BMJ Open

Protocol for the Endometriosis Research Queensland Study (ERQS): An integrated approach for research to improve diagnosis and stratify treatment.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-064073
Article Type:	Protocol
Date Submitted by the Author:	20-Apr-2022
Complete List of Authors:	Tanaka, Keisuke; Royal Brisbane and Women's Hospital, Department of Obstetrics & Gynaecology, Womens' and Newborn Services; University of Queensland Gilroy, Deborah; University of Queensland Subramaniam, Sugarniya; University of Queensland Lakshmi, Preethi; University of Queensland Bhadravathi Lokeshappa, Madhura; The University of Queensland Wallace, Leanne; University of Queensland Atluri, Sharat; The University of Queensland Schmidt, Bart; University of Queensland Ganter, Peter; University of Queensland Baartz, David; Royal Brisbane and Women's Hospital, Department of Obstetrics & Gynaecology, Womens' and Newborn Services Smith, Matthew; Royal Brisbane and Women's Hospital, Department of Obstetrics & Gynaecology, Womens' and Newborn Services Mortlock, Sally; University of Queensland Henders, Anjali k.; University of Queensland Khalil, Akram; Royal Brisbane and Women's Hospital, Department of Obstetrics & Gynaecology, Womens' and Newborn Services Montgomery, Grant ; The University of Queensland Institute for Molecular Bioscience McKinnon, Brett; The University of Queensland - Saint Lucia Campus, Institute for Molecular Bioscience Amoako, Akwasi ; Royal Brisbane and Women's Hospital, Department of Obstetrics & Gynaecology, Womens' and Newborn Services; University of Queensland
Keywords:	GYNAECOLOGY, REPRODUCTIVE MEDICINE, GENETICS

Protocol for the Endometriosis Research Queensland Study (ERQS): An integrated approach for research to improve diagnosis and stratify treatment.

KEISUKE TANAKA[1,2], DEBORAH GILROY[4], SUGARNIYA, SUBRAMANIAM[3], PREETHI LAKSHMI[4], MADHURA BHADRAVATHI LOKESHAPPA[4], LEANNE M. WALLACE[4], SHARAT ATLURI, BART SCHMIDT[2], PETER GANTER[2], DAVID BAARTZ[1], MATTHEW SMITH[1], SALLY MORTLOCK[3], ANJALI K. HENDERS[4], AKRAM KHALIL[1] GRANT W. MONTGOMERY[3] BRETT D. MCKINNON[3#], AND AKWASI A AMOAKO[1,2#],

AUTHOR AFFILIATIONS

1. Department of Obstetrics & Gynaecology, Womens' and Newborn Services, Royal Brisbane and Women's Hospital, Brisbane, Australia
2. Discipline of Obstetrics & Gynaecology, Faculty of Medicine, University of Queensland
3. Genomics of Reproductive Disorders, Institute for Molecular Bioscience, University of Queensland
4. Human Studies Unit, Institute for Molecular Bioscience, University of Queensland.

# These authors should be considered joint Senior Authors.

Corresponding Author: Akwasi Amoako: a.amoako@uq.edu.au

Word Count = 5865

Abstract

Introduction

Endometriosis is a common gynaecological disease associated with pelvic pain and subfertility. There are no non-invasive diagnostic tests, medical management requires suppression of estrogens and surgical removal is associated with risk. Endometriosis is a complex genetic disease with variants in at least 27 genetic regions associated with susceptibility. Previous research has implicated a variety of biological mechanisms in multiple cell types. Endometrial and endometriotic epithelial cells acquire somatic mutations at frequency higher than expected in normal tissue. Stromal cells have altered adhesive and capacity. Immune cells show altered cytotoxicity. Understanding the functional consequences of these genetic variations in each cell type requires tissue from patients collected in an integrated program of symptoms and clinical data collection genetic assessment.

Methods and analysis

The aims of this study are to collect tissue associated with endometriosis, chart the genetic architecture related to endometriosis in this tissue, isolate and propagate patient specific cellular models, understand the functional consequence of these genetic variants and understand how they interact with environmental factors in pathogenesis and treatment response.

We will collect patient information from online questionnaires prior to surgery and at 6 and 12 months post-surgery. Treating physicians will document detailed surgical data. During surgery we will collect blood, peritoneal fluid, endometrium and endometriotic tissue. Tissue will be used to isolate and propagate *in vitro* models of individual cells. Genome wide genotyping and gene expression data will be generated. Somatic mutations will be identified via whole genome sequencing.

Ethics and Dissemination

The study has been approved and will be monitored by the Metro North Human Research Ethics committee (HREC) and research activities at the University of Queensland (UQ) will be overseen by the UQ HREC with annual reports submitted. Research results will be published in peer-reviewed journals and the results presented at conferences where appropriate.

Strengths and limitations of the study:

A strength of this study is the breadth and size. It is expected that each activity will create meaningful results within themselves that will be more powerful once combined. Specific strengths include;

- Carefully worded patient questionnaires will ensure patients responses and the tissue collected corresponds to current episodes of endometriosis. Endometriosis is a disease characterized by high rates of recurrence, repeated surgery and multiple lesion removal. To associate endometriosis tissue with clinically relevant outcomes it is imperative tissue samples and symptom reporting are linked.
- Patient data collection is conducted online via email contact and with direct patient entry of data, increasing privacy, reducing time required for data entry and potential for data-entry operator errors.
- Patient samples are collected in line with harmonized World Endometriosis Research Foundation (WERF) protocols, ensuring high quality material for research projects and provide the opportunity to link with other centres and studies both nationally and worldwide.
- *In vitro* models created are underpinned with both genetic and extensive phenotype data that can inform their choice and application for laboratory studies and personalised treatment screening.
- The patient response to treatment will be acquired with a 6 and 12 month follow-up, providing pre and post medical management data.

The main limitation of the study is that it will currently be performed at a single tertiary centre, which may limit samples size and heterogeneity restricting patient cohort to more severe disease.

Introduction

Endometriosis affects approximately 10% of women of reproductive age leading to significant pelvic pain and subfertility[1]. It has been associated with an increased chance of developing ovarian cancer later in life[2]. It is characterized by the growth of endometrial epithelial and stromal cells in lesions outside the uterine cavity currently defined in three categories; superficial peritoneal (SUP), ovarian endometrioma (OMA) and deeply infiltrating endometriosis lesions (DIE), most severe form defined by infiltration of more than 5mm into the underlying tissue. Endometriosis is currently managed either medically via hormonal modulation, or surgically by laparoscopic excision. There are no non-invasive diagnostic tests. Hormonal modulation is inadequate for women wishing to conceive[3] and surgical excision can be difficult and accompanied by a high rate of recurrence, with up to 30% of patients having lesions return within 5 years[4]. Inadequacies of medical management places extraordinary strain on health care systems[5]. The critical need to improve endometriosis management has been recognized by the Australian Federal Government with the launch of a National Action Plan[6].

Originally proposed in 1928 Sampson's theory of retrograde menstruation, in which viable endometrial cells are refluxed back into the peritoneal cavity[7], is still the most commonly accepted theory for the origin of the lesions. However, up to 90% of women experience retrograde menstruation[8] thus additional factors must predispose some to an increased risk. Inherent differences have been observed in both the endometrial epithelial and stromal cells of women with endometriosis that leads to favourable growth profiles of these cells[9]. Endometrial mesenchymal stem cells (eMSC), the precursors cell of the endometrial stromal cells[10] are present in the menstrual efflux[11] and may initiate lesions, as are epithelial progenitor cells[12]. Macrophages[13] and dendritic cells[14] have altered phagocytosis, and natural killer (NK) cells show reduced cytotoxic function[15] in women with endometriosis. Increased levels of neutrophils[16] and cytotoxic T cells[17] in women with endometriosis have also been reported. What underpins the aberrant activity of each cell type and how it contributes to endometriosis progression and how these interact is not yet clear.

Endometriosis is a heritable disease with common genetic variants contributing approximately 51% to endometriosis risk[18]. Genome wide association studies (GWAS) are mapping genetic risk factors for endometriosis[19-21] and a recent study from eleven case-control data sets (17,045 endometriosis cases and 191,596 controls) replicated nine previously reported genomic regions and identified five novel regions[19]. Acquired, rare mutations may also play a role in endometriosis pathogenesis. Using exome sequencing researchers identified somatic changes

in cells from 79% of the 24 patients with DIE lesions[22] with lesions from five patients (21%) harbouring known somatic cancer driver mutations in *ARID1A*, *PIK3CA*, *KRAS*, and *PPP2R1A*. Importantly, even though lesions contain multiple cell types separation of epithelial and stromal cells found the cancer driver mutations of KRAS, ARID1A and PPP2R1A were confined to the epithelial cells. Evidence supports a cell-specific influence of genetic variants[23] on cellular behaviour and these studies underline the importance not only of inherited genetic risk, but also acquired mutations and their potential for differing contributions in individual cell types.

Currently, medical management of endometriosis focusses on reducing circulating estrogen[24]. This is often accompanied by unwanted side effects and in many cases, fails to alleviate primary symptoms, nor ablate lesions. Understanding the role of genetic variants in current hormonal treatments is warranted, as genetic architecture of individuals can be a predictive factor in treatment response. A genetic variant in a progesterone metabolizing enzyme correlates with an increased progesterone breakdown and reduced effectiveness of progesterone-based treatments[25]. While improved targeting of current therapeutics is a short-term goal, novel, targeted non-hormonal treatments are still needed to avoid systemic hormone modulation. A number of biological mechanisms, including the inflammatory response and the endocannabinoid system (ECS) play a role in endometriosis and could potentially be targets for non-hormonal treatments. The ECS affect mechanisms critical to endometriosis establishment and maintenance including cell migration, proliferation, survival and inflammation[26] and is significantly influenced by genetic variants[27-29].

The challenge now is to understand how these genetic variants, both common across the population and acquired rare mutations, influence cellular function and contribute to endometriosis pathogenesis, progression and response to treatment. Studying the consequences of genetic variants on biological and cellular behaviour requires well-characterized patient tissue that can be purified and studied in the laboratory. We have established tissue collection protocols, linked to the patient and surgical phenotyping data at the Royal Brisbane Women Hospital (RBWH) to undertake this study designed to improve diagnosis, understand the pathogenesis and stratify treatment response in endometriosis, through the implementation of the aims and objectives outlined below.

Aims

Identify underlying genetic variants in women with and without endometriosis, determine their relationship to biological variation and functional activity in individual cells, and identify their role in individual cells that lead to disease initiation, progression, symptoms and treatment response.

Objectives

- Collect sufficient, high-quality tissue involved in endometriosis pathogenesis from women both with and without endometriosis for subsequent clinical and laboratory-based studies to understand pathogenesis and cellular response to specific compounds
- Characterize the genetic variants related to endometriosis pathogenesis and treatment response in the variety of cell types from relevant tissue
- Identify gene and protein expression differences resulting from genetic variants in endometriosis related biological material which may relate to endometriosis pathogenesis and inform treatment approaches
- Isolate and propagate pure populations of endometriosis associated cells with known germline and somatic genetic variants
- Identify the consequence of germline and somatic genetic variations in endometriosis related cells from women with and without endometriosis
- Understand how these genetic variants interact with environmental factors to influence the prevalence and progress of the disease and how it may respond to treatment

Expected outcomes

It is expected that this project will establish a large, professionally processed and curated tissue resource with linked clinical and surgical information. These samples and information will greatly contribute to both current and future endometriosis projects. In the context of this project it will allow investigation into the germline and somatic mutations in endometrial tissue. It will provide an understanding of:

1. How both somatic and germline genetic variants influence gene expression and the functional activity of individual cell types associated with endometriosis.
2. How these genetic, biological and functional variations contribute to endometriosis pathogenesis.
3. How these genetic, biological and functional variations lead to differences in symptoms and treatment response.

