PRECISION-CARRS

Study Protocol

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

AIIMS	All India Institute of Medical Sciences
AKU	Aga Khan University
ApoA/B	Apolipoproteins A & B
BMI	Body Mass Index
BP	Blood Pressure
BRFSS	Behavioural Risk Factor Surveillance System
CAP	College of American Pathologists
CCC	CARRS Coordinating Center
CCDC	Centre for Chronic Disease Control
CEB	Census Enumeration Blocks
CHD	Coronary Heart Disease
CHF	Congestive Heart Failure
CKD	Chronic Kidney Failure
CMD	Cardiometabolic Diseases
CUPS	Chennai Urban Population Study
CURES	Chennai Urban Rural Epidemiology Study
CVD	Cardiovascular Disease
DM	Diabetes Mellitus
DMC	Delhi Municipal Corporation
EQ-5D	European Quality of Life 5
FPG	Fasting Plasma Glucose

HbA1c	Glycated Hemoglobin
HDL	High Density Lipoprotein
HINTS	Health Information National Trends Study
HTN	Hypertension
IDSP	Integrated Diseases Surveillance Project
IGIB	Institute of Genomics and Integrative Biology
IRB	Institutional Review Board
KAP	Knowledge, Attitudes, and Perceptions
LDL	Low Density Lipoprotein
LMCI	Low- and Middle-Income Countries
MDRF	Madras Diabetes Research Foundation
MI	Myocardial Infarction
MONICA	Multinational MONItoring of trends and determinants in CArdiovascular disease
MOP	Manual of Operations
NABL	National Accreditation Board for testing and calibration Laboratories
NCCD	National Center for Chronic Diseases
NCD	Non-Communicable Diseases
NDMC	New Delhi Municipal Corporation
NHLBI	National Heart, Lung and Blood Institute
NIH	National Institutes of Health
NPDCS	National Program for Prevention and Control of Diabetes, CVD and Stroke
PHFI	Public Health Foundation of India
PI	Principal Investigators
PVD	Peripheral Vascular Disease
QME	Quality Monitoring and Evaluation
SOP	Standard Operating Procedures
TG	Triglycerides
UA	Unstable Angina
VLDL	Very Low Density Lipoprotein
WHO	World Health Organization

I. BACKGROUND

I.i. Cardio-metabolic Diseases (CMD)

Cardio-metabolic diseases include diabetes mellitus (DM), cardiovascular disease (CVD), kidney disease, and common risk factors that underlie these conditions such as central obesity, insulin resistance, glucose intolerance, dyslipidemia, and hypertension. CMDs are growing public health problems worldwide. [1] They are among the top ten most costly diseases, but they have the advantage of being predictable through identification of distal and intermediate risk factors and also preventable through changes in lifestyle particularly through healthy eating habits and regular physical activity.[1]

I.ii. Cardio-metabolic Disease Burden

It is estimated that coronary heart disease (CHD), cerebrovascular disease, and diabetes together account for 30% of global mortality and 80% of these deaths occur in low- and middle-income countries (LMIC).[2-4] Diabetes commonly co-exists with obesity (both generalised and central), hypertension,[5] and lipid abnormalities (elevated triglycerides, low HDL-cholesterol, abnormal LDL-cholesterol sub-fractions) and is a central feature accelerating athero-thrombotic cardiovascular disease (CVD), it is the leading cause of adult-onset blindness, non-traumatic amputations and kidney failure worldwide. The addition of these inter-related risk factors and co-morbidities results in a multiplicative, rather than additive, amplification of risk of severe outcomes (such as CVD events, amputation, etc.) and mortality.[6]

In people of South Asian origin, diabetes, cardio-metabolic risk factors[7, 8] and events[9, 10] occur at younger ages and lower body mass indices (BMI) when compared to other ethnic groups,[10-23] and are rapidly increasing with socioeconomic and nutrition transitions.[4, 24-26] The South Asia region includes three of the top ten countries in the world in terms of total people living with diabetes (India, Pakistan, and Bangladesh)[27] and is the region with the highest number of diabetes-related deaths currently.[28] However, deaths are only the tip of the iceberg, beneath which are a large number of diabetes patients with complications and long term sequelae. A study by Bajaj et al. in 1997 estimated the prevalence of diabetic retinopathy in India to be 20.8%[29]; Mohan et al. in 1995 estimated that about 1.9% of the diabetic patients in India develop Nephropathy[29]; 23.9% Indians develop diabetic neuropathy (Ramachandran et al., 1988)[29]; about 18.5% go on to develop CHD (Mohan et al., 1995)[29]; about 3.9% develop peripheral vascular diseases resulting in amputation of 0.5% of these patients (Mohan et al., 1995).[29] Asian Indians, as a group, are projected to account for between 40-60% of the global CVD burden within the next 10-15 years.[30] Furthermore, 35% of CVD-related deaths in India occur in those between 35-64 years of age as compared to only 12% in the U.S.[31]

Previously, it was thought that CMDs were confined to affluent urban residents, but this paradigm is gradually shifting as these conditions are now increasingly prevalent in lower socioeconomic groups in South Asia,[32-34] and CVD is currently the leading cause of death in both urban *and* rural India.[26, 35] Projections suggest that India and Pakistan's national income losses over the decade up to 2015 will amount to US$ 267 billion due to cardiovascular and diabetes deaths alone;[36] of these, the economically active age range (25-64) will bear great morbidity and mortality resulting in loss of human capital and productivity, perpetuating poverty faced by many, and potentially stifling development.

I.iii. Surveillance for the Prevalence of CMDs and their Risk factors

Surveillance of risk factors and disease is an invaluable public health research tool for: monitoring population health status; guiding resource allocation and policy; identifying and prioritizing interventions for subpopulations at particular risk; identifying disparities in outcomes; planning and evaluation of health programs.[37,38] Given the elevated and growing cardio-metabolic risk in South Asia, [7, 10, 13, 19, 23, 25, 26, 39-42] the importance of surveillance cannot be underestimated. Current assessments of surveillance efforts in the subcontinent, and indeed most developing countries,[43, 44] suggest large data deficits, vast state-wise heterogeneity and variable data quality, limiting the value of existing figures.

I.iv. Existing Surveillance System and Gaps

Currently, the major source of population level estimates of CMD risk factors, morbidity and mortality in India and Pakistan has been ad hoc surveys. These surveys can generally be characterized as state-specific, with small, often highly variable sample sizes, varying and often low response rates, with use of different diagnostic criteria, and limited by problems of sample design, lack of standardization, frequent measurement errors and incomplete reporting of results.[45] Recent initiatives by the Government of India have attempted to address these deficiencies through setting up a National Program for Prevention and Control of Diabetes, CVD and Stroke (NPDCS), an Integrated Diseases Surveillance Project (IDSP) at multiple sites, and establishing two other NCD risk factor surveillance projects.[46, 47] However, these systems fall short by failing to provide critical CMD-specific incidence and mortality data, measures of diet and physical activity, secular trends in risk factors, health service utilization, health care costs, and quality of care. The Sample Registration System of India, for example, relies on medically certified deaths which account for just 15% of total mortality and coverage is limited to institutional deaths in urban areas.[48] Pakistan conducted a national health survey from 1990-94, but has not mobilized unified, national efforts to collect subsequent health data in keeping with the transitioning socio-political milieu. The region as a whole suffers from a fragmented, chaotic, public-private mix of health-care providers with little or non-existent documentation.

I.v. Arguments for Surveillance Models

Justification for establishment of a well-designed, integrated surveillance system lies not only in helping align resource allocation with actual needs, but also broader themes[49, 50] which include:

a. More extensive comprehension of the distribution and trends of determinants and disease outcomes, especially given the asymptomatic prolonged course of NCD risk factors, the risk of debilitating target organ damage and often fatal disease events, as well as ensuing health and socioeconomic burdens. This is accomplished by:

    1. Investigating the determinants of disease prevention through early risk factor detection and control, spanning the spectrum of awareness, knowledge, attitudes, and practice (lifestyle behaviors, health-seeking and utilization, as well as treatment adherence, and perceived quality of life);

    2. Capturing newly-diagnosed cases, events, recurrent disease and mortality as well as the distribution and determinants associated;

b. Dynamic integration of information from multiple sources, improving case detection,[51] quality of individual chronic care delivery, health information infrastructure and the opportunity to increase accountability through regular audits and evaluating efficacy of prevention and control strategies;[52-55]

c. Reducing long-term health expenditure through culmination of safe, effective prevention and care models lowering rates of target organ damage, first events, recurrent disease, disability and premature mortality.

