The PFAS Health Study

Phase II Research Protocols

Martyn Kirk,[1] Rosemary Korda,[1] Archie Clements,[2] Robyn Lucas,[1] Cate D'Este,[1] Bruce Armstrong,[3,4] Jochen Mueller,[5] Alison Jones,[6] Adrian Miller,[7] Cathy Banwell,[1] Philip Batterham,[8] Aparna Lal,[1] Jennifer Braunig,[5] Tambri Housen[1]

1. National Centre for Epidemiology and Population Health, Research School of Population Health, The Australian National University, Canberra, ACT, Australia.
2. Research School of Population Health, The Australian National University, Canberra, ACT, Australia.
3. Sydney School of Public Health, The University of Sydney, NSW, Australia
4. Sax Institute, Sydney, NSW, Australia
5. National Research Centre for Environmental Toxicology, The University of Queensland
6. University of Wollongong, NSW, Australia
7. Office of Pro Vice-Chancellor Indigenous Leadership, Charles Darwin University, Darwin, NT, Australia
8. Centre for Mental Health Research, Research School of Population Health, The Australian National University, Canberra, ACT, Australia.

Australian National University

For the Australian Government Department of Health

The PFAS Health Study Team

Role	Name	Organisation
Principal Investigator	Associate Professor Martyn Kirk	The Australian National University
Co-Investigator	Dr Rosemary Korda	The Australian National University
Co-Investigator	Professor Robyn Lucas	The Australian National University
Co-Investigator	Professor Cate D'Este	The Australian National University
Co-Investigator	Professor Archie Clements	The Australian National University
Co-Investigator	Emeritus Professor Bruce Armstrong	The University of Sydney
Co-Investigator	Professor Jochen Mueller	The University of Queensland
Co-Investigator	Professor Alison Jones	University of Wollongong
Co-Investigator	Professor Adrian Miller	Charles Darwin University
Co-Investigator	Associate Professor Cathy Banwell	The Australian National University
Associate Investigator	Associate Professor Philip Batterham	The Australian National University
Associate Investigator	Dr Aparna Lal	The Australian National University
Associate Investigator	Dr Jennifer Braunig	The University of Queensland
Associate Investigator	Dr Tambri Housen	The Australian National University
Research Officer	Ms Susan Trevenar	The Australian National University
Research Assistant	Ms Kayla Smurthwaite	The Australian National University
Research Assistant	Ms Liz Walker	The Australian National University

Study Management
Associate Professor Martyn Kirk
National Centre for Epidemiology and Population Health
Research School of Population Health
The Australian National University
ACT 2601 AUSTRALIA
Telephone: +61 2 6125 5609
Mobile: +61 4 2613 2181
Email: pfas.health.study@anu.edu.au

Contents

1. Synopsis

The Per and Poly-Fluoroalkyl Substances (PFAS) Health Study: Phase II will investigate the exposure to and potential health effects of PFAS in areas of known contamination in the communities of Williamtown, New South Wales (NSW),Oakey, Queensland (Qld) and Katherine, Northern Territory (NT), Australia.

In particular, this study will concentrate on the main chemical components of the firefighting foams used on Defence facilities in the townships of Williamtown, Oakey and Katherine, which are perfluorohexane sulfonic acid (PFHxS), perfluorooctanoic acid (PFOA) and perfluorooctane sulfonic acid (PFOS). The applied research study will build on the results of a systematic review of the health effects of these chemicals.

There are 4 components of Phase II of the study:

1. A focus group study to determine the concerns of individuals living in the vicinity of Williamtown, Oakey and Katherine have in relation to exposure to PFAS and their health.
2. A blood serum study to define the serum concentrations (mean and range) of PFAS in Williamtown, Oakey and Katherine residents living in the Investigation Areas and to compare the levels to those of people residing in non-contaminated areas in the townships and surrounding areas.
3. A cross-sectional survey to investigate the exposure and risk factors for high serum PFAS levels, including sociodemographic (e.g. age, sex, location) and other factors (e.g. duration of residence in the area, water source), and associations of high serum PFAS levels with common symptoms, signs and diagnosed illnesses in the Williamtown, Oakey and Katherine communities.
4. A data linkage study to examine whether sex-specific age adjusted rates of diseases potentially associated with PFAS are higher among people who have lived in the Investigation Areas of Williamtown, Oakey and Katherine, compared to those living outside the Investigation Areas and in the general Australian population.

2. Abbreviations, acronyms and terminology

AFFF—Aqueous film forming foam

AMSANT – Aboriginal Medical Services Alliance Northern Territory

ANU—Australian National University

DVA – Department of Veterans' Affairs

MBG—Model-based geostatistics

NCEPH—National Centre for Epidemiology and Population Health

NSW—New South Wales

NT – Northern Territory

NTDoH and MSHR – Northern Territory Department of Health and Menzies School of Health Research

PFAS—Per- and poly-fluoroalkyl substances

PFHxS—Perfluorohexane sulfonate

PFNA—Perfluorononanoic acid

PFOA—Perfluorooctanoic acid

PFOS—Perfluorooctanate sulfonate

Qld—Queensland

3. Introduction

Per- and poly-fluoroalkyl substances (PFAS), including perfluorooctane sulfonic acid (PFOS) and perfluorooctanoic acid (PFOA), are man-made chemicals. [1] They have a wide range of industrial uses, including in protective coatings on furniture and textiles and non-stick coatings on cookware, and as flame retardants in fire-fighting foams. [2] These chemicals are highly inert thermally, chemically and biologically, however, their toxicity may have been underestimated. [3] PFAS are environmentally persistent and can last for decades in water and soil, and they have been shown to accumulate in the environment (both soil and aquatic systems) of contaminated areas. [2, 4]

Human exposure to PFAS occurs primarily through consumption of contaminated food and water with some exposure through inhalation of contaminated dust and air, which results in higher PFAS levels in the blood than background levels found in the general community. [5, 6] Some long-chain PFAS are biochemically stabile and accumulate in the human body. [7, 8] Once absorbed through digestion and gas-exchange pathways, PFAS bind to plasma protein and accumulate in the blood stream. [9] Prenatal exposure to PFAS can occur through the placental transfer of maternal blood to the foetus in utero.[10] Additionally, infants may be exposed to PFAS through contaminated breastmilk. [11]

International research on the health effects of PFAS

There have been many epidemiological studies reported in the literature examining the health effects of exposure to PFAS. Although a diverse range of health outcomes have been investigated, findings have been inconsistent. Studies have covered people exposed to PFAS at different ages, including those in utero. Chemicals of interest to most researchers have been PFOA and PFOS, although many studies have investigated exposure to a wider range of PFAS, including perfluorohexane sulfonic acid (PFHxS) and perfluorononanoic acid (PFNA). Whilst most studies aimed to identify the health effects associated with exposure to each PFAS separately, researchers have also defined exposure through categorising PFAS based on principal component analysis. [12]

The health effects of exposure to PFAS have primarily been investigated through cross-sectional and cohort studies. A large number of studies have been conducted through existing longitudinal studies, including occupational cohorts and periodic national surveys, such as the National health and Nutrition Examination Survey conducted in the United States. Although many health outcomes have been investigated, a key focus has been pregnancy outcomes and reproductive health. Large birth cohort studies have been used to examine the association between maternal blood concentrations of PFAS chemicals and a range of prenatal effects, including foetal growth, miscarriage, birth weight, and birth defects. [13, 14] Reproductive health focused studies have

examined the association between PFAS exposure and semen quality in males, and the time-to-pregnancy and menstrual cycles of females. [15, 16] The hormonal effects of PFAS exposure have further been investigated, particularly in relation to the functioning of the thyroid gland. [17]

International research on PFAS exposure has investigated the metabolic, cardiovascular, respiratory and gastrointestinal conditions associated with exposure to PFAS. Each of these health outcomes have been investigated separately through epidemiological studies. Metabolic outcomes tested through these studies have largely focused on the function of the kidneys and liver, including concentrations of uric acid and liver enzymes. [18] Additional research has been completed on type 2 diabetes and obesity, specifically in relation to cholesterol levels associated with PFAS exposure. [19]

Cardiovascular conditions related to PFAS exposure have not been evaluated to the same extent as the metabolic outcomes. Hypertension and cardiovascular disease diagnoses, including stroke, have been the key health outcomes examined. [20] Respiratory and gastrointestinal research has been limited in comparison, with asthma and ulcerative colitis the main outcomes investigated. [12, 21] These medical conditions have further been investigated in relation to the immunological effects of exposure to PFAS. Additional research has been completed on the neurodevelopmental effects of PFAS, although the health outcomes investigated have been vast and the results conflicting.

The International Agency for Research on Cancer (IARC) reported that that there was 'limited evidence' in humans for the carcinogenicity of PFOA. A positive association was observed for cancer of the testis and kidney. [22] The IARC concluded overall that PFOA was possibly carcinogenic to humans. [23] Despite correlations between exposure and disease or PFAS compounds, the findings from epidemiological and biological studies have been inconsistent and subject to debate, specifically in regards to the relevance of animal studies to human health outcomes. [11]

PFAS in Williamtown, NSW, Oakey, Qld and Katherine, NT

In recent years, the environments of areas in Williamtown in NSW, Oakey in Qld and Katherine, NT have been identified as being contaminated due to Defence Force firefighting activities on nearby Defence bases. Use of aqueous film forming foam (AFFF) containing PFOS and PFHxS as the main component on the Royal Australian Airforce base in Williamtown, the Army Aviation Centre in Oakey and Royal Australian Airforce base in Katherine, NT has been associated with detection of increased PFAS levels in ground water, soil, biota and humans living or working in the three towns. [24, 25] Since the initial use of the firefighting foams in the late 1970s, members of these communities have been exposed to potentially concerning concentrations of PFAS through the contamination of the local environment. Manufacturers have largely phased out PFOS and PFOA from AFFF due to toxicological concerns and have switched to alternative chemicals with similar properties. These often have a different fluorination that makes them less

persistent. The Australian Defence Force now uses AFFF concentrates that do not contain PFOS and PFOA as active ingredients. However, due to their environmental persistence, PFAS from AFFF used decades earlier still remain in ground water, dust, sediment and soil in the local areas.

Rationale for examining the health effects of PFAS

The public health significance of exposure to these PFAS compounds is unclear. Risk assessments of contaminated areas have evaluated likely exposure, but have been equivocal about the health risk due to the uncertainty about health outcomes of exposure. To date, there have been no well-designed epidemiological studies examining health effects of these chemicals in Australia, as the affected communities are often small and the levels of exposure highly variable. Due to the limited scientific evidence regarding the risks of PFAS exposure, the residents of Williamtown, Oakey and Katherine have expressed concerns about the potential effects on their health and wellbeing. The decline in property and agricultural produce value across the three towns has further contributed to these concerns within the rural communities.

The Australian Government Department of Health established the PFAS Coordination Unit in 2016 to facilitate health-related work regarding PFAS in Williamtown and Oakey and in late 2017 extended this to include Katherine. This work included establishing a voluntary blood testing program to enable residents to have their blood tested for PFAS. The Departmental Taskforce ensured that there were adequate mental health services in the affected areas for residents dealing with issues relating to PFAS. The PFAS Coordination Unit contracted the Australian National University (ANU) to conduct Phase I of an epidemiological study of the potential health outcomes associated with living in contaminated areas. In Phase I, the ANU were to provide advice regarding the blood testing program, conduct a systematic review of the potential health effects of PFAS and develop a protocol for epidemiological studies in the Williamtown, Oakey and Katherine communities (Phase II).

To examine the health effects of PFAS on the communities of Williamtown, Oakey and Katherine, the PFAS study team proposes the following protocol:

1. A focus group study to examine concerns of individuals living in affected areas,
2. A blood serum study to define serum PFAS levels in residents and people residing in non-contaminated areas,
3. A cross-sectional survey to investigate sociodemographic and other risk factors associated with high serum PFAS, health problems and psychological distress, and
4. A data linkage study to examine rates of PFAS candidate diseases among people living in affected areas, compared to the general population.

In this report, we use the term PFAS to cover chemical compounds used in AFF, particularly PFOS, PFOA and PFHxS.

4. Component 1: Focus group study

4.1. Aims and objectives

The primary aim of the focus group discussion study is to gather a range of social and health-related experiences and views from current residents and workers exposed to PFAS from the three communities of Williamtown, Oakey and Katherine, particularly in the Investigation Areas as defined by the Department of Defence and the NSW Environmental Protection Agency (EPA) respectively. The focus groups are designed to elicit experiences of people living in the communities, which may include people outside of the Investigation Areas.

