STUDY PROTOCOL

Solan Surveillance Study

Study protocol: Version 1.2 | 12 November 2013

Solan Surveillance Study

Study Protocol

Version 1.2 | 12 November 2013 -

CENTRE FOR CHRONIC DISEASE CONTROL

Funding agency:

Medtronic Foundation

Sub-study: National Heart Lung and Blood Institute, National Institutes of Health, USA

Contents

Contents

List of Abbreviation

BMI	Body Mass Index
BP	Blood Pressure
BRFSS	Behavioural Risk Factor Surveillance System
COE-CARRS	Center of excellence- Center for Cardiometabolic Risk Reduction in South Asia
CDC	Centers for Disease Control and Prevention
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
EQ-5D	European Quality of Life 5
FPG	Fasting Plasma Glucose
HINTS	Health Information National Trends Study
HTN	Hypertension
IDSP	Integrated Diseases Surveillance Project
IRB	Institutional Review Board
KAP	Knowledge, Attitudes, and Perceptions
LMCI	Low- and Middle-Income Countries
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
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
UA	Unstable Angina
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. CMD are a growing public health problem worldwide which is gradually becoming a pandemic.[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]

Among Indians, 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] 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'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 CMD 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 India, [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 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] Thus India 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]

I.vi. Study partners and their strengths

The Investigators of our study are Drs. **Dorairaj Prabhakaran** (Principal Investigator, CCDC), **Nikhil Tandon** (AIIMS), and **Ajay V S** (PHFI). The study will be coordinated by the Centre for Chronic Disease Control, New Delhi, India. The investigators are currently leading another surveillance study in urban Delhi, Chennai and Karachi using a similar protocol – The CARRS Surveillance Study. The expertise and learning from implementing the CARRS Surveillance Study will be an added benefit in implementing the proposed study.

I.vii. Objectives of the study

Primary objectives

a. - To implement and evaluate a model surveillance system for CMDs and its risk factors in rural Solan, Himachal Pradesh

b. - To assess the prevalence of CMD risk factors and diseases among adults aged 20 years and above, permanently residing in well-defined rural communities in Solan, Himachal Pradesh

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

Secondary objective

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

b. - 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

II. Methodology

II.i. Study design

While the primary study design for the surveillance model is cross-sectional, a cohort study design will be used to follow-up the participants for four years subsequent to the cross-sectional study. The cross-sectional study will assess the prevalence of CMDs and their risk factors while the pilot cohort study will estimate the incidence of morbidity and mortality associated with CMDs.

The **Cross-sectional study** will be the initial study conducted in 2012-2013 which will also form the baseline for the cohort study conducted in subsequent years. A repeat cross-sectional survey will be conducted in 2016-2017 and by recruiting an independent sample in the same study site to estimate the trend (over 4 years) in prevalence of CMDs and their risk factors.

The **Cohort study** will follow the cross-sectional study and will be conducted at two years interval during 2014-2015 and 2016-2017. With informed consent, subjects enrolled into the cross-sectional study will be followed up as a cohort for four years to measure the outcomes of interest.

Outcomes of interest:

i. - Anthropometric Changes
  a. - Weight
  b. - Waist Circumference
  c. - Skinfold Thickness
  d. - Body Fat

ii. Development of new-onset intermediate risk factors
    a. Hypertension
    b. Diabetes mellitus
    c. Dyslipidemia
iii. Incident morbidity
    a. Stroke
    b. MI
    c. CHF
    d. Chronic Stable Angina
    e. Chronic Kidney Disease
iv. Complication
    a. PVD
    b. Retinopathy
    c. Nephropathy
    d. Neuropathy
    e. Amputation
v. Health service utilization and costs
    a. Hospitalization
    b. Outpatient use
vi. Mortality
    a. All cause
    b. CMD-specific

II.ii. Study site and settings

The Surveillance study will be conducted in Solan District of Himachal Pradesh. This will be a household survey wherein recruitment of participants and data collection will take place in the households.