Experiences with surveillance models in developed countries have varied according to the stage of health system maturity and economic development. The U.S., for example, has relied on nationally-representative surveys, focusing primarily on self-reported disease risk factors (National Health Interview Survey, Behavioral Risk Factor Surveillance System or BRFSS). In following trends in cardiovascular risk factors, only one survey (National Health and Nutrition Examination Surveys) now routinely collects laboratory samples.[49] Countries with socialized national health systems, like the UK and Canada, have publicly-funded, networked, routine data capturing registers, although use and auditing of these systems is inconsistent. Models in Australia and much of Europe are based on regular standardized quality of care evaluation, acquiring population characteristics as well as provision of performance indicators based on provider processes and patient outcomes.[49]

The lessons drawn from these experiences and the published literature[49, 50, 49] support the utilization of standardized models that are not reliant on self-reporting, such as the World Health Organization (WHO) STEPwise Approach to Surveillance.[50] An initial, uniform prototype of this nature can be used to overcome infrastructural deficits in low-resource settings, and the foundation created may help advocate for modernizing and scaling up surveillance efforts

towards an ideal system (networked, electronic health recording registers with data integrated from primary care, hospital, laboratory and home monitoring settings).[51, 56]

II. OVERVIEW OF CARRS COHORT

II.i. History of CARRS Cohort

Recognizing the importance of understanding CVD risk in this population, the NHLBI supported the establishment of the Center for cArdiometabolic Risk Reduction in South Asia cohort (CARRS) over 2009-2015.[64] Since then, the CARRS cohort has grown to follow a diverse population-based sample of n=21,864 urban adults age ≥ 20 years, representative of Delhi and Chennai, India, with ongoing follow-up for clinical ASCVD risk factors, clinical disease, and mortality. The cohort was recruited in two waves, CARRS-1 in 2010-11 and CARRS-2 in 2015-16, and has high retention rates (70-85% annual follow-up, 88% hybrid tele/in-person attendance even during COVID, and >95% have at least one follow-up or death in 10 years of follow up) with a biorepository of 360,000 stored samples.

II.ii Overview of Precision-CARRS

While a substantial proportion of CVD manifests in people with ASCVD risk factors,[65] many individuals with these risk factors do not develop CVD.[65] In addition, some individuals without conventional risk factors do develop CVD.[66,67] Motivating **Precision-CARRS** is our insufficient understanding of the factors driving these differences in risk and manifest disease.

Precision-CARRS will enhance the scientific value of the CARRS Cohort by mapping the natural history of CVD at the granular level of subclinical/clinical phenotypes, using targeted markers and untargeted multi-omics, and assessing the role of biological, environmental, behavioral, and social factors along both atherosclerotic and heart failure CVD pathways. We will perform comprehensive analyses in stored samples and freshly collect detailed subclinical/clinical CVD phenotypes during five more years of follow-up of participants living in Delhi and Chennai. We will accrue a total of 176,536 person-years of follow-up, >1,000 incident ASCVD and >900 events of HF (Stage C/D, 50% HFrEF and HFpEF) by 2025.

Precision-CARRS is comprised of **four inter-linked projects**, with **three complementary cores**. The synergistic projects will be conducted in the Delhi and Chennai CARRS participants, will build upon existing data, incorporate new data collection, have common endpoints of subclinical/clinical CVD, and share resources provided by all three cores.

The four projects will:

- **Project 1:** Estimate the prevalence, incidence, and predictors of subclinical and clinical vascular and myocardial disease, with exploration of pathophysiologic pathways.
- **Project 2:** Investigate the role of environmental exposures (air pollution/built environment) on the development of subclinical and clinical vascular and myocardial disease.
- **Project 3:** Investigate multi-omics determinants of subclinical and clinical vascular and myocardial disease.
- **Project 4:** Investigate spousal and behavioral influences on subclinical and clinical vascular and myocardial disease.

The three cores will:

- **Administrative and Field Coordination (AFC) Core**: Provide leadership, field and laboratory coordination and standardization, and communication services.
- **Cardiovascular Disease Phenotyping Core:** Be responsible for advanced subclinical and clinical CVD phenotyping of all living, consenting participants enrolled in Precision-CARRS.
- **Data Management and Analysis (DMA) Core:** Provide comprehensive data analytics support.

Precision-CARRS will unravel the natural history, pathophysiology, and causal factors of ASCVD and myocardial disease and pave the way for precision CVD diagnostics, prevention, and care for South Asians, who account for 20% of humanity, are a rapidly growing but understudied US population, and are at high CVD and diabetes risk even at younger ages, despite having a thin body phenotype.

II.iii. Study Partners

The CARRS Cohort is coordinated by Emory University of Atlanta, U.S.A. and the Centre for Chronic Disease Control (CCDC), New Delhi, India. The original CARRS Cohort included other key partners: All India Institute of Medical Sciences (AIIMS), New Delhi, India; Madras Diabetes Research Foundation (MDRF), Chennai, India; and Aga Khan University (AKU), Karachi, Pakistan. For the continued follow-up in Precision-CARRS, Emory and CCDC will continue to benefit from partners at AIIMS and MDRF and will also involve new colleagues at the Institute for Genomics and Integrative Biology (IGIB), New Delhi, India; Ashoka University, Haryana, India; and Sri Ramachandra Institute of Higher Education and Research, Chennai, India.

Our chosen network of partners is intimately familiar with the complexities of the region and appreciate the opportunity and need for context-specific, uniform and sustainable methods of capturing representative estimates of risk factor prevalence and outcomes**.** In addition, through scientific leadership, we aim to implement systems with capabilities for auditing and deriving cost indices, in order to model projected burdens and deliver comprehensive, effective response strategies.[57]

The Principal Investigators for the original CARRS cohort study were Drs. Dorairaj Prabhakaran (CCDC), K.M. Venkat Narayan (Emory), Viswanathan Mohan (MDRF), Nikhil Tandon (AIIMS), Masood Kadir (AKU), and Mohammed K. Ali (Emory). Additionally, for **Precision-CARRS**, Drs. K.M. Venkat Narayan (Emory, population sciences, cardiometabolic diseases, co-PI of CARRS Cohort from inception) and Arshed Quyyumi (Emory, laboratory and clinical sciences, cardiology) will serve as MPI's for the program project grant. The MPIs will be supported by Drs. Dorairaj Prabhakaran (CCDC, cardiologist-epidemiologist, Phenotyping Core Lead), Mohammed K. Ali (Emory, physician-epidemiologist, AFC Core Co-Lead), Howard Chang (Emory, Biostatistician, DMA Core Co-Lead), Yan Sun (Emory, NHLBI- funded 'omics researcher), and Shivani Patel (Emory, ESI cardiometabolic epidemiologist and social scientist). The Steering Committee additionally includes Drs. Sailesh Mohan (CCDC, AFC Core Co-Lead) and Dimple Kondal (CCDC, DMA Core Co-Lead).

In addition to the main investigators, Precision-CARRS is comprised of many investigators and staff from each of these institutions. Investigators/consultants from other institutions are also engaged in Precision-CARRS work.

II.iv. Objectives

Previous objectives (In CARRS)

a. To implement and evaluate a model surveillance system in three study sites: Delhi (India), Chennai (India) and Karachi (Pakistan).

b. To assess the prevalence of CMD risk factors and diseases among adults aged 20 years and above, permanently residing in well-defined urban communities in the three study sites.

c. Ascertain factors that influence knowledge, attitudes, and perceptions (KAP) of the sample population regarding CMD and their risk factors.

d. To derive cost and health-utilization indices which can be used to model projected burdens of CMD in order to formulate cost-effective and timely interventions.

Current objectives (In Precision-CARRS*):*

e. Estimate the prevalence, incidence, and predictors of subclinical and clinical vascular and myocardial disease, with exploration of pathophysiologic pathways, including the influence of environmental exposures, multi-omic determinants, and spousal and behavioral factors.

a. Determine the incidence of intermediate risk factors (in previously risk-free individuals), new-onset complications, and the associated morbidity and mortality.

III. METHODOLOGY

III.i. Study Design

CARRS recruited a representative sample of 21,864 adults aged ≥20 years in Delhi and Chennai in India over two waves of data collection (CARRS-1 in 2010-11 and CARRS-2 in 2015-16); of those enrolled, 19,275 gave blood specimens at baseline. CARRS-1 has completed five follow ups, and CARRS-2 has completed one follow up. The Cohort is followed biennially with bloods and has high retention. In CARRS-1, over 10 years, 95% have had at least one follow up or death. Due to the COVID-19 pandemic, we adopted a hybrid approach (telephonic plus selective in-person follow up) and have 88% retention using this approach.

For PRECISION-CARRS, two additional biennial follow-ups with bloods and simple measures will be conducted. In total, the CARRS Cohort will extend from 2010-2025 for a total of 15 years of overall follow-up

Outcomes of interest:

i. Anthropometric Changes
   a. Weight
   b. Waist circumference
   c. Body Fat
ii. Development of new-onset intermediate risk factors
   a. Hypertension
   b. Diabetes Mellitus
   c. Dyslipidemia
iii. Subclinical structural and functional cardiometabolic disease
   a. Arterial stiffness
   b. Arterial thickness
   c. Atherosclerosis
   d. Ventricular function
   e. Hepatic fat
iv. Incident morbidity
   a. Stroke
   b. Myocardial infarction
   c. CHF – preserved and reduce ejection fraction
   d. Chronic Stable Angina
   e. Chronic Kidney Disease

   f. Revascularization
v. Complication
   a. PVD
   b. Retinopathy
   c. Nephropathy
   d. Neuropathy
   e. Amputation
vi. Health service utilization and costs
   a. Hospitalization
   b. Outpatient use
vii. Mortality
   a. All cause
   b. CMD-specific

III.ii. Study Sites and Settings

The original cohort was established at three sites, two in India (Chennai and New Delhi) and one in Pakistan (Karachi). This was a household survey wherein recruitment of participants and data collection took place in households. These were metropolitan urban settings with large, growing populations (an estimated 4.5, 10 and 13.8 million people live in Chennai, Karachi, and New Delhi respectively) and represented archetypes of rapid socio-economic, demographic, epidemiologic, and nutrition/lifestyle transitions. These cities were characteristic of the endogenous regions in which they were situated and were home to both urban and semi-urban populations of varying socioeconomic status.