Methods and analysis

Project design and study sites

The project will be run as a collaboration between the Department of Obstetrics & Gynaecology (RBWH) and the Genomics of Reproductive Disorders laboratory (IMB, UQ). Sample collection will be performed at the Department of Obstetrics & Gynaecology, RBWH. Tissue and data collected will be processed at the IMB, UQ site. Experimental procedures and data analysis will be performed at the UQ IMB laboratory. Tissue collected and processed will be stored in locked University sites accessible only through identifiable keyed entry.

Patient and public involvement

The ERQS was designed based on our previous experience and informal discussion with patients and patient advocacy groups. All women recruited have been provided the opportunity to provide feedback, both informally and through dedicated sections on the patient questionnaires, on their clinical experience, workflow of the study and data collection.

Research Activities

Approach to provision of information to participants and consent

The recruitment will be continuous until the close of the study and under the direction of the project leader. Once patients have been identified as eligible for the study the main purpose and design of the study will be discussed directly with the patient by a research nurse or the treating physician at their original consultation. Once the project has been verbally explained by trained staff, the patient will also be provided with the participant information sheet and consent form that further outlines the details of the study and the procedures involved. If they choose to participate in the study, they will be asked to sign the consent form. They will be provided with a copy of the signed consent form for reference of the details of the study, the requirements of their involvement, as well as the phone number of the Department of Obstetrics & Gynaecology that can be used if further information is sought.

Both the participant information sheet and the consent form will state that participation is voluntary, and consent can be withdrawn at any time and that a decision to not participate will not affect treatment in anyway. The information sheet will outline how their data will be managed to maintain confidentiality and also detail the risks and benefits of participation in the study. If at any stage the patient withdraws consent any biological material and patient information collected will be destroyed. Subject data not included in the prior analysis will not be affected by the withdrawal from the study.

Inclusion criteria

Inclusion criteria will be women already selected for laparoscopic surgery for suspected endometriosis. Women of all age, ethnicity and sociodemographic background will be eligible for inclusion. Inclusion criteria for the additional lifestyle and environmental factors (LERF) questions is any participant that has already consented to the study and completed the baseline questionnaire.

Exclusion criteria

Exclusion criteria will be patients suffering from another or pre-existing inflammatory disease, pregnancy, malignancy or if surgery is performed due to an emergency.

Participant commitment

Engagement for the patient will include time required to complete the initial questionnaire, as well as the potential to contact the patient at least twice (at 6-month intervals) after the completion of surgery to provide details on the subsequent result of the treatment, as well as LERF, or additional details that might be required for this or other ethically approved projects. The participant commitment will include the completion of a series of online questionnaires both pre and post-surgery and the removal of biological material during laparoscopic surgery that is already planned for their treatment. Eligible patients will be requested to complete the LERF questions at 6 or 12 months follow up.

Participant recruitment strategies and timeframes

Initial patient contact will occur during an initial consultation in which the treating physician and patients will become known to each other. After the initial consultation the treating physician, under the supervision of the Chief Investigator will identify women that match the inclusion criteria for the study and that have been indicated and scheduled for laparoscopic surgery at the Department of Obstetrics & Gynecology, RBWH.

Once potential patients have been identified the treating physician will organize subsequent contact via a phone call, or other approved modes of telehealth contact by either the treating physician, or a dedicated research nurse. This may occur as part of either the patients planned telehealth treatment, or as an additional phone call. The treating physician, or dedicated research nurse will explain the project in detail and will ask the patient if they would like additional information.

A follow up phone call between the study doctor and the patient will be arranged to discuss the information that was provided. During this phone call the study doctor, or study nurse can go through the patient information sheet and informed consent forms and allow the patient to ask any questions they may have. If they agree the patient will be asked to sign the PICF by a process considered acceptable to Queensland Health.

Any electronically signed consent forms must be printed out on receipt by the study doctor or research nurse. The study doctor or research nurse is to write a file note on the consent form on the process used to obtain consent, sign and date the form. A copy of the completed consent form will be returned electronically to the patient for their records.

Recruited and consented participants will be sent either a link to an online questionnaire, or, if requested, a hardcopy version that has been based on the World Endometriosis Research Foundation (WERF) questionnaire. The approach used is based on the WERF questionnaire, although minor variations the language of the questions ensure participants answer questions relevant to their current episode of endometriosis. As endometriosis is a recurrent disease, we wish to ensure the data collected is linked to the tissues collected. If the patient reports a previous episode of endometriosis they are given the opportunity to answer the same set of questions again, describing and comparing their experience to their previous episode. The patients are also afforded the opportunity to complete a free form section describing their journey through endometriosis.

The online Questionnaire data will be collected using a "LimeSurvey" interface. This software is freeware but allows a user to install a version behind a secure firewall and does not rely on cloud-based storage. All data will be held independently within the Human Studies Unit (HSU), endometriosis database at the University of Queensland with the local site investigator approving access to site specific staff. Access is through personalised logins, passwords and mutifactor authentification for the research team. Collection of this data is done under informed consent allowing indefinite storage and use with other information collected and/or generated. Withdrawal of a participant may necessitate destruction of the data. This can be done although summary data may already have been used in previous publications. This questionnaire is to be completed in the participants own time and prior to surgery.

Day of Surgery

The women will arrive at the Department of Obstetrics & Gynaecology, RBWH for the scheduled surgical procedure. If a paper copy of the presurgical questionnaire was completed

patients will hand in their completed questionnaire. During surgical preparation and prior to the administration of anaesthesia, peripheral blood will be collected. Remaining biological samples, including eutopic endometrial biopsies, ectopic endometriotic lesions and peritoneal fluid/washing will be collected when clinically indicated during surgery and processed as described. Patients will receive the required standard of care and involvement in this study will not have an influence on the quality of treatment provided.

Collection of relevant patient symptom and phenotypical data

Patients that agree to the study and have provided informed consent will be asked to complete the background questionnaire that will capture data relevant for the study and include information on symptoms related to endometriosis based on the previously tabulated adapted version of the WERF Endometriosis phenome and biobanking harmonization project (EPhect) questionnaire. Modifications have been made to this questionnaire that does not change the content but does clarify the language to ensure the participant answers questions relevant to their current episode of endometriosis. The questionnaire contains the possibility to answer the same set of question, but in relation to a previous episode of endometriosis if the patient is willing. Additional questionnaires 6 and 12 months after treatment will elicit information regarding treatment response.

Collection of surgical and clinical data

Clinical data relevant to endometriosis pathology and symptoms, including but not limited to age, weight, BMI, gravidity and parity, ethnicity, previous use of medication and other gynaecological disorders will be collected through the completion of the WERF EPHect surgical phenotype data collection form that is to be completed by the surgeon at the time of surgery. To assist in entry of surgical and clinical data into the patient database an electronic interface has been built that mirrors the questions of the WERF surgical forms that allows streamlined data entry directly into the database and appended to the patient data.

Baseline Data Collection

Clinical and phenotypic data collection

- Demographic characteristics – Name, date of birth, Medicare number, Healthcare Identifier (HI), residential postcode, gender, ethnicity, language spoken at home, education and employment, marital status and history of smoking and alcohol use.
- Clinical presentation characteristics – Presenting symptoms (cyclical and non-cyclical signs and symptoms: dysmenorrhea, deep dyspareunia, dysuria, dyschezia, painful rectal

bleeding or haematuria, shoulder tip pain, catamenial pneumothorax, cyclical cough/haemoptysis/chest pain, cyclical scar swelling and pain, fatigue), pain history and symptoms, endometriosis history

- Pain-related questionnaires - Visual analogue scale (VAS) scores for dysmenorrhoea, non-menstrual pelvic pain, dyspareunia and dyschezia
- Patient reported outcomes measures (PROMs) characteristics – depression, anxiety, and Health-Related Quality of Life (QoL)
- Menstrual history including age at menarche, menstrual irregularities, last date of menstrual period, frequency, duration of flow, and menstrual cycle length will be recorded
- Fertility history, age at the start of each pregnancy, type of fertility treatment used for each pregnancy, and pregnancy outcome (if applicable), live births, type of delivery, and pregnancy complications will be captured
- Medical and surgical history – A detailed history of medical and surgical treatments, medication history, and primary diagnosis

Surgical Data Collection

The surgical data will be collected using the WERF-EPHect Standard Surgical Form [SSF][30] designed to capture all relevant visual information of the endometriosis lesion phenotype and the surgical treatment including:

- Current menstrual cycle, current hormonal treatment, pain or other medications used, dose and duration, side effects and other types of complementary interventions utilized. History of previous surgery for endometriosis
- Imaging (ultrasound and/or MRI) detection or pre-surgical mapping of endometriosis
- Surgical management – type of surgery performed, duration of the procedure, complications, additional conditions identified at surgery, confirmation of endometriosis
- Detailed visual information on intraoperative findings such as extent, exact location, and color of endometriotic lesions, with a specific focus on endometrioma dimensions and endometriotic nodules. It also includes information on the endometriosis fertility index
- Based on the EPHect SSF data, the phenotype characterization of endometriosis will be carried out into superficial peritoneal endometriosis, cystic ovarian endometriosis, or endometrioma and deep infiltrating endometriosis
- From the EPHect SSF data, rAFS (Revised American Fertility Society) score will be calculated and endometriosis cases will be grouped into Stage I/II or Stage III/IV

- Documentation of histopathology reports from samples excised at time of surgery for histopathological diagnosis and confirmation of endometriosis
- The soft copies of the labelled laparoscopic images will be uploaded to the electronic database

Collection of Biological tissue

Relevant biological samples from peripheral blood, peritoneal fluid/washing and tissue implicated in endometriosis pathogenesis, including endometrium and endometriotic lesions will be collected from women undergoing laparoscopic surgery and endometrial biopsies indicated as part of their designated treatment regime.

Blood samples will be collected from every participant on the day of surgery when intravenous access is sited. The remaining biological tissue, including peritoneal fluid, endometrial biopsies and ectopic lesions will be collected as clinically indicated during surgery. Endometrial biopsies will be performed only when clinically indicated. Ectopic lesions will be removed to surgically treat endometriosis, peritoneal fluid is removed and peritoneal cavity is washed with normal saline as a routine procedure. In the context of the surgery no additional surgical action is performed and therefore there is no additional risk to the patient. The treatment received is not altered by participation in the study. Biological material collected will be stored securely at appropriate storage facilities at the IMB UQ laboratory. In all cases, only approved research personnel will have access to the specimens.

- *Whole blood:* Whole blood to a total of 20ml will be collected in evacuated tubes per patient and separated into plasma, serum and buffy coat. Buffy coats will be used either for immune cell isolation and culture, or DNA isolation for genomic analysis.
- *Endometrial biopsy* Primary samples will be saved in RNAlater for subsequent analysis of somatic mutations, gene or protein expression. Any remaining sample will be flash frozen in liquid nitrogen or stored in cyro-protective media and used for gene or protein expression and the isolation and culture of endometrial and immune cells contained within endometrial biopsies. Any remaining sample(s) will be stored in *tissue tek* for the potential use of frozen sections.
- *Peritoneal fluid / washing***:** On arrival in the laboratory peritoneal fluid / washing will be immediately centrifuged to remove any immune cell content. The immune cell pellet will be kept in cyro-protective media for subsequent culture of immune cells, and the supernatant aliquoted for the analysis of pure peritoneal fluid. The aliquoted fluid will be

stored at -70°C to allow batch analysis of immune, hormonal, endocannabinoids or metabolomic components related to endometriosis pathogenesis and progression.

- *Ectopic lesions:* Endometrial ectopic tissue collected from either the SUP, OMA or DIE lesions will be carefully dissected from any surrounding tissue and both the lesion and surrounding tissue collected in either RNAlater for somatic mutation, gene or protein expression analysis, or either flash frozen or stored in cyro-protective media for subsequent isolation and culture of ectopic endometrial cells, or formalin fixed for production of slides that will be used for haematoxylin and eosin staining, histological confirmation of endometriosis and protein analysis through immunohistology.