II.iii. Sample size estimation

The proposed surveillance study will be implemented in 36 subcenter area in the Solan District. The study will be conducted in collaboration with the Government of Himachal Pradesh. As per the suggestions from the Department of Health & Family Welfare of Himachal Pradesh, six sub-centre areas adjoining each of the five Community Health Centres and the Civil Hospital of Solan district will be chosen for the study. Each sub-centre area has a population of 3000 subjects of all ages. Assuming that the age group of >20 years constitutes about 50% of the total population, about 54,000 people will be enrolled in this study i.e.1500 from each of the 36 sub-centers.

Using STEPwise methodology [59] we estimated the sample size required to capture CMD risk factor prevalence. 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 five age strata (20-30, 30-40, 40-50, 50-60 and 60 and above. Table-1 presents the cumulative subtotals of subjects required to observe appreciably consistent prevalence approximations for each of the commonly-known risk factors. Since, the population that we plan to cover is more than the highest required sample size (37,816 people), it will permit to reliably estimate one or more of the CMD risk factors for each of the gender and age strata identified above.

Beyond this, we will conduct follow-up surveys to collect pilot data on incidence of risk factors, CMDs complications, and CMD-specific mortality. Consent will be taken during the initial recruitment for the cross-sectional survey and only those participants who provide consent to be followed up for four years will be enrolled into the study. However we anticipate an overall 15% loss-to-follow-up by the end of the cohort study especially due to high probability of 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.

Table – 1 Sample size estimation

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	0.02	0.3	1.5	0.8	10	37816
Hypertension	1.96	0.02	0.3	1.5	0.8	10	**37816**
Diabetes	1.96	0.02	0.04	1.5	0.8	10	6915

Sub-Study:

With the support from NHLBI, blood samples will be collected in 4000 subjects using a multi-stage cluster random sampling technique considering equal proportions of males and females leading to a sample of 4000 participants.

Primary study design for original study is cross-sectional and a cohort study design will be used to follow-up the participants for four years subsequent to the cross-sectional study. Data collection for cross-section study stared in 2013 and is expected to be completed by May 2014.

The sub-study involves blood collection in a sub-sample of 4000 participants in this rural cohort at the baseline. For this, we plan to organize 100 blood sample collection camps in the clusters in six months to complete the proposed sub-study. Twelve phlebotomists will be recruited for blood sample collection from the camps and from households of selected participants. With this we will be able to complete the blood sample collection in 6 months."

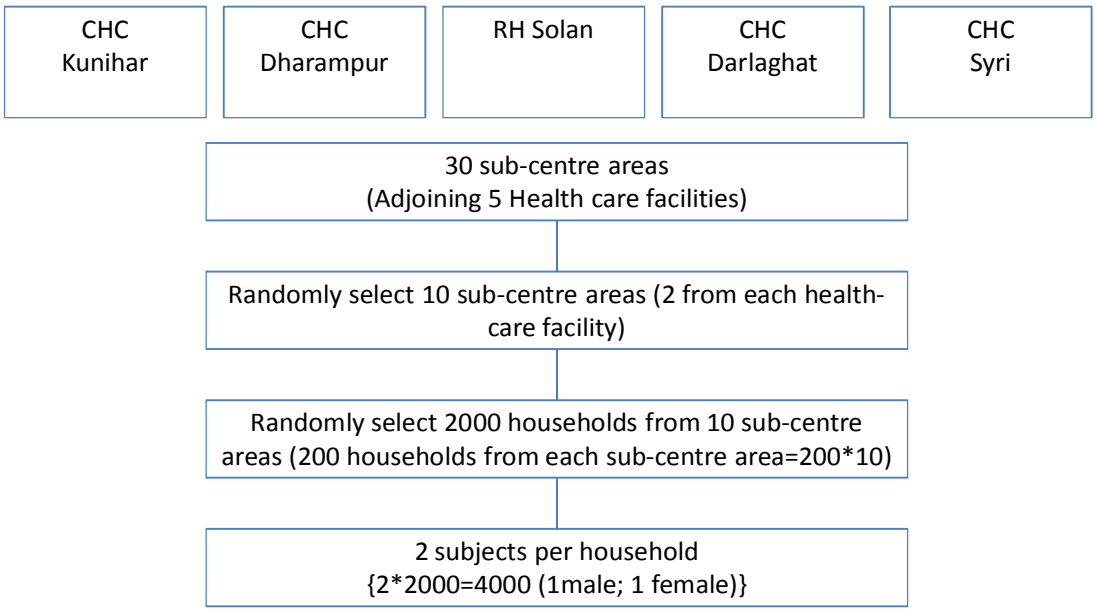
Sampling methodology for sub-study:

In Solan, 5 Government health care facilities (4 Community health centres and 1 regional hospital) were selected after recommendation from the health department, Solan. The six sub-centre areas adjoining each of the 5 healthcare facilities (i.e. 30 sub-centre areas) were selected based on proximity to the health facility.

The total sample size for the study is 4000 individuals. A multi-stage cluster random sampling technique will be used to capture a sample representative of the rural population. The 2 sub-centre areas will be selected randomly from each adjoining health care facility. From each sub-centre areas 200 households will be selected randomly. With informed consent, two subjects will be selected per household, one male and one female aged more than 20 years. The final sample for the study will be composed of equal proportions of males and females leading to a sample of 4000 participants (Figure 1).

Figure 1: Diagrammatic representation:

District Solan: 5 Health Care Facilities -



II.v. Study tools

Standardized sampling will enhance representativeness and reproducibility of results. Utilizing uniform tools and methods is a key component of this process and more easily replicable across multiple sites. We consulted validated questionnaires from the WHO Multinational MONItoring of trends and determinants in CArdiovascular disease (MONICA) study, WHO STEPwise Surveillance methodology, and existing questionnaires from regional and national surveys (accessible via partners at CDC), to construct a culturally appropriate and methodologically relevant questionnaire for India. Validated and new event (CVD or target organ damage) capturing tools was adapted to verify questionnaire-based follow-up survey incidence findings.

II.vi Surveillance indicators

A baseline cross-sectional survey of 54,000 participants in 2012-2013 will gather 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 will be comprehensive and encompass collection of data, anthropometric measures, and fasting glucose estimation using glucometers.

Demographic and social characteristics of the population will be collected including information regarding age, sex, marital status, religion, education, income, occupation, migration and household assets. Apart from this, meticulous collection of contact details of participants will help in following up the cohort and minimize loss to follow up. Validated questionnaires will be used to collect this information.

Several components of female reproductive history have been found to be associated with CMD and their risk factors. A detailed history on menarche, gestation, menopause and contraception will be taken through validated questionnaires to find any such association in our study population.

This study will assess both the distal and intermediate risk factors for CMD. Distal risk factors of interest are mostly related to lifestyle and include alcohol use, dietary habits, physical activity, stress and tobacco use. The intermediate risk factors will include study of prevalence of hypertension, obesity and diabetes. Tools that will be used to measure distal risk factors are validated questionnaires (food frequency questionnaire, physical activity questionnaire, etc.). The intermediate risk factors will be assessed through anthropometric measurements (Height / Weight / Waist Circumference/ Skinfold thickness / Body Fat), blood pressure measurement and through glucometer based fasting glucose estimation.

Prevalence of certain morbidity indicators of CMD will be studied. These include disability from Stroke, MI, CHF, Amputation, Chronic Stable Angina, CKD, revascularization, and other procedures and hospitalization related to CMD. This will be assessed through validated questionnaires and cross-checked with medical records (Table-3).

