BMJ Open

Biobanking for discovery of novel cardiovascular biomarkers using imaging-quantified disease burden: protocol for the longitudinal, prospective, BioHEART-CT cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-028649
Article Type:	Protocol
Date Submitted by the Author:	18-Dec-2018
Complete List of Authors:	Kott, Katharine; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Vernon, Stephen; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Hansen, Thomas ; Kolling Institute of Medical Research Yu, Christine; Kolling Institute of Medical Research Bubb, Kristen ; University of Sydney, Northern Clinical School, Faculty of Medicine and Health Coffey, Sean; Royal North Shore Hospital, Kolling Institute and Department of Cardiology; University of Otago, Dunedin School of Medicine Sullivan, David; The University of Sydney Yang, Jean ; University of Sydney, School of Mathematics and Statistics O'Sullivan, John; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Chow, Clara; The George Institute for Global Health and Westmead Hospital, Sydney Patel, Sanjay; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Chong, James ; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Celermajer, David; University of Sydney, Faculty of Medicine and Health; Royal Prince Alfred Hospital, Department of Cardiology Kritharides, Leonard; Concord Hospital, Cardiology Grieve, Stuart; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Figtree, Gemma; University of Sydney, Faculty of Medicine and Health; Kolling Institute of Medical Research,
Keywords:	Coronary heart disease < CARDIOLOGY, Cardiovascular imaging < RADIOLOGY & IMAGING, Ischaemic heart disease < CARDIOLOGY, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, MOLECULAR BIOLOGY, Clinical chemistry < PATHOLOGY

Biobanking for discovery of novel cardiovascular biomarkers using imaging-quantified disease burden: protocol for the longitudinal, prospective, BioHEART-CT cohort study

Katharine A. Kott*[1,2], Stephen T. Vernon*[1,2], Thomas Hansen[1], Christine Yu[1], Kristen J. Bubb[1], Sean Coffey[1,5], David Sullivan[2,3,6], Jean Y. Yang[3,4], John F. O'Sullivan[2,3,9], Clara K Chow[2,10,13], Sanjay Patel[2,3,7,9], James Chong[2,10,11,12], David S. Celermajer[2,7,9], Leonard Kritharides[2,11], Stuart M. Grieve[2,3,8], Gemma A. Figtree[1,2]

[1]Cardiothoracic and Vascular Health, Kolling Institute and Department of Cardiology, Royal North Shore Hospital, Northern Sydney Local Health District, Australia
[2]Sydney Medical School, Faculty of Medicine and Health, Univertsity of Sydney, Australia
[3]Charles Perkins Centre, University of Sydney, Australia
[4]School of Mathematics and Statistics, University of Sydney, Australia
[5]Dunedin School of Medicine, University of Otago, New Zealand
[6]Department of Biochemistry, Royal Prince Alfred Hospital, Sydney, Australia
[7]Department of Cardiology, Royal Prince Alfred Hospital, Sydney, Australia
[8]Department of Radiology, Royal Prince Alfred Hospital, Sydney, Australia
9The Heart Research Institute, Sydney, Australia
10Department of Cardiology, Westmead Hospital, Sydney, Australia
11ANZAC Research Institute & Cardiology Department, Concord Repatriation General Hospital, Sydney, Australia
12 Westmead Institute for Medical Research, Westmead Hospital, Sydney, Australia
13 Westmead Applied Research Centre, University of Sydney, Sydney, Australia

Corresponding author:

Professor Gemma Figtree MBBS DPhil (Oxon) FRACP FAHA

Cardiothoracic and Vascular Health, Level 12, Kolling Institute of Medical Research, Royal North Shore Hospital, St Leonards, NSW, 2065, Australia. Email: gemma.figtree@sydney.edu.au.

Abstract

Introduction: Coronary artery disease (CAD) persists as a major cause of morbidity and mortality worldwide despite intensive identification and treatment of traditional risk factors. Data emerging over the past decade shows a quarter of patients have disease in the absence of any known risk factor, and half have only one traditional risk factor. Improvements in quantification and characterisation of coronary atherosclerosis by CT coronary angiography (CTCA) can provide quantitative measures of subclinical atherosclerosis - enhancing the power of unbiased "omics" studies to unravel the missing biology of personal susceptibility, identify new biomarkers for early diagnosis, and to suggest new targeted therapeutics.

Methods & Analysis: BioHEART-CT is a longitudinal, prospective cohort study, aiming to recruit 5000 adult patients undergoing clinically-indicated CTCA. After informed consent, patient data, blood samples and CTCA imaging data are recorded. Follow-up for all patients is conducted one month after recruitment, and then annually for the life of the study. CTCA data provides volumetric quantification of total calcified and non-calcified plaque, and measures that characterise plaque vulnerability. Comprehensive molecular phenotyping will be performed using state-of-the-art genomics, metabolomics, proteomics, and immunophenotyping. Complex network and machine learning approaches will be applied to biological and clinical datasets to identify novel pathophysiological pathways and to prioritise new biomarkers. Discovery analysis will be performed in the first 1000 patients of BioHEART-CT, with validation analysis in the following 4000 patients. Outcome data will be utilised to build improved risk models for CAD.

Ethics & Dissemination: The study protocol has been approved by the human research ethics committee of North Shore Local Health District in Sydney, Australia. All findings will be published in peer-reviewed journals or at scientific conferences.

Registration: This study is registered with the Australia and New Zealand Clinical Trials Registry - ACTRN12618001322224.

Article Summary

Strengths and Limitations of this Study (max five bullet points)

- BioHEART-CT is a prospective, longitudinal cohort study assessing patients with suspected CAD undergoing CTCA across multiple Australian centres.
- Quantitative measures of coronary atherosclerosis from CTCA datasets will be used in conjunction with biological samples, improving the power to discover new mechanisms and markers for CAD.
- Samples stratified by imaging-quantified disease severity will be analysed by both candidate and unbiased “omics” approaches, utilising modern technologies and bioinformatic analysis to discover new biomarkers using proteomics, metabolomics, lipidomics, transciptomics, genomics and immunophenotyping.
- Longitudinal outcome data will be used to build risk models which can incorporate both traditional and novel risk factors and biomarkers.
- Limitations of the study include selection bias, as only patients with a clinical indication for CTCA are included, and the geographic isolation of Australia, which may result in the need for confirmation of results in multi-national studies.

Introduction

Cardiovascular disease has persisted as a major cause of human morbidity and mortality despite continual improvements in preventative therapies. In 2015 cardiovascular disease accounted for one-third of all deaths world-wide[1], and ischaemic heart disease (IHD) remains the leading cause of years of life lost for high and middle sociodemographic groups multi-nationally[2]. This large burden of disease puts significant stress on the health systems of individual countries. In Australia, cardiovascular disease remains the number one killer of Australians[3], with local health costs estimated to be $7.7 billion AUD per annum[4].

The standard modifiable cardiovascular risk factors (SMuRFs[5]) for atherosclerosis - hypertension, diabetes mellitus, hyperlipidaemia and smoking - were identified in epidemiological studies in the 1960s[6]. The identification and subsequent efforts to target these risk factors at community and primary care levels have led to a substantial reduction in mortality from cardiovascular disease[1]. However, substantial disease burden remains and importantly, individual variability in susceptibility to these risk factors is considerable. While continuing our efforts to tackle societal and modifiable risk factors, identifiying undiscovered mechanisms that lead to the development of atherosclerosis is critical. Such work will provide new biomarkers for early detection of subclinical atherosclerosis and open avenues for improved preventative and therapeutic strategies.

The importance of new detection mechanisms for subclinical atherosclerosis is highlighted by interrogating the “fine print” of clinical studies, including information not usually presented in the main tables of mansuscripts. Although data regarding the percentage of patients suffering myocardial infarction despite having no SMuRFs is often omitted, a large meta-analysis found that >50% of women and >60% of men presenting with their first ACS had 0 or 1 SMuRF[7], and similar proportions of SMuRFless patients have been identified in ACS cohort studies [5,8]. In a recent single centre Australian study, we showed that the proportion of patients presenting with ST-elevation myocardial infarction with 0 SMuRFs has risen from 11% to 27% over an 8 year period[5]. These patients have developed extensive CAD without having any of the red flags which would allow doctors to identify their risk of disease and intervene early. Although this is a relatively small piece of the overall IHD pie, the global burden of CAD is immense with an estimated prevalence of 470 million[9], which makes this 25-30% SMuRFless patient population important from a public health perspective. It is sobering to think that we have made no clear advances that allow us to identify risk and subclinical disease in this subgroup. Indeed, we also have limited information regarding how these patients respond to traditional secondary preventative strategies, and whether their long term outcomes and disease progression differs from those with traditional modifiable risk factors.

Further evidence that novel biological mechanisms contribute substantially to atherosclerosis can be seen in results from large scale genome-wide association studies which have reported that 66% of the identified genes conveying heritable cardiovascular risk are not associated with the traditional risk factors[10]. Many of these genes were related to inflammatory processes, and studies such as the CANTOS trial[11] confirm the importance of inflammation in CAD. However, some of these important genes have not yet been well characterised, and are not yet associated with any known pathway in atherosclerosis and CAD, highlighting the importance of ongoing discovery work.

To address this gap in our understanding of CAD pathophysiology, we have designed a unique cohort study with biobanked samples to facilitate investigation of novel cardiac risk factors and biomarkers. This biobank includes advanced, quantitative imaging measurements of coronary artery atherosclerosis to allow for accurate phenotyping of the patient groups. This overcomes a

weaknesses of previous genomics studies that have relied on the presence or absence of a coronary artery event to classify individuals as diseased or as a control, improving the statistical power of the subsequent analysis. The advanced imaging dataset and clinical data including baseline and follow-up are integrated with the collection of a broad range of biological samples for discovery work, utilising state-of-the-art molecular phenotyping platforms and computational bioinformatics.