For Precision-CARRS, living participants in Delhi and Chennai, India, will be approached to provide updated questionnaire, anthropometric, and biospecimen data, as well as participate in imaging studies.

III.iii. Sample Size Estimation

In CARRS, we applied the WHO STEPwise methodology [59] to estimate the sample size required to capture CMD risk factor prevalence precisely across the three study sites in South Asia. Utilizing risk factor prevalence estimates from previously published Indian studies and anticipating a response rate of 80 per cent (%) with a design effect factor of 1.5 (to account for cluster sampling), the sample size estimates were generated for males and females in three age strata (20-45, 45-60, 60 and above) in each urban setting. Table-1 presents the cumulative subtotals of subjects required to observe appreciably consistent prevalence approximations for each of the commonly-known risk factors. As shown, the highest required sample size (3,983 people) will permit each site to reliably estimate one or more of the CVD risk factors for each of the gender and age strata identified above.

Beyond this, we have been conducting follow-up surveys to collect pilot data on incidence of risk factors, CMDs complications, and CMD-specific mortality. Consent was taken during the initial recruitment for the cross-sectional surveys and only those participants who provided consent to be followed up for three years were enrolled into the study. Those participants who will choose to participate in **Precision-CARRS** will be reconsented. However, we anticipate an overall 15% loss-to-follow-up due to: potential migration of the young population (20-35 years), due to job opportunities or marriage (in case of females), and also because the study is community based.

Thus far, CARRS has recruited a representative sample of 21,864 adults aged ≥20 years in Delhi and Chennai, India over two waves (CARRS-1 in 2010-11 and CARRS-2 in 2015-16). The Cohort is followed biennially with bloods and has high retention.

Table 1: Sample Size Estimation (per site)

Risk factors	Level of Confidence	Margin of Error	Baseline levels of indicators	Design effect	Expected Response Rate	No. of age/sex Estimates	Sample size
Tobacco use	1.96	0.05	0.23	1.5	0.8	6	3062
Hypertension	1.96	0.05	0.36	1.5	0.8	6	3983
Diabetes	1.96	0.05	0.15	1.5	0.8	6	2204
Overweight (BMI $\geq$ 23)	1.96	0.05	0.65	1.5	0.8	6	3933

III.iv. Sampling Methodology

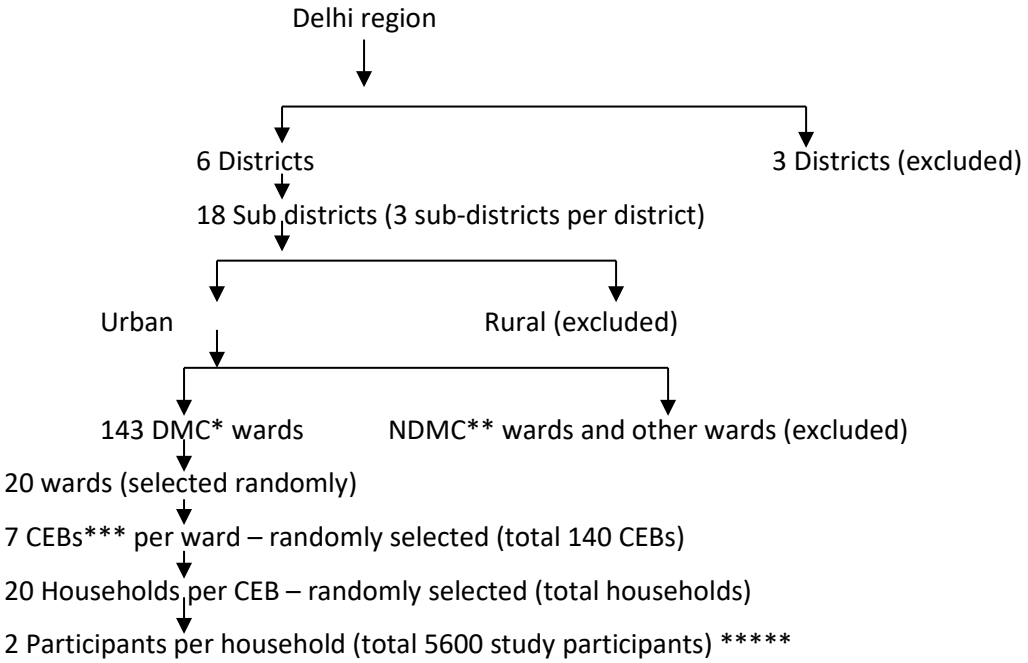
A multi-stage cluster random sampling technique was used to capture a sample representative of the urban population at the three sites. Each of the cities has its own distinctive municipal sub-divisions, encompassing municipal corporations, wards and census enumeration blocks (CEB) from which households were randomly selected. After the households were selected, a three step within-household sampling methodology (based on the 2002 Health Information National Trends Study [HINTS] conducted in the US[58]) was used to recruit study participants. With informed consent, two subjects were selected per household, one male and one female aged more than 20 years. However, in some households only one eligible participant was found. The final sample for the cross-sectional survey was composed of equal proportions of males and females in each of the three age strata (20-45 years, 45-60 years and >60 years) leading to a sample of approximately 4000 participants in each of the study sites (fig-1, fig-2 & fig-3).

In Precision-CARRS, no new sampling will occur. We will approach enrolled living participants regarding continued participation.

Thus, in 2022, participants from both Wave 1 and Wave 2 of the surveys in Delhi and Chennai, India will be approached and if they consent, will provide questionnaire, anthropometric, and biospecimen data, and will be assisted to present to a facility for additional imaging phenotyping measures.

Diagrammatic representation of city subdivisions for sampling of Cohorts 1 and 2

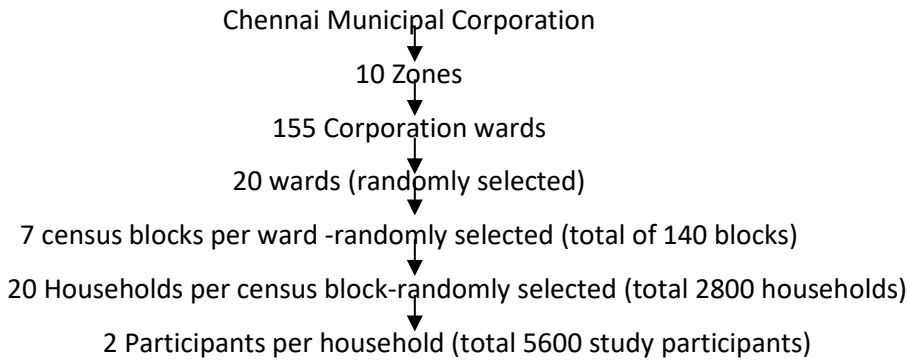
Figure 1: New Delhi Sampling Scheme



*Delhi Municipal Corporation; **New-Delhi Municipal Corporation, ***Census Enumeration Blocks

***** Recruitment will be stopped at 5500.

Figure 2: Chennai Sampling Scheme



III.v. Study Tools

Standardized sampling has enhanced the representativeness and reproducibility of results. We also utilized uniform tools and methods to ensure replicability across multiple sites. We utilize culturally appropriate and methodologically relevant questionnaires for South Asia which were originally derived from World Health Organization, NHANES, and other validated tools. Event modules and verbal autopsy forms will be adapted from PURE.

III.vi Surveillance Indicators

We completed a baseline cross-sectional survey of 16,287 participants (N=5,364 in Delhi; N=6,906 in Chennai; and N=4,017in Karachi) in 2010-2011 and a second baseline cross-sectional survey of 14,587 participants in 2015-16. We collected information encompassing broad categories such as: demographic and socio-economic characteristics of the population; presence of risk factors; previous or existing target organ damage (known nephropathy, angina, retinopathy, cataracts, peripheral vascular disease, previous stroke, previous therapeutic procedures such as amputation, revascularization procedure, peripheral endovascular procedure, dialysis, transplant, laser photocoagulation); quality of life, disability, health care utilization; quality, and cost(s) of care (as described in table-2). Surveys were comprehensive and encompassed collection of data, anthropometric measures, venous blood samples, urine samples and saliva.

Tools that were used to measure distal risk factors were validated questionnaires (food frequency questionnaire, physical activity questionnaire, etc.) and laboratory tests to confirm tobacco use in 15% of randomly selected participants through measurement of Cotinine in saliva. The intermediate risk factors were assessed through anthropometric measurements (Height / Weight / Waist Circumference/ Skinfold thickness / Body Fat), blood pressure measurement and through tests of bio-chemical parameters (laboratory estimation of FPG, HbA1c, Lipid profile, ApoA/B, serum urea and creatinine).

Prevalence of certain morbidity indicators of CMD were studied. These included disability from Stroke, MI, CHF, Amputation, Chronic Stable Angina, CKD, revascularization, and other procedures and hospitalizations related to CMD. These were assessed using questionnaires and cross-checked with medical records, where possible.

Starting in 2022**, for Precision-CARRS,** follow up examinations of both CARRS-1 and CARRS-2 in Delhi and Chennai, India, will include questions regarding health habits, marital status, home and work life, current and past medical history and medications, and early life experiences. Psychosocial questions will also be added to determine stress, depression, conflict resolution, and social support. Ambient air pollution will be measured as Daily average $PM_{2.5}$ at 1kmx1km spatial resolution. Built environment will be determined by the Vegetation Index as well as road networks,

grid-wise built-up area %, intersections, light intensity, and locations of polluting sources. Additional biomarker data will also be collected including advanced lipids, adipokines, as well as protein biomarkers of inflammation, thrombogenesis, myocardial infarction, and myocardial stress.

