COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author’s moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author’s reputation

For further information contact the University’s Copyright Service.

sydney.edu.au/copyright
Diabetes and lipid levels in Rural Andhra Pradesh, India - 2005 to 2014

Eshan Tahrir Affan

Thesis submitted in fulfilment of the requirements for the Degree of Master of Philosophy in Medicine,
The University of Sydney,
March 2015
Abstract

Background
Rural Andhra Pradesh is a developing region in India with reported high prevalence of diabetes mellitus and blood lipid levels. This is consistent with a worldwide transition of mortality and morbidity towards predominantly non-communicable conditions such as heart disease, stroke and diabetes. Detecting and monitoring these changes in less accessible regions of the world is a difficult task.

Methods
Three cross sectional studies examining cardiovascular risk factors in the Godavari region of rural Andhra Pradesh were done in 2005 (Andhra Pradesh Rural Health Initiative, APRHI), 2010 (Gates Grand Challenge 13, GC13) and 2014 (Systematic Medical Appraisal Referral and Treatment Health, SMART health). Diabetes prevalence from all three studies and blood lipid levels from 2005 and 2010 were compared to assess the trend. The data were further divided into a primary analysis including only fasting plasma glucose measurements and a secondary analysis which allowed for other methods of diabetes diagnosis. A systematic review and meta-analysis of the use of dried blood spots (DBS) for cardiometabolic risk factor analysis was also conducted.

Findings
Sixteen studies were included in the meta-analysis of the use of dried blood spots, 12 of which reported necessary data for haemoglobin A1c (HbA1c), one for triglycerides, two for both triglycerides and total cholesterol and one for HbA1c, total cholesterol and high density lipoprotein (HDL). Study sizes ranged from 30 to 613 participants. For HbA1c the summary regression (DBS = 0.9858Venous + 0.3809) showed close agreement between analyses based upon the venous and DBS sampling methods. The summary regression line for total cholesterol (DBS = 0.6807Venous + 1.151) indicates a requirement for moderate adjustment of values based upon analyses of DBS samples to obtain estimates equivalent to standard analyses based upon venous samples. For triglycerides, the summary regression for the three contributing studies showed a close association between the results obtained for the two methods (DBS = 0.9557Venous + 0.1427).

The primary analyses of dysglycaemia (diabetes and prediabetes) were based upon 3243 individuals from APRHI and 749 individuals from SMART health for whom fasting plasma glucose samples were available. The estimated prevalence of dysglycaemia was 53.7% (51.8 - 55.7, 95% CI) in 2005 and 62.0% (58.5 - 65.4, 95% CI) in 2014 (p<0.001). The primary difference in population characteristics across the two surveys was a more than one unit rise in mean BMI driven by an approximate 1.7kg rise in mean weight. For the secondary analyses there were 3333 individuals aged 40 to 85 included from the APRHI study in 2005; 2200 individuals included from the GC13 survey done in 2010; and 62 254 participants included from the SMART health survey in 2014. For the secondary analyses the estimated prevalence of dysglycaemia was 53.9% (52.0 - 55.9, 95% CI) in 2005, 50.5% (46.1 - 54.9, 95% CI) in 2010, and 41.3% (40.9 - 41.7, 95% CI) in 2014 with the data suggesting a decline across the three time points (p<0.001). There were significant increases observed between 2005, 2010, and 2014 for both body mass index (BMI) (p<0.001) and weight (p<0.001). Mean total cholesterol decreased from 4.6mM in 2005 to 3.4mM in 2010 while mean low density lipoprotein decreased from 2.9mM to 1.5mM during the same period.

Conclusions
The use of dried blood spots as a method of reporting HbA1c levels appears justified but further studies are required to confirm the suitability of blood lipid level measurements based upon dried blood spots. Every estimate of dysglycaemia was high suggesting that dysglycaemia is a major problem in this part of Andhra Pradesh. Further, the upward trend in dysglycaemia observed in the primary analyses and the corresponding adverse trend in obesity measures suggests that the problem is getting worse not better. The differences in observed trends between the primary and secondary analyses are almost certainly attributable to the variable
assay methods used across the data included in the secondary analyses. The steep changes in lipid levels are likely incorrect and have been biased by the study methods.
Table of Contents
Chapter 1 – Introduction and Literature Review
Chapter 2 – The Three Surveys
Chapter 3 – Comparability of HbA1c and Lipids Measured with Dried Blood Spots Versus Venous Samples: a systematic review and meta-analysis
Chapter 5 – Blood Lipids in Rural Andhra Pradesh – 2005 and 2010
Chapter 6 – Discussion and Conclusions
Appendix 1 – Assessing Potential Sources of Bias in GC13
Appendix 2 – APRHI Survey
Appendix 3 – GC13 Survey
Appendix 4 – SMARTHealth Survey
Acknowledgments

I would like to thank my supervisor Professor Bruce Neal for his continuous guidance and support throughout the project. Through the course of this study I have gained much experience and had much teaching in epidemiological methods and studies. His encouragement and perspective is what has led me to continue pursuing a research career, I feel I was always given just the right amount of autonomy and guidance which made my work enjoyable.