Objectives

The primary objectives of the BioHEART-CT study are:

1. To identify new biomarkers to assist in early CAD risk identification and stratification;
2. To identify new mechanistic pathways that may be targeted with novel therapeutic strategies to abrogate CAD risk.

The secondary objectives of the BioHEART-CT study include the following:

1. To determine the predictive value of a modified CTCA scoring system for CAD risk which incorporates plaque composition data;
2. To identify new risk factors for CAD based on integration of clinical information profiles, CTCA results and clinical outcomes;
3. To develop a novel risk scoring system incorporating clinical risk factors, novel biomarkers and CTCA scores to more accurarely predict CVD events.

Methods and Analysis

Study Type

BioHEART-CT is a multicentre, prospective cohort study of patients with suspected CAD, which brings together detailed clinical information with advanced imaging and molecular phenotyping.

Study Population

5000 patients will be recruited from participating hospitals' and associated imaging sites. All recruitment centres are large tertiary hospitals within the city of Sydney, Australia.

Inclusion criteria:

- Patients who have been referred for investigation of suspected CAD
- Age 18 years or older
- Willing and able to provide informed consent

Exclusion criteria:

- Patients highly dependent on medical care who are unable to provide informed consent
- Patients unwilling or unable to participate in ongoing follow-up

Study Protocol

Screening and Enrolment

Eligible patients having a CTCA are invited to participate in the study. Eligibility is confirmed according to the inclusion and exclusion criteria, the study is explained, and informed consent obtained.

Patient Data Collection

The baseline questionnaire is completed by research staff conducting a face-to-face patient interview. Data obtained includes the following information: demographic data including smoking and alcohol intake, past medical history, relevant family history, medication history, history of cardiac symptoms and indication for CTCA. Data is entered into secure, encrypted databases in a de-identified format. After recruitment all patients are assigned a unique study number and all samples, imaging data and demographic information is identified by this number. The master list including identifiable data is securely stored on an encrypted server and is available only to senior research staff. Relevant details are made available to research staff for follow-up purposes.

Follow-Up & Outcome Adjudication

Participants are contacted by phone at approximately 30 days and then annually to determine if any events have occurred. Information recorded for any potential outcome event includes the acuity, need for hospitalisation and indication for admission, hospital attended, treating physician, date of admission and details of the events. New medical or surgical diagnoses since last contact are also recorded. Participants in BioHEART also consent to review of their electronic medical records and data linkage through the Centre for Health Record Linkage (CHeReL), which will be used for independent medical record reviews.

Significant outcome events that trigger a file review to assess for major adverse cardiovascular events (MACE) or secondary outcomes are the following:

- Unstable angina requiring hospitalisation
- Heart failure requiring hospitalisation
- Acute myocardial infarction
- Coronary angiogram with or without percutaneous intervention
- Cardiac bypass graft surgery
- Aortic or mitral valve surgery
- Cerebrovascular accident / transient ischaemic attack
- All-cause mortality

All medical records will be reviewed by an independent adjudication committee to determine if an event should be included in the study results. Detailed criteria for the individual diagnoses included in the outcomes must be met.

MACE is defined as cardiovascular death, non-fatal myocardial infarction or non-fatal stroke. Exploratory outcomes include revascularisation, unstable angina or heart failure requiring hospitalisation.

CTCA Data Acquisition

CTCA scans are obtained on 256 slice scanner, with reconstructions created using appropriate software for the individual machine. Each study is to be protocolled by a radiologist with at least level two accreditation (RANZCR/RANZCP). All radiographers are trained in acquisition and workup of CTCA. Oral metoprolol, or ivabradine if beta blockers are contraindicated, are given to optimise the heart rate if required. Dosing is adjusted by clinical staff according to baseline heart rate and bodyweight. Oral nitroglycerine (600-800 micrograms) is given immediately prior to the scan, and iodinated contrast is injected intravenously. Prospective studies are performed if the heart rate is

sufficiently controlled, otherwise retrospective acquisition is used. Average dose is minimised at all participating sites in line with current recommendations[12,13].

CTCA Imaging Analysis

CT data is exported as thin DICOM data and is stored securely in a de-identified manner. Data is later analysed using a dedicated workstation to obtain the primary clinical scores at each site. CTCAs are analysed by standard anatomical arterial region as per the 17-segment model outlined by the Society of Cardiovascular Computed Tomography[14]. Each segment is scored according to degree of stenosis and composition of plaque (calcified, soft and mixed), see Figure 2. The segmental scores are aggregated into a modified GENSINI score which represents the total amount of calcified and non-calcified plaque present. Calcium scores are generated using the scanner-associated software by the Agaston method. Total raw calcium scores, raw calcium scores for each vessel, and age and sex-matched calcium percentage scores are recorded.

Further analysis will be performed at a central core laboratory using a Frontier System (syngo.via, Siemens Healthcare, Erlangen), employing a semi-automated soft plaque algorithm.

Biological Sample Collection

A 20-30mL blood sample is taken from an in-situ intravenous cannula, or if this is not available, by venesection, prior to CTCA and drug administration. Blood is collected into EDTA, serum and citrate tubes and stored at 4°C until processing. A lithium heparin tube is collected and processed immediately to extract peripheral blood mononuclear cells (PBMCs) from the buffy coat.

Sample Processing & Storage

Blood samples are processed by hospital pathology scientists experienced in laboratory procedures. Briefly, whole blood is set aside, then tubes are centrifuged at 1861 x *g* for 15 minutes at 4°C prior to aliquoting 500uL samples of serum, EDTA plasma, whole blood, buffy coat and erythrocytes. The citrate tube is centrifuged at 1861 x *g* for a second 15 minutes at 4°C, after which citrate-plasma samples are aliquoted. All aliquoted samples are frozen and stored at -80°C until analysis.

For PBMC isolation a standard gradient-separation protocol is used[15]. Briefly, blood from the lithium heparin tube is diluted 1:1 with Hanks' Balanced Salt Solution (HBSS). Diluted whole blood is then layered on top of Ficoll-Paque Plus and centrifgued (22°C, 1460 x *g* no brake, 20 min) . The buffy coat layer containing PBMCs is then further purified with two wash steps in HBSS using repeat centrifugation as above. PBMCs are then plated at a density of ~2.5 x 10^4 cells/cm^2 into 0.1% gelatin-coated flasks containing endothelial cell growth medium with 2% fetal bovine serum (EGM2 bulletkit, Lonza, Australia) and cultured in standard conditions for up to 21 days with regular monitoring for the spontaneous growth of Endothelial Progenitor Cells (EPCs). Excess PBMCs are immediately frozen in bovine serum (heat-inactivated, Gibco, Australia) containing 10% dimethyl sulfoxide and stored in liquid nitrogen.

Sample Analysis

Samples will be analysed by a variety of methods with the intent of identifying candidate biomarkers and novel metabolites involved in atherosclerosis pathophysiology. There will be two principal approaches used to identify these factors.

First is a candidate approach, investigating currently identified factors that are thought to be involved in human biological pathways that could predispose to atherosclerosis. These factors will be

assessed within the cohort by specific assays that relate to the biological pathways involved. The candidate areas that have been identified for initial assessment are: soluble factors released by inflammatory signalling within the atherosclerotic plaque, dysregulated redox signalling within the endothelium, disorders of coagulation which lead to hypercoagulable states and/or hypofibrinolysis, soluble platelet factors which indicate platelet dysfunction, and dysregulated signalling within endothelial cells as evidenced by assessment of cultured EPCs derived from individual patient samples. Individual assays have been designed to assess each candidate marker or pathway.

The second approach aims to investigate atherosclerosis biology in an unbiased manner by using "omics" techniques. These assessments will utilise metabolomics, proteomics, transcriptomics, genomics including SNP arrays and immunophenotyping techniques.

Proteomics for the cohort will be assessed utilising a clinical proteomics platform[16] which will give a complete description of the human peptidome in relevant samples, allowing for identification of novel or common factors worthy of further investigation. Metabolomics and lipidomics will be performed utilising six liquid chromatograpy-mass spectrometry platforms which will allow for identification and quantitation of a broad spectrum of metabolites, including amino acids, nucleotides, neutrotransmitters and lipid subtypes[17-21].

Immunophenotyping will be performed on PBMC samples via mass spectrometry time of flight (CyTOF) which allows for detailed leucocyte phenotyping on an individual basis[22]. These profiles can be associated with serum levels of established markers of inflammation such as VCAM-1, ICAM-1 and IL-6, as well as correlated with the balance of M1 and M2 phenotype monocytes which have a well established role in atherosclerosis[23].

Transcriptomic and genomic assays will be utilised on subsets of the cohort of particular interest, who for example may display particular resistance or vulnerability to CAD based on traditional risk factor profile and disease burden. Whole transcriptome and RNA sequencing will be performed with ribosomal depletion, and sequencing of the resulting reads and counting will be based on GENCODE. Whole genome profiling will be performed on individuals with extreme athero-susceptible, or athero-resilient phenotype, allowing for unbiased discovery of genetic variants that may be causally associated with the extreme phenotypes[24].

Data Analysis Plan

Statistical machine learning approaches will be applied to these biological and clinical datasets, to identify novel pathophysiological pathways, and to prioritise new markers based on likely causal roles. Discovery analysis will be performed in the first 1000 patients of BioHEART-CT, with validation in the following 4000 patients recruited, and in international collaborators studies (Figure 3.). Sample size planning for developing classifiers using high dimensional data is based on NIH Biometric online calculator assuming 1.3 fold standardised change, 50,000 molecular variables, and balanced extreme groups.

Biomarker and risk factor data analysis will depend on the specifics of the assay or investigation performed. Generally, continuous variables will be presented as means (with standard deviations) or medians (with inter-quartile ranges), categorical variables as proportions (%) with chi-squared or Fisher's exact test used to determine differences between groups. Hazard ratios for a one standard deviation increase in a biomarker, after log transformation to remove effects of outliers, on MACE will be obtained from weighted Cox regression models. All p-values reported will be two-sided, with the 5% threshold used to determine significance.