Laboratory studies and data generation

- *Cell isolation and culture:* The eutopic, ectopic and immune cells isolated from women undergoing surgery will be used for the isolation and culture of cells to produce *in vitro* models. Standard isolation procedures including size exclusion and fluorescence activated cell sorting (FACS) will be utilised to isolate cells of interest and establish *in vitro* models.
- *Genetic assays* Peripheral blood will be separated into serum and plasma. To determine genetics variants related to endometriosis risk, DNA will be extracted from buffy coats and genotyped on Illumina GSA chips using standard protocols assessed on genomic arrays to determine genome wide genetic variants. These arrays are developed to identify single nucleotide polymorphisms that commonly occur randomly across the population. To increase the number of assayed single nucleotide variants (SNP) imputation using data derived from the 1000 genomes project will be used to infer SNPs in linkage disequilibrium (LD).
- *Somatic mutational analysis:* Selected tissue, or isolated cells will be used for somatic mutational screening via either targeted sequencing assays, whole exome or whole genome sequencing.
- *Induced genetic mutations for in vitro analysis of somatic cells.* To assess the influence of individual, or multiple SNPs associated with endometriosis in the context of consistent genetic backgrounds in isolated endometrial cells specific alleles can be induced in cells via CRISPR-cas9 technology. We will introduce specific allelic variants in the *in vitro* cell models generated from isolated endometrial cells. We will determine their influence on gene expression and cell function using methods described below. These somatic cells will be used for *in vitro* experiments only.

- *Gene expression:* RNA will be isolated using standard procedures and either targeted gene expression via Real-time PCR, multiplexed targeting via Nanostring technology, or genome wide gene expression via RNA-sequencing or single cell RNA-sequencing, will be performed to determine gene and transcript expression levels.
- *Protein expression:* Cellular protein concentrations of the isolated tissue will be performed by both quantitative and semi-quantitative methods including Western blot analysis and immunohistochemistry. Secreted proteins will be measured via enzyme linked immune-absorbent assay using the standard manufacturers protocol.
- *Metabolomic profiling:* A non-targeted metabolomic profiling will be performed with Ultra high-performance liquid chromatography high resolution mass spectrometry (UHPLC-HRMS).
- *Functional Assay:* The influence of observed biological variations on the cellular functional behaviour either as individual or as dual cell cultures will be determined by numerous cellular assays that will include, but not be limited to viability, proliferation, apoptosis, migration and invasion.

Participant follow-up

Post-surgical treatment patients will be followed up to determine their response to treatment, including the reduction in symptoms as well as the presence of side effects. Follow-up questionnaires are to be provided to participants at 6-and 12-months post-surgery. These questionnaire capture details of the individual's response to treatment and will be either emailed via a personalized link to the questionnaire or mailed to individual participants by personnel authorized to access the address of the patients.

Collection of Lifestyle and Environmental Factors (LERF)

Participants who had surgery and completed the baseline questionnaire will be invited to answer LERF questions at the 6 or 12-month follow up. An initial assessment will be conducted with 500 participants invited to complete LERF questions.

The LERF questions collects information in relation to:

- Demographics
- Family
- Lifetime occupations (formatted to match Australian census data)
- General health (including mental health)

- Exposures to potential environmental toxins.

There are some questions regarding use of prescription and illicit drug as well as medical history information including questions pertaining to sexually transmitted diseases (STDs). For female participants there are specific questions relating to menstruation and pregnancy history. This will be administered to all consented individuals and will take approximately 40 minutes to complete. Participants are informed that they are not obliged to answer any questions that make them feel uncomfortable or they can skip questions they do not wish to answer.

Participant withdrawal:

Patients may withdraw their consent at any time during the project by contacting either the Department of Obstetrics and Gynaecology, or their treating physician whose details are provided on the patient information sheet. Any biological sample already collected, as well as health related data will be removed from the project and the laboratory. Subject data included in the prior analysis will not be affected by the withdrawal from the study.

Data management

Each patient identified for inclusion in the study will be assigned a unique study identification number (Study ID). This study number will be assigned upon collection of the signed consent at the Department of Obstetrics & Gynaecology, RBWH and matched with a hospital derived patient identification number (PID). This file will represent the only link between the study number and PID number and will remain accessible only to authorized members of the RBWH clinical team in a secure database with an audit trail function and accessible only via a username and password. Clinical and surgical data collected through the completion of the surgeons reporting questionnaire and the patient questionnaire respectively will be entered into a secure database by authorized staff.

Patient and surgical data, lifestyle and environmental data collected will be stored in locked, linked databases housed behind University/Hospital firewalls at IMB/UQ. Experimental procedures and data analysis will be performed at the UQ, IMB Laboratory. The biological material and the health-related data required for analysis will be provided to the UQ laboratory attached to the study ID only and all subsequent experimental data generated during the project will be stored attached to the study ID on secure servers behind UQ University firewall. This data will be considered anonymous to laboratory staff. Only authorized staff for the project will have access to the data. Transfer of any data will be via secure transfer protocols. Data

storage will be for the duration of the project, as well as and for subsequent projects based on information or material derived from this project and retained for a minimum of 5 years.

Electronic database

A dedicated database has been developed for the project using the open-source Drupal software. The data will be housed on a University server behind the University firewall. Treating physicians or research nurses who consent the patients will enter the patient data directly into the electronic database that will generate an ID. Subsequent patient, surgical and biological collections will linked to the study ID. Surgical data collected from the WERF EPHect surgical form will be assessed for consistency by key clinical surgical project leaders and the data entered directly into the Drupal database via a built in electronic interface via the relevant physician, study nurse or project leader.

The Study ID will also be used to generate personalized links to the patient's questionnaire, both before and after surgery, that will allow the patient answers to be directly entered into the Lime Survey database. The Drupal database and Lime Survey interface will function as a central repository that will allow real time data entry and reduce input error. Access to data will only be for authorized personnel.

Data Analysis plan

Data analysis will be performed using a combination of SPSS, GraphPad Prism, plink and R statistical and genomic packages. Analysis of patient and surgical data will be used to identify pre-surgical, risk factors that may facilitate an early diagnosis of endometriosis. Patient data collection matched to the current episode of endometriosis will be used in linear regression models to establish correlation between patient symptoms and lesion characteristics observed during surgery and facilitate subgroup discovery. This will include statistical comparisons between patient reported symptoms prior to surgery, as well as phenotypic descriptions of lesion observed during surgery. Follow up questionnaires will be used to assess patient response to their individualized treatments and whether linked pre-surgical data and surgical observation can subcategorize patients into groups that predict treatment response and disease progression.

Biological samples of peripheral blood, peritoneal fluid and eutopic and ectopic endometrium will be explored to identify biological mechanisms that underlie molecular signatures, including gene and transcriptomic expression, alternative splicing and common and acquired genetic mutations, proteomic and metabolic profiles. Molecular signatures generated with

collected biological samples will be matched to pre-surgery patient data, surgical data, follow up data and LERF data to discover correlations between molecular signature and clinically relevant patient outcomes.

In vitro models will be used to assess the functional influence of genotypes on individual cellular function, including gene and transcriptome expression, alternative splicing, common and rare genetic mutations, proteomic and metabolomic profiles. The interaction between different cell types and their influence on clinically relevant molecular signatures. These molecular signatures will also be identified, or induced, in individual *in vitro* models and utilized to determine potential personalization of current treatments to individual molecular profiles or identify novel treatments.

Additional genotyping data and functional genomics datasets aimed at identifying the causal genetic variants and biological mechanism that are associated with endometriosis susceptibility will be generated to potentially contribute to consortium meta-analysis in collaboration with other groups worldwide.

Power Calculations

It is assumed comparisons will be performed predominantly between two groups; women with and without endometriosis, and either positive or negative response to treatment. Analysis between two independent means will aim to achieve a moderate power of 80% to target small effect sizes of 0.2. It is estimated to achieve a significant difference a sample size of 394 sample per group, or 786 total would be required. It is intended for recruitment procedures to continue for three years to establish a database of sufficient samples that will satisfy power requirements. Over a 3-year period we plan to recruit approximately 300 patients per year to reach up to 900 patients by the end of the period. Based on first year response rates we predict an incomplete protocol from approximately 15% of patients each year, leaving approximately 750 – 800 patients over the 3-year collection period, providing the required sample numbers.

Research Team Training and Quality Controls

Research staff are all professionally trained in their relevant positions. They have also received study-specific training. This included training for treating physicians and study nurse on the appropriate recruitment of study participants, how to obtain informed consent, counseling of the study participant, data collection, interacting with the database interface and the follow up of patients with incomplete study protocols. Specific training has been provided on correct data entry, data management, data analysis and the maintenance of data sets.

Specific training has been provided for the collection and processing of biological samples including blood and its separation into constituents, aliquoting and recording. Standard operating procedures (SOP) based on best practice and the WERF biological processing protocols [31 32], including the storage of endometrial eutopic and ectopic lesions for either fresh frozen or cryopreserved process, as well as the separation of individual cells for establishing *in vitro* models have been developed.

Ethics and dissemination:

The study has been approved and will continue to be monitored by the Metro North Human Research Ethics committee (HREC) and has received site specific approval for the conduct of the study at the RBWH. Research activities at the University of Queensland with the data and samples generated also overseen by the UQ HREC. Annual reports will be submitted to both the Metro North and UQ HREC for regular updates on the study. Research results will be published in peer-reviewed journals and the results presented at conferences where appropriate. If requested by the journals, de-identified results may be submitted to recognized data repositories with controlled access policies to provide open access for research purposes.

Project closure procedure:

The total duration of the project is 4 years. Three years of tissue collection followed by an additional year to complete tissue and data analysis. There will be the potential to extend the project if deemed beneficial and ethics requirements are met. The project will include the production of *in vitro* models potentially useful to future projects and other researchers and samples will not be destroyed at the completion of the project. These models may be made conditionally available, subject to the receipt of appropriate ethical approval. Third party researchers will only have access to non-identifiable data and will not be able to match these cell models with patients.

Study status

The present study was conceptualized and designed in 2020. The first participant was enrolled in September 2020 and the study is currently ongoing with the number of collected sample reaching the desired number expected after the first year of recruitment. Baseline characteristics show positive recruitment and protocol completion rates. Characteristics of the of endometriosis patients, patients identified to have another non-endometrial pathology and no discernible endometrial pathology collected in the first year, are shown in **Table 1**.

Discussion

This protocol outlines the framework for a single site study investigating endometriosis and the development of the infrastructure that supports the collection of baseline data that will underpin the research. The infrastructure developed has the potential to be applied to additional sites if the study protocol was to be expanded. The study protocol has been designed to generate baseline data for each patient including data on patients reported symptoms, menstrual and fertility history and their previous treatment experience. It will also include surgical observations recorded during treatment and collected by physicians. The research infrastructure will ensure high quality baseline data is both generated, entered and stored in purpose-built database that increases ease of entry and security of the data and decreases user input error.

Collection of baseline patient and surgical data will be built upon with genotyping data from DNA isolated from peripheral blood that will allow functional investigation of genetic risk variants associated with endometriosis susceptibility. The collection of endometriosis relevant tissue will allow the generation of omics databases to investigate the functional influence of the genetic variants. It will also allow the creation of *in vitro* models of individual cell types with the accompanying genomics background.

This study represents a significant opportunity to collect deeply phenotyped endometriosis samples to generate cellular models and detailed integrated molecular datasets. It will be the foundation of endometriosis biorepositories of multiple tissue types including serum, plasma DNA, RNA, as well as the establishment of personalized *in vitro* models that can be screened for responses to treatments and compounds. The results from the analysis will improve our understanding of endometriosis pathogenesis, the clinical symptoms associated with subtypes and how these respond to various treatments. It will provide valuable datasets to contribute to global networks and International Consortia studying endometriosis and improve outcomes for endometriosis patients.

Funding;

G.W.M is supported by the National Health and Medical Research Council (NHMRC) fellowship (GNT1078399 and GNT1177194). G.W.M, A.A. K.T and S.M were supported by the Metro North Collaborative Research Grant (CRG-176-2019). K.T. is supported by RBWH and RBWH Foundation Post Graduate Scholarship. B.M. is supported by Endometriosis Australia (1427).

Author contributions

B.M, A.A and K.T conceived the study and wrote the protocol with assistance from GWM. AA and KT consented patients. AA, KT, BS, PG, AK and DB, MS and DG collected patient data. BM, SS, SA, KT, LMW collected and processed patient samples. MBL and PL build and maintained database and AKH oversaw database and sample processing.

