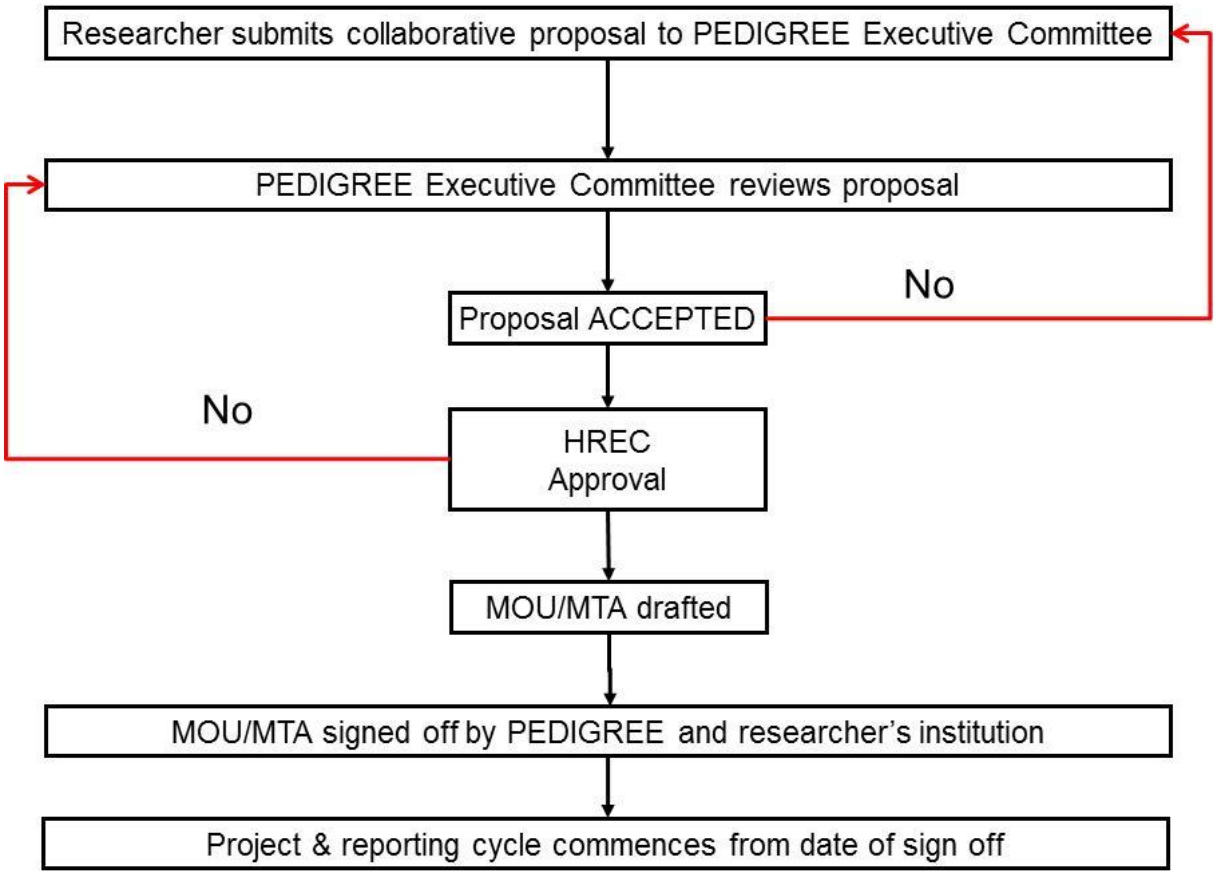
5 Procedure for accessing PEDIGREE data and biospecimens

The main steps to be taken by researchers interested in carrying out studies with PEDIGREE are delineated in the figure below.

Requirements

Researcher must read and agree to policies and write proposal in line with PEDIGREE guidelines.

Commencement of accepted projects is subject to gaining HREC approval for high risk projects plus institutional sign off of MOU / MTA and the payment of fees.



The PEDIGREE Coordinator will arrange for data extracts required for each approved project. If the data are de-identified and <4MB they can be sent via email. All other data must be collected in person or sent on CD to the Principal Investigator via registered post.

For projects that require contact with participants, or release of biospecimens, the Principal Investigator should arrange a meeting with the PEDIGREE Executive Committee and other relevant people to discuss how the project should proceed.

6 Data and Biospecimens

6.1 Ownership, Storage & archiving

PEDIGREE data and biospecimens are held in trust by the Cancer Epidemiology Centre (CEC) at Cancer Council Victoria and by the MEGA Epidemiology Centre of The School of Population Health at The University of Melbourne. Samples are held on behalf of CEC and

MEGA at the Genetic Epidemiology Laboratory (GEL), The Department of Pathology, The University of Melbourne. Release of samples is coordinated by the GEL.

To ensure the PEDIGREE database remains complete and the resource is used to maximal efficiency, collaborating investigators must submit all new data measured or collected to the central PEDIGREE database as agreed in the *Additional Terms of Agreement*. It will generally be expected that any new data measured or collected will be submitted, prior to publication, as a complete data set with a data dictionary of variables and valid values. Similarly, it is a condition of access to PEDIGREE biospecimens (including DNA) that any new biospecimens generated by researchers should be submitted for addition to the biospecimen repository. For biospecimens that have not been collected by PEDIGREE staff, details of the specimen collection must be lodged with the PEDIGREE coordinator.

These new data or biospecimens become the property of PEDIGREE (as well as that of the researcher(s)) and will be protected and accessed according to the "*PEDIGREE policies and procedures for access to and use of data and biospecimens*". An access exclusion period may be negotiated with the Principal Investigators of projects delivering new data and biospecimens to PEDIGREE.