The specific objectives of the focus group discussions are to:

1. Examine the range of experiences and opinions of people living, working or owning property in a PFAS affected area;
2. Understand residents' perceptions of health and other risks from exposure to PFAS in order to inform ways to assist affected residents (e.g. provide mental, social and health services or support);
3. Inform policy responses regarding risk communication relating to environmental threats, to reduce suffering and unnecessary anxiety; and
4. Inform the development of a questionnaire for a future cross-sectional survey of residents.

4.2. Study design

Focus group discussions are a valuable method for collecting qualitative data, as they enable discussion of public knowledge, and underlying attitudes and opinions. They are well suited to exploring a range of views on topics in the public domain although they are less appropriate for gathering information about highly personal and sensitive matters. Discussions may reveal concerns and issues that are often generated by the interaction within the group. As a consequence, focus group discussions are commonly used in health research and the development of social action programs often in conjunction with other research methods such as surveys.

A focus group study, such as this, is predicated on the understanding that health-related concerns are social in nature. The difficulties and concerns that residents may experience are health risks related to exposure and others partly due to the social context in which they live. For example, selling property, moving residence, relocating children and awareness of media-related information are all socially mediated interactions that may provoke feelings of stress, anxiety, or relief and have potential health risks.

4.3. Study population

The study population is residents of the townships of Williamtown, NSW, Oakey Qld and Katherine, NT. Residents living outside the Investigation Area, but in the townships will also have concerns about their health and the impact the contamination is having on their lives. Previous residents will not be actively sought to participate in the focus groups, but will not be excluded if they want to participate.

4.4. Study procedure

Recruitment

Between 4–5 focus groups will be recruited from Williamtown and Oakey each in the Investigation Areas, while slightly more focus groups may be required for Katherine taking into consideration the larger population (~10,000) and different Aboriginal groups giving potentially 12-14 groups in total. Focus groups are not expected to be representative and instead are used to gather an array of views on issues of community importance; in this case PFAS contamination. Ideally, a focus group consists of about 6-12 people although they may be larger.

As these group discussions will focus on a community issue, participants will be invited to take part through a range of community-based groups and media outlets, including local radio and newspapers. This may include displaying posters in shops and community centres advertising focus group discussion times and locations. We will consult with the PFAS Health Study Community Reference Panel (not yet established) and other local 'experts', such as local general practitioners, government representatives and community health workers on the best ways to recruit participants in each location.

We will aim to ensure that the groups will contain a mix of men and women with a range of interests (e.g. property owners, townspeople, farmers). In small communities, it is likely that some focus group participants may know each other. As the topic of the discussions is of considerable concern to the communities, it is expected that people will be keen to attend. It may be difficult to get less concerned residents and workers to attend the focus groups, but invitations will be made through local groups and community members. As focus groups are qualitative in nature there is less emphasis on representativeness. The venues selected for the meetings should be easily accessible for community members (e.g. local halls, community centres) and politically neutral. Each focus groups discussion will be conducted by at least two experienced researchers; one to moderate and the other to observe, take notes and assist.

Study conduct

The manner in which the focus group discussions are conducted will be clearly explained to participants (face-to-face at the time and through information sheets, see Appendix 1) who will be required to give consent (see Appendix 2). Before the discussions commence, participants will

be given a list of open-ended topics (by phone, postal mail or email) so that they are aware of what will be covered in the discussions. The focus group discussions will be conducted to minimise breaches in privacy and confidentiality. They will be audio-recorded and participants will be asked to provide a pseudonym at the beginning of the discussion that will be recorded and used by participants during the discussions. Any identifying information about participants will be deleted before sending recordings to transcribers. Staff from the transcribing service will sign a Confidentiality Deed. A brief one page questionnaire will be distributed to participants to collect basic socio-demographic data, such as age, gender, marital status, employment, number of children (see

Appendix 3). These data can be used to describe the composition of the groups, although it is not expected that group participants will be representative. When participants return the completed questionnaire at the end of the discussion they will also be given a $50 EFTPOS card as reimbursement for their time, which is standard practice where people have made significant contributions to a qualitative study.

Data will be stored on secure servers at the ANU for a minimum of five years following completion of the study. It is expected that the focus group discussions will be held between July and September 2017, and each will be no longer than 2 hours in length.

4.5. Data analysis

The discussions will be audio-recorded and professionally transcribed by a service previously used by the researchers. The transcriptions will be read several times by experienced qualitative researchers who will then develop a coding manual using ATLAS.ti software to assist in the management of data. The codes, reflecting the questions asked and concepts that arise during the discussions, will be used to build broader themes. This approach is based on a modification of Strauss and Corbin's grounded theory methodology. [26]

4.6. Ethical issues

Ethical issues relate to the risks and benefits of participating in a focus group. These usually include concerns about anonymity and protection of privacy and the potential harms associated with participation. The study will obtain ethics approval and oversight from the Human Research Ethics Committees of the ANU, Departments of Defence and Veteran's Affairs (DDVA), and Northern Territory Department of Health and Menzies School of Health Research (NTDoH and MSHR) and AMSANT.

Participants will be informed that they can withdraw while the discussion is taking place or choose not to answer any question that they perceive to be sensitive. Study investigators conducting the group discussions are trained and sensitive to potential issues. They will manage the group to avoid one person dominating the discussion and provide the opportunity for less forthright members to voice their opinions. If study participants become upset discussing the impact of living in an affected area on their lives more broadly, they will be referred to mental health support services recommended by the Primary Health Networks in the local area.

5. Component 2: Blood serum study

5.1. Objectives

The primary objective of this study will be to measure serum concentrations of PFAS in Williamtown, Oakey and Katherine residents of the Investigation Areas, and compare their levels to those of people residing in nearby non-contaminated areas, which will be sampled using a geostatistical framework. Through this, the study aims to determine whether there is evidence of elevated exposure to PFAS in the affected communities, and to further quantify the baseline levels of PFAS in the surrounding population through estimation of mean concentrations through individual testing.

5.2. Research questions

The specific research questions that the PFAS Health Study aims to answer in the Component 2: Blood serum study are:

1. Does the geographic distribution of blood PFAS levels correlate with known zones of contamination of groundwater and soil?
2. What are the mean serum concentrations of PFAS in Williamtown, Oakey and Katherine residents and how do these levels compare to those of people residing in non-contaminated areas?
3. How do serum concentrations vary by location and demographic factors, such as age, sex and length of residence, in the townships of Williamtown, Oakey and Katherine?
4. Do people who have voluntarily submitted blood serum under the Australian Government Department of Health's *Voluntary Blood Testing Program for PFAS* have higher PFAS concentration than those sampled randomly from the communities of Williamtown, Oakey and Katherine?
5. How do serum concentrations of PFAS in Williamtown, Oakey and Katherine residents correlate with other blood markers of disease risk, such as cholesterol and kidney function?

5.3. Hypotheses

We hypothesise that the PFAS levels in blood serum of residents of the Investigation Areas as defined by the Department of Defence will be higher than the PFAS levels in blood serum of residents in surrounding areas. Further, we would expect that the further away from the investigation area that a person resides, the lower the PFAS level will be in their blood serum.

5.4. Study design

The blood serum study is a cross-sectional study of residents in the three townships of Williamtown, Oakey and Katherine; particularly of people in the Investigation Areas of each township, along with people outside the Investigation Area. These people outside the Investigation Area may be residents of the township or wider rural area. The study will rely on testing for PFAS that is occurring among people who have lived in the Investigation Areas in the three towns under the Australian Government Department of Health's *Voluntary Blood Testing Program for PFAS*, as well as sampling that the study team will undertake. The blood serum study will link in with Component 3: The Cross-Sectional Survey that will examine potential exposure to PFAS along with health outcomes and concerns.

Factors associated with high levels within the communities would be investigated to better understand exposure risk, including individual's vicinity to the Royal Australian Airforce base in Williamtown, the Army Aviation Centre in Oakey and the Royal Australian Airforce base in Katherine. Results of the study would be presented in the form of a geospatial analysis, which would provide a visual representation of PFAS exposure within the communities, as determined by the geocoding of individual's serum concentrations of PFAS. The study will analyse the correlation between PFAS blood concentrations and other biomarkers of potential disease identified through a systematic review into health effects from PFAS that is currently being undertaken.

5.5. Study population

The study population is the residents of the Investigation Areas and surrounding areas. The 'exposed' group are those that lived or worked in the Investigation Areas of Williamtown, Oakey and Katherine, and the 'non-exposed' or reference population is a sample of households in areas surrounding the Investigation Areas. The 'non-exposed' in the Katherine area may require a reference population from another nearby community with similar demographic, as the Katherine Investigation Area encompasses the entire town south of the Katherine River.

5.6. Study procedure

A geostatistical sampling design will be used, which considers the expected range of spatial autocorrelation informed by the preliminary data from the initial voluntary blood testing results from the two communities. Using the cutting-edge geostatistical design—'grid plus close pairs design'—a primary grid will be overlaid on the study area (the three communities and surrounding areas) in Google Earth, with the household structures and their inhabitants lying in closest proximity to the grid nodes being selected. A secondary set of households located in near proximity to a random subset of those selected at the nodes of the grid (the close pairs) will also

be selected. This approach was shown by Diggle et al. [27] to be the most efficient survey design for estimating spatial variability in variables of interest. The grid is used to provide a statistically efficient sample that achieves optimal coverage of the study area and the close pairs are essential for estimating spatial autocorrelation parameters that are used in the generation of the risk maps.

5.7. Sample Size

The sample size will be limited by the number of individuals in the community Investigation Areas, but will likely be in the range 1,000 persons per study site for Williamtown and Oakey: 500 in the Investigation Areas and 500 outside the Investigation Area. For the Katherine Investigation Area the site encompasses a much larger population (~10,000 people), and thus could include up to 1,500 persons: 1,000 in the Investigation Area and 500 outside the Investigation Area. The study sample may include multiple members of households, making recruitment more feasible, but requiring adjustment for clustering in analysis. If numbers are small and/or response rates low we will consider increasing the number of individuals outside the Investigation Area to increase statistical power. We will use a geostatistical sampling design to enable estimation of spatial autocorrelation parameters, which involves sampling within specified geographic grids to select households. Multiple individuals from households will also be eligible to participate.

The sampling scheme will involve clustered (correlated) data and potential weighting to ensure that the sample correctly reflects the population of interest. This sampling strategy will increase the standard errors for estimates, thus reducing the 'effective' sample size. While the amount by which the sample size is impacted is unknown, a design effect of 1.2 is suggested. This means that the actual samples size should be divided by 1.2 to allow for the effect of the sampling scheme. For example, if the actual sample size is 500, the PFAS sampling design means that there will be the same precision and power as a sample of 500/1.2 ~= 415. To obtain the estimates of precision and power for this sample size, the estimates for a sample of 415, not 500 would be appropriate. Appendix 4 and 5 show estimates of precision for different sample sizes and prevalence.

A key consideration for the blood serum study is the detectable difference, also called the effect size in clinical trials, which refers to the difference that can be detected between groups with a specified sample size, significance level (α) and power ($1-\beta$). For an (effective) sample size of 400 participants per group, the study will have 80% power for a 5% significance level to detect a difference in mean PFAS (or log PFAS) of approximately 0.3 standard deviations for binary explanatory/exposure variables and correlation of 0.1 or more for continuous explanatory variables; and a difference between quartiles of PFAS of approximately 15% for binary explanatory/exposure variables and approximately 0.3 standard deviations for continuous

explanatory variables (Research Questions 1, 2, 3). For Research Question 4, this sample size will have 80% power, for a 5% significance level, to detect correlations between the blood markers of disease risk (e.g. cholesterol and kidney function) and PFAS level (or log PFAS) of 0.1 or more and differences in mean blood markers between quartiles of PFAS level of 0.3 standard deviations. Appendix 6 and Appendix 7 show the detectable difference in means and proportions respectively for varying sample sizes and prevalence.

5.8. Invitation to participate

Participants will be invited to participate when they voluntarily seek blood testing under the Australian Government Department of Health's *Voluntary Blood Testing Program* through their local general practitioner, or they will be sent an invitation to participate from the PFAS Health Study team. The study team will prepare letters of invitation addressed to the householder identified from the Electoral Roll, which will be followed up by a house visit or a phone call.

5.9. Blood collection

Blood samples collected under the Australian Government Department of Health's *Voluntary Blood Testing Program for PFAS* will be collected through residents' general practitioner. The consent can only be given in writing, on the Sonic Healthcare Australia pathology request form or a separate ANU consent form specific to the PFAS Health Study. Some participants may have had blood tested prior to the voluntary blood testing program. The Study team will seek consent from these people to include their results, and may request an additional test if appropriate. For participants randomly sampled from communities, the PFAS Health Study team will arrange for collection of blood specimens from the participant's home, or at a blood collection facility.