Ethics approval

Ethics approval has been granted by the Metro North Human ethics research council Ref No: HREC/2019/QRBW/56763 and the University of Queensland Human research ethics committee 2019/HE002744.

References.

1. Vercellini P, Viganò P, Somigliana E, et al. Endometriosis: pathogenesis and treatment. *Nature Reviews Endocrinology* 2014;10(5):261-75. doi: 10.1038/nrendo.2013.255
2. Grandi G, Toss A, Cortesi L, et al. The Association between Endometriomas and Ovarian Cancer: Preventive Effect of Inhibiting Ovulation and Menstruation during Reproductive Life. *BioMed Research International* 2015;2015:751571. doi: 10.1155/2015/751571
3. Dunselman GAJ, Vermeulen N, Becker C, et al. ESHRE guideline: management of women with endometriosis †. *Human Reproduction* 2014;29(3):400-12. doi: 10.1093/humrep/det457
4. Guo S-W. Recurrence of endometriosis and its control. *Human Reproduction Update* 2009;15(4):441-61. doi: 10.1093/humupd/dmp007
5. Simoens S, Dunselman G, Dirksen C, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum Reprod* 2012;27(5):1292-9. doi: 10.1093/humrep/des073 [published Online First: 2012/03/17]
6. Australian Government DoH. National Action Plan for Endometriosis. 2018
7. J.A. S. Peritoneal Endometriosis Due to the Menstrual Dissemination of Endometrial Tissue into the Peritoneal Cavity. *American Journal of Obstetrics & Gynecology* 1927;14:442-69.
8. Halme J, Hammond MG, Hulka JF, et al. Retrograde menstruation in healthy women and in patients with endometriosis. *Obstet Gynecol* 1984;64(2):151-4. [published Online First: 1984/08/01]
9. Burney RO, Giudice LC. Pathogenesis and pathophysiology of endometriosis. *Fertil Steril* 2012;98(3):511-9. doi: 10.1016/j.fertnstert.2012.06.029 [published Online First: 2012/07/24]
10. Barragan F, Irwin JC, Balayan S, et al. Human Endometrial Fibroblasts Derived from Mesenchymal Progenitors Inherit Progesterone Resistance and Acquire an Inflammatory Phenotype in the Endometrial Niche in Endometriosis. *Biol Reprod* 2016;94(5):118. doi: 10.1095/biolreprod.115.136010 [published Online First: 2016/04/15]
11. Gargett CE, Schwab KE, Deane JA. Endometrial stem/progenitor cells: the first 10 years. *Hum Reprod Update* 2016;22(2):137-63. doi: 10.1093/humupd/dmv051 [published Online First: 2015/11/11]

12. Nguyen HPT, Xiao L, Deane JA, et al. N-cadherin identifies human endometrial epithelial progenitor cells by in vitro stem cell assays. *Human Reproduction* 2017;32(11):2254-68. doi: 10.1093/humrep/dex289
13. Chuang PC, Wu MH, Shoji Y, et al. Downregulation of CD36 results in reduced phagocytic ability of peritoneal macrophages of women with endometriosis. *J Pathol* 2009;219(2):232-41. doi: 10.1002/path.2588 [published Online First: 2009/07/17]
14. Izumi G, Koga K, Takamura M, et al. Mannose receptor is highly expressed by peritoneal dendritic cells in endometriosis. *Fertil Steril* 2017;107(1):167-73.e2. doi: 10.1016/j.fertnstert.2016.09.036 [published Online First: 2016/10/30]
15. Thiruchelvam U, Wingfield M, O'Farrelly C. Natural Killer Cells: Key Players in Endometriosis. *Am J Reprod Immunol* 2015;74(4):291-301. doi: 10.1111/aji.12408 [published Online First: 2015/06/25]
16. Tariverdian N, Siedentopf F, Rücke M, et al. Intraperitoneal immune cell status in infertile women with and without endometriosis. *J Reprod Immunol* 2009;80(1-2):80-90. doi: 10.1016/j.jri.2008.12.005 [published Online First: 2009/04/21]
17. Furuya K, Murakami M, Makimura N, et al. Immunological and endocrinological studies on lymphocyte subpopulation and medical treatment for infertility in patients with endometriosis. *Mol Cell Endocrinol* 2003;202(1-2):195-9. doi: 10.1016/s0303-7207(03)00083-2 [published Online First: 2003/05/29]
18. Treloar SA, O'Connor DT, O'Connor VM, et al. Genetic influences on endometriosis in an Australian twin sample. sueT@qimr.edu.au. *Fertil Steril* 1999;71(4):701-10. doi: 10.1016/s0015-0282(98)00540-8 [published Online First: 1999/04/15]
19. Sapkota Y, Steinthorsdottir V, Morris AP, et al. Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nature Communications* 2017;8(1):15539. doi: 10.1038/ncomms15539
20. Fung JN, Rogers PA, Montgomery GW. Identifying the biological basis of GWAS hits for endometriosis. *Biol Reprod* 2015;92(4):87. doi: 10.1095/biolreprod.114.126458 [published Online First: 2015/02/20]
21. Nilufer R, Karina B, Paraskevi C, et al. Large-scale genome-wide association meta-analysis of endometriosis reveals 13 novel loci and genetically-associated comorbidity with other pain conditions. *bioRxiv* 2018:406967. doi: 10.1101/406967
22. Anglesio MS, Papadopoulos N, Ayhan A, et al. Cancer-Associated Mutations in Endometriosis without Cancer. *N Engl J Med* 2017;376(19):1835-48. doi: 10.1056/NEJMoa1614814 [published Online First: 2017/05/11]
23. Suda K, Nakaoka H, Yoshihara K, et al. Clonal Expansion and Diversification of Cancer-Associated Mutations in Endometriosis and Normal Endometrium. *Cell Rep* 2018;24(7):1777-89. doi: 10.1016/j.celrep.2018.07.037 [published Online First: 2018/08/16]
24. Barra F, Grandi G, Tantari M, et al. A comprehensive review of hormonal and biological therapies for endometriosis: latest developments. *Expert Opin Biol Ther* 2019;19(4):343-60. doi: 10.1080/14712598.2019.1581761 [published Online First: 2019/02/15]
25. Lazorwitz A, Aquilante CL, Oreschak K, et al. Influence of Genetic Variants on Steady-State Etonogestrel Concentrations Among Contraceptive Implant Users. *Obstetrics and gynecology* 2019;133(4):783-94. doi: 10.1097/AOG.0000000000003189
26. Di Marzo V, Izzo AA. Endocannabinoid overactivity and intestinal inflammation. *Gut* 2006;55(10):1373-76. doi: 10.1136/gut.2005.090472

27. Smith DR, Stanley CM, Foss T, et al. Rare genetic variants in the endocannabinoid system genes CNR1 and DAGLA are associated with neurological phenotypes in humans. *PLoS One* 2017;12(11):e0187926. doi: 10.1371/journal.pone.0187926 [published Online First: 2017/11/18]
28. Meyer HC, Lee FS, Gee DG. The Role of the Endocannabinoid System and Genetic Variation in Adolescent Brain Development. *Neuropsychopharmacology* 2018;43(1):21-33. doi: 10.1038/npp.2017.143
29. Chidambaran V, Pilipenko V, Spruance K, et al. Fatty acid amide hydrolase-morphine interaction influences ventilatory response to hypercapnia and postoperative opioid outcomes in children. *Pharmacogenomics* 2017;18(2):143-56. doi: 10.2217/pgs-2016-0147 [published Online First: 2016/12/16]
30. Becker CM, Laufer MR, Stratton P, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: I. Surgical phenotype data collection in endometriosis research. *Fertil Steril* 2014;102(5):1213-22. doi: 10.1016/j.fertnstert.2014.07.709 [published Online First: 2014/08/26]
31. Rahmioglu N, Fassbender A, Vitonis AF, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: III. Fluid biospecimen collection, processing, and storage in endometriosis research. *Fertil Steril* 2014;102(5):1233-43. doi: 10.1016/j.fertnstert.2014.07.1208 [published Online First: 2014/09/27]
32. Fassbender A, Rahmioglu N, Vitonis AF, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: IV. Tissue collection, processing, and storage in endometriosis research. *Fertil Steril* 2014;102(5):1244-53. doi: 10.1016/j.fertnstert.2014.07.1209 [published Online First: 2014/09/27]

Table 1: Patient characteristics collected in the first year of the ERQS

Characteristics		Endometriosis N (%)	Other endometrial pathology N (%)	No pathology N (%)	P value
Demographic		N = 168 (52.97%)	N = 67 (20.23%)	N = 88 (26.79%)	
BMI (kg/m^2), Mean ± SD (*Ordinary one-way ANOVA*)		27.07 ± 8.95 (n = 151)	31.19 ± 6.95 (n = 62)	27.33 ± 6.41 (n = 80)	*P = 0.002
BMI (kg/m^2) categories (*Chi Squared*)	Underweight (≤18.5)	5 (3.3%)	1 (1.6%)	2 (2.5%)	*P = 0.003
	Normal (18.5 – 24.9)	64 (42.4%)	12 (19.4%)	34 (42.5%)	
	Overweight (25 – 29.9)	47 (31.1%)	17 (27.4%)	19 (23.8%)	
	Obese (≥30)	35 (23.2%)	32 (51.6%)	25 (31.3%)	
	Unknown/No response	17	5	8	
Education level	Year 10, or equivalent	11 (7.5%)	5 (8.8%)	4 (6.2%)	P = 0.729
	Year 12	17 (11.6%)	6 (10.5%)	11 (13.6%)	
	Diploma	51 (34.9%)	22 (38.6%)	30 (37.0%)	
	TAFE	12 (8.2%)	3 (5.3%)	2 (2.5%)	
	University undergraduate	34 (23.3%)	12 (21.1%)	25 (30.9%)	
	University Postgraduate	21 (14.4%)	9 (15.8%)	8 (9.9%)	
	No response	22	10	7	
Employment	Government/ Private job	112 (77.2%)	46 (78.0%)	53 (65.4%)	P = 0.235
	Self employed	10 (6.9%)	2 (3.4%)	5 (6.2%)	
	Unemployed	9 (6.2%)	8 (13.6%)	13 (16.0%)	
	Student	8 (5.5%)	1 (1.7%)	7 (8.6%)	
	Unable to work due to chronic pain	6 (4.1%)	2 (3.4%)	3 (3.7%)	
	Unknown	23	8	7	
Marital status	Married	100 (67.6%)	43 (68.3%)	52 (64.2%)	P = 0.868
	Divorced/Separated	8 (5.4%)	4 (6.3%)	2 (2.5%)	
	In a relationship, not living together	12 (8.1%)	4 (6.3%)	9 (11.1%)	
	Single	27 (18.2%)	12 (19.0%)	18 (22.2%)	
	Widowed	1 (0.7%)	0 (0%)	0 (0%)	
	Unknown/No response	20	4	7	

BMJ Open

Protocol for the Endometriosis Research Queensland Study (ERQS): An integrated cohort study approach to improve diagnosis and stratify treatment.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-064073.R1
Article Type:	Protocol
Date Submitted by the Author:	17-Aug-2022
Complete List of Authors:	Tanaka, Keisuke; Royal Brisbane and Woman's Hospital Health Service District, Department of Obstetrics & Gynaecology, Womens' and Newborn Services; The University of Queensland Gilroy, Deborah; The University of Queensland Subramaniam, Sugarniya; The University of Queensland Lakshmi, Preethi; The University of Queensland Bhadravathi Lokeshappa, Madhura; The University of Queensland Wallace, Leanne; The University of Queensland Atluri, Sharat; The University of Queensland Schmidt, Bart; The University of Queensland Ganter, Peter; The University of Queensland Baartz, David; Royal Brisbane and Woman's Hospital Health Service District, Department of Obstetrics & Gynaecology, Womens' and Newborn Services Smith, Matthew; Royal Brisbane and Woman's Hospital Health Service District, Department of Obstetrics & Gynaecology, Womens' and Newborn Services Mortlock, Sally; Royal Brisbane and Women's Hospital Henders, Anjali k.; Royal Brisbane and Woman's Hospital Health Service District Khalil, Akram; Royal Brisbane and Woman's Hospital Health Service District, Department of Obstetrics & Gynaecology, Womens' and Newborn Services Montgomery, Grant ; The University of Queensland Institute for Molecular Bioscience McKinnon, Brett; The University of Queensland - Saint Lucia Campus, Institute for Molecular Bioscience Amoako, Akwasi ; Royal Brisbane and Women's Hospital, Department of Obstetrics & Gynaecology, Womens' and Newborn Services; The University of Queensland
Primary Subject Heading:	Reproductive medicine
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	GYNAECOLOGY, REPRODUCTIVE MEDICINE, GENETICS

Protocol for the Endometriosis Research Queensland Study (ERQS): An integrated cohort study approach to improve diagnosis and stratify treatment.