For testing the prognostic value of prioritised biomarkers, and assuming a C-statistic of 0.75 for traditional risk factors, we estimate that we will require 705 participants to detect a C-statistic of 0.85 with p=0.05, power of 80% and an event rate of 20%. Models developed with different sets of variables and biomarkers and C-statistics for one year risk will be determined and compared. The model's ability to discriminate and reclassify one year risk will be assessed using the integrated discrimination index and the net reclassification improvement.

Discussion

Three main advances have been applied in the design of BioHEART-CT that provide strategic advantage compared to previous studies which also used unbiased approaches to discover novel mechanisms and markers of atherosclerosis. First, this study includes quantification and characterisation of plaque (and the absence of plaque) by CT coronary angiography and advanced imaging algoritims, improving the power of the study compared to those that relied on CAD-related clinical events. Second, this study integrates high through-put, state-of-the-art multi-omic molecular phenotyping with comprehensive clinical data and follow up. And third, the study brings together a global team of biologists and bioinformaticians with complementary approaches and perspectives to enhance the likelihood of new discoveries.

In previous CAD studies, healthy controls were often defined as patients who had not had a cardiac event, or who had non-obstructive atherosclerosis on invasive coronary angiography which has limited utility in identifying extraluminal plaque. As CAD often remains silent until it is very advanced, those without a myocardial infarction could have quite extensive disease and still be considered a "healthy" control. In BioHEART-CT, we have the capability to use our CTCA subgroup to more precisely phenotype subjects, allowing the accurate creation of a group of healthy controls which truly have no atherosclerotic disease present in their coronary arteries. This group of healthy patients can be used as a comparator for those with various stages of disease, reducing confounding factors and enhancing the capacity to detect real differences in biology.

Additionally, CTCA phenotyping can be utilised to analyse plaque composition in detail. Using scoring systems such as GENSINI[25], severity scores can be generated for each individual. We have also developed a modified GENSINI score as outlined above, which gives different weight to the presence of older, calcified plaque and newer, more biologically active non-calcified plaque. Automated detection of soft plaque shows promise as a scalable, reproducible method to add value to existing CTCA analysis[26]. In identifying patients with high levels of active atherosclerotic plaque, we can create groups enriched for this biological feature, with and without traditional risk factors. This will better enable us to detect novel mechanisms of CAD in those with unexplained susceptibility.

Our unbiased approach to cardiovascular risk factor discovery, applying the advantages of high throughput multi-omics platforms, improved quantification of plaque and its characteristics, and advanced statistical bioinformatics, provides a high chance of discovering previously unidentified mechanisms and markers of disease. We have recently discussed the advantages of integrating multi-omics over isolated genomics studies for the complex disease process of atherosclerosis[27], and this will be particularly beneficial in such well-phenotyped groups.

Molecular markers found to be associated with atherosclerosis may be direct drivers of disease, or result from the plaque and dysregulated vascular physiology. The closer the marker is to the underlying biology, the more likely it will be to be useful in clinical practice, transforming our ability to identify early vascular disease and susceptibility above and beyond traditional risk calculators. We are well placed to take any markers and their surrounding pathways into preclinical cellular and

animal models to determine causality and/or mechanisms for association. This will help prioritise markers for testing in prospective clinical trials for their value in improving patient care and outcomes, as well as identifying potential novel signalling pathways involved in disease that may be therapeutic targets.

In conclusion, BioHEART-CT is a study created in response to patients who have had heart attacks and their frequent question from the cardiac catheter laboratory table- "Why me?". While it is imperative that we measure and treat well-established risk factors for atherosclerosis, this approach misses a substantial group of patients with CAD who are not identified as being at risk. We hypothesise that major mechanisms for atherosclerosis remain to be discovered and can be unravelled by comprehensive molecular characterisation, combined with powerful bioinformatic analyses. BioHEART-CT's platforms and team are well positioned for discovering new markers and mechanisms of disease, and to take these both back to the bench to unravel new biology, as well as through to prospective clinical trial to test prognostic and clinical utility.

Ethics and Dissemination

This study was approved by the Human Research Ethics Committee of the North Shore Local Health District, Sydney, Australia – approval number HREC/17/HAWKE/343, and has been registered as a clinical trial on the Australian New Zealand Clinical Trials Registry (ACTRN12618001322224).

Informed consent is obtained from participants prior to enrolment, and the data and samples are de-identified and managed entirely anonymously with the exception of the required information for follow-up phone calls. The most significant risk to the patients in this study is that of venesection, with a possible consequence of mild bruising or superficial infection. Participants can withdraw from the study at any time and this will not have any impact on their clinical care.

The results of this study will be published in peer-reviewed journals and presented at domestic and international scientific meetings. No identifiable information will be published.

Author's Contributions

KK and SV participated in study design, drafted the manuscript and coordinated manuscript revisions.
GF obtained the research funding and is the principal investigator of the study.
All other authors contributed to study design and revisions of the manuscript.

Author's Acknowledgements

The authors would also like to acknowledge in particular the extensive efforts of Luisa Osorio in supporting the BioHEART study. We would also like to extend thanks to the research coorindators and nurses for their hard work across all sites, as well as the nursing staff, radiologists and radiographers at the involved imaging facilties for their assistance. This project would not be possible without their support.

Funding Statement

GF is supported by NHMRC Practitioner Fellowship. SMG thanks the Parker Hughes Bequest for support. This study has received support from a combination of grants including from the Ramsay Teaching and Research Foundation, BioPlatforms Australia, the Vonwiller Foundation and Heart Research Australia.

Competing Interests Statement

None declared

References

1 Roth, G. A. *et al.* Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. *Journal of the American College of Cardiology* **70**, 1-25, doi:https://doi.org/10.1016/j.jacc.2017.04.052 (2017).

2 Wang, H. *et al.* Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **390**, 1084-1150, doi:10.1016/S0140-6736(17)31833-0 (2017).

3 *Australian Bureau of Statistics: Australia's leading causes of death*, <http://www.abs.gov.au/> (2016).

4 Welness, A. I. o. H. a. Australia's Health 2014.

5 Vernon, S. T. *et al.* Increasing proportion of ST elevation myocardial infarction patients with coronary atherosclerosis poorly explained by standard modifiable risk factors. *Eur J Prev Cardiol* **24**, 1824-1830, doi:10.1177/2047487317720287 (2017).

6 Mahmood, S. S., Levy, D., Vasan, R. S. & Wang, T. J. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet* **383**, 999-1008, doi:10.1016/S0140-6736(13)61752-3 (2014).

7 Khot, U. N. *et al.* Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA* **290**, 898-904, doi:10.1001/jama.290.7.898 (2003).

8 Wang, J. Y. *et al.* Cardiovascular Risk Factors and In-hospital Mortality in Acute Coronary Syndromes: Insights From the Canadian Global Registry of Acute Coronary Events. *Canadian Journal of Cardiology* **31**, 1455-1461, doi:10.1016/j.cjca.2015.04.007 (2015).

9 Vos, T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **390**, 1211-1259, doi:10.1016/S0140-6736(17)32154-2 (2017).

10 The, C. D. C. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nature genetics* **45**, 25-33, doi:10.1038/ng.2480 (2013).

11 Ridker, P. M. *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med* **377**, 1119-1131, doi:10.1056/NEJMoa1707914 (2017).

12 Harden, S. *et al.* The safe practice of CT coronary angiography in adult patients in UK imaging departments. *Clinical Radiology* **71**, 6 (2016).

13 Abbara, S. *et al.* SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography: A report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *Journal of Cardiovascular Computed Tomography* **10**, 435-449, doi:10.1016/j.jcct.2016.10.002 (2016).

14 Leipsic, J. *et al.* SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *J Cardiovasc Comput Tomogr* **8**, 342-358, doi:10.1016/j.jcct.2014.07.003 (2014).

15 Riedhammer, C., Halbritter, D. & Weissert, R. Peripheral Blood Mononuclear Cells: Isolation, Freezing, Thawing, and Culture. *Methods in molecular biology (Clifton, N.J.)* **1304**, 53-61, doi:10.1007/7651_2014_99 (2016).

16 Parker, B. L. *et al.* Multiplexed Temporal Quantification of the Exercise-regulated Plasma Peptidome. *Mol Cell Proteomics* **16**, 2055-2068, doi:10.1074/mcp.RA117.000020 (2017).

17 Wang, T. J. *et al.* 2-Aminoadipic acid is a biomarker for diabetes risk. *Journal of Clinical Investigation* **123**, 4309+ (2013).

18 Roberts, L. D. *et al.* beta-Aminoisobutyric acid induces browning of white fat and hepatic beta-oxidation and is inversely correlated with cardiometabolic risk factors. *Cell Metab* **19**, 96-108, doi:10.1016/j.cmet.2013.12.003 (2014).

19 Turer, A. T. *et al.* Increases in myocardial workload induced by rapid atrial pacing trigger alterations in global metabolism. *PLoS One* **9**, e99058, doi:10.1371/journal.pone.0099058 (2014).

20 Patsch, C. *et al.* Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat Cell Biol* **17**, 994-1003, doi:10.1038/ncb3205 (2015).

21 Hinson, J. T. *et al.* Integrative Analysis of PRKAG2 Cardiomyopathy iPS and Microtissue Models Identifies AMPK as a Regulator of Metabolism, Survival, and Fibrosis. *Cell Rep* **17**, 3292-3304, doi:10.1016/j.celrep.2016.11.066 (2016).

22 Lomax, A. J. *et al.* Immunotherapy-induced sarcoidosis in patients with melanoma treated with PD-1 checkpoint inhibitors: Case series and immunophenotypic analysis. *Int J Rheum Dis* **20**, 1277-1285, doi:10.1111/1756-185X.13076 (2017).

23 Guzik, T. J., Skiba, D. S., Touyz, R. M. & Harrison, D. G. The role of infiltrating immune cells in dysfunctional adipose tissue. *Cardiovasc Res* **113**, 1009-1023, doi:10.1093/cvr/cvx108 (2017).