Each blood sample should be collected in 2mL aliquot tubes for the epidemiological study. One sample is required per participant. Each blood sample should be clearly labelled with the date of collection, time of collection, ID number, initials of the participant and the pathology barcode. The details written on each blood sample need to be compared to the details provided on the participant's pathology request form. The blood samples will be stored upright in 9 x 9 cryogenic boxes in a freezer set to – 80 degrees Celsius. Each cryogenic box should be labelled with ANU PFAS Health Study.

Sonic Healthcare Australia will test all blood samples for PFAS, cholesterol, creatinine and other tests, depending on the findings of the systematic review. Following testing, Sonic Pathology will ship specimens all blood serum samples to the ANU for future analyses.

5.10. Data analysis

All analyses will be undertaken separately for Williamtown, Oakey and Katherine communities, and will be adjusted for the sampling scheme, including correlation of outcomes within households, where appropriate.

The geospatial analysis aims to quantify environmental risk factors for elevated blood PFAS concentration, and identify any clusters of high PFAS exposure within the Williamtown, Oakey and Katherine communities, to potentially indicate the difference in risk from living and working in the Investigation Areas compared to living and working elsewhere. Households will be geo-referenced using a global positioning system. We will use the blood sample data in conjunction with the prospectively collected survey data to predict the distribution of blood concentrations of PFAS across the two communities and surrounding areas. We will apply a spatial prediction method, model-based geostatistics (MBG). [28] A key advantage of this study will be that we will have high-quality data collected for the specific purpose of undertaking a spatial analysis – often studies use data collected for a different purpose, resulting in inefficient spatial predictions with areas of sparse data and high prediction uncertainty.

For Research Question 1, the outcome will be maps of the distribution of PFAS blood concentration and associated uncertainty, and quantification of environmental and demographic factors associated with PFAS exposure. Note that no maps showing individual households in a way that will allow identification of study participants will be published or shared outside the study team. In addition, the results will allow the analyses of correlation between PFAS blood concentrations and other potential biomarkers of disease, along with demographic and exposure information from the cross sectional survey.

PFAS blood concentration level is likely to be highly right skewed. While preliminary / exploratory data analyses will use actual PFAS level (potentially log transformed), primary analyses will involve comparison of characteristics (or outcomes) by quantiles of PFAS (quintiles or quartiles, depending on the distribution of PFAS and sample sizes), with contrast between lowest versus highest PFAS quantile of particular interest.

Analyses for Research Questions 2–4 will involve PFAS as the outcome of interest. We will undertake multivariable regression (linear for log transformed PFAS value or multinomial for PFAS quantile, with the lowest quantile as the reference group) to examine the relationship between characteristics of interest (e.g. age; sex; location of residence inside or outside of the Investigation Area; participation in the *Voluntary Blood Testing Program for PFAS* versus randomly selected for the study) and PFAS blood level / quantile. Location or residence or proximity to Investigation Area site will be informed by contamination site mapping and the geo-spatial analysis.

For Research Question 5, PFAS is the exposure of interest, and outcomes are the relevant blood markers of disease risk, such as cholesterol and kidney function. We will examine the association between PFAS and outcomes using non-parametric correlation and multiple linear regression, using both actual PFAS blood level (likely log transformed) and PFAS quantile, adjusted for age and sex.

5.11. Ethical issues

Privacy and confidentiality

All participants will be informed about the nature of the blood testing study, the risks of participation, and the options for non-completion and include a statement of consent for collection of the blood specimen, inclusion of data for blood testing analysis, and storage of the blood serum.

The data will be stored on secure servers at the ANU with access restricted to authorized personnel. All personnel associated with the survey, including market research company staff, will sign a Confidentiality Deed.

Ethics

The PFAS Health Study: Blood serum study will be submitted to the ANU, DDVA, NTDoH and MSHR and AMSANT Human Health Ethics Committees for approval.

6. Component 3: Cross-sectional survey

6.1. Introduction

This third component study is a cross sectional survey of residents in the townships of Williamtown, NSW, Oakey, Qld and Katherine, NT. In the survey, we will assess three main outcomes of (1) self-reported associated health problems, (2) levels of psychological distress and concern, and (3) PFAS levels arising from the blood serum study in Component 2. We will explore associations with three different main exposures: (1) whether a person lived or worked in an Investigation Area, (2) assessment of intake of possible sources of PFAS from food, water and the environment, and for health problems and distress, (3) blood testing results arising from Component 2.

The survey will ascertain health-related concerns and likely levels of exposure to PFAS. People living in Investigation Areas and nearby uncontaminated areas will be surveyed at or after the time of the blood collection process for Component Study 2. The survey will include the collection of demographic data; current and historical residential and employment information; water source; consumption of locally-caught fish and home grown produce; other potential sources of PFAS contamination; information about current and past health conditions, including reproductive history; and lifestyle and health behaviours.

The cross sectional survey will rely on a questionnaire derived from the information provided by residents who have participated in the focus groups, along with literature on adverse effects of PFAS and other relevant survey instruments. The research team will refer to the Department of Defence Water Usage and Community Survey questionnaires for both Williamtown, Oakey, Katherine, the C-8 Health Project Baseline Questionnaire, the Airservices Australia's Aviation Rescue and Fire Fighting staff questionnaire and the Swedish Kallinge Study for potential questions to be include in the survey.

6.2. Objectives

The objective of this cross sectional survey are to:

1. Identify likely exposures over time from consumption of water and food, and potential exposure to dust and other potential sources of PFAS of survey participants and development of exposure metric, informed by geo-spatial analysis of PFAS blood serum in Component 2 and by international exposure assessment tools (questionnaires)
2. Assess health concerns and self-reported health outcomes of residents and people who have lived or worked in the Investigation Areas and compare them with people who have

not lived or worked in the Investigation Areas in Williamtown, NSW, Oakey, Qld and Katherine, NT.

3. Measure levels of psychological distress in residents and workers and compare these to those not living or working in the Investigation Areas
4. Expanding the analysis of risk factors associated with PFAS blood serum levels to incorporate additional information from the survey, including sociodemographic (e.g. age, sex, location) and other factors (e.g. duration of residence in the area, water source) within the Williamtown, Oakey and Katherine communities; both in and out of the Investigation Areas.

6.3. Research questions

The specific research questions that this cross-sectional survey will address are:

1. What are the main potential sources of exposure to PFAS through occupation, food, waters, or other factors In Williamtown, NSW, Oakey, Qld or Katherine, NT?
2. What are the main concerns regarding health problems associated with living or working in the Investigation Area in Williamtown, NSW, Oakey, Qld or Katherine, NT?
3. What are the main self-reported health outcomes associated with living in or working in the Investigation Area in Williamtown, NSW, Oakey, Qld or Katherine, NT?
4. What are the current levels of psychological distress and how do these relate to PFAS blood results and location of residence or work?
5. What are the main risk factors for higher than background level serum PFAS concentration regarding sociodemographic and other factors?

6.4. Hypotheses

We hypothesise that residents of the Investigation Areas will have higher levels of exposure to PFAS. In addition, we hypothesise that residents of the Investigation Areas will have higher levels of concerns, psychological distress and self-reported health outcomes than residents living in surrounding areas. Further, we would expect that working with PFAS, living in the Investigation Area, eating fresh produce and drinking water would be risk factors for high PFAS blood serum level.

6.5. Study design

This cross-sectional survey will provide information about the health experience and potential exposure to PFAS of residents of Investigation Areas in Williamtown, Oakey and Katherine, as well as workers on the Defence Bases in those areas, both past and present. Additionally, a

sample of residents living in nearby uncontaminated areas will also be surveyed for comparison of health outcome measures and potential exposure to PFAS.

6.6. Study population

The study population is all current and former residents in the Investigation Areas, and all current and former Defence Personnel who were exposed to PFAS chemicals while working at either the Royal Australian Airforce Base at Williamtown, the Army Aviation Centre at Oakey or the Royal Australian Airforce Base at Katherine along with residents sampled from surrounding areas. There is a potential for a mismatch in time between surveying people in Investigation Areas, compared to those outside. We will attempt to take this into account in analysis. There are approximately 2,500 households containing 15,000 people in the Investigation Areas at the three communities.

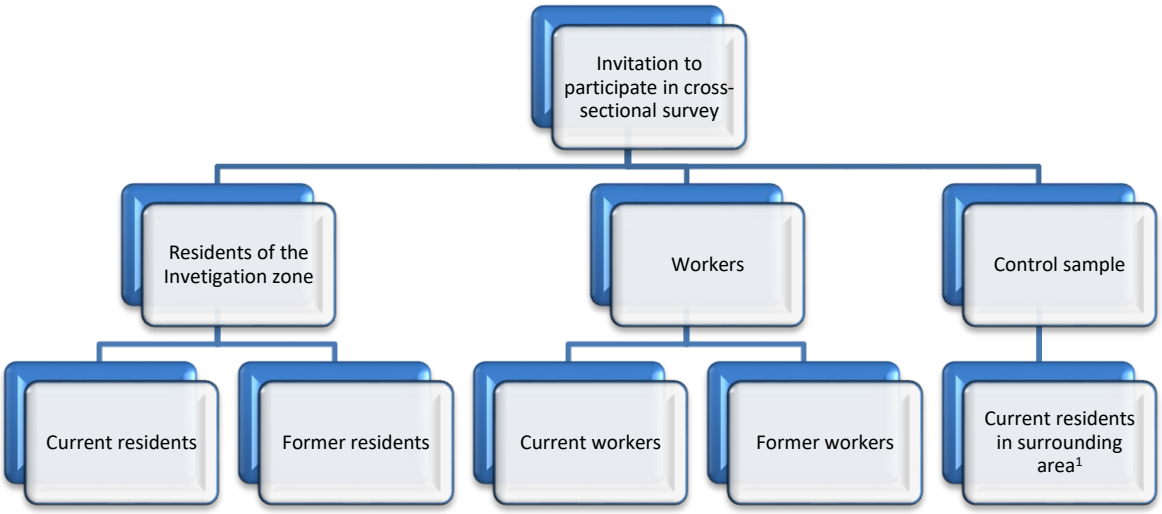
6.7. Reference population

The reference population for the cross-sectional survey is residents of Williamtown, NSW, Oakey, Qld and Katherine, NT, and areas surrounding them. We will also compare results from the survey for specific health outcomes (e.g. self-assessed overall health, psychological distress) to national data collected by the Australiana Bureau of Statistics, such as the National Health Survey.

6.8. Study procedure

A schematic overview of the study is shown in Figure 1, which is explained in more detail in the sections below.

Figure 1. Schema showing the conduct of the cross-sectional survey



[1] This group will be asked to complete the survey at the time of blood sample collection, either online, or face-to-face using a mobile data collection, or on a paper questionnaire. Residents and workers may also complete the survey at the time of the blood collection, if they have not taken part in the Voluntary Blood Testing Program prior to the survey being launched.

6.9. Invitation to participate

To recruit people to the cross-sectional survey, the study team will send out invitations informing them that they have been selected for the study. The sample selection will occur in conjunction with Component 2: blood serum study to ensure that data can be used in both studies. PFAS Health Study staff will collect the data either face-to-face at the time of blood collection or over the phone using an online database. Participants who have provided blood under the Australian Government Department of Health's Voluntary Blood Testing Program for PFAS will be contacted by PFAS Health Study staff to complete the questionnaire, either over the phone, online or by paper. Invitees will be asked to complete the survey whether or not they have given or will give blood for PFAS testing.

6.10. Sample size

For an (effective) sample size of 400 participants per group will have 80% power, for a 5% significance level, to detect: 1) for health continuous outcomes - differences in mean outcomes between quartiles of PFAS level of 0.3 standard deviations; 2) for binary health outcomes – differences in between quartiles of PFAS of approximately 15% for binary explanatory / exposure variables and approximately 0.3 standard deviations for continuous explanatory variables Appendix 6 and Appendix 7 show the detectable difference in means and proportions

respectively for varying sample sizes and prevalence (Research Objectives 2, 3, 4); 3) for PFAS quantile outcome - differences between quartiles of PFAS of approximately 15% for binary explanatory/exposure variables and approximately 0.3 standard deviations for continuous explanatory variables.