We will arrange transportation of participants in groups to a core facility for detailed subclinical structural and functional cardiovascular disease (CVD) and metabolic (e.g., hepatic fat) imaging at one time-point. These tests will include echocardiography, carotid ultrasound, a CT scan of the heart and liver, as well as arterial stiffness testing and ECG monitoring.

In a subsample of cases and controls (n=3,000) defined by different atherosclerotic cardiovascular disease (ASCVD) and myocardial disease phenotypes and controls, integrative 'omics analyses will be conducted on stored and newly-collected bio-specimens.

In a separate subsample of spousal dyads (n=1,500 dyads = 3,000 adults), we will employ use of mobile health (mHealth) devices (e.g., wearable wristbands and mobile phones, etc.) for a 7-day period that will be loaned to and collected from participants at one time point to assess measures of sleep, physical activity, and stress.

Additional sub-studies will be approved on an ongoing basis after considering the burden to participants and the scientific advances offered by new data.

Table 2: Summary of the Surveillance Indicators

			Currently Available or Measured from Stored Sera							**Newly Measured from Fresh Samples**	
Domain	**Variables**	**Measurement**	**CARRS-1 N=10,205**				**CARRS-2 N=9070**			**Precision-CARRS P01 N=14,000**	
		Midpoint Year of measurement.	2011	2012	2013	2016	2017	2015	2018	2022	2024
Questionnaires or wearables											
Socio-demographic	Contact information, age, sex, income, education, occupation, social caste, place of birth, marital status, household size	Questionnaire	A					A		F	F
Health behaviors: questionnaires	Dietary habits, physical activity, alcohol use, tobacco use, sleep duration and quality, COVID	Modified FFQ,[122] IPAQ, Questionnaires	A				A	A		F	F
Healthcare services	Awareness and risk factor control, access to health care services, utilization of services, health insurance, cost of treatment	Questionnaires	A				A	A		F	F
Health behaviors: wearables	†Dynamic measures using wearables in subsample	†Actiwatch[171]								F	F
Family History	Family history of HTN; DM; Dyslipidemia; Heart disease; Stroke	Self-report	A				A	A		F	F
Physician diagnosis of disease	Previously diagnosed HTN; DM; Dyslipidemia; Heart disease; Stroke; COPD	Self-report / medical records	A	A	A	A	A	A	A	F	F
Medication use	Cholesterol, glucose, lipid, and BP-lowering medication	Self-report	A				A	A		F	F
Psychosocial	Stress, Depression, Health Related QoL, Partner conflict resolution, stress,	PHQ-9[172] and EQ-5D[173] Negotiation Subscale of the	A				A	A		F	F

	social support †Smartphone-based ecological momentary assessment	Conflict Tactics Scale-2; Daily Inventory of Stressful Events (DISE[174]); Medical Outcomes Study Social Support Survey questionnaires †Periodic daily stress appraisals[175]									
Marital characteristics	Age at marriage, family structure, marriage duration, consanguinity, arranged vs. not arranged	Questionnaires								F	F
Environmental											
Ambient Air Pollution	Daily average $PM_{2.5}$ at 1kmx1km spatial resolution	B-attenuation monitors; Gravimetric Samplers[141]	A	A	A	A	A	A	A	F	F
Built Environment	Vegetation index, Road Networks, Gridwise built-up area %, Intersections, Light intensity, Locations of polluting sources	Satellite observations and land use maps[119]	A	A	A	A	A	A	A	F	F
CVD Risk Factors											
Anthropometry	Weight, Height, Waist and hip circumference	Tanita BC-418, Seca-213 Stadiometer, Tape measure[107]	A	A	A	A	A	A	A	F	F
Blood pressure	Systolic and Diastolic BP	Automated Omron HEM-7080[107]	A	A	A	A	A	A	A	F	F
Glucose and Insulin	Fasting, 30-min, 2-hour blood glucose; HbA1c, Fasting, 30-min, 2-hr insulin; C-peptide	Hexokinase;[107] HPLC[107] ECL immunoassay[129]	A		A		A		A	F	F
Lipid Markers	Total, HDL, LDL Cholesterol; Triglycerides	Direct; Friedewald Equation;[107] Martin/Hopkins Equation if triglyceride >400 mg/dl;[176] Enzymatic methods Rate immunone-phelometry[107]	A		A		A		A	F	F
Renal function	Serum and urine creatinine; albumin:creatinine ratio	Jaffe kinetic[107] Spot urine[114]	A		A		A		A	F	F
	Cystatin C	Immunoturbidimetric[177]	S					S		F	F
Advanced lipids	Apolipoproteins A1 and B; Lpa; Homocysteine	Abbott labs	S					S		F	F
Adipokines	Leptin, adiponectin	Elisa assay[178]	S					S		F	F
Hepatic Steatosis	Hepatic fat deposition	CT Scan[179]									F
		Protein Biomarkers (Supported by Abbott Laboratories)									
Inflammation	Hs-CRP; suPAR	Abbott Labs	S					S		F	F
Thrombogenesis											
MI	High sensitivity troponin-I;	Abbott Labs	S					S		F	F
Myocardial stress	B-type natriuretic peptide	Abbott Labs								F	F
		† Multi-omics (Supported by IGIB, India)									
Genomics	Genotypes	GWAS Array[180]	S†					S†			

Epigenomics	DNA methylation	EPIC Beadchip[181]	S†	S†	F†
Metabolomics	Circulating Metabolites	Liquid chromatography[182] Mass spectrometry[183]	S†	S†	F†

A=Available currently**, S=**To be measured from stored samples**, F=**To be measured at follow-up**. †** Measures will be in a sub-sample of 3,000 individuals**.** *HTN=hypertension, DM=Diabetes Mellitus, COPD=Chronic Obstructive Pulmonary Disease, BP=Blood Pressure, HbA1c=Glycated Hemoglobin A1c, QoL=Quality of life, HDL=High Density Lipoprotein, LDL=Low Density Lipoprotein, Lpa=lipoprotein a; hs-CRP=High Sensitivity C-reactive protein, suPAR= soluble urokinase plasminogen activator receptor, MI=myocardial ischemia, ECL= electrochemiluminesence CIMT=Carotid Intima Media Thickness, ECL= electrochemiluminesence, MDCT= Multi-detector computed tomography scanner; * Echocardiography not currently available in CARRS but added in **Precision-CARRS**.*

BP = blood pressure; FPG = fasting plasma glucose; HbA_1c = glycated hemoglobin; TC = total cholesterol; HDL = high-density lipoprotein cholesterol; LDL = Low density lipoprotein; VLDL = Very low density lipoprotein; ApoA/B = apolipoproteins A & B; BMI = body mass index; MI = myocardial infarction; CHF = congestive heart failure; CKD = chronic kidney disease; CMD = cardio-metabolic disease; EQ-5D = European Quality of Life 5 Dimensions questionnaire.

III.vii. Study of Cost Burden

The study has previously utilized a **bottom-up**, **cost-of-illness approach** to collect important data on:

- ***Health-service utilization and treatment patterns*** (assessing the patterns of health-seeking, and treatment regimens that are used widely)
- ***Direct costs*** ('inputs' encompassing out-patient and in-patient care, pharmacotherapies, therapeutic procedures, and transportation to and from health care facilities in a given timeframe);
- ***Indirect costs*** ('lost outputs' representing the value of economic productivity lost by society on account of temporary or permanent absence, disability, or premature mortality);
- ***Health-related quality of life*** (including health utilities); and
- ***Health outcomes*** (including incidence of new-onset risk factors, CVD events, morbidity, and mortality),

These data may be used to model **societal burdens** of disease and potential avoidable mortality, disability, and costs. Direct medical and non-medical costs were ascertained from respondents through standardized questionnaires applied in other low- and middle-income countries.

Where possible, data regarding caregiver time and costs were ascertained and will be included in analyses. The study will not collect data regarding foregone opportunities of children or other household members.

III.viii. Knowledge, Attitudes, and Care Practices

There is very little currently published regarding the knowledge, perceptions, attitudes, and care practices of people with cardio-metabolic risk factors and diseases in South Asia. The surveillance study serves as a robust platform for investigating these aspects of cardio-metabolic diseases. Data collected in the questionnaires and medication assessments can and have been used to assess self-care and perceptions of chronic diseases in the CARRS cohort.

III.ix. Study Outcomes

We will conduct follow-up surveys to detect anthropometric changes, incidence of new intermediate risk factors or changes in participants with pre-existing risk factors (either or multiple of hypertension, diabetes, and dyslipidemia), morbidity (underlying target organ damage, health consequences, and disability), and mortality associated with CMDs during the follow up assessments of CARRS-1 and CARRS-2 participants in Delhi and Chennai, India, as well as both cohorts combined. We will also assess the incidence of subclinical cardiovascular disease in both cohorts combined. CMDs and their complications will be adjudicated using standard definitions (described in the MOP) and will be coded using MedDRA / ICD codes.