During any agreed access exclusion period, data generated by a specific project or biospecimens collected for a specific project will not be available to another researcher applying for data access from PEDIGREE without the permission of the Principal Investigator of the study which collected the data. Where the Principal Investigator provides permission to release the data or biospecimens, approval of the relevant HREC must also be obtained.

Any paper-based records (including questionnaires and consent) should be stored with the main PEDIGREE record collections at either the CEC or MEGA. The storage of paper based records should be arranged with the PEDIGREE Coordinator. Biospecimens shall be stored in the manner agreed in the *Additional Terms of Agreement*.

Any PEDIGREE data and the analysis thereof (paper and electronic) prepared for the purposes of publication should be retained by the institution housing the Principal Investigator that conducted the study. These files should be stored in accordance with NHMRC guidelines and the relevant privacy legislation. Where it is not possible to store data or perform analyses within the Principal Investigator's institution, it can be stored by the CEC or MEGA on behalf of PEDIGREE. Data and analyses to be stored by PEDIGREE should be submitted on CD-R in duplicate to the PEDIGREE Coordinator.

6.2 Charges associated with obtaining data

Charges may be levied to cover the costs associated with:

- Obtaining participants' permission to provide personal details to a researcher.
- Extracting and preparing data for the project
- Data analysis

Charges may be levied to cover the cost of supplying biospecimens; such costs include personnel required to select samples from storage, aliquot of samples, transport of samples and consumables.

7 Reporting

Each Project's Principal Investigator is required to submit an annual report on the anniversary of the signing of the MOU/MTA for the project's duration. The annual report should include a progress report, any changes to protocols not already notified, details of publications and presentations, the names of active investigators including any new investigators and provisions for ensuring the privacy and confidentiality of information. The PEDIGREE Coordinator will notify the Principal Investigator when reports are due.

The Principal Investigator will also be asked to submit a final report on completion of the project. All projects are deemed completed after three years from the date of a signed MOU/MTA unless a new application is approved.

8 PEDIGREE: memorandum of understanding

Between

Cancer Epidemiology Centre, Cancer Council Victoria & Centre for MEGA Epidemiology, The University of Melbourne

AND

(INSERT INSTITUTION)

The research covered by this memorandum is restricted to the project entitled:

Project title

The principal investigator for this project will be:

Name of PI

The data items agreed to be provided for this project includes measures of:

A

B

C

D

E

F

Approval of the project protocol has been obtained from the following HREC(s):

(name of HREC)

The principal collaborators agree:

To complete the project by {*insert date*} at which time, if not renewed, the MOU will expire

To abide by the *PEDIGREE Data Access & Publication Guidelines* and *Additional Terms of Agreement*

To conduct this research as described in the *Collaborators application* AND advise the PEDIGREE Executive Committee of any changes to this application

To collect, store and use data and biospecimens consistent with PEDIGREE's HREC requirements and in accordance with relevant privacy legislation

That all costs of the substudy, and charges for PEDIGREE biospecimens and the cost of obtaining data be covered by the budget for the research proposal

Not to disclose or release data or biospecimens obtained or collected as part of The Project to third parties

That new data or biospecimens become the property of PEDIGREE (as well as that of the researcher(s)) and will be protected and accessed according to the "PEDIGREE policies and procedures for access to and use of data and biospecimens". An access exclusion period may be negotiated with the Principal Investigators of projects delivering new data and biospecimens to PEDIGREE

To register all publications and presentations that arise from this research with the PEDIGREE Publications Coordinator

To provide a progress report every 12 months for the duration of this project and a final report at project conclusion

To acknowledge PEDIGREE, and its relevant components, as the source of the data in all publications, presentations and in the media

To provide copies of all data files and statistical codes to the PEDIGREE Co-ordinator on completion of the studies

To ensure that all quantitative and qualitative data collected as part of a substudy will be provided to PEDIGREE for secure electronic archiving. These data may be used in future projects, and the project leader who collected the data will always be invited to participate in these future analyses, and

That hard copies of data will be the responsibility of the project leader and must be stored and disposed of in accordance with HREC and NHMRC guidelines http://www.nhmrc.gov.au/publications/synopses/r39syn.htm

PEDIGREE agrees not to undertake work, either by itself, or in collaboration with other researchers, that would compromise The Project for the length of this MOU.

Prof Graham Giles on behalf of the Cancer Epidemiology Centre, The Cancer Council Victoria	Name of Principal Collaborator on behalf of (institution)
_____ *signature* *Date:* _____/_____/_____	_____ *signature* *Date:* _____/_____/_____

8.1 Possible additional terms of agreement

The principal collaborators agree:

To meet the costs involved in preparing and shipping biospecimens or data and in collecting and extracting data.

To store and lodge data collected as part of The Project with PEDIGREE. {*Insert details of format of file and timelines*}

To store and lodge biospecimens collected as part of The Project with PEDIGREE. {*Insert details of protocol to be followed and timelines*}

That the information shall not be matched, in whole or in part, with any other information for the purposes of attempting to identify individuals, nor shall any other attempt to identify an individual be made.

That once supplied, de-identified data cannot ever be matched to the PEDIGREE main database.

That the source of funding for the data supplied shall be acknowledged as follows – {*insert acknowledgement*}

The Cancer Epidemiology Centre, The Cancer Council Victoria agrees:

Not to disclose or release data or biospecimens generated by The Project to another researcher without the permission of the Principal Investigator, until the end of the agreed exclusion period

To archive the Project data and analysis once it has been accepted for publication.