24 Warren, C. R. *et al.* Induced Pluripotent Stem Cell Differentiation Enables Functional Validation of GWAS Variants in Metabolic Disease. *Cell Stem Cell* **20**, 547-557 e547, doi:10.1016/j.stem.2017.01.010 (2017).

25 Gensini, G. G. A more meaningful scoring system for determining the severity of coronary heart disease. *American Journal of Cardiology* **51**, 606, doi:10.1016/S0002-9149(83)80105-2 (1983).

26 Tesche, C. *et al.* Coronary CT angiography derived morphological and functional quantitative plaque markers correlated with invasive fractional flow reserve for detecting hemodynamically significant stenosis. *Journal of Cardiovascular Computed Tomography* **10**, 199-206, doi:10.1016/j.jcct.2016.03.002 (2016).

27 Vernon, S. T. *et al.* Utilizing state-of-the-art “omics” technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease. *Microcirculation* **0**, e12488, doi:doi:10.1111/micc.12488.

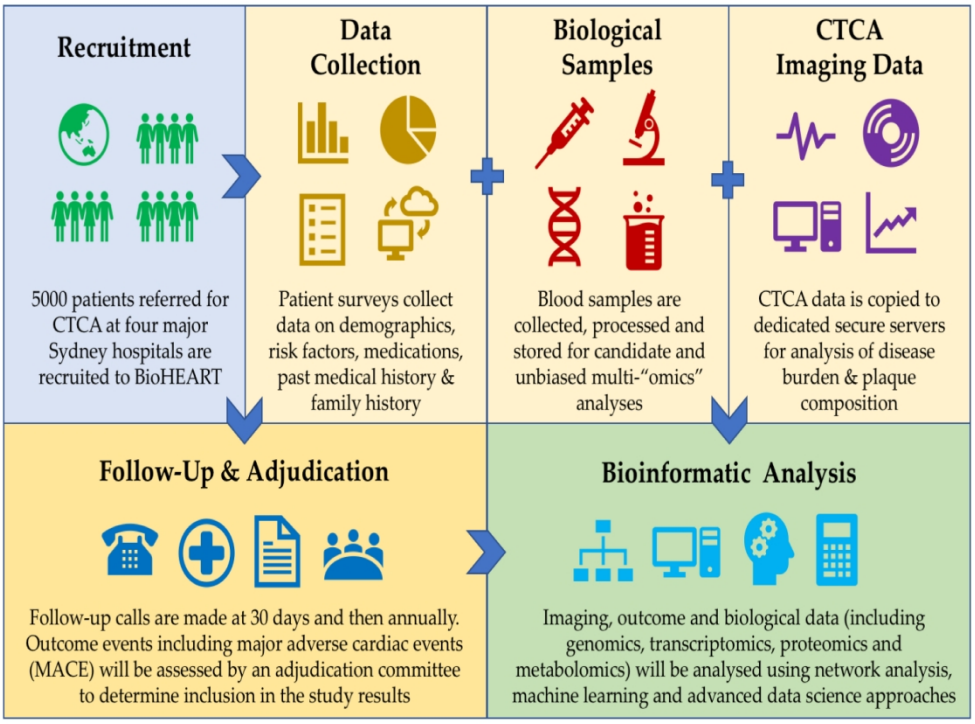


Figure 1: BioHEART-CT Study Design

359x269mm (110 x 110 DPI)

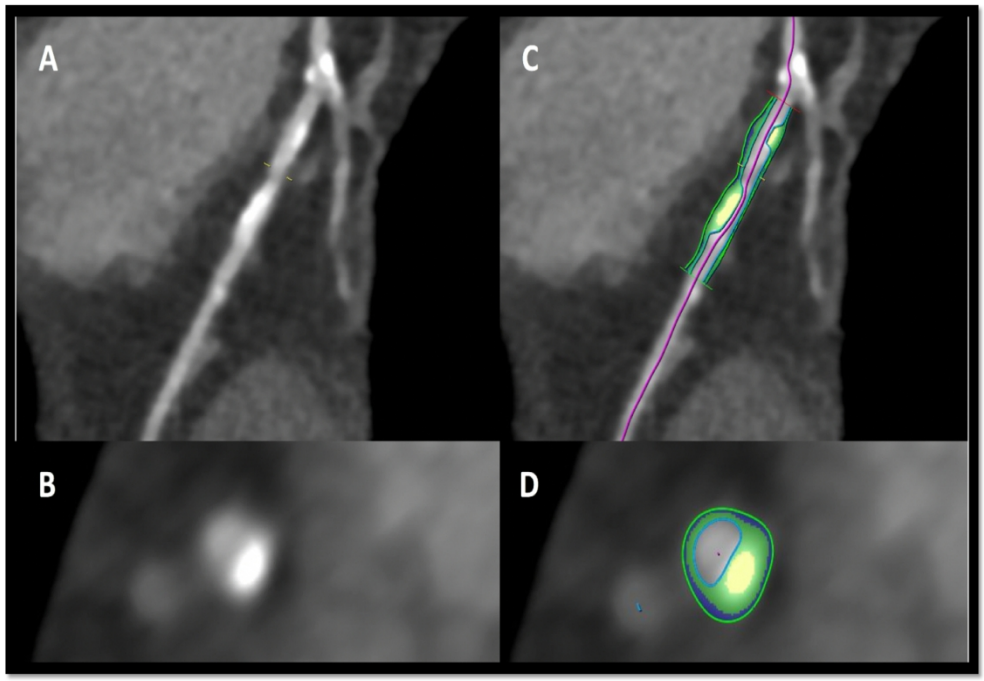


Figure 2: Example of semi-automated analysis of a complex plaque using syngo.via running on a Frontier system (Siemens Healthcare, Erlangen). Panels A & B show a coronary artery lesion with a large calcified component longitudinally and in cross-section. Panels C & D show classification of the calcific (yellow) and fibrous (green) components of this lesion

359x269mm (110 x 110 DPI)

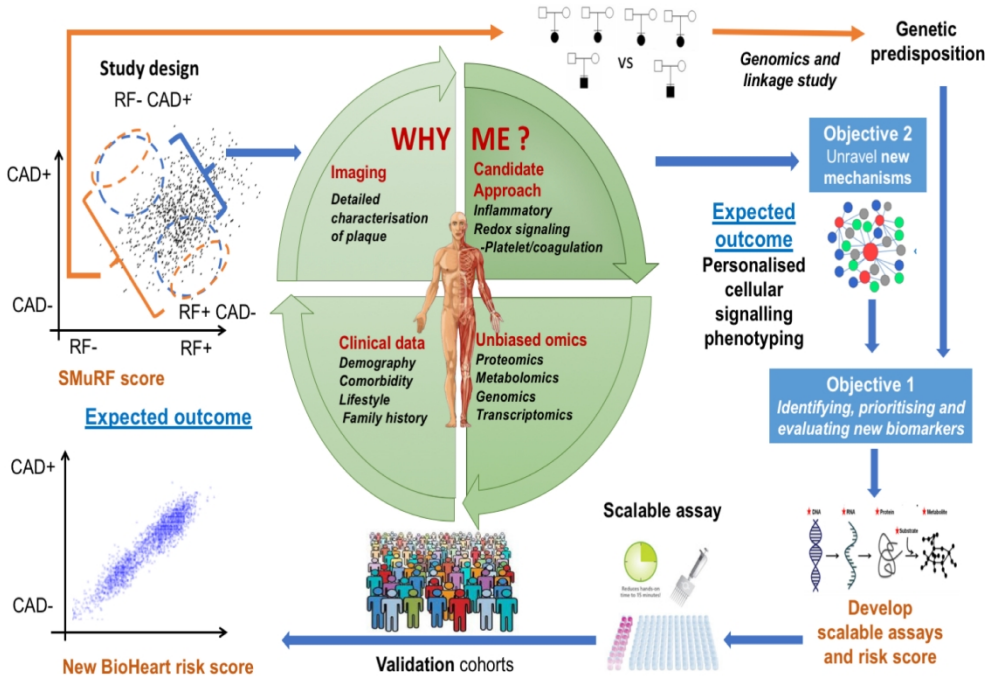


Figure 3: Overview of the Analytic Approach.
Precise phenotype characteristisation including advanced plaque quantification and analysis of CTCA imaging and clinical data will be utilised in candidate and unbiased omics based assessments. Machine learning approaches will be applied to: identify new biomarkers, biological mechanisms of atherosclerosis, and new risk scores.