Table 3: Outcome Indicators in Follow-up Surveys

Outcome	Measures	Methods
Anthropometric Changes	Height / Weight / Waist Circumference / Skinfold Thickness / Body Fat	Clinical Examination, Stadiometer, Weighing Machine, Tape Measure, Calipers, Bio-electrical impedance
Intermediate risk factors (during the third follow-up only)	Hypertension	Clinical Examination, BP measurement
	Diabetes	Laboratory estimation of FPG, HbA_1c
	Dyslipidemia	Laboratory estimation of serum TC, HDL, Triglycerides, ApoA/B
Incident Morbidity	Stroke / MI / CHF / Chronic Stable Angina/ Revascularization, Peripheral Arterial Disease	Follow-up surveys; Symptoms, Echocardiograms, Medical records of documented events, admissions or procedures; Rose Angina Questionnaire, serum urea and creatinine for CKD.
	CKD/ Dialysis / Renal transplantation	
	Amputation/diabetic retinopathy/ Procedures /Revascularization/ Hospitalization	
Subclinical Cardiometabolic and Myocardial Disease	Pulse wave velocity, Augmentation Index, CIMT, Carotid Plaque, CAC, and Systolic and Diastolic Function	SphygmoCor, Ultrasound, and Echocardiograms
Mortality	All cause	Follow-up surveys; Death Certificates; Verbal Autopsy (PURE)
	CVD-specific; Diabetes-specific	

BP = blood pressure; FPG = fasting plasma glucose; HbA_1c = glycated hemoglobin; TC = total cholesterol; HDL = high-density lipoprotein cholesterol; LDL = Low density lipoprotein; VLDL = Very low density lipoprotein; ApoA/B = apolipoproteins A & B; BMI = body mass index; MI = myocardial infarction; CHF = congestive heart failure; CKD = chronic kidney disease; CMD = cardio-metabolic disease; EQ-5D = European Quality of Life 5 Dimensions questionnaire.

III.x. Biological Sample Collection and Storage

Biological samples to be collected in future assessment rounds include up to 35 ml of blood, 100 ml of urine, and 10 ml of saliva. Depending on the study site's experience with previous community-based studies, blood samples (fasting), urine sample (first morning void) and the sample of saliva (fasting) will be collected at the participant's residence or by organizing camps/clinics. These will be transported in proper cold chain to the laboratories at the field site, where the analysis will be done (described in table-2 and 3). The sample will then be stored in cryo-vials for future analysis. The challenges of collection, handling, transfer, and storage of specimen samples can be offset by many of the quality assurance plans. Maintaining the condition of samples from collection to long-term storage facilities is crucial to surveillance research, so we will integrate procedures and policies that ensure clear labeling, temperature control, pre-emptive solutions to power failure, and safety from biological hazards for all field staff. The methods of analysis of the biological samples at each of the three study sites used in the past are given in table-4.

Table 4: Biological Samples and their Methods of Analysis

Clinical parameter	Laboratory parameter	Methods used for analysis		
		Chennai	Delhi	Karachi (previous)
Diabetes	Fasting plasma glucose	Hexokinase/Kinetic*	Same as that of Chennai	Oxygen rate method
	HbA1c	HPLC*		HPLC
Dyslipidemia	Total cholesterol	CHOD-POD/End point*		Enzymatic Colorimetric method

	HDL	Direct*		Enzymatic Colorimetric method
	LDL	Friedewald Formula		Enzymatic Colorimetric method
	VLDL	Calculation		Calculation
	Triglycerides	GPO-PAP/End point*		Enzymatic Colorimetric method
	Apo A, Apo B	Immuno-turbidimetric*		**Not done**
Kidney disease	Serum urea	Urease GLDH/ Kinetic*		BUN: Enzymatic conductivity rate method
	Serum creatinine	Jaffe Kinetic*		Modified Jaffe's Method
	Microalbuminuria	Immuno-turbidimetric*		Rate nephelometry
Tobacco exposure	Salivary cotinine#	Enzyme Immunoassay		Not done

*Auto-analyzer; # In a sub-sample

In addition, in **Precision-CARRS**, participants will be invited to participate in further testing that is designed to assess early heart disease and stroke.

Participants will be taken to AIIMS in Delhi for all measures or Sri Ramachandra Institute in Chennai for Cardiac Echocardiography, Carotid ultrasound, and CT scan of the heart and liver or Dr. Mohan's Specialties Diabetes Center in Chennai for Arterial Stiffness and EKG monitoring. Additionally, a subset may participate in neuroimaging, as well. Participants will be transported in a study van and have the following tests performed:

1. **Echocardiography:** will be done to assess the function of the heart muscle and heart valves using ultrasound. The ultrasound probe will be placed on the chest while the participant is lying down on the exam table and turned to one side. The technician will move the probe around the chest to obtain pictures of the heart. This test will take about 15 minutes to perform.

2. **Carotid ultrasound**: will be done to assess hardening of the arteries (blood vessels) in the neck and to look for early plaque (waxy substance) deposits. Presence of these deposits increases the risk of stroke in the future and predicts future heart attacks. An ultrasound probe will be placed on both sides of the neck by a trained technician to obtain pictures of the carotid arteries (a large artery in the neck supplying blood to the brain). This test will take about 20 minutes to perform.

3. **CT scan of the heart and liver:** will be done to measure calcium deposits in the arteries of the heart and assessing fat deposits in the liver. Participants will be taken to the radiology department and a CT scan picture, similar to an x-ray of the heart and abdomen will be obtained. This test will take about 5 minutes.

4. **Arterial Stiffness Testing:** will be done to measure the stiffness of the blood vessels. While lying down on an exam table, a probe will be placed over the participant's wrist to measure how well the blood vessel moves. In addition, 3 patches (similar to EKCG patches) will be placed on the body to record the electrical activity of the heart. A cuff will also be placed on the thigh and a probe over the artery in the neck to measure the speed of the pulse movement in the arteries. This test will take 5 to 10 minutes to perform.

5. **ECG monitoring:** will be done to look at the heart's electrical activity. This will be done using a patch placed on the skin that the participant may wear for up to 72 hours. The participant will be given a pre-paid envelope to mail the patch back.

6**. Neuroimaging** (subset): may be done in a small subset of participants >50 years old to take pictures of the brain using MRI to quantify brain volumetrics, white matter hyperintensities, and microstructure, cerebrovascular integrity, and functional brain networks. The procedures include structural 3D T1-weighted (T1w) imaging, T2-weighted (T2w) FLAIR image, multi-shell diffusion MRI; arterial spin labeling, MR spectroscopy for brain metabolites, resting-state functional MRI, and whole brain 3D GRE sequence for quantitative susceptibility mapping. The tests take about 1.5 hours to complete.

III.xi. Sub-Studies

Precision-CARRS encourages and will help facilitate additional sub-studies that will utilize the CARRS Cohort and enhance the overall goals of the study. Our current sub-studies solely add additional questions to the existing questionnaire for a sub-sample of the cohort and, therefore, do not change how the study is being carried out. Each sub-study request will be reviewed internally prior to approval – consideration will be given to participant burden as well as scientific value being added by the sub-study. Additional ethical review approval will be the responsibility of the proposing investigators and must be obtained prior to any data collection. Participants will consent for sub-study participation at the time of consent for the main questionnaire.

Current sub-studies include:

1. Time/Risk preferences: 3500 participants needed in Delhi and Chennai.
2. Food Insecurity: 2000 participants needed in Delhi only.
3. Early Life Exposure: 1000 participants ≥ age 40 in Delhi and Chennai (500/site).
4. Multi-Morbidity and Frailty: 2300 participants ≥ age 40 in Delhi and Chennai (1750/site).
5. Multi-Modal Neuropsychological Examination Assessments: 400 participants ≥ age 50 in Delhi and Chennai (200/site).
6. Davos Alzheimer Collaborative Pilot: 300 participants ≥ age 50 in Delhi and Chennai (150/site).
7. Detailed Reproductive History Pilot: 300 female participants in Delhi and Chennai (150/site).

IV. DATA MANAGEMENT

IV.i. Data Collection

Data has previously been collected on paper questionnaires and transferred to a database that is managed by CCDC and Emory. The interviewers were rigorously trained to minimize errors and constantly supervised. Quality assurance and monitoring and evaluation are discussed in section VII.

For **Precision-CARRS**, there will be two follow-up survey data collection (questionnaires, anthropometric [height, weight, hip, waist] and blood pressure measurements) visits at participant households over the 5-year study. Living CARRS participants residing in Delhi and Chennai will be contacted via telephone by a study participant navigator to schedule a visit time by the study team. The field team interviewers will provide detailed information about the risks, benefits, and importance of the questionnaires, anthropometric measurements, bio-specimens, storage, and analysis. Participants will be given the opportunity to ask clarifying questions and asked whether they would like to participate, and if so, to sign an informed consent document. Interviewers will administer the questionnaire via a tablet-based application. Please note, some participants may be asked to perform digital cognitive tests using an iPad. Following questionnaire administration, interviewers will conduct an anthropometric examination and measure blood pressure.