The sample size will be determined by those participating in Component 2: Blood Serum Study, which is approximately 1,000 people in each for Williamtown and Oakey (500 in the Investigation Areas and 500 outside) and approximately 1,500 people for Katherine (1,000 in the Investigation Area and 500 outside). A key consideration for the cross-sectional study is the detectable difference, also called the effect size in clinical trials, which refers to the difference in outcomes that can be detected between groups with a specified sample size, significance level (α) and power ($1-\beta$). Detectable differences for the cross-sectional study have been estimated assuming a significance level of 0.05 (5%) and a power of 0.8 (80%), and comparison of two groups, with groups defined as quantiles (e.g. quintiles or quartiles) of PFAS level, or between those in the PFAS Investigation Area versus not in the PFAS Investigation Area. For comparison of quantiles of PFAS level, the most powerful comparison is expected to be between the highest versus lowest quantiles. The sample sizes in the figures and tables are the total number of (effective) participants in two groups. If the highest and lowest quintiles of PFAS level are compared, the (effective) sample size presented is the sum of the numbers in the highest and lowest quartile, i.e. 2/5 or 0.4 of the total sample so that the total sample size of exposed individuals will be the (effective) displayed sample size divided by 0.4 (or equivalently multiplied by 2.5).

Figure 2 shows the detectable difference in means between groups, for varying (effective) sample sizes, in terms of standard deviations. As an example, if comparing highest and lowest quintiles of PFAS level with a total (effective) sample size of 200 (100 in each quintile, or 500 in total), the study would have 80% power, with a 5% significance level, to detect a difference in continuous outcome (e.g. cholesterol level) of 0.4 standard deviations. Alternatively if there was an (effective) sample size of 400 participants in the PFAS exposure zone and 400 comparison individuals outside of the exposure zone (i.e. a total (effective) sample size of 800), the study would have 80% power, with a 5% significance level, to detect a difference in continuous outcome (e.g. cholesterol level) of 0.2 standard deviations.

A table of the details of the precision for varying sample sizes is provided in Appendix 6.

Figure 2: Detectable difference in mean outcome between groups

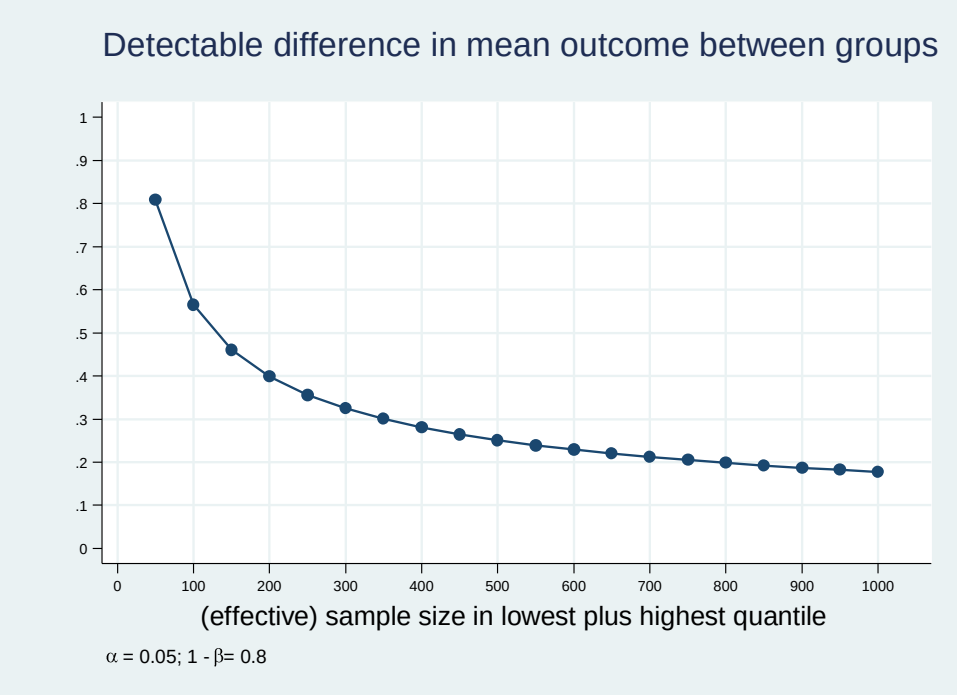
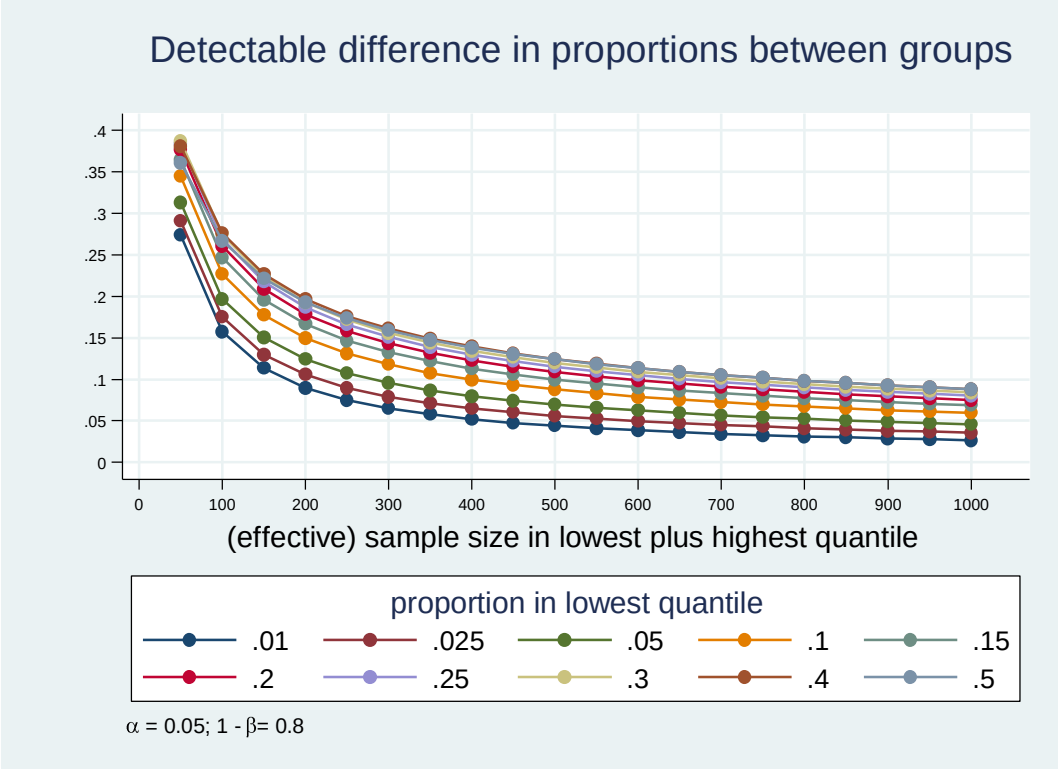


Figure 3 shows the detectable difference in proportions between groups, for a range of proportions in the lowest risk quantile (or group) and varying (effective) sample sizes. As an example, if comparing the prevalence of an outcome between the highest and lowest quintiles of PFAS level with a prevalence of the outcome in the quintile with the lowest risk of 0.01 (1%), a total (effective) sample size of 200 (100 in each quintile, or 500 in total), the study would have 80% power, with a 5% significance level, to detect a difference in prevalence of 0.05 (5%); i.e. a prevalence of 0.01 + 0.05 or 0.06 (6%) in the comparison quintile. Alternatively if there was an (effective) sample size of 400 participants in the PFAS exposure zone and 400 comparison individuals outside of the exposure zone (i.e. a total (effective) sample size of 800), the study would have 80% power, with a 5% significance level, to detect a difference in proportions of approximately 10%, assuming a prevalence in the group with the lowest risk of approximately 40%.

A table of the details of the precision for varying proportions and sample sizes is provided in Appendix 5.

Figure 3: Detectable difference in proportions between groups



6.11. Respondent participation

Response rates to surveys have been declining over time throughout the world. In this instance, we expect the response rate to be relatively high due to the relevance of the survey to participants. We estimate that the response rate to the initial approach will be around 30%.

To maximize participation in the survey, we will:

- Employ invitation letters followed by face-to-face visits of PFAS Health Study staff to boost participation in the survey.
- Promote the survey through the Community Walk in Sessions, run by the Department of Defence, at both Williamtown, Oakey and Katherine.
- Prepare media releases indicating that the survey is underway when the invitations are sent out. These releases may be issued both in print and through local radio stations.
- Study team members will also visit both communities in the lead-up to the survey to provide feedback on the Focus Group Discussions as well as provide information about the cross-sectional survey.
- Contact community reference groups and other important local stakeholders to encourage participation.
- Study team members will speak to legal firms representing local residents to request that they encourage clients to participate in the survey.
- Reminders will be sent via post/email to potential participants.

6.12. Survey conduct

The survey will employ online data collection using online forms and mobile data collection using tablets at the time of blood collection. Members of the study sample from the surrounding areas will be given either a paper questionnaire to complete, or PFAS Health Study staff will collect data online using an online form at the time of blood sample collection for the Blood Serum Study.

6.13. Survey instrument

The survey will collect a range of information for the previous 10 years, including:

- Residence and address history, with respect to the Investigation Area
- Exposure (water supply to the property, use of bores and bore water, consumption of locally caught fish, local beef and other meats, and home grown produce, occupational)
- Health conditions, including a reproductive history
- Lifestyle and health behaviours
- Level of distress Kessler 10 Psychological Distress scale (K-10)/Distress Questionnaire (DQ-5) plus Patient Health Questionnaire (PHQ-2) and Generalised Anxiety Disorder scale (GAD-2)
- Perceived economic impact for the household
- Perception of dissemination of health information by the Australian Government (risk communication)

- Demographics characteristics (age, sex, Indigenous status, marital status, country of birth, language spoken at home, educational attainment, occupation, children's age, children's sex, household income, number of residents in household)
- Whether participants are happy to be contacted in the future for further health studies relating to exposure to PFAS, including data linkage studies.

The questionnaire will be developed based on the literature and findings from the Component 1 Focus Groups Study. The questionnaire will be tested among study team members to ensure that the questionnaire is functional and easily understood. Following this, the survey will be piloted using 10 randomly selected households from within each investigation area. Where possible, relevant questions will be identical to those used in other surveys of the reference population (i.e. ABS National Health Survey, C-8 Community Follow-up Study, Department of Defence Water Use and Community Surveys for Williamtown, Oakey and Katherine residents) to enable comparison of results.

6.14. Data analysis

All analyses will be undertaken separately for Williamtown, Oakey and Katherine communities, and will be adjusted for the sampling scheme, including correlation of outcomes within households, where appropriate. We refer to quantiles of PFAS, with the particular type of quantile (e.g. quintile, quartile, etc.) to be determined based on the distribution of PFAS level and the samples size (to ensure an appropriate number of observations in each quantile).

An exposure metric will be developed, using data obtained on a range of exposures including area of residence, consumption of water and food, and exposure to dust and other potential sources of PFAS exposure of survey participants, informed by geo-spatial analysis of PFAS blood serum in Component 2 and by international exposure assessment tools. Exposures included in the exposure metric will be informed by the qualitative study. There will then be three different measures of PFAs 'exposure': area of residence (living inside or outside of the Investigation Area); quantile of blood PFAS level; and exposure based on the developed exposure metric.

For Research Objectives 2, 3, and 4 multiple regression (linear or logistic as appropriate) will be undertaken to examine the relationship between outcomes (health concerns, self-reported health outcomes and psychological distress) and exposures, adjusted for relevant covariates (e.g. sex, age). Different models will be generated for each of the three different definitions/measures of exposure described above.

Analyses for Research Objective 5 will consider PFAS exposure as the outcome variable, and expand on the analyses undertaken for Research Question 2 of the Blood Serum Study to include additional risk factors for FPAS quantile obtained from the cross sectional survey data. We will

undertake multivariable multinomial regression of PFAS quantile (with the lowest quantile as the reference group) to examine the relationship between characteristics of interest (e.g. age; sex; location of residence inside or outside of the Investigation Area; and PFAS blood level quantile. Location or residence or proximity to the Investigation Area will be informed by contamination site mapping and the geo-spatial analysis.

6.15. Ethical issues

Privacy and confidentiality

The first page of the survey will outline the nature of the survey, the risks of participation, and the options for non-completion and include a statement of consent for collection of information.

At the end of the survey, the respondent will be asked if they are willing to participate in future research. If they indicate yes, they will be asked to provide their full name, Medicare, Defence or DVA) number, date of birth, and contact details. This will be used in future studies of the association between exposure to PFAS and disease (subject to appropriate approvals).

The data will be stored on secure servers at the ANU with access restricted to authorized personnel. All personnel associated with the survey, including staff collecting data in the field, will sign a Confidentiality Deed.

Ethics

The PFAS Health Study: Cross-sectional survey will be submitted to the ANU, DDVA, and NTDoH and MSHR and AMSANT Human Health Ethics Committees for approval.

7. Component 4: Data linkage study

7.1. Aims and objectives

The aim of this study is to examine whether sex-specific age adjusted rates of health outcomes that are known to be associated, or possibly associated, with PFAS exposure are higher among people who have lived in Investigation Areas of Williamtown, Oakey and Katherine compared with those who have in lived in similar but uncontaminated areas and in the general population of the relevant State.