I would also like to thank my co-supervisor Associate Professor Clara Chow for her advice, guidance and provision of the Andhra Pradesh Rural Health Initiative dataset. I am particularly grateful for Dr Devarsetty Praveen who I worked with day to day to formulate new directions, carry out statistical analyses and produce publications based on all three datasets. Also I thank Associate Professor David Pereis and Professor Anushka Patel who along with Dr Devarsetty Praveen provided the SMARTHealth dataset and advised on the development of the publications.

Also I am indebted to Dr Jason Wu, Dr Stephane Hertier and Dr Laurent Billot who provided much support and teaching in statistical analysis methods for survey data.

Finally I would like to thank all of my family and friends who were persistent with me and provided invaluable support throughout these years.
Chapter 1 – Introduction and Literature Review

Changing disease patterns

We are currently passing through a notable transition in the history of human health. The most significant causes of mortality for hundreds of years are regressing. Previously, infectious diseases, poor pregnancy care and childhood undernutrition were the significant conditions plaguing quality of life and premature mortality. Changes in this are due to improvements in basic sanitation and technological advances in health care mainly during the 20th and 21st centuries. With this transformation there are a new wave of diseases now responsible for the majority of global mortality. Non-communicable diseases now account for over two thirds of global mortality. This includes chronic diseases such as ischaemic heart disease, diabetes and cerebrovascular disease. The Global Burden of Disease 2010 study was paramount to exposing the changing impacts of these chronic diseases worldwide. The extensive data generated has provided numerous examples of the epidemiological transition to increasing global morbidity and mortality from chronic, non-communicable diseases. This is true when viewing the mortality and morbidity attributed to either the specific diseases or their risk factors.

In 1990, ischaemic heart disease and stroke was responsible for one fifth of all mortality and that has increased to one quarter in 2010. Similarly, deaths due to diabetes have doubled from 1.3 million worldwide in 1990. When measured as disability adjusted life years (DALY) lost, ischaemic heart disease has progressed from 4th to 1st, stroke from 5th to 3rd and diabetes mellitus from 21st to 14th. Meanwhile lower respiratory tract infections fell from 1st to 2nd, diarrhoea from 2nd to 4th, tuberculosis from 8th to 13th, pre-term birth complications from 3rd to 8th and protein-energy malnutrition from 9th to 20th. Regarding years lost to disability (YLD), the Global Burden of Disease study also shows that the main contributors were mental and behavioural disorders, musculoskeletal disorders and diabetes mellitus. These results underline the general transition in global health patterns due to demographic changes, changes in cause of mortality and changes in the causes of disability. It seems these changes are here to stay and will only continue to drive the transition in the sources of disease burden. A World Health Organisation (WHO) forecasting study used simulations that predicted deaths from communicable diseases will decline by a further 50% and mortality due to non-communicable diseases and injuries will more than double.

Another notable epidemiological transition is towards a higher fraction of the total burden of disease belonging to disability rather than premature mortality. The causes of disability and mortality are not necessarily congruent. Between the top 25 contributors to YLD and the equivalent list for years of life lost, only diabetes, road injury, ischaemic heart disease, chronic obstructive pulmonary disease, tuberculosis and diarrhoea are found in both groups. Much of this change is attributed to rising rates of diabetes and it has been suggested that in today’s age what ails the population the greatest is not always what kills them.

Certainly these changes in disease patterns are global but it appears that South Asia is facing an appreciable amount of the consequences. Non-communicable diseases have become increasingly prevalent as they are responsible for almost half of the burden of disease in South Asia. The rise of cardiovascular diseases in low and middle level income countries has been largely unnoticed when compared to the direct threats of infectious diseases, although 80% of deaths due to cardiovascular disease now occur in these countries.