Table 5: Implementation of Study Tools

Year	Schedule	Household proforma	Demographic and residential details	Tobacco, alcohol, diet, physical activity and stress	Female reproductive health	Medical history	Co-morbidities	Complications	EQ-5D	Respiratory disease	Family history	Treatment history and expenditure	Anthro-pometry/BP	Blood	Saliva	Urine		Subclinical Cardiovascular measures, CT Scans and EKG	Verbal Autopsy/ Death Certificates
CROSS SECTIONAL STUDY – I																			
2010-2012	Participant recruitment	X																	
	Data collection		X	X	X	X	X	X	X	X	X	X	X	X	X	X			
COHORT – 1 (Follow-up)																			
2011-2012	Follow-up - 1			X		X	X	X	X			X	X						X
2013	Follow-up - 2			X		X	X	X	X			X	X	X		X			X
2014	Follow-up - 3			X		X	X	X	X			X	X						X
2015	Follow-up-4			X		X	X	X	X			X	X	X		X			X
2017	Follow-up-5		X	X		X	X	X	X	X	X	X	X	X	X	X			X
CROSS SECTIONAL STUDY – II																			
2014	Participant recruitment	X																	
	Data collection		X	X	X	X	X	X	X	X	X	X	X	X	X	X			
COHORT – 2 (Follow-up)																			
2015	Follow up-1			X		X	X	X	X			X	X						x
2018	Follow up-2		X	X	X	X	X	X	X	X	X	X	X	X	X	X			X
PRECISION-CARRS																			
2022	Data Collection		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X
2024	Follow-up 1		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X

IV.ii. Data Entry

For **Precision-CARRS**, REDCap will be used for data collection and management. A data field specification list will be created along with a coding list for designing the database. The database will be programmed to have automated in-built checks for, logic, clinical reasonable ranges, absolute and relative values, context, and structure. This will control the quality of data. The database will be piloted and necessary restructuring done before finalizing it.

Data will be entered in the entry forms of the database at the respective study sites (MRDF, CCDC). There will be regular checks by site coordinators to minimize errors, missing values, and outliers. Data entry completion and errors will be continuously monitored by the Data Management and Analysis (DMA) Core. Final data cleaning and de-linking of participant identification information will be done before conducting analyses.

IV.iii. Data Management:

In Precision-CARRS, a Data Management and Analysis (DMA) core will be established to create workflows and protocols for consistent data cleaning, indexing, harmonization, monitoring, and documentation. The DMA Core will also assist in the dissemination of research findings, analytic tools, and data visualization to a variety of academic, public health, and stakeholder audiences. The DMA Core will also provide expertise in using biostatistics, bioinformatics, and machine learning method applied to CVD epidemiology, epigenetics, metabolomics, predictive modeling, and multilevel exposome research. In addition, the DMA core will support the development of new and innovative analytic methods and tools motivated by complexities of research design, exposure assessment, and phenotyping measurements.

The data manager and statisticians will re-check all data sent from the sites for outliers, coding errors and missing values. The data will be checked for missing values, outliers, and inconsistency by running do-files. Queries will be generated for inappropriate data and query rates will be calculated for sites, interviewers and outcome. A decision log will be used to document all issues related errors and queries (using excel sheets and emails).

IV.iv. Data Analysis

(a) Quantitative Analysis

Data analyses that have been conducted and published can be found on the CARRS website and are indexed in the National Library of Medicine. All analyses confirm to standard practices of data management, analytical approaches, and appropriate reporting.

For Precision-CARRS, the DMA Core will perform the following data analysis support:

Descriptive Analyses:

- Provide descriptive statistics (including outcome prevalence/incidence and exposure distributions) accounting for study-design, sampling and response weights using cluster-sampling selection probability, differential non-response rates at the household and individual levels, and random selection of two individuals in the same household. Standard error will be based on the Taylor linearization method.
- Record the number, timing, pattern, and reason for missing data. For all analyses, we will assume missing data to be missing at random a priori. Potential bias due to loss to follow-up will be examined in a sensitivity analysis with inverse probability of attrition weights that vary across visits. Multiple imputation will be used to account for missing exposures

Time-to-event Analysis of CVD Events and Prevalence Analysis of Subclinical CVD Measures.

- Cox proportional hazard models will be used to analyze incident CVD events (e.g., non-fatal MI, CVD death, heart failure). Log-binomial models will be used to analyze for dichotomized subclinical outcomes (e.g., CIMT, carotid plaque, CAC), as well as linear models for continuous measures. Primary exposures are longitudinal conventional and metabolic risk factors and pathway-specific protein biomarkers When appropriate, models will consider (1) clustering due to households/neighborhoods using random effects that account for exchangeable, spatial, or genetic relatedness (2) non-proportional hazard assumptions using baseline hazard

stratification (e.g., by age group or geographical areas), (3) competing risks (e.g., non-CVD death events) using sub-distribution hazard models (4) non-linear associations using parametric splines or categorical levels, and (5) different exposure windows for time-varying exposures and lagged associations. All analyses will be stratified by sex.

Longitudinal Models for CVD Risk Factors

- Mixed effect models will used to analyze longitudinal conventional and metabolic risk factors as a function of follow-up time to extract trajectory features as potential predictors of clinical and subclinical CVD outcomes
- Mixed effect models will be used to analyze longitudinal CVD risk factors and behaviors and estimate associations with time-varying spousal covariates

4. Development of Prediction Equations and Algorithms

- Based on fitted Cox proportional hazard and log-binomial models, we will evaluate prediction performance of incident CVD events and subclinical measures, along with identification of optimal cut-points for protein biomarkers, and assess prediction improvement with the addition of identified epigenetic and metabolic biomarkers and spousal influences
- Apply multiple supervised machine learning algorithms and ensemble modeling to predict CVD outcomes, environmental exposures, and to identify complex interactions with omics data

Additional Cross-cutting Data Analytic Needs.

- Coordinate a consistent set of sensitivity analyses, include examinations of alternate outcome definitions, different set of confounder controls, and removal of records due to known issues in data quality or extreme values.
- Implement methods to enhance inference from observational studies: negative exposure control to detect residual confounding, assessing impacts of unmeasured confounders, and estimating causally-oriented effects from observation studies using propensity score matching, instrumental variable, mediation analysis, and marginal structural model for longitudinal data.

.

V. REPORTING AND PUBLICATION

The investigators have formed a Publications, Presentations, and Ancillary (PP&A) Studies Committee led by Drs. Mohammed K. Ali and Viswanathan Mohan, and have developed a suitable policy to protect the rights regarding ownership of study materials and data. The investigators will be the custodians responsible for assigning subgroup analyses and publication under the oversight of this subcommittee. Reported outcomes of interest will include: age-standardized prevalence and mean levels of CMD risk factors; incidence rates for emergence of risk factors, subclinical disease, disease events, and mortality; levels of awareness; health service utilization; quality of life and determinants of human behaviors. The PP&A policies are intended to support productivity and ensure high-quality, impactful products that advance our understanding of CMD are disseminated, and credit and recognition is inclusively and diversely shared.

VI. QUALITY ASSURANCE

VI.i. Quality Assurance Strategies

Quality assurance strategies have been and will continue to be applied throughout the duration of the study using a framework which comprehensively considers each phase of the study and applies inter-related themes to every level of the study (as shown in Table 6).

Table 6: Quality Assurance Strategies

		Phase			
		Design and Planning	**Pilot Testing**	**Data Collection**	**Data Analysis**
Levels	**Institutions**	• Critical review of protocols (IRB)* • Develop a common manual of operations • Coordination of timelines, activities	• Assess fluidity and feasibility of field operations	• Monitoring field activities	• Audit and evaluate validity of findings prior to publication • Internal peer reviews prior to publication
	Investigators	• Certification • Pre-situational analysis	• Audit results after pilot is completed	• Monitoring	• Validity checks • Review results
	Field Personnel	• Extensive training • Objective evaluation • Easy-to-carry SOPs **	• Evaluate all field and documenting techniques	• Random checks, re-training	
	Survey Questionnaires	• Peer-review • Translation into local language(s) • Internal consistency estimates and reliability exercises	• Establish clarity and face validity in small field sample	• Regular checks to assess completeness	• Identify and discard compromised or inadequately com-pleted questionnaires
	Measuring Equipment	• Central procurement • Central training • Calibration guidelines and checks	• Evaluate calibration techniques, acceptability of use in field	• Regular calibration of tools; replace as and when required	
	Specimens	• Central procurement of kits and equipment • Specific protocols for each biochemical assay • Training (labelling, handling, storage)	• Evaluate adherence to protocols, labelling, processing, storage and handling • Interim analysis to detect outliers	• Random checks • External temperature gauge labels to monitor sample temperature	• Stored samples for future investigation • Identify and discard compromised samples
	Laboratory	• Laboratory selection and identification of reference laboratory • NABL or CAP certification *** • Central procurement before distribution • Develop internal and external quality assessment protocols and schedule of regularity	• Evaluate procedural fluidity • Evaluate intra- and inter-laboratory variability • Interim analysis to detect outliers	• Internal quality checks and calibration • Assess & evaluate intra- and inter-laboratory coefficients of variation • Regular external validation – lyophilized samples from reference laboratory	• Regular external validation – lyophilized samples from reference laboratory • Calibration and Internal quality checks
	Communication	• Establish reporting structures • Establish data transfer plans	• Assess agility of transfers		
	Documentation	• Develop checklists, logbooks	• Assess recording legibility		• Audit logbooks for response rates and

	Phase			
	Design and Planning	**Pilot Testing**	**Data Collection**	**Data Analysis**
	• Training in appropriate and legible documentation			field activity indicators
Data Storage & Confidentiality	• Establish data back-up and protection policies • Training of all staff	• Assess accessibility, simplicity and flexibility of software	• Locked and password-protected data storage • Active back-up	• De-identified datasets • Limited access to personal identifiers
Data Entry	• Establish protocols, consistent data cleaning methods and verification systems	• Assess variability	• Interim analyses to identify duplicate entries • Decision log to document issues	• Reporting on outliers • Validity checks • Track database errors

**IRB = institutional review board; **SOP = standard operating procedures; ***NABL = National Accreditation Board for Testing and Calibration Laboratories, Department of Science and Technology, Government of India; CAP = College of American Pathologists, Northfield, IL, USA*

Validated questionnaires will be used to design the survey questionnaire and these along with the data collection forms will be peer-reviewed to ensure construct validity. Further, the questionnaire will be pilot tested in a small sample of population to ensure face validity. Rigorous training and familiarity with tools will minimize intra- and inter-observer variability. Key considerations throughout the process are consistency, strong leadership, communication, and adherence to clearly defined roles.