The specific objectives of the study are:

1. From the systematic review of the epidemiological literature, compile a list of outcomes that have been associated with PFAS exposure, hereafter referred to as candidate outcomes, and that can be measured with administrative data.
2. Estimate the relative rates of candidate outcomes in relation to community PFAS exposure. By relative rates we mean rates of candidate outcomes in those who have lived in Investigation Areas compared with those who have not lived an Investigation Areas defined in several different ways (e.g. lived in surrounding non-Investigation Areas; lived in comparable areas of the same State as defined by socioeconomic status (SES) and remoteness; lived anywhere in the same State and adjusted for socioeconomic status and remoteness in addition to age and sex).
3. Estimate relative rates for control conditions, i.e., selected common outcomes not known or thought to be associated with PFAS, and compare these to the corresponding relative rates for candidate outcomes.

7.2. Hypotheses

If living in the Investigation Areas where there is PFAS contamination (Williamtown, Oakey and Katherine) increases the risk of candidate diseases, after taking into account differences in age and sex between residents in the Investigation Areas and the comparison populations (Investigation Areas see objective 2 above) we would expect:

- Rates of candidate outcomes to be higher in residents who have lived in contaminated areas (relative rates > 1)
- Rates of control conditions to be no higher in residents who have lived in contaminated areas (relative rates=1).

7.3. Study design

The study design proposed is a cohort study, involving linking Medicare, Defence and DVA registration data to routinely collected health data, including hospital, cancer and death data.

7.4. Study population

The study population will include all people registered with Medicare, Defence or DVA sometime between 1984 and 2015 [1] who had been registered with an address in the Investigation Areas of Williamtown, Oakey or Katherine (exposed group) at any time during that period and a random sample of people on the register with addresses outside these areas (the unexposed) defined by having:

- lived in surrounding non-Investigation Areas;
- lived in comparable areas of the same State as defined by socioeconomic status (SES) and remoteness; and
- lived anywhere in the same State and adjusted for socioeconomic status and remoteness in addition to age and sex.

The unexposed will be sampled from addresses in areas with similar area characteristics with respect to rurality (based on the ARIA+ scores) and socioeconomic profile (based on SEIFA score).

7.5. Study procedure

A final list of outcomes to be investigated will be based on the systematic review carried out in Phase I of the Study. The list will be as inclusive as possible, including all health outcomes that have be shown to be linked or possibly linked to PFAS in humans, and that can be identified though routinely collected health data collections (see Table 1 for examples).

Table 1. Examples of potential health outcomes for inclusion in PFAS data linkage study

Heath outcome	**Source**
Kidney cancer [ICD code C64]	Australian cancer database
Testicular cancer [ICD code C62]	Australian cancer database
Ulcerative colitis [ICD code K51]	Admitted patient data collections
Low birth weight	National perinatal database

Medicare, Defence and Department of Veterans' Affairs register data, which include name, sex, date of birth and addresses of people registered, are held by the Australian Institute of Health and Welfare (AIHW), and will be used to identify the study population. Using these data will allow estimation of duration of residence in different places, including Investigation Areas, although

[1] Note that while the population of interest is anyone who has resided in the investigation areas since PFAS were in regular use on the Defence bases, i.e. since the 1970's, there are no databases dating back to that time that capture names and addresses of all residents (both adults and children), nor any from which diseases of relevance might be identified before 1982 or deaths from diseases of relevance easily identified before 1980.

Medicare start and end dates may be unreliable. The Study Team may also need to consider linking other data, such as electoral roll or Defence Force medical records, to improve estimation of duration of residence or exposure.

Health outcomes could be ascertained from the following data collections:

- Australian Cancer Database (ACD), held by the AIHW
- Admitted Patient Data Collections (APDC), including the NSW APDC and Queensland Hospital APDC (QHAPDC), held by respective state health departments and data linkage centres;
- National Perinatal Data Collection (NPDC), held at the AIHW.

The National Death Index (NDI), held at AIHW, will be used to identify deaths for censoring in the analysis. The NDI contains records of all deaths occurring in Australia since 1980. The data are provided by the Registries of Births, Deaths and Marriages, the Australian Bureau of Statistics and the National Coroners Information System. All data are available for health research subject to ethics clearances and data custodian approvals. (Further information on the National Death Index, including data request procedures, can be found at the AIHW National Death Index.

Formal guidelines for integrating Commonwealth data for research projects have been established by the National Statistical Service. The Data Integration Services Centre at the AIHW, which is a Commonwealth-accredited data integration authority, is likely to be the appointed authority for this project. They would facilitate access to the data, perform the linkage and provide secure storage of the data.

Data will be linked probabilistically based on relevant variables in the various datasets, including full name, sex, date of birth and address. Importantly, a separation principle is in place. The separation principle means that no one working with the data can view both the linking (identifying) information (such as name, address, or date of birth) together with the merged analysis (content) data (that is, health information) in an integrated dataset.

7.6. Data analysis

For each outcome, we will use indirect standardisation to generate standardised incidence ratios (SIRs) and 95% confidence intervals. The indirect approach will be used because of the likely small number of events for each outcome. To do this, we will first calculate the age-sex- period specific rates (number of diagnoses/person years) for each outcome in the unexposed. We will apply these rates to the exposed to generate the expected number of cases. The SIR is the total number of observed cases in the exposed divided by the expected number in the exposed. An SIR>1 means rates are higher in the exposed than the unexposed, an SIR <1 means rates are lower in the exposed than the unexposed, and a SIR equal to 1 means there is no difference in rates

between the exposed and unexposed. Rates in the exposed will also be compared to rates in the general population, by generating SIRs using age-sex rates for the general population as the standard rates, which will be based on published data (e.g. use AIHW data cubes for numerators and ABS population data for denominators to generate rates). The caveat to this approach is that there are likely to be other factors that differ between the exposed group and the reference populations that will result in residual confounding of the standardised incidence ratios, which cannot be quantified.

7.7. Ethical issues

Ethics applications

While the AIHW acts as a custodian of state and territory registry data for the purposes of producing national cancer statistics, cancer registries retain ownership of their jurisdiction's data at all times. Thus, multiple ethics and data custodian approvals are required, from the different jurisdictions, as well as from the institutions involved in the research. Ethics approvals will be sought from:

1. ANU Human Research Ethics Committee
2. AIHW Ethics Committee
3. ACT Human Research Ethics Committee
4. Departments of Defence and Veteran's Affairs Human Research Ethics Committee
5. Northern Territory Department of Health and Menzies School of Health Research Human Research Ethics Committee
6. NSW Population and Health Services Research Ethics Committee
7. SA Health Human Research Ethics Committee
8. Human Research Ethics Committee (Tasmania) Network
9. WA Department of Health Human Research Ethics Committee[2]

Data custodian approvals will be sought from state and territories as required for cancer and other data.

Privacy and waiver of consent issues

The study is compliant with all Australian Privacy Principles (APP) except APP6 (use or disclosure of personal information). As this project is to be conducted without consent, which would breach APP6, a waiver of consent pursuant to section 95 of the Privacy Act 1988 will be sought on the basis of the large number of people involved, the lack of current address information for most of them, the high degree of privacy protection afforded by application of the separation principle

[2] Note institutional ethics committee approvals are sufficient to access QLD, Vic and NT data.

and additional measures relating to data access and use that are aimed at minimisation of linked data records' re-identification risk.

Secure data management

The study team will adhere to strict guidelines to ensure security of data. Individual-level identifying information, required to link data, will be performed by the AIHW, a Commonwealth-accredited data integration authority. Details on the accreditation criteria, which have been fulfilled by the AIHW, can be found on the Australian Government National Statistical Service website:

http://nss.gov.au/nss/home.nsf/pages/Data%20Integration%20-%20AIHW%20accreditation%20application%20and%20audit%20summary/#CIV

Only de-identified data will be available to the researchers at the ANU who will be carrying out the analyses. These data will be made available through the Secure Unified Research Environment (SURE). SURE is a high-powered computing environment, which is a remote-access data research laboratory for analysing routinely collected health data. It is allows researchers to log in remotely and securely to analyse data from sources such as cancer registries and death registries. SURE was developed by the Sax Institute, as part of the Population Health Research Network. SURE is accessed via AARNET (the Australian Academic and Research Network) or the internet using an encrypted connection from researchers' local computers, which must meet security requirements. All users must complete training on privacy, ethics, information security and statistical disclosure control and sign a deed that sets out the terms and conditions for using SURE. Further details can be found at the Sax Institute

8. Quality assurance, monitoring and safety

The ANU is proactive and responsible in its approach to risk management. The Research Office within the College of Medicine Biology and the Environment oversees all population health research within ANU. The Research Office oversees the application of research proposals and financial accountability for the conduct of research. The ANU has human and animal research ethics committees that function in accordance with National Health and Medical Research guidance. The Research Office ensures that all funded research is approved by the appropriate ethics committee and complies with University policies.

The study will have ethics approval from the Human Research Ethics Committee at the ANU and all other relevant committees. This Committee will ensure that research is conducted according to the National Health and Medical Research Council's National Statement on Ethical Conduct in Human Research.

9. Governance

The Chief Investigator, Associate Professor Martyn Kirk, will oversee all aspects of the PFAS Health Study, in close consultation with other Chief Investigators. The PFAS Health Study team includes national and international experts in population health and will ensure that the study is of high quality. The PFAS Health Study team will include 4–6 representatives from the Williamtown, Oakey and Katherine communities to ensure engagement and acceptability to the community.

The PFAS Health Study team will provide monthly reports on progress and issues to the Department of Health to ensure that the study is meeting proposed timelines.

10. Timelines

The PFAS Health Study will be conducted over a three and a half year period between June 2017 and December 2020. The timeline has been modified to reflect the addition of the Katherine study area, but this has not affected the completion date of the Phase II study. The inclusion of Katherine has extended the length of the Blood Serum Study and Cross-sectional Survey by six months, with delivery of both reports due in June 2020. Delivery of the Data Linkage Study is unaffected and due for delivery in December 2020.

It should be noted that approximately 22% of the Katherine community identify as Aboriginal and Torres Strait Islander as per the 2016 census. This inclusion of Aboriginal and Torres Strait Islander participants in the PFAS Health Study requires additional ethics components (PartD.) for submission to the Human Research Ethics Committee of Northern Territory Department of Health and Menzies School of Health Research (EC 00153).

Part D. requires the PFAS Health Study ethics submission to address how the study will benefit or, have an impact on Aboriginal and Torres Strait Islander and their communities. Furthermore, letters of support are required from relevant community authorities stating they are aware of the aims and methods of the proposed research. These additional aspects may require potential partnerships with Aboriginal Medical/Health Services and could take time to establish.

11. Finance and resources

The study will be funded by the Australian Government Department of Health.

12. Dissemination of results and publication policy

The results from each component of the PFAS Health Study will be detailed in a report provided to the Australian Government Department of Health. The study team will also prepare articles for publication in peer-reviewed journals. All reports and publications will acknowledge funding from the Department and input from the community and other experts. Authorship of peer reviewed articles will be determined in accordance with the Uniform Requirements for Manuscripts Submitted to Biomedical Journals produced and endorsed by the International Committee of Medical Journal Editors, specifically those considerations set out under heading IIA 'Ethical Considerations in the Conduct and Reporting of Research: Authorship and Contributorship'.

Results will be summarised in a format suitable for lay-people. The form and timing of the communication will be determined by the Study team in consultation with the Department. Findings will be communicated to national and international media through a coordinated media release between the Australian National University and the Australian Government Department of Health.

The report for each component of the study will be made available to community members via the study webpage. Additionally, after the acceptance of each component report the study team will hold a community consultation in each of the two communities at which time the findings of each report will be presented, and the next stages of the study will be reported.

The first report will be due in April 2018, this will present the results of the focus group discussions from Williamtown and Oakey, and will give feedback to the Department and the communities on the issues that were raised by participants. Community consultations will be held in both townships in May 2018. These community consultations will be particularly important, as people who did not attend the focus group discussions will be able to provide feedback about other relevant issues not picked up during the group discussions. At this time the study team will be able to discuss the systematic review, the types of questions that will be included in the Cross-sectional Survey and the timing of both the Cross-sectional Survey and the Blood Serum Study. A separate focus group discussion report will be delivered for Katherine to the Department in August 2018.

The second and third reports will be finalised by June 2020 and it will present the findings of the Cross-sectional Survey and Blood Serum Study. Community consultations will be held in each township, providing details of the findings for both studies. At this time the study team will also provide a progress report on the final study—the Data Linkage Study.