359x269mm (110 x 110 DPI)

BMJ Open

Biobanking for discovery of novel cardiovascular biomarkers using imaging-quantified disease burden: protocol for the longitudinal, prospective, BioHEART-CT cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-028649.R1
Article Type:	Protocol
Date Submitted by the Author:	28-Jun-2019
Complete List of Authors:	Kott, Katharine; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Vernon, Stephen; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Hansen, Thomas ; Kolling Institute of Medical Research Yu, Christine; Kolling Institute of Medical Research Bubb, Kristen ; University of Sydney, Northern Clinical School, Faculty of Medicine and Health Coffey, Sean; Royal North Shore Hospital, Kolling Institute and Department of Cardiology; University of Otago, Dunedin School of Medicine Sullivan, David; The University of Sydney Yang, Jean ; University of Sydney, School of Mathematics and Statistics O'Sullivan, John; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Chow, Clara; The George Institute for Global Health and Westmead Hospital, Sydney Patel, Sanjay; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Chong, James ; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Celermajer, David; University of Sydney, Faculty of Medicine and Health; Royal Prince Alfred Hospital, Department of Cardiology Kritharides, Leonard; Concord Hospital, Cardiology Grieve, Stuart; University of Sydney, Sydney Medical School, Faculty of Medicine and Health Figtree, Gemma; University of Sydney, Faculty of Medicine and Health; Kolling Institute of Medical Research,
Primary Subject Heading:	Cardiovascular medicine
Secondary Subject Heading:	Patient-centred medicine, Radiology and imaging
Keywords:	Coronary heart disease < CARDIOLOGY, Cardiovascular imaging < RADIOLOGY & IMAGING, Ischaemic heart disease < CARDIOLOGY, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, MOLECULAR BIOLOGY, Clinical chemistry < PATHOLOGY

Biobanking for discovery of novel cardiovascular biomarkers using imaging-quantified disease burden: protocol for the longitudinal, prospective, BioHEART-CT cohort study

Katharine A. Kott*[1,2], Stephen T. Vernon*[1,2], Thomas Hansen[1], Christine Yu[1], Kristen J. Bubb[1], Sean Coffey[1,5], David Sullivan[2,3,6], Jean Y. Yang[3,4], John F. O'Sullivan[2,3,9], Clara K Chow[2,10], Sanjay Patel[2,3,7,9], James Chong[2,10], David S. Celermajer[2,7,9], Leonard Kritharides[2,11], Stuart M. Grieve[2,3,8], Gemma A. Figtree[1,2]

[1]Cardiothoracic and Vascular Health, Kolling Institute and Department of Cardiology, Royal North Shore Hospital, Northern Sydney Local Health District, Australia
[2]Sydney Medical School, Faculty of Medicine and Health, University of Sydney, Australia
[3]Charles Perkins Centre, University of Sydney, Australia
[4]School of Mathematics and Statistics, University of Sydney, Australia
[5]Dunedin School of Medicine, University of Otago, New Zealand
[6]Department of Biochemistry, Royal Prince Alfred Hospital, Sydney, Australia
[7]Department of Cardiology, Royal Prince Alfred Hospital, Sydney, Australia
[8]Department of Radiology, Royal Prince Alfred Hospital, Sydney, Australia
[9]The Heart Research Institute, Sydney, Australia
[10]Westmead Institute for Medical Research & Cardiology Department, Westmead Hospital, Sydney, Australia
[11]ANZAC Research Institute & Cardiology Department, Concord Repatriation General Hospital, Sydney, Australia
*The first two authors contributed equally to this publication

Corresponding author:

Professor Gemma Figtree MBBS DPhil (Oxon) FRACP FAHA

Cardiothoracic and Vascular Health, Level 12, Kolling Institute of Medical Research, Royal North Shore Hospital, St Leonards, NSW, 2065, Australia. Email: gemma.figtree@sydney.edu.au.

Word Count: 3834

Abstract

Introduction: Coronary artery disease (CAD) persists as a major cause of morbidity and mortality worldwide despite intensive identification and treatment of traditional risk factors. Data emerging over the past decade shows a quarter of patients have disease in the absence of any known risk factor, and half have only one risk factor. Improvements in quantification and characterisation of coronary atherosclerosis by CT coronary angiography (CTCA) can provide quantitative measures of subclinical atherosclerosis - enhancing the power of unbiased "omics" studies to unravel the missing biology of personal susceptibility, identify new biomarkers for early diagnosis, and to suggest new targeted therapeutics.

Methods & Analysis: BioHEART-CT is a longitudinal, prospective cohort study, aiming to recruit 5000 adult patients undergoing clinically-indicated CTCA. After informed consent, patient data, blood samples and CTCA imaging data are recorded. Follow-up for all patients is conducted one month after recruitment, and then annually for the life of the study. CTCA data provides volumetric quantification of total calcified and non-calcified plaque, which will be assessed using established and novel scoring systems. Comprehensive molecular phenotyping will be performed using state-of-the-art genomics, metabolomics, proteomics, and immunophenotyping. Complex network and machine learning approaches will be applied to biological and clinical datasets to identify novel pathophysiological pathways and to prioritise new biomarkers. Discovery analysis will be performed in the first 1000 patients of BioHEART-CT, with validation analysis in the following 4000 patients. Outcome data will be utilised to build improved risk models for CAD.

Ethics & Dissemination: The study protocol has been approved by the human research ethics committee of North Shore Local Health District in Sydney, Australia. All findings will be published in peer-reviewed journals or at scientific conferences.

Registration: This study is registered with the Australia and New Zealand Clinical Trials Registry - ACTRN12618001322224.

Article Summary

Strengths and Limitations of this Study (max five bullet points)

- BioHEART-CT is a prospective, longitudinal cohort study assessing patients with suspected CAD undergoing CTCA across multiple Australian centres.
- Quantitative measures of coronary atherosclerosis from CTCA datasets will be used in conjunction with biological samples, improving the power to discover new mechanisms and markers for CAD.
- Samples stratified by imaging-quantified disease severity will be analysed by both candidate and unbiased “omics” approaches, utilising modern technologies and bioinformatic analysis to discover new biomarkers using proteomics, metabolomics, lipidomics, transciptomics, genomics and immunophenotyping.
- Longitudinal outcome data will be used to build risk models which can incorporate both traditional and novel risk factors and biomarkers.
- Limitations of the study include selection bias, as only patients with a clinical indication for CTCA are included, and the geographic isolation of Australia, which may result in the need for confirmation of results in multi-national studies.

Introduction

Cardiovascular disease has persisted as a major cause of human morbidity and mortality despite continual improvements in preventative therapies. In 2015 cardiovascular disease accounted for one-third of all deaths world-wide[1], and ischaemic heart disease (IHD) remains the leading cause of years of life lost for high and middle sociodemographic groups multi-nationally[2]. This large burden of disease puts significant stress on the health systems of individual countries. In Australia, cardiovascular disease remains the number one killer of Australians[3], with local health costs estimated to be $7.7 billion AUD per annum[4].

The standard modifiable cardiovascular risk factors (SMuRFs[5]) for atherosclerosis - hypertension, diabetes mellitus, hyperlipidaemia and smoking - were identified in epidemiological studies in the 1960s[6]. The identification and subsequent efforts to target these risk factors at community and primary care levels have led to a substantial reduction in mortality from cardiovascular disease[1]. However, substantial disease burden remains and importantly, individual variability in susceptibility to these risk factors is considerable. Conditions such as impaired glucose tolerance and pre-hypertension may also contribute to cardiovascular risk. While continuing our efforts to tackle societal and modifiable risk factors, identifiying undiscovered mechanisms that lead to the development of atherosclerosis is critical. Such work will provide new biomarkers for early detection of subclinical atherosclerosis and open avenues for improved preventative and therapeutic strategies that may be relevant to those with and without known risk factors.

The importance of new detection mechanisms for subclinical atherosclerosis is highlighted by interrogating the “fine print” of clinical studies, including information not usually presented in the main tables of mansuscripts. Although data regarding the percentage of patients suffering myocardial infarction despite having no SMuRFs is often omitted, a large meta-analysis found that >50% of women and >60% of men presenting with their first ACS had 0 or 1 SMuRF[7], and similar proportions of SMuRFless patients have been identified in ACS cohort studies [5,8]. In a recent single centre Australian study, we showed that the proportion of patients presenting with ST-elevation myocardial infarction with 0 SMuRFs has risen from 11% to 27% over an 8 year period[5]. These patients have developed extensive CAD without having any of the red flags which would allow doctors to identify their risk of disease and intervene early. Although this is a relatively small piece of the overall IHD pie, the global burden of CAD is immense with an estimated prevalence of 470 million[9], which makes this 25-30% SMuRFless patient population important from a public health perspective. It is sobering to think that we have made no clear advances that allow us to identify risk and subclinical disease in this subgroup. Indeed, we also have limited information regarding how these patients respond to traditional secondary preventative strategies, and whether their long term outcomes and disease progression differs from those with traditional modifiable risk factors.

Further evidence that novel biological mechanisms contribute substantially to atherosclerosis can be seen in results from large scale genome-wide association studies which have reported that 66% of the identified genes conveying heritable cardiovascular risk are not associated with the traditional risk factors[10]. Many of these genes were related to inflammatory processes, and studies such as the CANTOS trial[11] confirm the importance of inflammation in CAD. However, some of these important genes have not yet been well characterised, and are not yet associated with any known pathway in atherosclerosis and CAD, highlighting the importance of ongoing discovery work.

To address this gap in our understanding of CAD pathophysiology, we have designed a unique cohort study with biobanked samples to facilitate investigation of novel cardiac risk factors and biomarkers.

This biobank includes advanced, quantitative imaging measurements of coronary artery atherosclerosis to allow for accurate phenotyping of the patient groups. This overcomes a weaknesses of previous genomics studies that have relied on the presence or absence of a coronary artery event to classify individuals as diseased or as a control, improving the statistical power of the subsequent analysis. The advanced imaging dataset and clinical data including baseline and follow-up are integrated with the collection of a broad range of biological samples for discovery work, utilising state-of-the-art molecular phenotyping platforms and computational bioinformatics.

Objectives

The primary objectives of the BioHEART-CT study are:

1. To identify new biomarkers to assist in early CAD risk identification and stratification;
2. To identify new mechanistic pathways that may be targeted with novel therapeutic strategies to abrogate CAD risk.

The secondary objectives of the BioHEART-CT study include the following:

1. To determine the predictive value of a modified CTCA scoring system for CAD risk which incorporates plaque composition data;
2. To identify new risk factors for CAD based on integration of clinical information profiles, CTCA results and clinical outcomes;
3. To develop a novel risk scoring system incorporating clinical risk factors, novel biomarkers and CTCA scores to more accurarely predict CVD events.

Methods and Analysis

Patient and Public Involvement

The initial concept for the BioHEART-CT study was prompted by patients who had no known cardiovascular risk factors who presented with ACS and wanted to know “why me?” The ensuing discussions with the patients and their familes reinforced the public interest in research in this area. Informal consultation with patient groups through cardiovascular non-profit organisations helped build the initial framework for the study, and patient feedback regarding the recruitment process was sought and was positive in nature. When results of interest to the general public are available, summaries will be posted to the Australia and New Zealand Clinical Trials Registry. Information about this registry is included on the patient information sheet.