VI.ii. Anticipated Challenges in Quality Assurance and Suggested Solutions

(a) *Representativeness of Sample*

Given the heterogeneity in socioeconomic status, cultural, nutritional, and linguistic characteristics of the sample populations derived from distinct sites, the investigators anticipate challenges in capturing samples that are representative of the wider South Asia region, and that inherent differences may limit generalizations in deductions made. This will be acknowledged in all products.

(b) *Response Rates and Loss to follow-up*

Further concerns include encountering low response rates as well as loss to follow-up of participants. Suggested solutions are: (a) engaging communities to be sampled in advance and obtaining approval from local authorities; (b) leveraging our institutional reputations; (c) methodical application of training, regular surveyor motivation and encouraging positive, professional and respectful behavior towards respondents; (d) advance scheduling of visits that are convenient for respondents, and providing a list of subsequent visits; (e) storing and backing up contact details of all respondents surveyed and supplemental contact information (for relatives, employers) in an accessible format, while appealing to those surveyed to notify the network surveillance site office of any change of contact; and (f) identifying those at risk of mobility and customizing our approach accordingly. Our investigators have extensive experience in recruiting and maintaining population cohorts (e.g., Industrial Health Study, New Delhi Birth Cohort Study, CUPS, and CURES). Furthermore, with continued follow up, differential loss-to-follow up or mortality can cause biases in those retained in the cohort – these will be addressed by examining the influences of these in analyses and acknowledging potential biases.

VII. MONITORING AND EVALUATION

Going forward, the **Precision-CARRS** Administrative and Field Coordination (AFC) Core will be responsible for quality assurance of the study, will monitor all phases of the study and will conduct formative, process, and outcome evaluations. This will help to maintain the timeline of the study.

VIII. ETHICAL CONSIDERATIONS

All subjects in Precision-CARRS will be asked to provide informed consent for follow-up data collection. All informed consent procedures, assessments, and interviews will be completed in Hindi, Tamil, or English, the most common local languages in Delhi and Chennai. The consent forms will describe each study component, the respondent burden, the potential risks and benefits, and provide names and contact details of individuals at CCDC, AIIMS, or MDRF who can be contacted for additional details, clarifications, or grievances. Field data collectors will read the informed consent forms to the respondent if requested, offer the opportunity to have any questions answered, and then hand it to the respondent for their signature. Written or thumb-print consent for illiterate participants (in the presence of a family member or witness) will be taken prior to enrollment or any further procedures taking place. Verbal consent will also be obtained and documented before each testing period. A copy of the consent form will be given to the respondent. Participants will be informed that their care providers will be notified of the assessment results and that they are free to withdraw at any time, without any consequence to existing care.

All data collection will use well-established techniques, including trained staff to administer survey questionnaires and conduct physical assessments. For a small portion of the questionnaire, a subsample of participants' voices may be recorded and stored on a secure server, where access will be restricted to the research team only as part of the cognitive assessments. The collection of bio-specimens is minimally invasive, requiring only venipuncture, and all efforts will be made to reduce discomfort. Data collection will be conducted at participant households or nearby neighborhood camps by trained staff from CCDC and MDRF. Laboratory analyses of blood specimens to estimate cardio-metabolic indices and pathophysiological biomarkers will involve standard operating procedures and quality control protocols. Data will be entered into secure databases. All analyses will involve accessing de-identified data. Subclinical and clinical CVD phenotyping procedures at neighborhood camps and our core facilities (ie. AIIMS) will follow well-defined, standardized procedures and are minimally invasive and do not cause more than minimal discomfort. Trainings will be held for each component of the study to ensure standardized and optimal techniques.

IX. TIMELINE

The first baseline cross-sectional survey (CARRS-1) began in 2010 with a completion time of twelve months. The participants were then followed up as a cohort with annual follow-up surveys for four subsequent rounds of data collection. A baseline cross sectional survey was conducted on newly recruited participants in 2015 and they were followed up as a cohort with a follow-up survey in 2018. In 2022, additional measures will be added to the follow-ups of both surveys combined for **Precision-CARRS** for participants in Delhi and Chennai, India. An additional follow-up of **Precision-CARRS** participants in India will take place in 2024-2026. Data analysis and reporting will take place after the collation of data at the end of the cross-sectional survey and thereafter at the end of annual follow-ups in each study site. Monitoring will be an ongoing process and the system of evaluation will be inbuilt into the study.

REFERENCES

1. Kraushaar L, Kramer A. Are we losing the battle against cardiometabolic disease? The case for a paradigm shift in primary prevention. BMC Public Health. 2009;9(1):64.
2. Mathers CD, Lopez AD, Murray CJL. The Burden of Disease and Mortality by Condition: Data, Methods, and Results for 2001. In: Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL, eds. *Global Burden of Disease and Risk Factors,*. New York: Oxford University Press; 2006.
3. International Diabetes Federation. The human, social & economic impact of diabetes. Available at: http://www.idf.org/home/index.cfm?node=41. Accessed February 2008.
4. Economic Intelligence Unit. *The silent epidemic: An economic study of diabetes in developed and developing countries.* New York, London, Hong Kong: The Economist June 2007.
5. Kempler P. Learning from large cardiovascular clinical trials: classical cardiovascular risk factors. *Diabetes research and clinical practice.* 2005;68 Suppl1:S43-47.
6. Stamler J, Vaccaro O, Neaton JD, Wentworth D. Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care.* 1993;16(2):434-444.
7. Ramachandran A, Snehalatha C, Latha E, Satyavani K, Vijay V. Clustering of cardiovascular risk factors in urban Asian Indians. *Diabetes care.* 1998;21(6):967-971.
8. Pradeepa R, Mohan V. The changing scenario of the diabetes epidemic: implications for India. *Indian J Med Res.* 2002;116:121-132.
9. Hughes LO, Raval U, Raftery EB. First myocardial infarctions in Asian and white men. *BMJ.* 1989;298(6684):1345-1350.
10. McKeigue PM, Ferrie JE, Pierpoint T, Marmot MG. Association of early-onset coronary heart disease in South Asian men with glucose intolerance and hyperinsulinemia. *Circulation.* 1993;87(1):152-161.
11. Ghaffar A, Reddy KS, Singhi M. Burden of non-communicable diseases in South Asia. *BMJ.* 2004;328(7443):807-810.
12. Mohan V, Sandeep S, Deepa R, Shah B, Varghese C. Epidemiology of type 2 diabetes: Indian scenario. *Indian J Med Res.* 2007;125(3):217-230.
13. Enas EA, Mehta J. Malignant coronary artery disease in young Asian Indians: thoughts on pathogenesis, prevention, and therapy. Coronary Artery Disease in Asian Indians (CADI) Study. *Clin Cardiol.* 1995;18(3):131-135.
14. Radha V, Mohan V. Genetic predisposition to type 2 diabetes among Asian Indians. *Indian J Med Res.* 2007;125(3):259-274.
15. Chambers JC, Elliott P, Zabaneh D, Zhang W, Li Y, Froguel P, Balding D, Scott J, Kooner JS. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet.* 2008;40(6):716-718.
16. Yajnik CS. Early life origins of insulin resistance and type 2 diabetes in India and other Asian countries. *J Nutr.* 2004;134(1):205-210.
17. Ramachandran A. Epidemiology of diabetes in India--three decades of research. *J Assoc Physicians India.* 2005;53:34-38.
18. Lev-Ran A. Human obesity: an evolutionary approach to understanding our bulging waistline. *Diabetes Metab Res Rev.* 2001;17(5):347-362.
19. Joshi P, Islam S, Pais P, Reddy S, Dorairaj P, Kazmi K, Pandey MR, Haque S, Mendis S, Rangarajan S, Yusuf S. Risk factors for early myocardial infarction in South Asians compared with individuals in other countries. *Jama.* 2007;297(3):286-294.
20. Diaz VA, Mainous AG, 3rd, Baker R, Carnemolla M, Majeed A. How does ethnicity affect the association between obesity and diabetes? *Diabet Med.* 2007;24(11):1199-1204.
21. McKeigue PM, Marmot MG, Syndercombe Court YD, Cottier DE, Rahman S, Riemersma RA. Diabetes, hyperinsulinaemia, and coronary risk factors in Bangladeshis in east London. *Br. Heart J.* 1988;60(5):390-396.
22. McKeigue PM, Shah B, Marmot MG. Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians. *Lancet.* 1991;337(8738):382-386.
23. Vardan S, Mookherjee S, Sinha AK. Special features of coronary heart disease in people of the Indian sub-continent. *Indian Heart J.* 1995;47(4):399-407.
24. International Diabetes Federation. Diabetes Atlas 3rd Edition. Available at: www.eatlas.idf.org/media. Accessed February 2008.