KEISUKE TANAKA[1,2], DEBORAH GILROY[4], SUGARNIYA, SUBRAMANIAM[3], PREETHI LAKSHMI[4], MADHURA BHADRAVATHI LOKESHAPPA[4], LEANNE M. WALLACE[4], SHARAT ATLURI, BART SCHMIDT[2], PETER GANTER[2], DAVID BAARTZ[1], MATTHEW SMITH[1], SALLY MORTLOCK[3], ANJALI K. HENDERS[4], AKRAM KHALIL[1] GRANT W. MONTGOMERY[3] BRETT D. MCKINNON[3#], AND AKWASI A AMOAKO[1,2#],

AUTHOR AFFILIATIONS

1. Department of Obstetrics & Gynaecology, Womens' and Newborn Services, Royal Brisbane and Women's Hospital, Brisbane, Australia
2. Discipline of Obstetrics & Gynaecology, Faculty of Medicine, University of Queensland
3. Genomics of Reproductive Disorders, Institute for Molecular Bioscience, University of Queensland
4. Human Studies Unit, Institute for Molecular Bioscience, University of Queensland.

# These authors should be considered joint Senior Authors.

Corresponding Author: Akwasi Amoako: a.amoako@uq.edu.au

Word Count = 5865

Abstract

Introduction

Endometriosis is a common gynaecological disease associated with pelvic pain and subfertility. There are no non-invasive diagnostic tests, medical management requires suppression of estrogens and surgical removal is associated with risk. Endometriosis is a complex genetic disease with variants in at least 27 genetic regions associated with susceptibility. Previous research has implicated a variety of biological mechanisms in multiple cell types. Endometrial and endometriotic epithelial cells acquire somatic mutations at frequency higher than expected in normal tissue. Stromal cells have altered adhesive and capacity. Immune cells show altered cytotoxicity. Understanding the functional consequences of these genetic variations in each cell type requires tissue from patients collected in an integrated program of symptoms and clinical data collection genetic assessment.

Methods and analysis

The aims of this study are to collect tissue associated with endometriosis, chart the genetic architecture related to endometriosis in this tissue, isolate and propagate patient specific cellular models, understand the functional consequence of these genetic variants and understand how they interact with environmental factors in pathogenesis and treatment response.

We will collect patient information from online questionnaires prior to surgery and at 6 and 12 months post-surgery. Treating physicians will document detailed surgical data. During surgery we will collect blood, peritoneal fluid, endometrium and endometriotic tissue. Tissue will be used to isolate and propagate *in vitro* models of individual cells. Genome wide genotyping and gene expression data will be generated. Somatic mutations will be identified via whole genome sequencing.

Ethics and Dissemination

The study has been approved and will be monitored by the Metro North Human Research Ethics committee (HREC) and research activities at the University of Queensland (UQ) will be overseen by the UQ HREC with annual reports submitted. Research results will be published in peer-reviewed journals and the results presented at conferences where appropriate.

Strengths and limitations of the study:

Strengths of the study include;

- Carefully worded patient questionnaires that ensures patients' responses and the tissue collected corresponds to specific episodes of endometriosis.
- Patient data collection is conducted online via email contact and with direct patient entry of data and will be collected both pre and post surgery
- Patient samples are collected in line with harmonized World Endometriosis Research Foundation (WERF) protocols.
- *In vitro* models are underpinned with both genetic and phenotype data.

The main limitation of the study is;

- It is currently performed at a single tertiary centre, limiting sample size and heterogeneity.

Introduction

Endometriosis affects approximately 10% of women of reproductive age leading to significant pelvic pain and subfertility[1]. It has been associated with an increased chance of developing ovarian cancer later in life[2]. It is characterized by the growth of endometrial epithelial and stromal cells in lesions outside the uterine cavity currently defined in three categories; superficial peritoneal (SUP), ovarian endometrioma (OMA) and deeply infiltrating endometriosis lesions (DIE), most severe form defined by infiltration of more than 5mm into the underlying tissue. Endometriosis is currently managed either medically via hormonal modulation, or surgically by laparoscopic excision. There are no non-invasive diagnostic tests. Hormonal modulation is inadequate for women wishing to conceive[3] and surgical excision can be difficult and accompanied by a high rate of recurrence, with up to 30% of patients having lesions return within 5 years[4]. Inadequacies of medical management places extraordinary strain on health care systems[5]. The critical need to improve endometriosis management has been recognized by the Australian Federal Government with the launch of a National Action Plan[6].

Originally proposed in 1928 Sampson's theory of retrograde menstruation, in which viable endometrial cells are refluxed back into the peritoneal cavity[7], is still the most commonly accepted theory for the origin of the lesions. However, up to 90% of women experience retrograde menstruation[8] thus additional factors must predispose some to an increased risk. Inherent differences have been observed in both the endometrial epithelial and stromal cells of women with endometriosis that leads to favourable growth profiles of these cells[9]. Endometrial mesenchymal stem cells (eMSC), the precursors cell of the endometrial stromal cells[10] are present in the menstrual efflux[11] and may initiate lesions, as are epithelial progenitor cells[12]. Macrophages[13] and dendritic cells[14] have altered phagocytosis, and natural killer (NK) cells show reduced cytotoxic function[15] in women with endometriosis. Increased levels of neutrophils[16] and cytotoxic T cells[17] in women with endometriosis have also been reported. What underpins the aberrant activity of each cell type and how it contributes to endometriosis progression and how these interact is not yet clear.

Endometriosis is a heritable disease with common genetic variants contributing approximately 51% to endometriosis risk[18]. Genome wide association studies (GWAS) are mapping genetic risk factors for endometriosis[19-21] and a recent study from eleven case-control data sets (17,045 endometriosis cases and 191,596 controls) replicated nine previously reported genomic regions and identified five novel regions[19]. Acquired, rare mutations may also play a role in endometriosis pathogenesis. Using exome sequencing researchers identified somatic changes

in cells from 79% of the 24 patients with DIE lesions[22] with lesions from five patients (21%) harbouring known somatic cancer driver mutations in *ARID1A*, *PIK3CA*, *KRAS*, and *PPP2R1A*. Importantly, even though lesions contain multiple cell types separation of epithelial and stromal cells found the cancer driver mutations of KRAS, ARID1A and PPP2R1A were confined to the epithelial cells. Evidence supports a cell-specific influence of genetic variants[23] on cellular behaviour and these studies underline the importance not only of inherited genetic risk, but also acquired mutations and their potential for differing contributions in individual cell types.

Currently, medical management of endometriosis focusses on reducing circulating estrogen[24]. This is often accompanied by unwanted side effects and in many cases, fails to alleviate primary symptoms, nor ablate lesions. Understanding the role of genetic variants in current hormonal treatments is warranted, as genetic architecture of individuals can be a predictive factor in treatment response. A genetic variant in a progesterone metabolizing enzyme correlates with an increased progesterone breakdown and reduced effectiveness of progesterone-based treatments[25]. While improved targeting of current therapeutics is a short-term goal, novel, targeted non-hormonal treatments are still needed to avoid systemic hormone modulation. A number of biological mechanisms, including the inflammatory response and the endocannabinoid system (ECS) play a role in endometriosis and could potentially be targets for non-hormonal treatments. The ECS affect mechanisms critical to endometriosis establishment and maintenance including cell migration, proliferation, survival and inflammation[26] and is significantly influenced by genetic variants[27-29].

The challenge now is to understand how these genetic variants, both common across the population and acquired rare mutations, influence cellular function and contribute to endometriosis pathogenesis, progression and response to treatment. Studying the consequences of genetic variants on biological and cellular behaviour requires well-characterized patient tissue that can be purified and studied in the laboratory. We have established tissue collection protocols, linked to the patient and surgical phenotyping data at the Royal Brisbane Women Hospital (RBWH) to undertake this study designed to improve diagnosis, understand the pathogenesis and stratify treatment response in endometriosis, through the implementation of the aims and objectives outlined below.

Aims

Identify underlying genetic variants in women with and without endometriosis, determine their relationship to biological variation and functional activity in individual cells, and identify their role in individual cells that lead to disease initiation, progression, symptoms and treatment response.

Objectives

- Collect sufficient, high-quality tissue involved in endometriosis pathogenesis from women both with and without endometriosis for subsequent clinical and laboratory-based studies to understand pathogenesis and cellular response to specific compounds
- Characterize the genetic variants related to endometriosis pathogenesis and treatment response in the variety of cell types from relevant tissue
- Identify gene and protein expression differences resulting from genetic variants in endometriosis related biological material which may relate to endometriosis pathogenesis and inform treatment approaches
- Isolate and propagate pure populations of endometriosis associated cells with known germline and somatic genetic variants
- Identify the consequence of germline and somatic genetic variations in endometriosis related cells from women with and without endometriosis
- Understand how these genetic variants interact with environmental factors to influence the prevalence and progress of the disease and how it may respond to treatment

Expected outcomes

It is expected that this project will establish a large, professionally processed and curated tissue resource with linked clinical and surgical information. These samples and information will greatly contribute to both current and future endometriosis projects. In the context of this project it will allow investigation into the germline and somatic mutations in endometrial tissue. It will provide an understanding of:

1. How both somatic and germline genetic variants influence gene expression and the functional activity of individual cell types associated with endometriosis.
2. How these genetic, biological and functional variations contribute to endometriosis pathogenesis.
3. How these genetic, biological and functional variations lead to differences in symptoms and treatment response.

Methods and analysis

Project design and study sites

The project will commence on September 2019 and is initially planned to run for 5 years with 4 years of sample and data collection finishing in December 2023 with an additional year of data analysis to be completed by December 2024. A decision on whether to extend the trial will be made if it is considered worthwhile and if can be secured. The trial will be run as a collaboration between the Department of Obstetrics & Gynaecology (RBWH) and the Genomics of Reproductive Disorders laboratory (IMB, UQ). Sample collection will be performed at the Department of Obstetrics & Gynaecology, RBWH. Tissue and data collected will be processed at the IMB, UQ site. Experimental procedures and data analysis will be performed at the UQ IMB laboratory. Tissue collected and processed will be stored in locked University sites accessible only through identifiable keyed entry.

Patient and public involvement

The ERQS was designed based on our previous experience and informal discussion with patients and patient advocacy groups. All women recruited have been provided the opportunity to provide feedback, both informally and through dedicated sections on the patient questionnaires, on their clinical experience, workflow of the study and data collection.

Research Activities

Approach to provision of information to participants and consent

The recruitment will be continuous until the close of the study and under the direction of the project leader. Once patients have been identified as eligible for the study the main purpose and design of the study will be discussed directly with the patient by a research nurse or the treating physician at their original consultation. Once the project has been verbally explained by trained staff, the patient will also be provided with the participant information sheet and consent form that further outlines the details of the study and the procedures involved. If they choose to participate in the study, they will be asked to sign the consent form. They will be provided with a copy of the signed consent form for reference of the details of the study, the requirements of their involvement, as well as the phone number of the Department of Obstetrics & Gynaecology that can be used if further information is sought.