The final report will be delivered December 2020. Final community consultations will be held in each township when this report has been accepted by the Department. These final consultations will provide the details of the last report and provide an overall picture of the findings from the entire study. Study team members will also discuss potential future studies.

13. Status of these protocols

The protocols in this report represent the PFAS Health Study team plans and may not reflect final study protocols. Study conduct may change over time depending on various factors, such as feasibility, availability of data, ethical requirements, and the availability of new information. Prior to the commencement of all studies, the Study team will prepare final protocols to guide study conduct. These final protocols will be made available to community members via the study webpage and communicated through community consultations.

14. References

1. Kissa E: **Fluorinated surfactants and repellents**, vol. 97: CRC Press; 2001.
2. Buck RC, Franklin J, Berger U, Conder JM, Cousins IT, de Voogt P, Jensen AA, Kannan K, Mabury SA, van Leeuwen SPJ: **Perfluoroalkyl and polyfluoroalkyl substances in the environment: Terminology, classification, and origins**. *Integrated Environmental Assessment and Management* 2011, **7**(4):513-541.
3. Grandjean P, Clapp R: **Perfluorinated Alkyl Substances: Emerging Insights Into Health Risks**. *New Solut* 2015, **25**(2):147-163.
4. Kannan K, Tao L, Sinclair E, Pastva SD, Jude DJ, Giesy JP: **Perfluorinated Compounds in Aquatic Organisms at Various Trophic Levels in a Great Lakes Food Chain**. *Archives of Environmental Contamination and Toxicology* 2005, **48**(4):559-566.
5. Trudel D, Horowitz L, Wormuth M, Scheringer M, Cousins IT, Hungerbuhler K: **Estimating consumer exposure to PFOS and PFOA**. *Risk Anal* 2008, **28**(2):251-269.
6. Fromme H, Tittlemier SA, Volkel W, Wilhelm M, Twardella D: **Perfluorinated compounds--exposure assessment for the general population in Western countries**. *Int J Hyg Environ Health* 2009, **212**(3):239-270.
7. Steenland K, Fletcher T, Savitz DA: **Epidemiologic evidence on the health effects of perfluorooctanoic acid (PFOA)**. *Environmental Health Perspectives* 2010, **118**(8):1100-1108.
8. Olsen GW, Burris JM, Ehresman DJ, Froehlich JW, Seacat AM, Butenhoff JL, Zobel LR: **Half-life of serum elimination of perfluorooctanesulfonate, perfluorohexanesulfonate, and perfluorooctanoate in retired fluorochemical production workers**. *Environmental health perspectives* 2007:1298-1305.
9. Faniband M, Lindh CH, Jönsson BA: **Human biological monitoring of suspected endocrine-disrupting compounds**. *Asian journal of andrology* 2014, **16**(1):5.
10. Chen MH, Ha EH, Wen TW, Su YN, Lien GW, Chen CY, Chen PC, Hsieh WS: **Perfluorinated compounds in umbilical cord blood and adverse birth outcomes**. *PloS one* 2012, **7**(8):e42474.
11. United States Department of Health and Human Services: **Toxicological profile for perfluoralkyls**. In.: Agency for Toxic Substances and Disease Registry; 2015: 574.
12. Smit LA, Lenters V, Hoyer BB, Lindh CH, Pedersen HS, Liermontova I, Jonsson BA, Piersma AH, Bonde JP, Toft G *et al*: **Prenatal exposure to environmental chemical contaminants and asthma and eczema in school-age children**. *Allergy* 2015, **70**(6):653-660.
13. Bach CC, Bech BH, Nohr EA, Olsen J, Matthiesen NB, Bonefeld-Jorgensen EC, Bossi R, Henriksen TB: **Perfluoroalkyl Acids in Maternal Serum and Indices of Fetal Growth: The Aarhus Birth Cohort**. *Environ Health Perspect* 2016, **124**(6):848-854.
14. Stein CR, Savitz DA, Dougan M: **Serum levels of perfluorooctanoic acid and perfluorooctane sulfonate and pregnancy outcome**. *Am J Epidemiol* 2009, **170**(7):837-846.
15. Louis GM, Chen Z, Schisterman EF, Kim S, Sweeney AM, Sundaram R, Lynch CD, Gore-Langton RE, Barr DB: **Perfluorochemicals and human semen quality: the LIFE study**. *Environ Health Perspect* 2015, **123**(1):57-63.
16. Lum KJ, Sundaram R, Barr DB, Louis TA, Buck Louis GM: **Perfluoroalkyl Chemicals, Menstrual Cycle Length, and Fecundity: Findings from a Prospective Pregnancy Study**. *Epidemiology* 2017, **28**(1):90-98.
17. Tsai MS, Lin CY, Lin CC, Chen MH, Hsu SH, Chien KL, Sung FC, Chen PC, Su TC: **Association between perfluoroalkyl substances and reproductive hormones in adolescents and young adults**. *International journal of hygiene and environmental health* 2015, **218**(5):437-443.

18. Steenland K, Woskie S: **Cohort mortality study of workers exposed to perfluorooctanoic acid**. *Am J Epidemiol* 2012, **176**(10):909-917.
19. Su TC, Kuo CC, Hwang JJ, Lien GW, Chen MF, Chen PC: **Serum perfluorinated chemicals, glucose homeostasis and the risk of diabetes in working-aged Taiwanese adults**. *Environ Int* 2016, **88**:15-22.
20. Shankar A, Xiao J, Ducatman A: **Perfluorooctanoic acid and cardiovascular disease in US adults**. *Archives of internal medicine* 2012, **172**(18):1397-1403.
21. Steenland K, Zhao L, Winquist A, Parks C: **Ulcerative colitis and perfluorooctanoic acid (PFOA) in a highly exposed population of community residents and workers in the mid-Ohio valley**. *Environ Health Perspect* 2013, **121**(8):900-905.
22. International Agency for Research on Cancer: **Perfluorooctanoic Acid**. In: *IARC Monographs on the evaluation of carcinogenic risks to humans.* vol. 110: World Health Organization; 2016.
23. Chang ET, Adami HO, Boffetta P, Cole P, Starr TB, Mandel JS: **A critical review of perfluorooctanoate and perfluorooctanesulfonate exposure and cancer risk in humans**. *Critical reviews in toxicology* 2014, **44 Suppl 1**:1-81.
24. AECOM Services Pty Ltd: **Stage 2C Environmental Investiagtion - Preliminary Ecological Risk Assessment, Army Aviation Centre Oakey**. In. Edited by Department of Defence; 2016: 2.
25. AECOM Services Pty Ltd: **Preliminary Ecological Risk Assessment December 2016 - RAAF Base Williamtown Stage 2B Environmental Investigation** In. Edited by Department of Defence; 2016: 1.
26. Strauss A, Corbin J: **Grounded theory procedures and techniques**. Thousand Oaks, CA, USA: Sage Publications Inc.; 1990.
27. Diggle P, Ribeiro P: **Geostatistical design in Model-based Geostatistics**. In: *Springer Series in Statistics.* edn.: Springer; 2007.
28. Magalhães RJS, Clements AC, Patil AP, Gething PW, Brooker S: **The applications of model-based geostatistics in helminth epidemiology and control**. *Advances in parasitology* 2011, **74**:267.

15. Appendices

15.1. Appendix 1 Focus group discussions: DRAFT participation information sheet

PFAS Health Study

Participant Information Sheet for Focus Group Discussions

Project Title:

The PFAS Health Study: A Focus Group Study

The Australian Government Department of Health is funding the study.

Researcher Team	**Contact Details**
Professor Martyn Kirk	E: martyn.kirk@anu.edu.au
Associate-Professor Cathy Banwell	E: cathy.banwell@anu.edu.au
Dr Tambri Housen	E: tambri.housen@anu.edu.au
Ms Susan Trevenar	E: susan.trevenar@anu.edu.au
Ms Kayla Smurthwaite	E: kayla.smurthwaite@anu.edu.au

This Participant Information Sheet/Consent Form tells you about the research project. It explains the processes involved with taking part. Knowing what is involved will help you decide if you want to take part in the research. Please read this information carefully. Ask questions about anything that you don't understand or want to know more about.

Participation in this research is entirely voluntary; there is no obligation to take part in the study, and if you choose not to participate there will be no negative consequences.

If you decide you want to take part in the research project, you will be asked to sign the consent section.

You will be given a copy of this Participant Information and Consent Form to keep.

Project Title: The PFAS Health Study: A Focus Group Study

This study is being conducted by researchers from the National Centre for Epidemiology and Population Health in the Research School of Population Health at The Australian National University. Professor Martyn Kirk is the primary investigator of the PFAS Health Study. Associate-Professor Cathy Banwell will lead the focus group discussion study with Dr Tambri Housen, Ms Sue Trevenar, and Ms Kayla Smurthwaite.

General Outline of the Study:

The focus group discussion study is part of a broader study concerning the health and related risks of living in a PFAS Investigation Area. The focus groups are an opportunity for residents to express concerns about their health and the social impacts of living in the area and to shape the direction of the broader study. Anyone living in a PFAS Investigation Area is invited to participate in a discussion group. We expect to hold five group discussions in your area in public locations, such as town halls or community centres. The groups will consist of up to 12 people who will be invited to contribute to a general discussion. It is likely that the discussion will last between one and two hours.

The Australian Government has commissioned this study. No identifiable personal information will be provided to the Australian Government in the course of this study.

Participant Involvement:

Participation in the study is entirely voluntary; there is no obligation to take part in the study, and if you choose not to participate there will be no detriment to your career or future health care. Participants are free to withdraw from the research at any time without penalty and without providing a reason. If this occurs, the researchers will dispose of any data already collected from you. However, it may not be possible to remove statements that you have made as part of the general discussion. At the group discussion individuals will be asked to sign a consent form presented to them at the time.

We are asking all focus group attendees for their consent to collect their discussion via audio-recording so we can accurately record everything everyone tells us. As people talk quickly it is difficult to write everything down and we do not want to miss anything anyone tells us. Your

contribution to the discussion will be confidential beyond the group in which you participated. Your name will not be recorded anywhere on the recorder and if it is mentioned it will not be transcribed. If anyone does not wish to be recorded, the person may withdraw from the discussion.

During the focus group discussions, participants will be asked to discuss the following topics related to living in a PFAS Investigation Area:

- Health concerns
- Risk perception and management related to potential PFAS exposure
- Stress related to financial concerns due to living in the area
- Social issues
- Practical issues – where to live, moving, schooling, work, replacement of belongings, rebuilding house - time costs, other barriers
- The response to the PFAS situation by government, media, other
- And other issues that participants raise

Participants will be asked to fill out a short form collecting demographic information. With consent, the focus group discussions will be recorded. The discussions will last about an hour and the total time will be about two hours.

After the study the discussion material will be transcribed, collated and analysed and will then contribute to the findings from the broader study. The findings of the broader study will be disseminated to participants, to the general public and published in academic papers. The group discussion transcripts will not be available to individual participants.

Risks of Participating:

These discussions may raise some feelings of distress as they concern potential threats to health and well-being. The Australian Government has funded dedicated mental health and counselling services to provide support during this time. If you should become distressed, free counselling services are available and can be accessed through your local GP, the local primary health network or through Support Now. If you are a currently serving member of the ADF, you can access services through your usual Defence Health Centre.

A small token of our appreciation, a $50 EFTPOS voucher, will be offered on completion of the discussion.

The focus group discussions provide residents with an opportunity to express concerns and describe experiences related to their health and their social circumstances. The findings from the focus groups will be used to design a survey to be conducted in PFAS affected communities in 2018 and will contribute to the development of policy related to PFAS contamination. These will be used to inform the development of a survey questionnaire to be sent to current and past residents. The findings, with other parts of the study findings, will be presented in a report to the Australian Government Department of Health and to the general public and may be presented at scientific meetings and conferences, and published in academic books and journals. Information will be presented in such a way that individuals cannot be identified.

Confidentiality:

We will not be discussing whether you participated or not with other people. Only members of the research team will have access to the data. Your privacy is important to us. The identity of participants will not be collected except as a signature on the consent forms that are stored separately from data. We also ask that focus group members maintain the confidentiality of group discussions, and that participants in focus groups should refrain from making statements of a confidential nature or that are defamatory of any person. We ask that participants use pseudonyms. It is possible that transcripts from the focus group discussions may be subpoenaed as part of legal actions related to PFAS litigations. However, participants in focus groups will be anonymous, in the situation a participant's name is mentioned during interview, it will not be transcribed. Your participation will not affect your position at work, or your use of any local or state government service. It is entirely voluntary and there are no consequences for non-participation. The information you provide will not be linked to a name or phone number. Your data will be stored securely on ANU servers for five years and then destroyed. It will not be used in future studies.