Study Type

BioHEART-CT is a multicentre, prospective cohort study of patients with suspected CAD, which brings together detailed clinical information with advanced imaging and molecular phenotyping. Recruitment commenced in 2017 and is ongoing, currently with expansion into a multi-centre study.

Study Population

5000 patients will be recruited from participating hospitals’ and associated imaging sites. All recruitment centres are large tertiary hospitals within the city of Sydney, Australia.

Inclusion criteria:

- Patients who have been referred for investigation of suspected CAD
- Age 18 years or older

- Willing and able to provide informed consent

Exclusion criteria:

- Patients highly dependent on medical care who are unable to provide informed consent
- Patients unwilling or unable to participate in ongoing follow-up

Study Protocol

Screening and Enrolment

Eligible patients having a CTCA are invited to participate in the study. Eligibility is confirmed according to the inclusion and exclusion criteria, the study is explained, and informed consent obtained (Figure 1).

Patient Data Collection

The baseline questionnaire is completed by research staff conducting a face-to-face patient interview. Data obtained includes the following information: demographic data including smoking and alcohol intake, past medical history, relevant family history, medication history, history of cardiac symptoms and indication for CTCA. Data is entered into secure, encrypted databases in a de-identified format. After recruitment all patients are assigned a unique study number and all samples, imaging data and demographic information is identified by this number. The master list including identifiable data is securely stored on an encrypted server and is available only to senior research staff. Relevant details are made available to research staff for follow-up purposes.

Follow-Up & Outcome Adjudication

Participants are contacted by phone at approximately 30 days and then annually to determine if any events have occurred. Information recorded for any potential outcome event includes the acuity, need for hospitalisation and indication for admission, hospital attended, treating physician, date of admission and details of the events. New medical or surgical diagnoses since last contact are also recorded. Participants in BioHEART also consent to review of their electronic medical records and data linkage through the Centre for Health Record Linkage (CHeReL), which will be used for independent medical record reviews.

Significant outcome events that trigger a file review to assess for major adverse cardiovascular events (MACE) or secondary outcomes are the following:

- Unstable angina requiring hospitalisation
- Heart failure requiring hospitalisation
- Acute myocardial infarction
- Coronary angiogram with or without percutaneous intervention
- Cardiac bypass graft surgery
- Aortic or mitral valve surgery
- Cerebrovascular accident / transient ischaemic attack
- All-cause mortality

All medical records will be reviewed by an independent adjudication committee to determine if an event should be included in the study results. Detailed criteria for the individual diagnoses included in the outcomes must be met.

MACE is defined as cardiovascular death, non-fatal myocardial infarction or non-fatal stroke. Exploratory outcomes include revascularisation, unstable angina or heart failure requiring hospitalisation.

CTCA Data Acquisition

CTCA scans are obtained on 256 slice scanner, with reconstructions created using appropriate software for the individual machine. Each study is to be protocolled by a radiologist with at least level two accreditation (RANZCR/RANZCP). All radiographers are trained in acquisition and workup of CTCA. Oral metoprolol, or ivabradine if beta blockers are contraindicated, are given to optimise the heart rate if required. Dosing is adjusted by clinical staff according to baseline heart rate and bodyweight. Oral nitroglycerine (600-800 micrograms) is given immediately prior to the scan, and iodinated contrast is injected intravenously. Prospective studies are performed if the heart rate is sufficiently controlled, otherwise retrospective acquisition is used. Average dose is minimised at all participating sites in line with current recommendations[12,13].

CTCA Imaging Analysis

CT data is exported as thin DICOM data and is stored securely in a de-identified manner. Data is later analysed using a dedicated workstation to obtain the primary clinical scores at each site. CTCAs are analysed by standard anatomical arterial region as per the 17-segment model outlined by the Society of Cardiovascular Computed Tomography[14]. Each segment is scored according to degree of stenosis and composition of plaque (calcified, soft and mixed). The segmental scores are aggregated into a modified GENSINI score (Figure 2) which represents the total amount of calcified and non-calcified plaque present. Calcium scores are generated using the scanner-associated software by the Agaston method. Total raw calcium scores, raw calcium scores for each vessel, and age and sex-matched calcium percentage scores are recorded.

Further analysis will be performed at a central core laboratory using a Frontier System (syngo.via, Siemens Healthcare, Erlangen), employing a semi-automated soft plaque algorithm (Figure 3).

Biological Sample Collection

A 20-30mL blood sample is taken from an in-situ intravenous cannula, or if this is not available, by venesection, prior to CTCA and drug administration. Blood is collected into EDTA, serum and citrate tubes and stored at 4°C until processing. A lithium heparin tube is collected and processed immediately to extract peripheral blood mononuclear cells (PBMCs) from the buffy coat.

Sample Processing & Storage

Blood samples are processed by hospital pathology scientists experienced in laboratory procedures. Briefly, whole blood is set aside, then tubes are centrifuged at 1861 x *g* for 15 minutes at 4°C prior to aliquoting 500uL samples of serum, EDTA plasma, whole blood, buffy coat and erythrocytes. The citrate tube is centrifuged at 1861 x *g* for a second 15 minutes at 4°C, after which citrate-plasma samples are aliquoted. All aliquoted samples are frozen and stored at -80°C until analysis.

For PBMC isolation a standard gradient-separation protocol is used[15]. Briefly, blood from the lithium heparin tube is diluted 1:1 with Hanks' Balanced Salt Solution (HBSS). Diluted whole blood is then layered on top of Ficoll-Paque Plus and centrifgued (22°C, 1460 x *g* no brake, 20 min) . The buffy coat layer containing PBMCs is then further purified with two wash steps in HBSS using repeat centrifugation as above. PBMCs are then plated at a density of ~2.5 x 10^4 cells/cm^2 into 0.1% gelatin-coated flasks containing endothelial cell growth medium with 2% fetal bovine serum (EGM2

bulletkit, Lonza, Australia) and cultured in standard conditions for up to 21 days with regular monitoring for the spontaneous growth of Endothelial Progenitor Cells (EPCs). Excess PBMCs are immediately frozen in bovine serum (heat-inactivated, Gibco, Australia) containing 10% dimethyl sulfoxide and stored in liquid nitrogen.

Sample Analysis

Samples will be analysed by a variety of methods with the intent of identifying candidate biomarkers and novel metabolites involved in atherosclerosis pathophysiology. There will be two principal approaches used to identify these factors. Table 1 provides a summary of all variables collected and planned assays.

TABLE 1 Summary of variables and planned assays

Clinical Variables	Disease Quantitation Variables	Established Markers	Unbiased -omics Approaches	Candidate Approaches
Demographics Standard modifiable cardiovascular risk factors (SMuRFs) - Hypertension - Hyperlipidaemia - Diabetes - Smoking Other medical history BMI Family history of IHD Current medications History of cardiac symptoms	Coronary artery calcium scores (CACS) GENSINI score Modified GENSINI score (see figure 2) Frontier System semi-automated plaque analysis (syngo.via, Siemens Healthcare, Erlangen)	Cardiac - Troponin - NT-BNP Inflammatory - CRP - VCAM-1 - ICAM-1 - IL-6	Metabolomics Proteomics Lipidomics Transcriptomics Genomics Immunophenotyping	Redox signalling dysregulation Endothelial cell signalling dysregulation Soluble factors released by inflammatory plaque Apopotosis signalling Mitochondrial function Angiogenesis potential Disorders of overall coagulation Soluble platelet factors Alpha-gal antibodies Heavy metal toxicities Galectin-3

First is a candidate approach, investigating currently identified factors that are thought to be involved in human biological pathways that could predispose to atherosclerosis. These factors will be assessed within the cohort by specific assays that relate to the biological pathways involved. The candidate areas that have been identified for initial assessment are: soluble factors released by inflammatory signalling within the atherosclerotic plaque, dysregulated redox signalling within the endothelium, disorders of coagulation which lead to hypercoagulable states and/or hypofibrinolysis, soluble platelet factors which indicate platelet dysfunction, and dysregulated signalling within

endothelial cells as evidenced by assessment of cultured EPCs derived from individual patient samples. Individual assays have been designed to assess each candidate marker or pathway.

The second approach aims to investigate atherosclerosis biology in an unbiased manner by using "omics" techniques. These assessments will utilise metabolomics, proteomics, transcriptomics, genomics including SNP arrays and immunophenotyping techniques. The cells utilised for transcriptomcs will be patient-derived EPCs and PBMCs.

Proteomics for the cohort will be assessed utilising a clinical proteomics platform[16] which will give a complete description of the human peptidome in relevant samples, allowing for identification of novel or common factors worthy of further investigation. Metabolomics and lipidomics will be performed utilising six liquid chromatograpy-mass spectrometry platforms which will allow for identification and quantitation of a broad spectrum of metabolites, including amino acids, nucleotides, neutrotransmitters and lipid subtypes[17-21].

Immunophenotyping will be performed on PBMC samples via mass spectrometry time of flight (CyTOF) which allows for detailed leucocyte phenotyping on an individual basis[22]. These profiles can be associated with serum levels of established markers of inflammation such as VCAM-1, ICAM-1 and IL-6, as well as correlated with the balance of M1 and M2 phenotype monocytes which have a well established role in atherosclerosis[23].

Transcriptomic and genomic assays will be utilised on subsets of the cohort of particular interest, who for example may display particular resistance or vulnerability to CAD based on traditional risk factor profile and disease burden. Whole transcriptome and RNA sequencing will be performed with ribosomal depletion, and sequencing of the resulting reads and counting will be based on GENCODE. Whole genome profiling will be performed on individuals with extreme athero-susceptible, or athero-resilient phenotype, allowing for unbiased discovery of genetic variants that may be causally associated with the extreme phenotypes[24].