Both the participant information sheet and the consent form will state that participation is voluntary, and consent can be withdrawn at any time and that a decision to not participate will not affect treatment in anyway. The information sheet will outline how their data will be managed to maintain confidentiality and also detail the risks and benefits of participation in

the study. If at any stage the patient withdraws consent any biological material and patient information collected will be destroyed. Subject data not included in the prior analysis will not be affected by the withdrawal from the study.

Inclusion criteria

Inclusion criteria will be women already selected for laparoscopic surgery for suspected endometriosis. Women of all age, ethnicity and sociodemographic background will be eligible for inclusion. Inclusion criteria for the additional lifestyle and environmental factors (LERF) questions is any participant that has already consented to the study and completed the baseline questionnaire.

Patients will be considered either endometriosis positive, or endometriosis negative based on either the surgical findings or histopathological examination of excised tissue. Both results will be collected and entered into the database. Subsequent analysis based on endometriosis status will be defined by the values in these fields and will be stated as such in the resulting data and outcomes.

Exclusion criteria

Exclusion criteria will be patients suffering from another or pre-existing inflammatory disease, pregnancy, malignancy or if surgery is performed due to an emergency.

Participant commitment

Engagement for the patient will include time required to complete the initial questionnaire, as well as the potential to contact the patient at least twice (at 6-month intervals) after the completion of surgery to provide details on the subsequent result of the treatment, as well as LERF, or additional details that might be required for this or other ethically approved projects. The participant commitment will include the completion of a series of online questionnaires both pre and post-surgery and the removal of biological material during laparoscopic surgery that is already planned for their treatment. Eligible patients will be requested to complete the LERF questions at 6 or 12 months follow up.

Participant recruitment strategies and timeframes

Initial patient contact will occur during an initial consultation in which the treating physician and patients will become known to each other. After the initial consultation the treating physician, under the supervision of the Chief Investigator will identify women that match the

inclusion criteria for the study and that have been indicated and scheduled for laparoscopic surgery at the Department of Obstetrics & Gynecology, RBWH.

Once potential patients have been identified the treating physician will organize subsequent contact via a phone call, or other approved modes of telehealth contact by either the treating physician, or a dedicated research nurse. This may occur as part of either the patients planned telehealth treatment, or as an additional phone call. The treating physician, or dedicated research nurse will explain the project in detail and will ask the patient if they would like additional information.

A follow up phone call between the study doctor and the patient will be arranged to discuss the information that was provided. During this phone call the study doctor, or study nurse can go through the patient information sheet and informed consent forms and allow the patient to ask any questions they may have. If they agree the patient will be asked to sign the PICF by a process considered acceptable to Queensland Health.

Any electronically signed consent forms must be printed out on receipt by the study doctor or research nurse. The study doctor or research nurse is to write a file note on the consent form on the process used to obtain consent, sign and date the form. A copy of the completed consent form will be returned electronically to the patient for their records.

Recruited and consented participants will be sent either a link to an online questionnaire, or, if requested, a hardcopy version that has been based on the World Endometriosis Research Foundation (WERF) questionnaire. The approach used is based on the WERF questionnaire, although minor variations the language of the questions ensure participants answer questions relevant to their current episode of endometriosis. As endometriosis is a recurrent disease, we wish to ensure the data collected is linked to the tissues collected. If the patient reports a previous episode of endometriosis they are given the opportunity to answer the same set of questions again, describing and comparing their experience to their previous episode. The patients are also afforded the opportunity to complete a free form section describing their journey through endometriosis.

The online Questionnaire data will be collected using a "LimeSurvey" interface. This software is freeware but allows a user to install a version behind a secure firewall and does not rely on cloud-based storage. All data will be held independently within the Human Studies Unit (HSU), endometriosis database at the University of Queensland with the local site investigator approving access to site specific staff. Access is through personalised logins, passwords and

mutifactor authentification for the research team. Collection of this data is done under informed consent allowing indefinite storage and use with other information collected and/or generated. Withdrawal of a participant may necessitate destruction of the data. This can be done although summary data may already have been used in previous publications. This questionnaire is to be completed in the participants own time and prior to surgery.

Day of Surgery

The women will arrive at the Department of Obstetrics & Gynaecology, RBWH for the scheduled surgical procedure. If a paper copy of the presurgical questionnaire was completed patients will hand in their completed questionnaire. During surgical preparation and prior to the administration of anaesthesia, peripheral blood will be collected. Remaining biological samples, including eutopic endometrial biopsies, ectopic endometriotic lesions and peritoneal fluid/washing will be collected when clinically indicated during surgery and processed as described. Patients will receive the required standard of care and involvement in this study will not have an influence on the quality of treatment provided.

Collection of relevant patient symptom and phenotypical data

Patients that agree to the study and have provided informed consent will be asked to complete the background questionnaire that will capture data relevant for the study and include information on symptoms related to endometriosis based on the previously tabulated adapted version of the WERF Endometriosis phenome and biobanking harmonization project (EPhect) questionnaire. Modifications have been made to this questionnaire that does not change the content but does clarify the language to ensure the participant answers questions relevant to their current episode of endometriosis. The questionnaire contains the possibility to answer the same set of question, but in relation to a previous episode of endometriosis if the patient is willing. Additional questionnaires 6 and 12 months after treatment will elicit information regarding treatment response.

Collection of surgical and clinical data

Clinical data relevant to endometriosis pathology and symptoms, including but not limited to age, weight, BMI, gravidity and parity, ethnicity, previous use of medication and other gynaecological disorders will be collected through the completion of the WERF EPHect surgical phenotype data collection form that is to be completed by the surgeon at the time of surgery. To assist in entry of surgical and clinical data into the patient database an electronic

interface has been built that mirrors the questions of the WERF surgical forms that allows streamlined data entry directly into the database and appended to the patient data.

Baseline Data Collection

Clinical and phenotypic data collection

- Demographic characteristics – Name, date of birth, Medicare number, Healthcare Identifier (HI), residential postcode, gender, ethnicity, language spoken at home, education and employment, marital status and history of smoking and alcohol use.
- Clinical presentation characteristics – Presenting symptoms (cyclical and non-cyclical signs and symptoms: dysmenorrhea, deep dyspareunia, dysuria, dyschezia, painful rectal bleeding or haematuria, shoulder tip pain, catamenial pneumothorax, cyclical cough/haemoptysis/chest pain, cyclical scar swelling and pain, fatigue), pain history and symptoms, endometriosis history
- Pain-related questionnaires - Visual analogue scale (VAS) scores for dysmenorrhoea, non-menstrual pelvic pain, dyspareunia and dyschezia
- Patient reported outcomes measures (PROMs) characteristics – depression, anxiety, and Health-Related Quality of Life (QoL)
- Menstrual history including age at menarche, menstrual irregularities, last date of menstrual period, frequency, duration of flow, and menstrual cycle length will be recorded
- Fertility history, age at the start of each pregnancy, type of fertility treatment used for each pregnancy, and pregnancy outcome (if applicable), live births, type of delivery, and pregnancy complications will be captured
- Medical and surgical history – A detailed history of medical and surgical treatments, medication history, and primary diagnosis

Surgical Data Collection

The surgical data will be collected using the WERF-EPHect Standard Surgical Form [SSF][30] designed to capture all relevant visual information of the endometriosis lesion phenotype and the surgical treatment including:

- Current menstrual cycle, current hormonal treatment, pain or other medications used, dose and duration, side effects and other types of complementary interventions utilized. History of previous surgery for endometriosis
- Imaging (ultrasound and/or MRI) detection or pre-surgical mapping of endometriosis

- Surgical management – type of surgery performed, duration of the procedure, complications, additional conditions identified at surgery, confirmation of endometriosis
- Detailed visual information on intraoperative findings such as extent, exact location, and color of endometriotic lesions, with a specific focus on endometrioma dimensions and endometriotic nodules. It also includes information on the endometriosis fertility index
- Based on the EPHect SSF data, the phenotype characterization of endometriosis will be carried out into superficial peritoneal endometriosis, cystic ovarian endometriosis, or endometrioma and deep infiltrating endometriosis
- From the EPHect SSF data, rAFS (Revised American Fertility Society) score will be calculated and endometriosis cases will be grouped into Stage I/II or Stage III/IV
- Documentation of histopathology reports from samples excised at time of surgery for histopathological diagnosis and confirmation of endometriosis
- The soft copies of the labelled laparoscopic images will be uploaded to the electronic database

Collection of Biological tissue

Relevant biological samples from peripheral blood, peritoneal fluid/washing and tissue implicated in endometriosis pathogenesis, including endometrium and endometriotic lesions will be collected from women undergoing laparoscopic surgery and endometrial biopsies indicated as part of their designated treatment regime.

Blood samples will be collected from every participant on the day of surgery when intravenous access is sited. The remaining biological tissue, including peritoneal fluid, endometrial biopsies and ectopic lesions will be collected as clinically indicated during surgery. Endometrial biopsies will be performed only when clinically indicated. Ectopic lesions will be removed to surgically treat endometriosis, peritoneal fluid is removed and peritoneal cavity is washed with normal saline as a routine procedure. In the context of the surgery no additional surgical action is performed and therefore there is no additional risk to the patient. The treatment received is not altered by participation in the study. Biological material collected will be stored securely at appropriate storage facilities at the IMB UQ laboratory. In all cases, only approved research personnel will have access to the specimens.

- *Whole blood:* Whole blood to a total of 20ml will be collected in evacuated tubes per patient and separated into plasma, serum and buffy coat. Buffy coats will be used either for immune cell isolation and culture, or DNA isolation for genomic analysis.

- *Endometrial biopsy* Primary samples will be saved in RNAlater for subsequent analysis of somatic mutations, gene or protein expression. Any remaining sample will be flash frozen in liquid nitrogen or stored in cyro-protective media and used for gene or protein expression and the isolation and culture of endometrial and immune cells contained within endometrial biopsies. Any remaining sample(s) will be stored in *tissue tek* for the potential use of frozen sections.
- *Peritoneal fluid / washing*: On arrival in the laboratory peritoneal fluid / washing will be immediately centrifuged to remove any immune cell content. The immune cell pellet will be kept in cyro-protective media for subsequent culture of immune cells, and the supernatant aliquoted for the analysis of pure peritoneal fluid. The aliquoted fluid will be stored at -70°C to allow batch analysis of immune, hormonal, endocannabinoids or metabolomic components related to endometriosis pathogenesis and progression.
- *Ectopic lesions:* Endometrial ectopic tissue collected from either the SUP, OMA or DIE lesions will be carefully dissected from any surrounding tissue and both the lesion and surrounding tissue collected in either RNAlater for somatic mutation, gene or protein expression analysis, or either flash frozen or stored in cyro-protective media for subsequent isolation and culture of ectopic endometrial cells, or formalin fixed for production of slides that will be used for haematoxylin and eosin staining, histological confirmation of endometriosis and protein analysis through immunohistology.

Laboratory studies and data generation

- *Cell isolation and culture:* The eutopic, ectopic and immune cells isolated from women undergoing surgery will be used for the isolation and culture of cells to produce *in vitro* models. Standard isolation procedures including size exclusion and fluorescence activated cell sorting (FACS) will be utilised to isolate cells of interest and establish *in vitro* models.
- *Genetic assays* Peripheral blood will be separated into serum and plasma. To determine genetics variants related to endometriosis risk, DNA will be extracted from buffy coats and genotyped on Illumina GSA chips using standard protocols assessed on genomic arrays to determine genome wide genetic variants. These arrays are developed to identify single nucleotide polymorphisms that commonly occur randomly across the population. To increase the number of assayed single nucleotide variants (SNP) imputation using data derived from the 1000 genomes project will be used to infer SNPs in linkage disequilibrium (LD).