Table 2: Summary of the Surveillance Indicators

Indicators	Measures	Methods
Demographic and Social Characteristics	Age / Sex / Marital Status / Religion	Questionnaires
	Education / Income / Occupation	
	Household assets	
	Contact Details (and supplemental contacts)	
Female Reproductive history	Menarche/ Gestational history (PIH, GD), Menopause (surgical / physiological / whether on HRT) / Contraception	Questionnaires
Distal risk factors	Alcohol Use / Dietary Habits / Physical Activity / Stress	Questionnaire
	Tobacco use	Questionnaire
Intermediate risk factors	Hypertension	Clinical Examination, BP measurement
	Obesity	Anthropometry (Height / Weight / Waist Circumference / Skinfold thickness / Body Fat)
	Diabetes	Glucometer (For everyone) and Fasting Blood Glucose (for 4000 participants in Sub-study) *
	Dyslipidemia*	Laboratory estimation of Lipid Profile (for 4000 participants in Sub-study)
Morbidity	Stroke / MI / CHF / Chronic Stable Angina	Questionnaires including medication history; Medical records of documented events, procedures, serum urea and creatinine and albumin for CKD
	CKD/ Dialysis / Renal transplantation	
	Amputation/diabetic retinopathy	
	Procedures, Revascularization, Hospitalization	
Quality of Life	Mobility, self care, usual activities, pain/discomfort, anxiety/depression (related to CMD and their risk factors)	EQ-5D
Health services, Quality of Care and Health care costs	Awareness and Risk Factor Control	One-to-one Interview, Questionnaire,
	Access to Health Care Services	One-to-one Interview, Questionnaire
	Utilization of Services	One-to-one Interview, Questionnaire
	Health Insurance / Coverage	One-to-one Interview, Questionnaire
	Costs of Treating CMDs and CMR	One-to-one Interview, Questionnaire
Genetic and metabolomic studies (for 4000 participants in Sub-study)*	To identify genetic determinants of CMDs and their risk factors	Genome-wide association studies (GWAS)

BP = blood pressure; 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.

** In the selected sub-sample population of 4000 participants, blood sample will be collected.*

II.vii. Study of cost burden

The study will utilize 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),

that may be used to model **societal burdens** of disease and potential avoidable mortality, disability, and costs with addition of proven interventions. Direct medical and non-medical costs will be ascertained from respondents through standardized questionnaires applied in other low- and middle-income countries. Data will also be collected from providers in different settings (private and government-funded clinics) to provide corroborated data for sensitivity analyses. Indirect costs will be estimated using a human capital approach which is based on data regarding absence from work and unit costs derived from national wage rates (for different occupation classes).

Lost production = total absence X wage rate of workers

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

II.viii. Qualitative exploration

II.viii.a. Background

There is very little currently published regarding the knowledge, perceptions, attitudes, and care practices of people with cardio-metabolic risk factors and diseases in India. The surveillance study serves as a robust platform for investigating these qualitative aspects of cardio-metabolic diseases.

II.viii.b. Study questions

1. - Identify valid indicators for assessment of the economic burden of CMD and their risk factors in different socioeconomic backgrounds and regions.
2. - Analyse the temporal trends and peculiarities of the socio-economic effects, knowledge, attitudes, perceptions, and practices of persons with risk factors (DM, HTN, Smokers)
3. - Knowledge, attitudes, perceptions, and practices of persons with cardiovascular (chronic stable angina, previous MI or UA, previous revascularization, PVD, previous stroke) and chronic kidney diseases

4. - Factors that influence self-care and health-seeking in India amongst those with risk factors and vascular diseases
    i. - Continuity of care
    ii. - Normal versus alternative forms of health care
    iii. - Trends and peculiarities associated with self-care and health-seeking in this region of the world
5. - The patient's experience with and without existence of complications (cardiac, eye, renal, foot, and neurological illnesses that result from diabetes and hypertension)

II.viii.c. Methodology

- Careful selection of cases
- Cases will be participants who are known patients of any of the CMDs with or without complications
- Sampling methodology - purposeful sampling (grounded theory) - 10-25 cases per site, each subsequent case will be decided based on the interview and analysis of the existing samples
- One-to- one interview
- Identification of issues through preliminary interviews which will determine the direction and the necessity of further interviews and selection of cases

II.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 and diabetes), morbidity (underlying target organ damage, health consequences, and disability) and mortality associated with CMDs during the follow-up study of the cohort for four subsequent years. CMDs and their complications will be diagnosed using standard definition and will be coded using MedDRA / ICD codes. The outcomes of interest are described in Table-3