Data Analysis Plan

Statistical machine learning approaches will be applied to these biological and clinical datasets, to identify novel pathophysiological pathways, and to prioritise new markers based on likely causal roles. Discovery analysis will be performed in the first 1000 patients of BioHEART-CT, with validation in the following 4000 patients recruited, and in international collaborators studies (Figure 4). Sample size planning for developing classifiers using high dimensional data is based on NIH Biometric online calculator assuming 1.3 fold standardised change, 50,000 molecular variables, and balanced extreme groups.

Biomarker and risk factor data analysis will depend on the specifics of the assay or investigation performed. Generally, continuous variables will be presented as means (with standard deviations) or medians (with inter-quartile ranges), categorical variables as proportions (%) with chi-squared or Fisher's exact test used to determine differences between groups. Hazard ratios for a one standard deviation increase in a biomarker, after log transformation to remove effects of outliers, on MACE will be obtained from weighted Cox regression models. All p-values reported will be two-sided, with the 5% threshold used to determine significance.

For testing the prognostic value of prioritised biomarkers, and assuming a C-statistic of 0.75 for traditional risk factors, we estimate that we will require 705 participants to detect a C-statistic of 0.85 with p=0.05, power of 80% and an event rate of 20%. Models developed with different sets of variables and biomarkers and C-statistics for one year risk will be determined and compared. The

model's ability to discriminate and reclassify one year risk will be assessed using the integrated discrimination index and the net reclassification improvement.

Discussion

Three main advances have been applied in the design of BioHEART-CT that provide strategic advantage compared to previous studies which also used unbiased approaches to discover novel mechanisms and markers of atherosclerosis. First, this study includes quantification and characterisation of plaque (and the absence of plaque) by CT coronary angiography and advanced imaging algoritims, improving the power of the study compared to those that relied on CAD-related clinical events. Second, this study integrates high through-put, state-of-the-art multi-omic molecular phenotyping with comprehensive clinical data and follow up. And third, the study brings together a global team of biologists and bioinformaticians with complementary approaches and perspectives to enhance the likelihood of new discoveries.

In previous CAD studies, healthy controls were often defined as patients who had not had a cardiac event, or who had non-obstructive atherosclerosis on invasive coronary angiography which has limited utility in identifying extraluminal plaque. As CAD often remains silent until it is very advanced, those without a myocardial infarction could have quite extensive disease and still be considered a "healthy" control. In BioHEART-CT, we have the capability to use our CTCA subgroup to more precisely phenotype subjects, allowing the accurate creation of a group of healthy controls which truly have no atherosclerotic disease present in their coronary arteries. This group of healthy patients can be used as a comparator for those with various stages of disease, reducing confounding factors and enhancing the capacity to detect real differences in biology.

Additionally, CTCA phenotyping can be utilised to analyse plaque composition in detail. Using scoring systems such as GENSINI[25], severity scores can be generated for each individual. We have also developed a modified GENSINI score as outlined above, which gives different weight to the presence of older, calcified plaque and newer, more biologically active non-calcified plaque. Automated detection of soft plaque shows promise as a scalable, reproducible method to add value to existing CTCA analysis[26]. In identifying patients with high levels of active atherosclerotic plaque, we can create groups enriched for this biological feature, with and without traditional risk factors. This will better enable us to detect novel mechanisms of CAD in those with unexplained susceptibility.

Our unbiased approach to cardiovascular risk factor discovery, applying the advantages of high throughput multi-omics platforms, improved quantification of plaque and its characteristics, and advanced statistical bioinformatics, provides a high chance of discovering previously unidentified mechanisms and markers of disease. We have recently discussed the advantages of integrating multi-omics over isolated genomics studies for the complex disease process of atherosclerosis[27], and this will be particularly beneficial in such well-phenotyped groups.

Molecular markers found to be associated with atherosclerosis may be direct drivers of disease, or result from the plaque and dysregulated vascular physiology. The closer the marker is to the underlying biology, the more likely it will be to be useful in clinical practice, transforming our ability to identify early vascular disease and susceptibility above and beyond traditional risk calculators. We are well placed to take any markers and their surrounding pathways into preclinical cellular and animal models to determine causality and/or mechanisms for association. This will help prioritise markers for testing in prospective clinical trials for their value in improving patient care and outcomes, as well as identifying potential novel signalling pathways involved in disease that may be therapeutic targets.

In conclusion, BioHEART-CT is a study created in response to patients who have had heart attacks and their frequent question from the cardiac catheter laboratory table- "Why me?". While it is imperative that we measure and treat well-established risk factors for atherosclerosis, this approach misses a substantial group of patients with CAD who are not identified as being at risk. We hypothesise that major mechanisms for atherosclerosis remain to be discovered and can be unravelled by comprehensive molecular characterisation, combined with powerful bioinformatic analyses. BioHEART-CT's platforms and team are well positioned for discovering new markers and mechanisms of disease, and to take these both back to the bench to unravel new biology, as well as through to prospective clinical trial to test prognostic and clinical utility.

Ethics and Dissemination

This study was approved by the Human Research Ethics Committee of the North Shore Local Health District, Sydney, Australia – approval number HREC/17/HAWKE/343, and has been registered as a clinical trial on the Australian New Zealand Clinical Trials Registry (ACTRN12618001322224).

Informed consent is obtained from participants prior to enrolment, and the data and samples are de-identified and managed entirely anonymously with the exception of the required information for follow-up phone calls. The most significant risk to the patients in this study is that of venesection, with a possible consequence of mild bruising or superficial infection. Participants can withdraw from the study at any time and this will not have any impact on their clinical care.

The results of this study will be published in peer-reviewed journals and presented at domestic and international scientific meetings. No identifiable information will be published.

Author's Contributions

KK and SV participated in study design, drafted the manuscript and coordinated manuscript revisions.
GF obtained the research funding and is the principal investigator of the study.
All other authors (TH, CY, KB, SC, DS, JY, JO, CC, SP, JC, DS, LK, SG) contributed to study design and revisions of the manuscript.

Author's Acknowledgements

The authors would also like to acknowledge in particular the extensive efforts of Luisa Osorio in supporting the BioHEART study. We would like to thank the patients and their families who advised us on the importance of the research and the basic study design, and to all the patients who have participated in the trial. We would also like to extend thanks to the research coorindators and nurses for their hard work across all sites, as well as the nursing staff, radiologists and radiographers at the involved imaging facilties for their assistance. This project would not be possible without their support. We would also like to acknowledge the Vonwiller Foundation for their support and advice on this project.

Funding Statement

GF is supported by NHMRC Practitioner Fellowship. SMG thanks the Parker Hughes Bequest for support. This study has received support from a combination of grants including from the Ramsay

Teaching and Research Foundation, BioPlatforms Australia, the Vonwiller Foundation and Heart Research Australia.

Competing Interests Statement

None declared

References

1 Roth, G. A. *et al.* Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. *Journal of the American College of Cardiology* **70**, 1-25, doi:https://doi.org/10.1016/j.jacc.2017.04.052 (2017).

2 Wang, H. *et al.* Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **390**, 1084-1150, doi:10.1016/S0140-6736(17)31833-0 (2017).

3 *Australian Bureau of Statistics: Australia's leading causes of death*, <http://www.abs.gov.au/> (2016).

4 Welness, A. I. o. H. a. Australia's Health 2014.

5 Vernon, S. T. *et al.* Increasing proportion of ST elevation myocardial infarction patients with coronary atherosclerosis poorly explained by standard modifiable risk factors. *Eur J Prev Cardiol* **24**, 1824-1830, doi:10.1177/2047487317720287 (2017).

6 Mahmood, S. S., Levy, D., Vasan, R. S. & Wang, T. J. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet* **383**, 999-1008, doi:10.1016/S0140-6736(13)61752-3 (2014).

7 Khot, U. N. *et al.* Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA* **290**, 898-904, doi:10.1001/jama.290.7.898 (2003).

8 Wang, J. Y. *et al.* Cardiovascular Risk Factors and In-hospital Mortality in Acute Coronary Syndromes: Insights From the Canadian Global Registry of Acute Coronary Events. *Canadian Journal of Cardiology* **31**, 1455-1461, doi:10.1016/j.cjca.2015.04.007 (2015).

9 Vos, T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **390**, 1211-1259, doi:10.1016/S0140-6736(17)32154-2 (2017).

10 The, C. D. C. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nature genetics* **45**, 25-33, doi:10.1038/ng.2480 (2013).

11 Ridker, P. M. *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med* **377**, 1119-1131, doi:10.1056/NEJMoa1707914 (2017).

12 Harden, S. *et al.* The safe practice of CT coronary angiography in adult patients in UK imaging departments. *Clinical Radiology* **71**, 6 (2016).

13 Abbara, S. *et al.* SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography: A report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *Journal of Cardiovascular Computed Tomography* **10**, 435-449, doi:10.1016/j.jcct.2016.10.002 (2016).

14 Leipsic, J. *et al.* SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *J Cardiovasc Comput Tomogr* **8**, 342-358, doi:10.1016/j.jcct.2014.07.003 (2014).

15 Riedhammer, C., Halbritter, D. & Weissert, R. Peripheral Blood Mononuclear Cells: Isolation, Freezing, Thawing, and Culture. *Methods in molecular biology (Clifton, N.J.)* **1304**, 53-61, doi:10.1007/7651_2014_99 (2016).

16 Parker, B. L. *et al.* Multiplexed Temporal Quantification of the Exercise-regulated Plasma Peptidome. *Mol Cell Proteomics* **16**, 2055-2068, doi:10.1074/mcp.RA117.000020 (2017).

17 Wang, T. J. *et al.* 2-Aminoadipic acid is a biomarker for diabetes risk. *Journal of Clinical Investigation* **123**, 4309+ (2013).

18 Roberts, L. D. *et al.* beta-Aminoisobutyric acid induces browning of white fat and hepatic beta-oxidation and is inversely correlated with cardiometabolic risk factors. *Cell Metab* **19**, 96-108, doi:10.1016/j.cmet.2013.12.003 (2014).

19 Turer, A. T. *et al.* Increases in myocardial workload induced by rapid atrial pacing trigger alterations in global metabolism. *PLoS One* **9**, e99058, doi:10.1371/journal.pone.0099058 (2014).