Privacy Notice:
The ANU Privacy Policy contains information about how you can

- Have access or seek correction to your personal information; and
- Complain about a breach of an Australian Privacy Principle (APP) by ANU and how ANU will handle the complaint.

Questions:

If you have any questions, do not hesitate to contact us (the researchers who are conducting the discussions) by email or phone.

Dr Cathy Banwell

T: 6125 0016

Cathy.Banwell@anu.edu.au

Dr Tambri Housen

(02) 6125 0460

Tambri.Housen@anu.edu.au

Concerns or complaints:

The Australian National University Human Research Ethics Committee and the DDVA Human Research Ethics Committee (ANU HREC protocol 2017/816 and DDVA HREC protocol 024-17). If you have concerns regarding the way this research was conducted please do not hesitate to contact the researchers or the following:

Executive Officer DDVA HREC CP3-6-037 PO Box 7911 Canberra BC ACT 2610 T: (02) 62663807 E: ddva.hrec@defence.gov.au	Human Research Ethics Officer The Australian National University Office of Research Integrity Chancelry 10B, T: (02) 6125 3427 E: Human.Ethics.Officer@anu.edu.au

No member of the research team will receive a personal financial benefit from involvement in this research project (other than their ordinary wages).

15.2. Appendix 2 Focus group discussions: DRAFT consent form

Participant Written Consent Form

CONSENT FORM

Title **The PFAS Health Study: Focus Group Discussions**

I, ... give my consent to participate in the project mentioned above on the following basis:

I have had explained to me the aims of this research project, how it will be conducted and my role in it.

I understand the risks involved as described in the Participant Information Sheet.

I am cooperating in this project on condition that:

- the information I provide will be kept confidential
- the information will be used only for this project. The research results will be made available to me at my request and any published reports of this study will preserve my anonymity
- I have been given a copy of the 'Departments of Defence and Veterans' Affairs Human Research Ethics Committee (DDVA HREC) Guidelines for Volunteers'.

I understand that:

- there is no obligation to take part in this study
- I am free to withdraw at any time

I have been given a copy of the participant information sheet and consent form, signed by me and by the principal investigator Martyn Kirk to keep.

_____Signature of participant

_____Name in full

_____Date

Signature of Principal Investigator

Cathy Banwell — Name in full

30 January 2018 — Date

Concerns or complaints to:

The Australian National University Human Research Ethics Committee and the DDVA Human Research Ethics Committee (ANU HREC protocol 2017/816 and DDVA HREC protocol 024-17). If you have concerns regarding the way this research was conducted please do not hesitate to contact the researchers or the following:

Executive Officer DDVA HREC CP3-6-037 PO Box 7911 Canberra BC ACT 2610 T: (02) 62663807 E: ddva.hrec@defence.gov.au	Human Research Ethics Officer The Australian National University Office of Research Integrity Chancelry 10B, T: (02) 6125 3427 E: Human.Ethics.Officer@anu.edu.au

15.3. Appendix 3 Focus group discussions: DRAFT questionnaires

PFAS Health Study

Questionnaire for Focus Group participants

1. Sex

☐ Male ☐ Female ☐ Other

2. Age

☐ 25-29 ☐ 30-34 ☐ 35-39 ☐ 40-44

☐ 45-49 ☐ 50-54 ☐ 55-59 ☐ 60+

3. What is your highest completed level of education?

☐ Incomplete secondary ☐ Completed secondary

☐ Certificate or diploma ☐ Bachelor degree or above

4. Partnership status

☐ Single (Never Married) ☐ Single (Separated/Divorced/ Widowed)

☐ Married ☐ Cohabiting/De Facto

5. What is your employment status?

☐ Not employed ☐ Retired ☐ Employed (casual))

☐ Employed (part-time) ☐ Employed (full-time)

6. What is your current job?

……

7. Do you have any children living with you?

☐ Yes ☐ No

8. If you have children living with you, what are their ages?

9. Did you own or rent your home?

☐ Own ☐ Rent

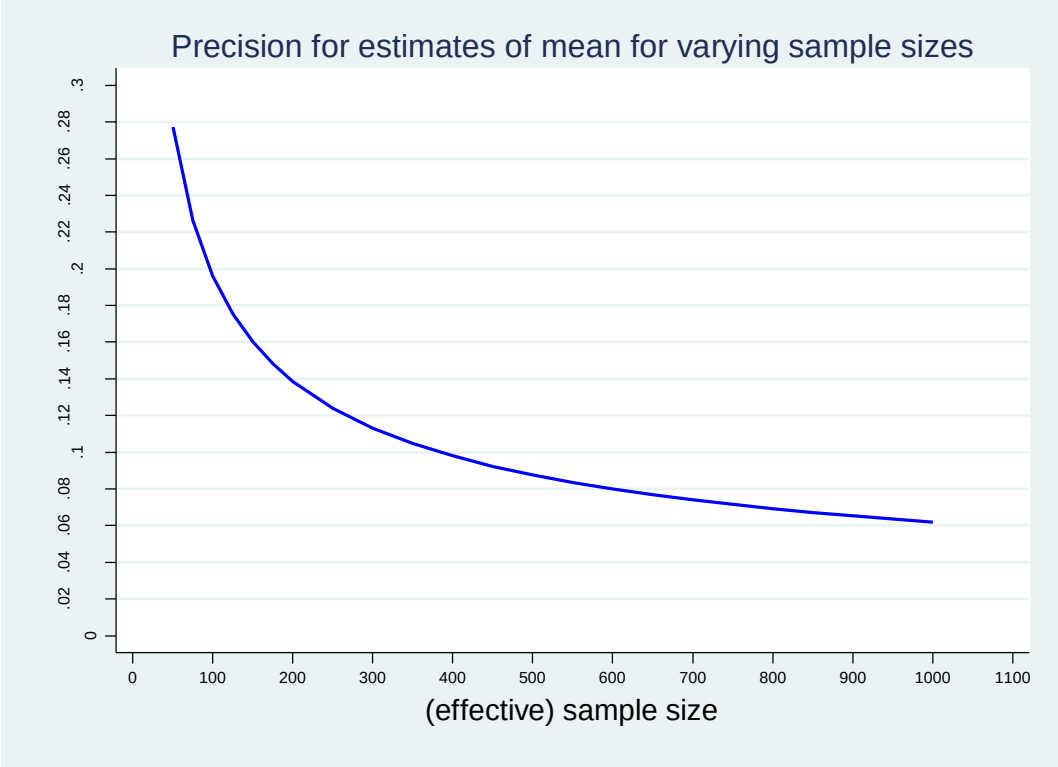
10. Where do you live?

☐ **In town** ☐ **Outskirts of town** ☐ **Rural property**

11. How long have you lived in this area? _____Years

15.4. Appendix 4 Precision for estimate of means: Components 2 & 3

Figure 4: Precision for estimates of means, for effective samples sizes ranging from 50 to 1000



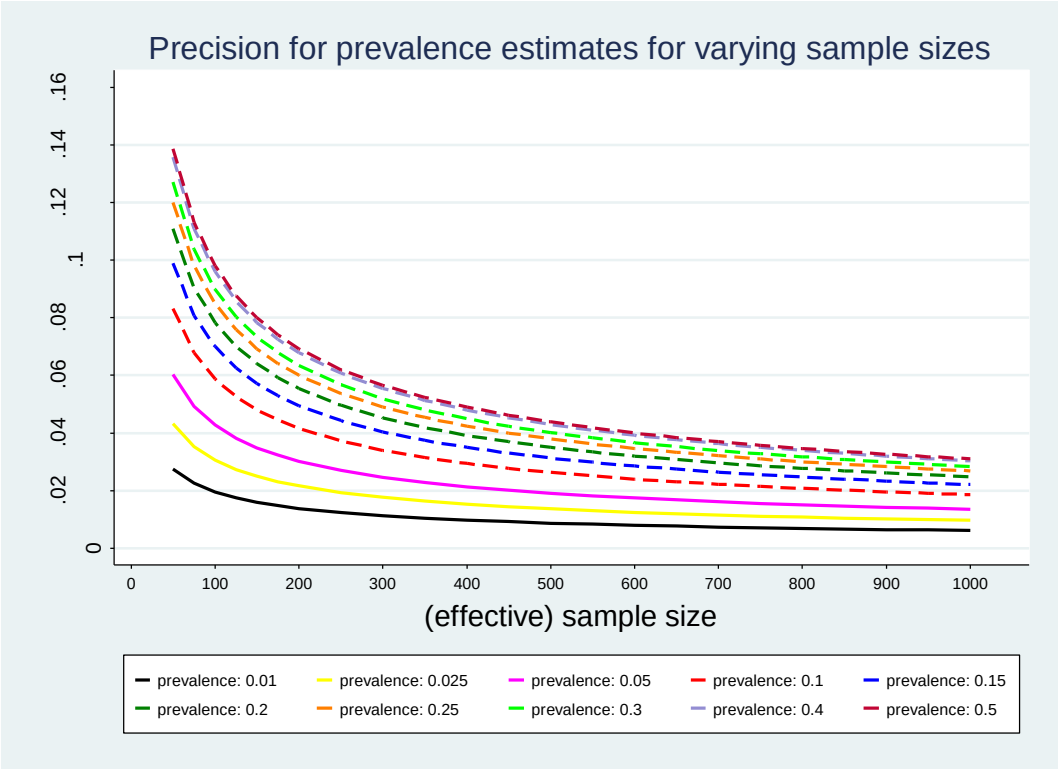
Precision for estimates of means
For 95% Confidence Intervals

Effective Sample size	Precision of mean
50	0.277
75	0.226
100	0.196
125	0.175
150	0.160
175	0.148
200	0.139
250	0.124
300	0.113
350	0.105
400	0.098
450	0.092
500	0.088

550	0.084
600	0.080
650	0.077
700	0.074
750	0.072
800	0.069
850	0.067
900	0.065
950	0.064
1000	0.062

15.5. Appendix 5 Precision for prevalence estimates: Components 2 & 3

Figure 5: Precision for prevalence estimates of 1% to 50%, for effective samples sizes ranging from 50–1000.



Precision for prevalence estimates
For 95% Confidence Intervals

Prevalence	Effective sample size	Precision
0.01	50	0.028
0.01	75	0.023
0.01	100	0.020
0.01	125	0.017
0.01	150	0.016
0.01	175	0.015
0.01	200	0.014
0.01	250	0.012
0.01	300	0.011
0.01	350	0.010
0.01	400	0.010
0.01	450	0.009
0.01	500	0.009
0.01	550	0.008
0.01	600	0.008
0.01	650	0.008
0.01	700	0.007

Prevalence	Effective sample size	Precision
0.01	750	0.007
0.01	800	0.007
0.01	850	0.007
0.01	900	0.007
0.01	950	0.006
0.01	1000	0.006
0.025	50	0.043
0.025	75	0.035
0.025	100	0.031
0.025	125	0.027
0.025	150	0.025
0.025	175	0.023
0.025	200	0.022
0.025	250	0.019
0.025	300	0.018
0.025	350	0.016
0.025	400	0.015
0.025	450	0.014
0.025	500	0.014
0.025	550	0.013
0.025	600	0.012
0.025	650	0.012
0.025	700	0.012
0.025	750	0.011
0.025	800	0.011
0.025	850	0.010
0.025	900	0.010
0.025	950	0.010
0.025	1000	0.010
0.05	50	0.060
0.05	75	0.049
0.05	100	0.043
0.05	125	0.038
0.05	150	0.035
0.05	175	0.032
0.05	200	0.030
0.05	250	0.027
0.05	300	0.025
0.05	350	0.023
0.05	400	0.021
0.05	450	0.020
0.05	500	0.019
0.05	550	0.018
0.05	600	0.017
0.05	650	0.017
0.05	700	0.016
0.05	750	0.016
0.05	800	0.015
0.05	850	0.015
0.05	900	0.014
0.05	950	0.014

Prevalence	Effective sample size	Precision
0.05	1000	0.014
0.1	50	0.083
0.1	75	0.068
0.1	100	0.059
0.1	125	0.053
0.1	150	0.048
0.1	175	0.044
0.1	200	0.042
0.1	250	0.037
0.1	300	0.034
0.1	350	0.031
0.1	400	0.029
0.1	450	0.028
0.1	500	0.026
0.1	550	0.025
0.1	600	0.024
0.1	650	0.023
0.1	700	0.022
0.1	750	0.021
0.1	800	0.021
0.1	850	0.020
0.1	900	0.020
0.1	950	0.019
0.1	1000	0.019
0.15	50	0.099
0.15	75	0.081
0.15	100	0.070
0.15	125	0.063
0.15	150	0.057
0.15	175	0.053
0.15	200	0.049
0.15	250	0.044
0.15	300	0.040
0.15	350	0.037
0.15	400	0.035
0.15	450	0.033
0.15	500	0.031
0.15	550	0.030
0.15	600	0.029
0.15	650	0.027
0.15	700	0.026
0.15	750	0.026
0.15	800	0.025
0.15	850	0.024
0.15	900	0.023
0.15	950	0.023
0.15	1000	0.022
0.2	50	0.111
0.2	75	0.091
0.2	100	0.078
0.2	125	0.070