25. Gupta R, Kumar P. Global Diabetes Landscape - Type 2 Diabetes Mellitus in South Asia: Epidemiology, Risk Factors, and Control. *Insulin.* 2008;3(2):78-94.

26. Gupta R, Joshi P, Mohan V, Reddy KS, Yusuf S. Epidemiology and causation of coronary heart disease and stroke in India. *Heart.* 2008;94(1):16-26.

27. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care.* 2004;27(5):1047-1053.

28. International Diabetes Federation. *Diabetes Atlas 3rd Edition.* Brussels 2006.

29. Amos AF, McCarty DJ, Zimmet P. The Rising Global Burden of Diabetes and its Complications: Estimates and Projections to the Year 2010. Diabetic Medicine. 1997;14(S5):S7-S85

30. Gaziano TA, Reddy KS, Paccaud F, Horton S, Chaturvedi V. Cardiovascular Disease. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, Jha P, Mills A, Musgrove P, eds. *Disease Control Priorities in Developing Countries, 2nd ed.* New York: Oxford University Press; 2006:645-662.

31. Lopez AD, Murray CC. The global burden of disease, 1990-2020. *Nat Med.* 1998;4(11):1241-1243.

32. Reddy KS, Prabhakaran D, Jeemon P, Thankappan KR, Joshi P, Chaturvedi V, Ramakrishnan L, Ahmed F. Educational status and cardiovascular risk profile in Indians. *Proceedings of the National Academy of Sciences.* 2007;104(41):16263-16268.

33. Misra A, Pandey RM, Devi JR, Sharma R, Vikram NK, Khanna N. High prevalence of diabetes, obesity and dyslipidaemia in urban slum population in northern India. *Int J Obes Relat Metab Disord.* 2001;25(11):1722-1729.

34. Gupta R, Gupta VP, Sarna M, Prakash H, Rastogi S, Gupta KD. Serial epidemiological surveys in an urban Indian population demonstrate increasing coronary risk factors among the lower socioeconomic strata. *J Assoc Physicians India.* 2003;51:470-477.

35. Tripathi A, Jeemon P, Ajay V, Prabhakaran D, Reddy K. *Cardiovascular Diseases & its Risk Profile.* World Health Organization Country Office, New Delhi: IC Health & WHO; August 2007.

36. Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. *Lancet.* 2007;370(9603):1929-1938.

37. National Heart Lung and Blood Institute. *Report of the Working Group on Heart Disease Epidemiology*: NIH Publication 79-1667; June 1979.

38. World Health Organization. Introduction: Rationale for Surveillance of Chronic Diseases Risk Factors. *Chronic diseases and health promotion: STEPwise approach to surveillance (STEPS).* Vol 2008. Geneva: http://www.who.int/chp/steps/Part1_Section1.pdf

39. Balarajan R. Ethnic differences in mortality from ischaemic heart disease and cerebrovascular disease in England and Wales. *BMJ.* 1991;302(6776):560-564.

40. Bhatnagar D. The metabolic basis of increased coronary risk attributed to people from the Indian subcontinent. *Current Science.* 1998;74(12):1087-1094.

41. Dhawan J, Bray CL, Warburton R, Ghambhir DS, Morris J. Insulin resistance, high prevalence of diabetes, and cardiovascular risk in immigrant Asians. Genetic or environmental effect? *Br Heart J.* 1994;72(5):413-421.

42. Gupta M, Brister S. Is South Asian ethnicity an independent cardiovascular risk factor? *The Canadian journal of cardiology.* 2006;22(3):193-197.

43. Eiseman E, Fossum D. *The Challenges of Creating a Global Health Resource Tracking System*. Santa Monica, CA: RAND Corporation; 2005.

44. Heathfield H, Pitty D, Hanka R. Evaluating information technology in health care: barriers and challenges. *Bmj.* 1998;316(7149):1959-1961.

45. Gupta R. Recent Trends in Coronary Heart Disease Epidemiology in India. *Indian Heart J.* 2008;Suppl B:B4-B18.

46. Nongkynrih B, Acharya A, Ramakrishnan L, Ritvik, Anand K, Shah B. Profile of biochemical risk factors for non communicable diseases in urban, rural and periurban Haryana, India. *J Assoc Physicians India.* 2008;56:165-170.

47. Reddy KS, Prabhakaran D, Chaturvedi V, Jeemon P, Thankappan KR, Ramakrishnan L, Mohan BV, Pandav CS, Ahmed FU, Joshi PP, Meera R, Amin RB, Ahuja RC, Das MS, Jaison TM. Methods for establishing a surveillance system for cardiovascular diseases in Indian industrial populations. *Bull World Health Organ.* 2006;84(6):461-469.

48. Sethi R. Sample Registration System in India. September 02, 2008. Available at: http://www.searo.who.int/LinkFiles/2007_India1.pdf.

49. Narayan KM, Benjamin E, Gregg EW, Norris SL, Engelgau MM. Diabetes translation research: where are we and where do we want to be? *Annals of internal medicine.* 2004;140(11):958-963.

50. Goff DC, Jr., Brass L, Braun LT, Croft JB, Flesch JD, Fowkes FG, Hong Y, Howard V, Huston S, Jencks SF, Luepker R, Manolio T, O'Donnell C, Robertson RM, Rosamond W, Rumsfeld J, Sidney S, Zheng ZJ. Essential features of a surveillance system to support the prevention and management of heart disease and stroke: a scientific statement from the American Heart Association Councils on Epidemiology and Prevention, Stroke, and Cardiovascular Nursing and the Interdisciplinary Working Groups on Quality of Care and Outcomes Research and Atherosclerotic Peripheral Vascular Disease. *Circulation.* 2007;115(1):127-155.

51. Morris AD, Boyle DIR, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, MacDonald TM. The diabetes audit and research in Tayside Scotland (darts) study: electronic record linkage to create a diabetes register. *Bmj.* 1997;315(7107):524-528.

52. Morris AD, Boyle DI, McMahon AD, Greene SA, MacDonald TM, Newton RW. Adherence to insulin treatment, glycaemic control, and ketoacidosis in insulin-dependent diabetes mellitus. The DARTS/MEMO Collaboration. Diabetes Audit and Research in Tayside Scotland. Medicines Monitoring Unit. *Lancet.* 1997;350(9090):1505-1510.

53. Evans JMM, Newton RW, Ruta DA, MacDonald TM, Stevenson RJ, Morris AD. Frequency of blood glucose monitoring in relation to glycaemic control: observational study with diabetes database. *Bmj.* 1999;319(7202):83-86.

54. Evans JM, McNaughton D, Donnan PT, MacDonald TM. Pharmacoepidemiological research at the Medicines Monitoring Unit, Scotland: data protection and confidentiality. *Pharmacoepidemiol Drug Saf.* 2001;10(7):669-673.

55. Evans JM, MacDonald TM, Leese GP, Ruta DA, Morris AD. Impact of type 1 and type 2 diabetes on patterns and costs of drug prescribing: a population-based study. *Diabetes Care.* 2000;23(6):770-774.

56. Joshy G, Simmons D. Diabetes information systems: a rapidly emerging support for diabetes surveillance and care. *Diabetes Technol Ther.* 2006;8(5):587-597.

57. Clifford GD, Blaya JA, Hall-Clifford R, Fraser HS. Medical information systems: a foundation for healthcare technologies in developing countries. *Biomed Eng Online.* 2008;7:18.

58. Rizzo L, Brick JM, Park I. A minimally intrusive method for sampling persons in Random Digit Dial surveys. Public Opinion Quarterly. 2004;68(2):267-74.

59. World Health Organization. STEPwise approach to surveillance (STEPS). Available at: http://www.who.int/chp/steps/manual/en/index.html. Accessed September 28, 2008.

60. SAS I. SAS/STAT Users Guide Version 6. Cary, NC: SAS Institute; 1990.

61. StataCorp LP. STATA Data Analysis and Statistical Software. Available at: http://www.stata.com/.

62. Neaton JD, Duchene AG, Svendsen KH, Wentworth D. An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. Statistics in Medicine. 1990;9(1-2):115-24.

63. Day BS, Fayers BP, Harvey BD. Double Data Entry: What Value, What Price? Controlled Clinical Trials. 1998;19(1):15-24.

64. Carrsprogram.org [Internet]. [cited 2021 Jan 7]. Available from: http://www.carrsprogram.org/

65. Lloyd-Jones DM, Wilson PWF, Larson MG, Beiser A, Leip EP, D'Agostino RB, Levy D. Framingham risk score and prediction of lifetime risk for coronary heart disease. The American Journal of Cardiology. 2004 Jul 1;94(1):20–24.

66. Jeppesen J, Hein HO, Suadicani P, Gyntelberg F. Low Triglycerides–High High-Density Lipoprotein Cholesterol and Risk of Ischemic Heart Disease. Arch Intern Med. 2001 Feb 12;161(3):361–366.

67. Kavousi M, Desai CS, Ayers C, Blumenthal RS, Budoff MJ, Mahabadi A-A, Ikram MA, Lugt A van der, Hofman A, Erbel R, Khera A, Geisel MH, Jöckel K-H, Lehmann N, Hoffmann U, O'Donnell CJ, Massaro JM, Liu K, Möhlenkamp S, Ning H, Franco OH, Greenland P. Prevalence and Prognostic Implications of Coronary Artery Calcification in Low-Risk Women: A Meta-analysis. JAMA. 2016 Nov 22;316(20):2126–2134.