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	Hypertension	Clinical Examination, BP measurement
	Diabetes	Glucometer (For everyone) and Fasting Blood Glucose (for 4000 participants in Sub-study)
	Dyslipidemia	Laboratory estimation for serum total cholesterol, Low density lipoprotein cholesterol, high density lipoprotein cholesterol, and triglycerides (for 4000 participants in Sub-study)
Incident Morbidity	Stroke / MI / CHF / Chronic Stable	Follow-up surveys; Medical records of

Outcome	Measures	Methods
	Angina	documented events, admissions or procedures; Rose Angina Questionnaire
	CKD/ Dialysis / Renal transplantation	
	Amputation/diabetic retinopathy/ Procedures /Revascularization/ Hospitalization	
Mortality	All cause	Follow-up surveys; Death Certificates; Verbal Autopsy (WHO)
	CVD-specific; Diabetes-specific	

BP = blood pressure; 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.

II.x. Biological sample collection and storage

It would be a unique opportunity to include bio-chemical investigations in a subsample of 4000 participants in this rural cohort at the baseline. These participants will be randomly selected and 15 ml of the blood will be collected at visit 2 of the cross-sectional study. This proposal seeks support from the NHLBI for including blood sample collection and investigations for blood fats (Total cholesterol, LDL Cholesterol, HDL Cholesterol and Triglycerides) as a sub-study to the main-study. Left over blood samples will be stored for Genetic and metabolomic studies in future in order to identify genetic determinants of CMDs and their risk factors.

III. Data management

III.i. Data collection

Data will be collected through paper based, interviewer administered questionnaires in the participant's house. For the cross-sectional survey this will be accomplished through two visits (first visit for questionnaire based data collection followed by anthropometry, BP measurement and fasting glucose measurement in the second visit). In a subsample of 4000 participants blood sample will also be collected during the second visit. Thereafter one follow-up visit for each participant will be done at an interval of two calendar years. The implementation process of the study tools are described in table-4. The interviewers will be rigorously to minimise errors and constantly supervised. Quality assurance will be through monitoring and evaluation discussed in section V.

Table 4: 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 (glucometer for every one)	Blood (Fasting Blood Glucose & lipid profile for 4000 participants in Sub-study)	Verbal Autopsy/ Death Certificates
Cross Sectional Study-I																
2012-13	Participant recruitment & Data collection	X	X	X	X	X	X	X	X	X	X	X	X	X		
***Cohort* (Follow-up)**																
2014-2015	Follow-up – 1			X		X	X	X	X			X	X	X	X	X
2016-2017	Follow-up -2			X		X	X	X	X			X	X	X		X
Cross Sectional Study-II																
2016-2017	Participant recruitment & Data collection	X	X	X	X	X	X	X	X	X	X	X	X	X		

EQ-5D: European Quality of life five dimensions Questionnaire

III.ii. Data entry

A database application will be created using Microsoft access having a user friendly data entry screen in English. 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 such as ranges, absolute and relative values, context and structure. This will control the quality of data. The database will be assessed during the pilot study and necessary restructuring done before finalising it.

Data will be entered in the entry forms of the database centrally. Double data entry will be done for 25% of the forms to cross-check errors. There will be regular checks by site coordinator to minimise errors, missing values and outliers. Any errors found will be corrected by referring to the filled questionnaire or if required by a revisit to the participant. The data entry will be monitored and checked at CCDC. Final data cleaning and de-linking of participant identification information will be done at CCDC before conducting the analyses.

III.iii. Data analysis

(a) Quantitative Analysis

Data analysis will be staggered to correspond to follow completion of each survey. All statistical analysis of quantitative data will be done using Statistical Analysis Software (SAS, version 9.1, SAS Institute, Cary, N.C)[60] or STATA (Statacorp, TX).[61]A probability of $p<0.05$ will be considered statistically significant for all tests. Assumptions of normally-distributed data will be assessed using plots and tests of normality. Non-normal variables will be transformed or categorized as required. All data will be presented before and after adjustment for confounding and testing for interactions.