- *Somatic mutational analysis:* Selected tissue, or isolated cells will be used for somatic mutational screening via either targeted sequencing assays, whole exome or whole genome sequencing.
- *Induced genetic mutations for in vitro analysis of somatic cells.* To assess the influence of individual, or multiple SNPs associated with endometriosis in the context of consistent genetic backgrounds in isolated endometrial cells specific alleles can be induced in cells via CRISPR-cas9 technology. We will introduce specific allelic variants in the *in vitro* cell models generated from isolated endometrial cells. We will determine their influence on gene expression and cell function using methods described below. These somatic cells will be used for *in vitro* experiments only.
- *Gene expression:* RNA will be isolated using standard procedures and either targeted gene expression via Real-time PCR, multiplexed targeting via Nanostring technology, or genome wide gene expression via RNA-sequencing or single cell RNA-sequencing, will be performed to determine gene and transcript expression levels.
- *Protein expression:* Cellular protein concentrations of the isolated tissue will be performed by both quantitative and semi-quantitative methods including Western blot analysis and immunohistochemistry. Secreted proteins will be measured via enzyme linked immune-absorbent assay using the standard manufacturers protocol.
- *Metabolomic profiling:* A non-targeted metabolomic profiling will be performed with Ultra high-performance liquid chromatography high resolution mass spectrometry (UHPLC-HRMS).
- *Functional Assay:* The influence of observed biological variations on the cellular functional behaviour either as individual or as dual cell cultures will be determined by numerous cellular assays that will include, but not be limited to viability, proliferation, apoptosis, migration and invasion.

Participant follow-up

Post-surgical treatment patients will be followed up to determine their response to treatment, including the reduction in symptoms as well as the presence of side effects. Follow-up questionnaires are to be provided to participants at 6-and 12-months post-surgery. These questionnaire capture details of the individual's response to treatment and will be either emailed via a personalized link to the questionnaire or mailed to individual participants by personnel authorized to access the address of the patients.

Collection of Lifestyle and Environmental Factors (LERF)

Participants who had surgery and completed the baseline questionnaire will be invited to answer LERF questions at the 6 or 12-month follow up. An initial assessment will be conducted with 500 participants invited to complete LERF questions.

The LERF questions collects information in relation to:

- Demographics
- Family
- Lifetime occupations (formatted to match Australian census data)
- General health (including mental health)
- Exposures to potential environmental toxins.

There are some questions regarding use of prescription and illicit drug as well as medical history information including questions pertaining to sexually transmitted diseases (STDs). For female participants there are specific questions relating to menstruation and pregnancy history. This will be administered to all consented individuals and will take approximately 40 minutes to complete. Participants are informed that they are not obliged to answer any questions that make them feel uncomfortable or they can skip questions they do not wish to answer.

Participant withdrawal:

Patients may withdraw their consent at any time during the project by contacting either the Department of Obstetrics and Gynaecology, or their treating physician whose details are provided on the patient information sheet. Any biological sample already collected, as well as health related data will be removed from the project and the laboratory. Subject data included in the prior analysis will not be affected by the withdrawal from the study.

Data management

Each patient identified for inclusion in the study will be assigned a unique study identification number (Study ID). This study number will be assigned upon collection of the signed consent at the Department of Obstetrics & Gynaecology, RBWH and matched with a hospital derived patient identification number (PID). This file will represent the only link between the study number and PID number and will remain accessible only to authorized members of the RBWH clinical team in a secure database with an audit trail function and accessible only via a username and password. Clinical and surgical data collected through the completion of the surgeons

reporting questionnaire and the patient questionnaire respectively will be entered into a secure database by authorized staff.

Patient and surgical data, lifestyle and environmental data collected will be stored in locked, linked databases housed behind University/Hospital firewalls at IMB/UQ. Experimental procedures and data analysis will be performed at the UQ, IMB Laboratory. The biological material and the health-related data required for analysis will be provided to the UQ laboratory attached to the study ID only and all subsequent experimental data generated during the project will be stored attached to the study ID on secure servers behind UQ University firewall. This data will be considered anonymous to laboratory staff. Only authorized staff for the project will have access to the data. Transfer of any data will be via secure transfer protocols. Data storage will be for the duration of the project, as well as and for subsequent projects based on information or material derived from this project and retained for a minimum of 5 years.

Electronic database

A dedicated database has been developed for the project using the open-source Drupal software. The data will be housed on a University server behind the University firewall. Treating physicians or research nurses who consent the patients will enter the patient data directly into the electronic database that will generate an ID. Subsequent patient, surgical and biological collections will linked to the study ID. Surgical data collected from the WERF EPHect surgical form will be assessed for consistency by key clinical surgical project leaders and the data entered directly into the Drupal database via a built in electronic interface via the relevant physician, study nurse or project leader.

The Study ID will also be used to generate personalized links to the patient's questionnaire, both before and after surgery, that will allow the patient answers to be directly entered into the Lime Survey database. The Drupal database and Lime Survey interface will function as a central repository that will allow real time data entry and reduce input error. Access to data will only be for authorized personnel.

Data Analysis plan

Data analysis will be performed using a combination of SPSS, GraphPad Prism, plink and R statistical and genomic packages. Analysis of patient and surgical data will be used to identify pre-surgical, risk factors that may facilitate an early diagnosis of endometriosis. Patient data collection matched to the current episode of endometriosis will be used in linear regression models to establish correlation between patient symptoms and lesion characteristics observed

during surgery and facilitate subgroup discovery. This will include statistical comparisons between patient reported symptoms prior to surgery, as well as phenotypic descriptions of lesions observed during surgery. Follow up questionnaires will be used to assess patient response to their individualized treatments and whether linked pre-surgical data and surgical observation can subcategorize patients into groups that predict treatment response and disease progression.

Biological samples of peripheral blood, peritoneal fluid and eutopic and ectopic endometrium will be explored to identify biological mechanisms that underlie molecular signatures, including gene and transcriptomic expression, alternative splicing and common and acquired genetic mutations, proteomic and metabolic profiles. Molecular signatures generated with collected biological samples will be matched to pre-surgery patient data, surgical data, follow up data and LERF data to discover correlations between molecular signature and clinically relevant patient outcomes.

In vitro models will be used to assess the functional influence of genotypes on individual cellular function, including gene and transcriptome expression, alternative splicing, common and rare genetic mutations, proteomic and metabolomic profiles. The interaction between different cell types and their influence on clinically relevant molecular signatures. These molecular signatures will also be identified, or induced, in individual *in vitro* models and utilized to determine potential personalization of current treatments to individual molecular profiles or identify novel treatments.

Additional genotyping data and functional genomics datasets aimed at identifying the causal genetic variants and biological mechanism that are associated with endometriosis susceptibility will be generated to potentially contribute to consortium meta-analysis in collaboration with other groups worldwide.

Patient samples with accompanying missing data releavnt for an individual analysis, will be removed from the analysis. Similarly, patients that has been lost to follow up will be excluded only from analyses that require a complete dataset.

Power Calculations

It is assumed comparisons will be performed predominantly between two groups. This group could be defined as either women with and without endometriosis, or as either positive or negative response to treatment, as well as many other possibilities. Analysis between two independent means from patient-derived variables (including patient response and

experimentally measured) will aim to achieve a moderate power of 80% to target small effect sizes of 0.2. It is estimated to achieve a significant difference a sample size of 394 samples per group, or 786 total would be required. It is intended for recruitment procedures to continue for four years to establish a database of sufficient samples that will satisfy power requirements. Over a 4-year period we plan to recruit approximately 300 patients per year to reach up to 1,200 patients by the end of the period. Based on first year response rates we predict an incomplete protocol from approximately 15% of patients each year, leaving approximately 1,00 – 1,100 patients over the 4-year collection period, providing the required sample numbers.

Research Team Training and Quality Controls

Research staff are all professionally trained in their relevant positions. They have also received study-specific training. This included training for treating physicians and study nurse on the appropriate recruitment of study participants, how to obtain informed consent, counseling of the study participant, data collection, interacting with the database interface and the follow up of patients with incomplete study protocols. Specific training has been provided on correct data entry, data management, data analysis and the maintenance of data sets.

Specific training has been provided for the collection and processing of biological samples including blood and its separation into constituents, aliquoting and recording. Standard operating procedures (SOP) based on best practice and the WERF biological processing protocols [31 32], including the storage of endometrial eutopic and ectopic lesions for either fresh frozen or cryopreserved process, as well as the separation of individual cells for establishing *in vitro* models have been developed.

Ethics and dissemination:

The study has been approved and will continue to be monitored by the Metro North Human Research Ethics committee (HREC) and has received site specific approval for the conduct of the study at the RBWH. Research activities at the University of Queensland with the data and samples generated also overseen by the UQ HREC. Annual reports will be submitted to both the Metro North and UQ HREC for regular updates on the study. Research results will be published in peer-reviewed journals and the results presented at conferences where appropriate. If requested by the journals, de-identified results may be submitted to recognized data repositories with controlled access policies to provide open access for research purposes.

Project closure procedure:

The total duration of the project is 5 years. Four years of tissue collection followed by an additional year to complete tissue and data analysis. There will be the potential to extend the project if deemed beneficial and ethics requirements are met. The project will include the production of *in vitro* models potentially useful to future projects and other researchers and samples will not be destroyed at the completion of the project. These models may be made conditionally available, subject to the receipt of appropriate ethical approval. Third party researchers will only have access to non-identifiable data and will not be able to match these cell models with patients.

Study status

The present study was conceptualized and designed in 2020. The first participant was enrolled in September 2020 and the study is currently ongoing with the number of collected sample reaching the desired number expected after the first year of recruitment. Baseline characteristics show positive recruitment and protocol completion rates. Characteristics of the of endometriosis patients, patients identified to have another non-endometrial pathology and no discernible endometrial pathology collected in the first year, are shown in **Table 1**.

Table 1: Patient characteristics collected in the first year of the ERQS

Characteristics		Endometriosis N (%)	Other endometrial pathology N (%)	No pathology N (%)	P value
Demographic		N = 168 (52.97%)	N = 67 (20.23%)	N = 88 (26.79%)	
BMI (kg/m^2), Mean ± SD (*Ordinary one-way ANOVA*)		27.07 ± 8.95 (n = 151)	31.19 ± 6.95 (n = 62)	27.33 ± 6.41 (n = 80)	*P = 0.002
BMI (kg/m^2) categories (*Chi Squared*)	Underweight (≤18.5)	5 (3.3%)	1 (1.6%)	2 (2.5%)	*P = 0.003
	Normal (18.5 – 24.9)	64 (42.4%)	12 (19.4%)	34 (42.5%)	
	Overweight (25 – 29.9)	47 (31.1%)	17 (27.4%)	19 (23.8%)	
	Obese (≥30)	35 (23.2%)	32 (51.6%)	25 (31.3%)	
	Unknown/No response	17	5	8	
Education level	Year 10, or equivalent	11 (7.5%)	5 (8.8%)	4 (6.2%)	P = 0.729
	Year 12	17 (11.6%)	6 (10.5%)	11 (13.6%)	
	Diploma	51 (34.9%)	22 (38.6%)	30 (37.0%)	
	TAFE	12 (8.2%)	3 (5.3%)	2 (2.5%)	
	University undergraduate	34 (23.3%)	12 (21.1%)	25 (30.9%)	
	University Postgraduate	21 (14.4%)	9 (15.8%)	8 (9.9%)	
	No response	22	10	7	
Employment	Government/ Private job	112 (77.2%)	46 (78.0%)	53 (65.4%)	P = 0.235
	Self employed	10 (6.9%)	2 (3.4%)	5 (6.2%)	
	Unemployed	9 (6.2%)	8 (13.6%)	13 (16.0%)	
	Student	8 (5.5%)	1 (1.7%)	7 (8.6%)	

	Unable to work due to chronic pain	6 (4.1%)	2 (3.4%)	3 (3.7%)	
	Unknown	23	8	7	
Marital status	Married	100 (67.6%)	43 (68.3%)	52 (64.2%)	P = 0.868
	Divorced/Separated	8 (5.4%)	4 (6.3%)	2 (2.5%)	
	In a relationship, not living together	12 (8.1%)	4 (6.3%)	9 (11.1%)	
	Single	27 (18.2%)	12 (19.0%)	18 (22.2%)	
	Widowed	1 (0.7%)	0 (0%)	0 (0%)	
	Unknown/No response	20	4	7	

Discussion

This protocol outlines the framework for a single site study investigating endometriosis and the development of the infrastructure that supports the collection of baseline data that will underpin the research. The infrastructure developed has the potential to be applied to additional sites if the study protocol was to be expanded. The study protocol has been designed to generate baseline data for each patient including data on patients reported symptoms, menstrual and fertility history and their previous treatment experience. It will also include surgical observations recorded during treatment and collected by physicians. The research infrastructure will ensure high quality baseline data is both generated, entered and stored in purpose-built database that increases ease of entry and security of the data and decreases user input error.