Prevalence	Effective sample size	Precision
0.2	150	0.064
0.2	175	0.059
0.2	200	0.055
0.2	250	0.050
0.2	300	0.045
0.2	350	0.042
0.2	400	0.039
0.2	450	0.037
0.2	500	0.035
0.2	550	0.033
0.2	600	0.032
0.2	650	0.031
0.2	700	0.030
0.2	750	0.029
0.2	800	0.028
0.2	850	0.027
0.2	900	0.026
0.2	950	0.025
0.2	1000	0.025
0.25	50	0.120
0.25	75	0.098
0.25	100	0.085
0.25	125	0.076
0.25	150	0.069
0.25	175	0.064
0.25	200	0.060
0.25	250	0.054
0.25	300	0.049
0.25	350	0.045
0.25	400	0.042
0.25	450	0.040
0.25	500	0.038
0.25	550	0.036
0.25	600	0.035
0.25	650	0.033
0.25	700	0.032
0.25	750	0.031
0.25	800	0.030
0.25	850	0.029
0.25	900	0.028
0.25	950	0.028
0.25	1000	0.027
0.3	50	0.127
0.3	75	0.104
0.3	100	0.090
0.3	125	0.080
0.3	150	0.073
0.3	175	0.068
0.3	200	0.064
0.3	250	0.057
0.3	300	0.052

Prevalence	Effective sample size	Precision
0.3	350	0.048
0.3	400	0.045
0.3	450	0.042
0.3	500	0.040
0.3	550	0.038
0.3	600	0.037
0.3	650	0.035
0.3	700	0.034
0.3	750	0.033
0.3	800	0.032
0.3	850	0.031
0.3	900	0.030
0.3	950	0.029
0.3	1000	0.028
0.4	50	0.136
0.4	75	0.111
0.4	100	0.096
0.4	125	0.086
0.4	150	0.078
0.4	175	0.073
0.4	200	0.068
0.4	250	0.061
0.4	300	0.055
0.4	350	0.051
0.4	400	0.048
0.4	450	0.045
0.4	500	0.043
0.4	550	0.041
0.4	600	0.039
0.4	650	0.038
0.4	700	0.036
0.4	750	0.035
0.4	800	0.034
0.4	850	0.033
0.4	900	0.032
0.4	950	0.031
0.4	1000	0.030
0.5	50	0.139
0.5	75	0.113
0.5	100	0.098
0.5	125	0.088
0.5	150	0.080
0.5	175	0.074
0.5	200	0.069
0.5	250	0.062
0.5	300	0.057
0.5	350	0.052
0.5	400	0.049
0.5	450	0.046
0.5	500	0.044
0.5	550	0.042

Prevalence	Effective sample size	Precision
0.5	600	0.040
0.5	650	0.038
0.5	700	0.037
0.5	750	0.036
0.5	800	0.035
0.5	850	0.034
0.5	900	0.033
0.5	950	0.032
0.5	1000	0.031

15.6. Appendix 6 Detectable differences in means: Components 2 & 3

Detectable Differences in Means
80% power, 5% significance level

Total (effective) sample size	(effective) sample size per group	Detectable difference between mean (in standard deviations)
50	25	0.81
100	50	0.57
150	75	0.46
200	100	0.40
250	125	0.36
300	150	0.32
350	175	0.30
400	200	0.28
450	225	0.26
500	250	0.25
550	275	0.24
600	300	0.23
650	325	0.22
700	350	0.21
750	375	0.20
800	400	0.20
850	425	0.19
900	450	0.19
950	475	0.18
1000	500	0.18

15.7. Appendix 7 Detectable differences in proportions: Components 2 & 3

Detectable Differences in Proportions
80% power, 5% significance level

Total (effective) sample size	(effective) sample size per group	Prevalence in group with lowest risk	Detectable difference in proportions	Prevalence in group with lowest risk
50	25	0.01	0.274	0.284
100	50	0.01	0.157	0.167
150	75	0.01	0.113	0.123
200	100	0.01	0.090	0.100
250	125	0.01	0.075	0.085
300	150	0.01	0.065	0.075
350	175	0.01	0.058	0.068
400	200	0.01	0.052	0.062
450	225	0.01	0.048	0.058
500	250	0.01	0.044	0.054
550	275	0.01	0.041	0.051
600	300	0.01	0.038	0.048
650	325	0.01	0.036	0.046
700	350	0.01	0.034	0.044
750	375	0.01	0.033	0.043
800	400	0.01	0.031	0.041
850	425	0.01	0.030	0.040
900	450	0.01	0.029	0.039
950	475	0.01	0.028	0.038
1000	500	0.01	0.027	0.037
50	25	0.025	0.291	0.316
100	50	0.025	0.175	0.200
150	75	0.025	0.130	0.155
200	100	0.025	0.105	0.130
250	125	0.025	0.090	0.115
300	150	0.025	0.079	0.104
350	175	0.025	0.071	0.096
400	200	0.025	0.065	0.090
450	225	0.025	0.060	0.085
500	250	0.025	0.056	0.081
550	275	0.025	0.053	0.078
600	300	0.025	0.050	0.075
650	325	0.025	0.047	0.072
700	350	0.025	0.045	0.070
750	375	0.025	0.043	0.068
800	400	0.025	0.041	0.066
850	425	0.025	0.040	0.065
900	450	0.025	0.038	0.063
950	475	0.025	0.037	0.062
1000	500	0.025	0.036	0.061
50	25	0.05	0.313	0.363
100	50	0.05	0.197	0.247

Total (effective) sample size	(effective) sample size per group	Prevalence in group with lowest risk	Detectable difference in proportions	Prevalence in group with lowest risk
150	75	0.05	0.150	0.200
200	100	0.05	0.124	0.174
250	125	0.05	0.108	0.158
300	150	0.05	0.096	0.146
350	175	0.05	0.087	0.137
400	200	0.05	0.080	0.130
450	225	0.05	0.074	0.124
500	250	0.05	0.070	0.120
550	275	0.05	0.066	0.116
600	300	0.05	0.062	0.112
650	325	0.05	0.059	0.109
700	350	0.05	0.057	0.107
750	375	0.05	0.055	0.105
800	400	0.05	0.052	0.102
850	425	0.05	0.051	0.101
900	450	0.05	0.049	0.099
950	475	0.05	0.047	0.097
1000	500	0.05	0.046	0.096
50	25	0.1	0.344	0.444
100	50	0.1	0.227	0.327
150	75	0.1	0.178	0.278
200	100	0.1	0.150	0.250
250	125	0.1	0.131	0.231
300	150	0.1	0.118	0.218
350	175	0.1	0.108	0.208
400	200	0.1	0.100	0.200
450	225	0.1	0.093	0.193
500	250	0.1	0.088	0.188
550	275	0.1	0.083	0.183
600	300	0.1	0.079	0.179
650	325	0.1	0.076	0.176
700	350	0.1	0.073	0.173
750	375	0.1	0.070	0.170
800	400	0.1	0.067	0.167
850	425	0.1	0.065	0.165
900	450	0.1	0.063	0.163
950	475	0.1	0.061	0.161
1000	500	0.1	0.059	0.159
50	25	0.15	0.365	0.515
100	50	0.15	0.247	0.397
150	75	0.15	0.196	0.346
200	100	0.15	0.167	0.317
250	125	0.15	0.147	0.297
300	150	0.15	0.133	0.283
350	175	0.15	0.122	0.272
400	200	0.15	0.113	0.263
450	225	0.15	0.106	0.256
500	250	0.15	0.100	0.250
550	275	0.15	0.095	0.245

Total (effective) sample size	(effective) sample size per group	Prevalence in group with lowest risk	Detectable difference in proportions	Prevalence in group with lowest risk
600	300	0.15	0.090	0.240
650	325	0.15	0.087	0.237
700	350	0.15	0.083	0.233
750	375	0.15	0.080	0.230
800	400	0.15	0.077	0.227
850	425	0.15	0.075	0.225
900	450	0.15	0.073	0.223
950	475	0.15	0.071	0.221
1000	500	0.15	0.069	0.219
50	25	0.2	0.377	0.577
100	50	0.2	0.260	0.460
150	75	0.2	0.209	0.409
200	100	0.2	0.179	0.379
250	125	0.2	0.158	0.358
300	150	0.2	0.143	0.343
350	175	0.2	0.132	0.332
400	200	0.2	0.123	0.323
450	225	0.2	0.115	0.315
500	250	0.2	0.109	0.309
550	275	0.2	0.103	0.303
600	300	0.2	0.099	0.299
650	325	0.2	0.095	0.295
700	350	0.2	0.091	0.291
750	375	0.2	0.088	0.288
800	400	0.2	0.085	0.285
850	425	0.2	0.082	0.282
900	450	0.2	0.080	0.280
950	475	0.2	0.077	0.277
1000	500	0.2	0.075	0.275
50	25	0.25	0.384	0.634
100	50	0.25	0.269	0.519
150	75	0.25	0.218	0.468
200	100	0.25	0.187	0.437
250	125	0.25	0.166	0.416
300	150	0.25	0.151	0.401
350	175	0.25	0.139	0.389
400	200	0.25	0.130	0.380
450	225	0.25	0.122	0.372
500	250	0.25	0.115	0.365
550	275	0.25	0.110	0.360
600	300	0.25	0.105	0.355
650	325	0.25	0.101	0.351
700	350	0.25	0.097	0.347
750	375	0.25	0.093	0.343
800	400	0.25	0.090	0.340
850	425	0.25	0.087	0.337
900	450	0.25	0.085	0.335
950	475	0.25	0.082	0.332
1000	500	0.25	0.080	0.330

Total (effective) sample size	(effective) sample size per group	Prevalence in group with lowest risk	Detectable difference in proportions	Prevalence in group with lowest risk
50	25	0.3	0.387	0.687
100	50	0.3	0.275	0.575
150	75	0.3	0.223	0.523
200	100	0.3	0.193	0.493
250	125	0.3	0.172	0.472
300	150	0.3	0.156	0.456
350	175	0.3	0.144	0.444
400	200	0.3	0.135	0.435
450	225	0.3	0.127	0.427
500	250	0.3	0.120	0.420
550	275	0.3	0.114	0.414
600	300	0.3	0.109	0.409
650	325	0.3	0.105	0.405
700	350	0.3	0.101	0.401
750	375	0.3	0.097	0.397
800	400	0.3	0.094	0.394
850	425	0.3	0.091	0.391
900	450	0.3	0.089	0.389
950	475	0.3	0.086	0.386
1000	500	0.3	0.084	0.384
50	25	0.4	0.381	0.781
100	50	0.4	0.276	0.676
150	75	0.4	0.227	0.627
200	100	0.4	0.197	0.597
250	125	0.4	0.176	0.576
300	150	0.4	0.161	0.561
350	175	0.4	0.149	0.549
400	200	0.4	0.139	0.539
450	225	0.4	0.131	0.531
500	250	0.4	0.125	0.525
550	275	0.4	0.119	0.519
600	300	0.4	0.114	0.514
650	325	0.4	0.109	0.509
700	350	0.4	0.105	0.505
750	375	0.4	0.102	0.502
800	400	0.4	0.098	0.498
850	425	0.4	0.095	0.495
900	450	0.4	0.093	0.493
950	475	0.4	0.090	0.490
1000	500	0.4	0.088	0.488
50	25	0.5	0.361	0.861
100	50	0.5	0.267	0.767
150	75	0.5	0.221	0.721
200	100	0.5	0.193	0.693
250	125	0.5	0.174	0.674
300	150	0.5	0.159	0.659
350	175	0.5	0.148	0.648
400	200	0.5	0.138	0.638
450	225	0.5	0.131	0.631

Total (effective) sample size	(effective) sample size per group	Prevalence in group with lowest risk	Detectable difference in proportions	Prevalence in group with lowest risk
500	250	0.5	0.124	0.624
550	275	0.5	0.118	0.618
600	300	0.5	0.113	0.613
650	325	0.5	0.109	0.609
700	350	0.5	0.105	0.605
750	375	0.5	0.102	0.602
800	400	0.5	0.098	0.598
850	425	0.5	0.096	0.596
900	450	0.5	0.093	0.593
950	475	0.5	0.090	0.590
1000	500	0.5	0.088	0.588