Descriptive data analyses will be performed for all variables. Means, standard deviations, quartiles, and median levels of risk factors for each gender and age group will be reported. Comparisons across gender, age groups, and time periods will also be done using Chi-square and logistic regression for categorical variables, and t-tests and linear regressions for continuous variables. Evaluation of predictors of incidence and mortality will employ Cox's proportional hazards models. Response rates and percent retention will be evaluated by comparing the number of eligible individuals approached, number agreeing to participate, and successful follow-up over duration of surveillance studies.

(b) Qualitative Data Analysis

All interviews will be de-identified prior to analysis. The NVIVO version 10 program will be used to manipulate textual data for analysis. Analysis of the textual data will follow the grounded theory methodology whereby key themes are identified inductively from the textual data. These themes will then be compared using structured comparisons to identify specific issues.

IV. Reporting and publication

At defined points over the course of the four years, formal analysis of quantitative and qualitative data will be reviewed for validity and subsequent publication. The investigators will be the custodians responsible for assigning subgroup analyses and publication under the oversight of this study. Reported outcomes of interest will include: age-standardized prevalence and mean levels of CMD risk factors; incidence rates for emergence of risk factors, disease events, and mortality; levels of awareness; health care costs; health service utilization; quality of life and determinants of human behaviors (qualitative measures).

V. Quality assurance

V.i. Quality assurance strategies

Quality assurance strategies will 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-5).

Table 5: Quality assurance strategies

		Phase			
		Design and Planning	**Pilot Testing**	**Data Collection**	**Data Analysis**
Levels	**CCDC**	• 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 (for the 4000 participants in Sub-study)	• 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	• Random checks • External temperature gauge labels to monitor sample temperature	• Stored samples for future investigation • Identify and discard compromised samples

	Phase			
	Design and Planning	**Pilot Testing**	**Data Collection**	**Data Analysis**
		to detect outliers		
Laboratory	• Central procurement • Develop internal and external quality assessment protocols and schedule of regularity	• Evaluate procedural fluidity • Evaluate intra- -laboratory variability • Interim analysis to detect outliers	• Internal quality checks and calibration • Regular external validation – lyophilized samples from reference laboratory	• Assess intra-laboratory coefficients of variation
Communication	• Establish reporting structures • Establish data transfer plans	• Assess agility of transfers		
Documentation	• Develop checklists, logbooks • Training in appropriate and legible documentation	• Assess recording legibility		• Audit logbooks for response rates and 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*

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.

V.ii. Anticipated challenges in quality assurance and suggested solutions

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 Surveillance Study, CARRS-Surveillance Study, New Delhi Birth Cohort Study, CUPS, and CURES).

VI. Monitoring and evaluation

A Quality Monitoring and Evaluation (QME) sub-committee will be responsible for quality assurance of the study, will monitor all phases of the study and will conduct formative, process and outcome evaluation. The QME sub-committee will include the principal investigator, site coordinators and external evaluators. Progress of the study will be monitored through regular appraisal by the QME sub-committee.

VI.i. Procedures for evaluating the process of the study

VI.i.1. Sampling*:*

a. - After the list of households from all the sub-centres are prepared, random checks will be done especially for areas that have been undergoing continuous development to ensure that all households are enlisted.

b. - Cross-check the household number generated with the number of the household in which participants were interviewed.

VI.i.2. Data collection**:**

a. - Initial shadow monitoring of all interviewers will be done for practical training and to resolve issues immediately. This will also help to identify weak personnel who may require intermittent supervision.

b. - As the study progresses, random checks for about 5-10% of the participants will be done every three months by site coordinators and project managers who will verify some of the answers in the filled questionnaires.

c. - There will be regular documentation of the response rates from participants especially for the follow-up of cohort to minimise loss to follow-up. The QME will also be responsible to audit these rates and other documented field activity indicators such that immediate corrective measure can be taken.