Collection of baseline patient and surgical data will be built upon with genotyping data from DNA isolated from peripheral blood that will allow functional investigation of genetic risk variants associated with endometriosis susceptibility. The collection of endometriosis relevant tissue will allow the generation of omics databases to investigate the functional influence of the genetic variants. It will also allow the creation of *in vitro* models of individual cell types with the accompanying genomics background.

This study represents a significant opportunity to collect deeply phenotyped endometriosis samples to generate cellular models and detailed integrated molecular datasets. It will be the foundation of endometriosis biorepositories of multiple tissue types including serum, plasma DNA, RNA, as well as the establishment of personalized *in vitro* models that can be screened for responses to treatments and compounds. The results from the analysis will improve our understanding of endometriosis pathogenesis, the clinical symptoms associated with subtypes and how these respond to various treatments. It will provide valuable datasets to contribute to global networks and International Consortia studying endometriosis and improve outcomes for endometriosis patients.

Funding;

G.W.M is supported by the National Health and Medical Research Council (NHMRC) fellowship (GNT1078399 and GNT1177194). G.W.M, A.A. K.T and S.M were supported by the Metro North Hospital and Health Service (CRG-176-2019). K.T. is supported by RBWH and RBWH Foundation Post Graduate Scholarship. B.M. is supported by Endometriosis Australia (N/A).

Author contributions

B.M, A.A and K.T conceived the study and wrote the protocol with assistance from GWM. AA and KT consented patients. AA, KT, BS, PG, AK and DB, MS and DG collected patient data. BM, SS, SA, KT, LMW collected and processed patient samples. MBL and PL build and maintained database and AKH oversaw database and sample processing. SM assisted with power calculations and performed data analysis.

Ethics approval

Ethics approval has been granted by the Metro North Human ethics research council Ref No: HREC/2019/QRBW/56763 and the University of Queensland Human research ethics committee 2019/HE002744.

Competing Interest Statement

No competing interests

References.

1. Vercellini P, Viganò P, Somigliana E, et al. Endometriosis: pathogenesis and treatment. *Nature Reviews Endocrinology* 2014;10(5):261-75. doi: 10.1038/nrendo.2013.255
2. Grandi G, Toss A, Cortesi L, et al. The Association between Endometriomas and Ovarian Cancer: Preventive Effect of Inhibiting Ovulation and Menstruation during Reproductive Life. *BioMed Research International* 2015;2015:751571. doi: 10.1155/2015/751571
3. Dunselman GAJ, Vermeulen N, Becker C, et al. ESHRE guideline: management of women with endometriosis †. *Human Reproduction* 2014;29(3):400-12. doi: 10.1093/humrep/det457
4. Guo S-W. Recurrence of endometriosis and its control. *Human Reproduction Update* 2009;15(4):441-61. doi: 10.1093/humupd/dmp007
5. Simoens S, Dunselman G, Dirksen C, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum Reprod*

2012;27(5):1292-9. doi: 10.1093/humrep/des073 [published Online First: 2012/03/17]
6. Australian Government DoH. National Action Plan for Endometriosis. 2018
7. J.A. S. Peritoneal Endometriosis Due to the Menstrual Dissemination of Endometrial Tissue into the Peritoneal Cavity. *American Journal of Obstetrics & Gynecology* 1927;14:442-69.
8. Halme J, Hammond MG, Hulka JF, et al. Retrograde menstruation in healthy women and in patients with endometriosis. *Obstet Gynecol* 1984;64(2):151-4. [published Online First: 1984/08/01]
9. Burney RO, Giudice LC. Pathogenesis and pathophysiology of endometriosis. *Fertil Steril* 2012;98(3):511-9. doi: 10.1016/j.fertnstert.2012.06.029 [published Online First: 2012/07/24]
10. Barragan F, Irwin JC, Balayan S, et al. Human Endometrial Fibroblasts Derived from Mesenchymal Progenitors Inherit Progesterone Resistance and Acquire an Inflammatory Phenotype in the Endometrial Niche in Endometriosis. *Biol Reprod* 2016;94(5):118. doi: 10.1095/biolreprod.115.136010 [published Online First: 2016/04/15]
11. Gargett CE, Schwab KE, Deane JA. Endometrial stem/progenitor cells: the first 10 years. *Hum Reprod Update* 2016;22(2):137-63. doi: 10.1093/humupd/dmv051 [published Online First: 2015/11/11]
12. Nguyen HPT, Xiao L, Deane JA, et al. N-cadherin identifies human endometrial epithelial progenitor cells by in vitro stem cell assays. *Human Reproduction* 2017;32(11):2254-68. doi: 10.1093/humrep/dex289
13. Chuang PC, Wu MH, Shoji Y, et al. Downregulation of CD36 results in reduced phagocytic ability of peritoneal macrophages of women with endometriosis. *J Pathol* 2009;219(2):232-41. doi: 10.1002/path.2588 [published Online First: 2009/07/17]
14. Izumi G, Koga K, Takamura M, et al. Mannose receptor is highly expressed by peritoneal dendritic cells in endometriosis. *Fertil Steril* 2017;107(1):167-73.e2. doi: 10.1016/j.fertnstert.2016.09.036 [published Online First: 2016/10/30]
15. Thiruchelvam U, Wingfield M, O'Farrelly C. Natural Killer Cells: Key Players in Endometriosis. *Am J Reprod Immunol* 2015;74(4):291-301. doi: 10.1111/aji.12408 [published Online First: 2015/06/25]
16. Tariverdian N, Siedentopf F, Rücke M, et al. Intraperitoneal immune cell status in infertile women with and without endometriosis. *J Reprod Immunol* 2009;80(1-2):80-90. doi: 10.1016/j.jri.2008.12.005 [published Online First: 2009/04/21]
17. Furuya K, Murakami M, Makimura N, et al. Immunological and endocrinological studies on lymphocyte subpopulation and medical treatment for infertility in patients with endometriosis. *Mol Cell Endocrinol* 2003;202(1-2):195-9. doi: 10.1016/s0303-7207(03)00083-2 [published Online First: 2003/05/29]
18. Treloar SA, O'Connor DT, O'Connor VM, et al. Genetic influences on endometriosis in an Australian twin sample. sueT@qimr.edu.au. *Fertil Steril* 1999;71(4):701-10. doi: 10.1016/s0015-0282(98)00540-8 [published Online First: 1999/04/15]
19. Sapkota Y, Steinthorsdottir V, Morris AP, et al. Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nature Communications* 2017;8(1):15539. doi: 10.1038/ncomms15539

20. Fung JN, Rogers PA, Montgomery GW. Identifying the biological basis of GWAS hits for endometriosis. *Biol Reprod* 2015;92(4):87. doi: 10.1095/biolreprod.114.126458 [published Online First: 2015/02/20]
21. Nilufer R, Karina B, Paraskevi C, et al. Large-scale genome-wide association meta-analysis of endometriosis reveals 13 novel loci and genetically-associated comorbidity with other pain conditions. *bioRxiv* 2018:406967. doi: 10.1101/406967
22. Anglesio MS, Papadopoulos N, Ayhan A, et al. Cancer-Associated Mutations in Endometriosis without Cancer. *N Engl J Med* 2017;376(19):1835-48. doi: 10.1056/NEJMoa1614814 [published Online First: 2017/05/11]
23. Suda K, Nakaoka H, Yoshihara K, et al. Clonal Expansion and Diversification of Cancer-Associated Mutations in Endometriosis and Normal Endometrium. *Cell Rep* 2018;24(7):1777-89. doi: 10.1016/j.celrep.2018.07.037 [published Online First: 2018/08/16]
24. Barra F, Grandi G, Tantari M, et al. A comprehensive review of hormonal and biological therapies for endometriosis: latest developments. *Expert Opin Biol Ther* 2019;19(4):343-60. doi: 10.1080/14712598.2019.1581761 [published Online First: 2019/02/15]
25. Lazorwitz A, Aquilante CL, Oreschak K, et al. Influence of Genetic Variants on Steady-State Etonogestrel Concentrations Among Contraceptive Implant Users. *Obstetrics and gynecology* 2019;133(4):783-94. doi: 10.1097/AOG.0000000000003189
26. Di Marzo V, Izzo AA. Endocannabinoid overactivity and intestinal inflammation. *Gut* 2006;55(10):1373-76. doi: 10.1136/gut.2005.090472
27. Smith DR, Stanley CM, Foss T, et al. Rare genetic variants in the endocannabinoid system genes CNR1 and DAGLA are associated with neurological phenotypes in humans. *PLoS One* 2017;12(11):e0187926. doi: 10.1371/journal.pone.0187926 [published Online First: 2017/11/18]
28. Meyer HC, Lee FS, Gee DG. The Role of the Endocannabinoid System and Genetic Variation in Adolescent Brain Development. *Neuropsychopharmacology* 2018;43(1):21-33. doi: 10.1038/npp.2017.143
29. Chidambaran V, Pilipenko V, Spruance K, et al. Fatty acid amide hydrolase-morphine interaction influences ventilatory response to hypercapnia and postoperative opioid outcomes in children. *Pharmacogenomics* 2017;18(2):143-56. doi: 10.2217/pgs-2016-0147 [published Online First: 2016/12/16]
30. Becker CM, Laufer MR, Stratton P, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: I. Surgical phenotype data collection in endometriosis research. *Fertil Steril* 2014;102(5):1213-22. doi: 10.1016/j.fertnstert.2014.07.709 [published Online First: 2014/08/26]
31. Rahmioglu N, Fassbender A, Vitonis AF, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: III. Fluid biospecimen collection, processing, and storage in endometriosis research. *Fertil Steril* 2014;102(5):1233-43. doi: 10.1016/j.fertnstert.2014.07.1208 [published Online First: 2014/09/27]
32. Fassbender A, Rahmioglu N, Vitonis AF, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: IV. Tissue collection, processing, and storage in endometriosis research. *Fertil Steril* 2014;102(5):1244-53. doi: 10.1016/j.fertnstert.2014.07.1209 [published Online First: 2014/09/27]

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	An integrated cohort study approach
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2	NA
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4	Introduction section
Objectives	3	State specific objectives, including any prespecified hypotheses	6	Objectives section
Methods				
Study design	4	Present key elements of study design early in the paper	7	Project design and study sites
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7	Project design and study site
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8 11 14	Inclusion criteria Baseline datacollection Participant follow up
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case		NA - will only be available after diagnosis
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	13	Outcomes defined in Laboratory studies and data generation
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	13	Outcomes defined in Laboratory studies and data generation
Bias	9	Describe any efforts to address potential sources of bias	NA	
Study size	10	Explain how the study size was arrived at	17	Power calculations

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	16	Data analysis plan
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	16	Data analysis plan
		(*b*) Describe any methods used to examine subgroups and interactions	16	Data analysis plan
		(*c*) Explain how missing data were addressed	16	Data analysis plan
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	16	Data analysis plan
		(*e*) Describe any sensitivity analyses	NA	
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	24	Table 1
		(b) Give reasons for non-participation at each stage	24	Table 1
		(c) Consider use of a flow diagram	NA	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	24	Table 1
		(b) Indicate number of participants with missing data for each variable of interest	NA	
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	24	Table 1
		Case-control study—Report numbers in each exposure category, or summary measures of exposure		
		Cross-sectional study—Report numbers of outcome events or summary measures		
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	24	Table 1
		(*b*) Report category boundaries when continuous variables were categorized	24	Table 1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA	

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA	
Discussion				
Key results	18	Summarise key results with reference to study objectives	19	Discussion
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	3	Strenght and limitations of the st
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	NA	
Generalisability	21	Discuss the generalisability (external validity) of the study results	NA	
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	20	Funding

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.