20 Patsch, C. *et al.* Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat Cell Biol* **17**, 994-1003, doi:10.1038/ncb3205 (2015).

21 Hinson, J. T. *et al.* Integrative Analysis of PRKAG2 Cardiomyopathy iPS and Microtissue Models Identifies AMPK as a Regulator of Metabolism, Survival, and Fibrosis. *Cell Rep* **17**, 3292-3304, doi:10.1016/j.celrep.2016.11.066 (2016).

22 Lomax, A. J. *et al.* Immunotherapy-induced sarcoidosis in patients with melanoma treated with PD-1 checkpoint inhibitors: Case series and immunophenotypic analysis. *Int J Rheum Dis* **20**, 1277-1285, doi:10.1111/1756-185X.13076 (2017).

23 Guzik, T. J., Skiba, D. S., Touyz, R. M. & Harrison, D. G. The role of infiltrating immune cells in dysfunctional adipose tissue. *Cardiovasc Res* **113**, 1009-1023, doi:10.1093/cvr/cvx108 (2017).

24 Warren, C. R. *et al.* Induced Pluripotent Stem Cell Differentiation Enables Functional Validation of GWAS Variants in Metabolic Disease. *Cell Stem Cell* **20**, 547-557 e547, doi:10.1016/j.stem.2017.01.010 (2017).

25 Gensini, G. G. A more meaningful scoring system for determining the severity of coronary heart disease. *American Journal of Cardiology* **51**, 606, doi:10.1016/S0002-9149(83)80105-2 (1983).

26 Tesche, C. *et al.* Coronary CT angiography derived morphological and functional quantitative plaque markers correlated with invasive fractional flow reserve for detecting hemodynamically significant stenosis. *Journal of Cardiovascular Computed Tomography* **10**, 199-206, doi:10.1016/j.jcct.2016.03.002 (2016).

27 Vernon, S. T. *et al.* Utilizing state-of-the-art “omics” technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease. *Microcirculation* **0**, e12488, doi:doi:10.1111/micc.12488.

Figure Legends

Figure 1: BioHEART-CT Study Design

Figure 2: Semi-quantitative plaque analysis incorporating the established GENSINI scoring system[25] and adding an additional multiplier for plaque composition to create a modified GENSINI score.

Figure 3: Example of semi-automated analysis of a complex plaque using syngo.via running on a Frontier system (Siemens Healthcare, Erlangen). Panels A & B show a coronary artery lesion with a large calcified component longitudinally and in cross-section. Panels C & D show classification of the calcific (yellow) and fibrous (green) components of this lesion.

Figure 4: Overview of the Analytic Approach.

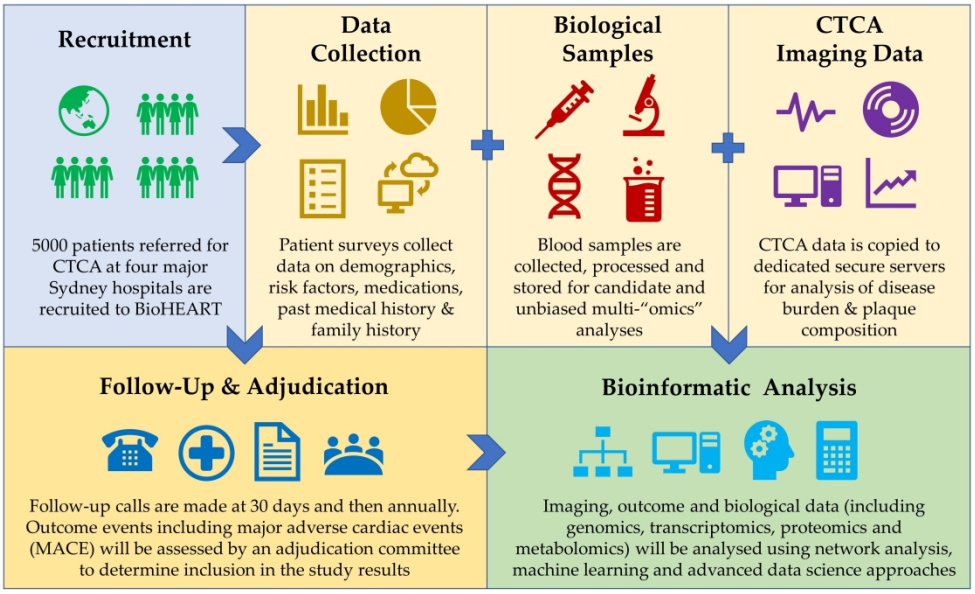
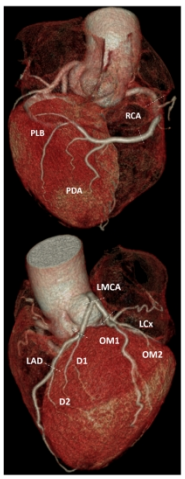


Figure 1: BioHEART-CT Study Design

358x220mm (300 x 300 DPI)

Stenosis Location	Multiplication Factor – Right Dominant (Left Dominant)
RCA - Proximal / Mid / Distal / - PDA - PLB	 x1 x1 x0.5
LMCA	x5
LAD - Proximal - Mid - Distal - D1 - D2	 x2.5 x1.5 x1.0 x1.0 x0.5
LCx - Proximal - Distal - OM1 - OM2 - (Left PDA) - (Left PLB)	 x2.5 (x3.5) x1 (x2) x1 x0.5 (x1) (x0.5)
Ramus	x1



Percent Stenosis	Score
0	0
1-25%	1
26-50%	2
51-75%	4
76-90%	8
91-99%	16
100%	32

Plaque Composition	Multiplication Factor
Calcified	x1
Mixed	x2
Non-calcified (soft)	x3

Figure 2: Semi-quantitative plaque analysis incorporating the established GENSINI scoring system25 and adding an additional multiplier for plaque composition to create a modifided GENSINI score.

210x148mm (300 x 300 DPI)

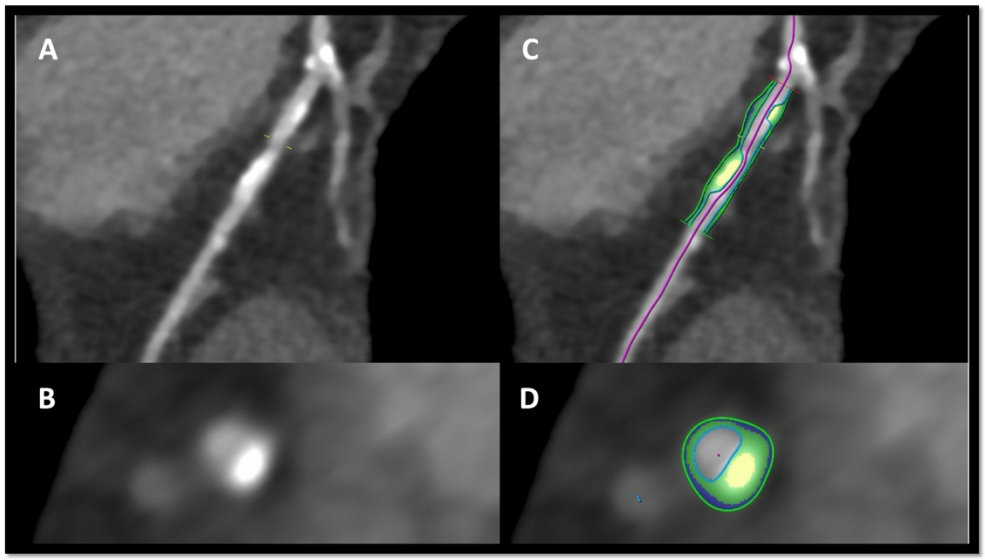


Figure 3: Example of semi-automated analysis of a complex plaque using syngo.via running on a Frontier system (Siemens Healthcare, Erlangen). Panels A & B show a coronary artery lesion with a large calcified component longitudinally and in cross-section. Panels C & D show classification of the calcific (yellow) and fibrous (green) components of this lesion.

358x220mm (300 x 300 DPI)

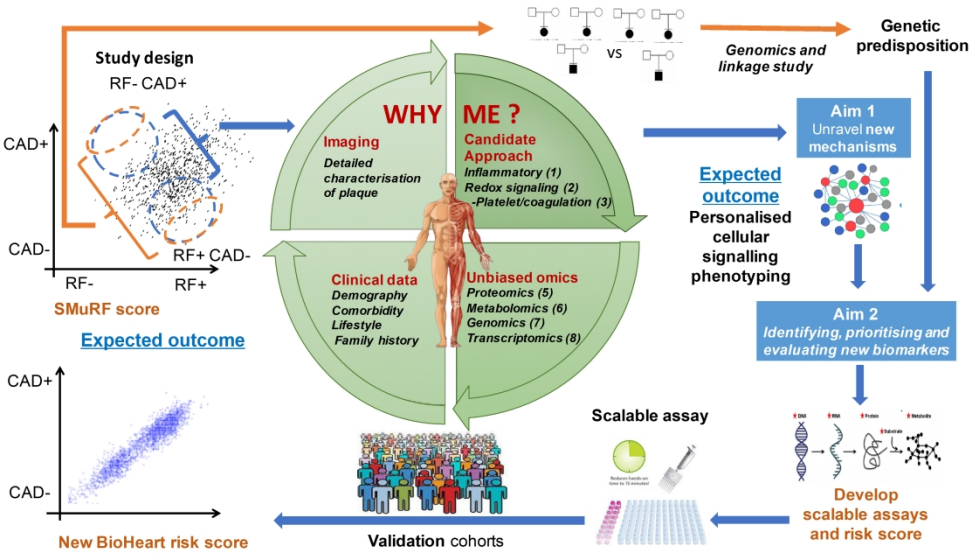


Figure 4: Overview of the Analytic Approach.

358x220mm (300 x 300 DPI)