VI.i.3. Data entry:

a. - Project Manager will be responsible for checking all forms for completion or for any obvious errors before data entry.

b. - In-built checks in the data entry software – logic checks, context checks and ranges will be incorporated for outliers.

c. - The outcomes of follow-up study for all participants will be validated by physicians and intra and inter observer variability will be checked by the Project Manager and / or an external evaluator before data is entered

d. - Expected error rates: Every time an error is located it will be given a mark of one, once all participants' and field data are checked (all fields, all rows, all columns), the error marks will be summed and using denominators of total fields, context fields and outcome fields, error rates for all fields, context fields and outcome will be respectively generated. The error rates are usually expressed as errors per 10,000 fields[62] [2]. Error rates vary for studies and different studies use different rates as acceptable limits[62, 63] [1, 2]. Since it is a large study with a very large number of data fields, we use the error rates suggested by Neaton et al. [2], 10 errors per 10,000 fields or 0.001. However, for context or participant identification / demographic fields and for outcome fields "zero tolerance" will be used (acceptable error rate=zero) [1, 2].

Error rate – all fields (0.001) -

Error rate – context fields (Zero) -

Error rate – outcome of cohort study (Zero) -

e. - Any errors found will be corrected by referring to the filled questionnaire or if required by a revisit to the participant.

a. - The data manager and statisticians will re-check all data entered for outliers, coding errors and missing values.

b. - 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 interviewers and outcome.

c. - Expected query rates: Every time a query is located it will be given a mark of one, once all data are checked, the query marks will be summed and using denominators of total fields per interviewer will be generated. Acceptable query rates will be:

Query rate – interviewer (0.001) -

Query rate – outcome for cohort study (zero) -

d. - Measures to be taken if the query and error rates are higher than the pre-decided value:

- High query/error rate for interviewer – all data for the particular interviewer will be checked against paper forms at the study site.
- If there is any error in the outcome for the cohort study, first the data will be cross-checked against paper forms and if required the participant will be re-visited. If none of these can correct the error, the error field will be dropped as missing.

e. - Decision log will be used to document all issues related errors and queries (using excel sheets and emails)

f. - Monitoring data storage and confidentiality procedures: This will be done by the PIs and external evaluators at the CCDC after the complete collation of data and before using the data for analysis.

VI.ii. Procedures for evaluating the outcome of the study

1. - After the study is complete an independent evaluation will be done by the PIs and the external evaluators to ascertain if the aims and objectives of the study are fulfilled. This will be done through review of all the preceding evaluation processes, and also through review of the findings and results of the study.

VII. Ethical considerations

Informed consent for interview, anthropometric measurements and blood collection will be taken from all individuals before enrolling them into the study. The study anticipates no major risk to participants except needle pricks for collection of blood sample during the follow-up and cross-sectional surveys. There could be rare instances of bleeding from the pricking site, however trained staff will minimise even the rarest of such possibility.

Supplemental Informed consent will be sought from the sub-study participants for blood sample collection and storage of left over blood for Genetic and metabolomic studies to identify genetic determinants of CMDs and their risk factors. Participation in this sub-study may involve some added risks or discomforts. These include: the risk of bleeding, bruising, and infection at the site of drawing blood. These risks will be minimized by having the blood drawn under sterile conditions by experienced personnel. Strict confidentiality of the information provided by the participants will be maintained and all possible source of identification of participants will be delinked before using the data for analysis.

VIII. Timeline

The Surveillance Study will be conducted over a period of five years. The baseline cross-sectional survey will begin in 2012 with estimated completion time of twelve months. The participants will be then followed up as a cohort with biannual follow-up surveys for four subsequent years. 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. Monitoring will be an ongoing process and the system of evaluation will be inbuilt into the study. The details of the timeline are provided in table-6.

Table 6: Timeline for Surveillance Study

Surveillance Study	Timeline	

Activities	2012		2013		2014		2015		2016		2017	
Household and participant recruitment												
Selection of households and recruitment of study participants for both the cross-sectional and cohort studies												
Data Collection												
Cross-sectional surveys												
Follow-up survey for incident events and mortality												
First follow-up survey for incident events and mortality												
Second follow-up survey for incident events and mortality												
Data entry and data cleaning												
Analysis												
Data analysis and reporting												
Audits and reporting												
Monitoring and evaluation												
Monitoring												
Process evaluation												
Outcome evaluation												
Obtain Support for Expansion of Surveillance Model												

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.