The combination of poverty and changing cardiovascular disease risk factors is responsible for half of the adult burden of disease in South Asia. The fact that one quarter of the world population belongs to South Asia and that half of those are below the poverty line outlines the significance of the oncoming epidemic of cardiovascular diseases. The outcome is evident in India potentially affecting the working age population due to the highest numbers of productive years of life lost from cardiovascular conditions. This is a result of the excessive deaths due to cardiovascular disease amongst the 35-64 years age group, with 9.2 million life years lost in the year 2000 alone. The severity of the situation is clear since this figure is predicted to increase to 17.9 million years by 2030 which is 940% greater than the corresponding loss of life years in the United States due to cardiovascular causes. It is also noted that India has the highest absolute number of people with diabetes, a contributor to the statistic that 52% of cardiovascular deaths occur below the age of 70, compared to only 23% in high income countries.
Andhra Pradesh is a large state in southern India with a population of 84 million, approximately 70% of whom live in rural areas. The Godavari region consists of the East Godavari and West Godavari districts, with a total area of 18,509 km² and a population of 9 million. The region includes the Godavari river delta making it very fertile with a large part of the local economy dependent on fisheries and agriculture, including the majority of the rice production in the state. This is a rural population susceptible to rapid changes in lifestyle and diet where already there are high levels of diabetes, blood lipids, hypertension, smoking and obesity. Most recently 13.2% were found to have diabetes, 27% hypertension, 7.2% with hypercholesterolaemia and 22.3% overweight.\textsuperscript{13, 14}

**Drivers of changing disease patterns**

Non-communicable diseases are increasingly being linked to poverty and socioeconomic disparity. These are effects of globalization, urbanisation, demographic and lifestyle changes which are occurring particularly rapidly in developing regions of the world. This leads to increasing income gaps and a growing lower middle class susceptible to lifestyle related chronic diseases. The wider impact on the countries involved is not trivial. From the economic perspective China is estimated to lose the equivalent of $558 billion USD while India is estimated to lose $237 billion USD, due to preventable cardiovascular diseases and diabetes over the next ten years.\textsuperscript{15}

A review article in 2004 drew some attention to the then neglected area of chronic disease in India.\textsuperscript{11} The author presents a dichotomy of globalization, which is primarily responsible for the accelerating changes in developing countries by importing diets and lifestyles from high income countries but at the same time also facilitates the prevention of cardiovascular disease through a well-established knowledge base of risk factor mitigation. In particular, these epidemics of non-communicable diseases are driven by rapid shifts in demographics and lifestyle due to urbanisation, globalisation and industrialisation. These changes have been seen in high income countries previously but they occurred over the time span of an entire century rather than the few decades the current transitions are compressed into. Also, the rate of urbanisation in areas of uncorrected poverty leads to increasing income disparity and will leave the most poor especially vulnerable.

Considering the health changes associated with urbanisation and socioeconomic status, it has been shown that over the previous 20 years the rates of obesity in developing countries have tripled and further and that being poorer is coupled to obesity.\textsuperscript{16} The results have been that some families of low socioeconomic status in developing regions, such as rural India, have children who are initially underweight but grow to become adults who are overweight. The presumption is that intrauterine growth restriction is conferring a predisposition to obesity in adulthood. When this possibility is combined with rapid childhood weight gain due to overconsumption of cheap, energy-dense food and a lack of physical exercise there exists a platform conducive to insulin resistance and the metabolic syndrome.

More recently in the East and West Godavari regions of rural Andhra Pradesh there has been a focus on industrialisation with the development of oil and gas refineries. This is likely to fuel increased economic development in the area and potentially further urbanisation\textsuperscript{17} with which lifestyle and diet changes will follow. It is hypothesised that there will be an increase in cardiovascular risk factors including diabetes, total cholesterol levels, low density lipoprotein levels and body mass index. This would be consistent with the global changes observed over the last 10 to 20 years and is expected of a region undergoing demographic, social and economic changes conducive to increasing cardiovascular disease.

**Changing cardiovascular risk factors**

Chronic diseases account for 60% of all deaths and 80% of these are now in low or middle income countries. Interestingly the deaths due to chronic diseases in low and middle income countries are also occurring in younger age groups when compared to developed countries.\textsuperscript{5} The global burden of disease can be assessed by looking at the most significant risk factors for mortality and morbidity. In order to undertake informed primary prevention and interventions the risk factors with highest prevalence must be targeted. The top three leading risk factors contributing to global disease burden in 2010 were hypertension, tobacco smoking including second hand smoke and household air pollution from solid fuels. This is compared to 1990, where the leading risks were childhood underweight status, household air pollution from solid fuels and smoking including...
second hand smoke\(^5\). Hypertension, high body mass index (BMI), high fasting plasma glucose, alcohol use and dietary risks have all become more prominent contributors to the global burden of disease during the period 1990 to 2010. Although, when seen as absolute percentage prevalence, most of these risk factors have actually decreased, except high BMI and high fasting plasma glucose.

Examining the link between risk factors and the diseases impacting the global burden of disease sheds some light on where interventions are best focused. ‘Metabolic mediators of the effects of body-mass index, overweight and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1.8 million participants’ collated the relevant data from 1948 and 2005 and provided the appropriate hazard ratios\(^9\). The factors linking obesity to cardiovascular disease are clear. A BMI above 25 kg/m\(^2\) and over 30 kg/m\(^2\) are both significantly associated with an increased risk for coronary heart disease and stroke when compared to the widely accepted normal range of 20 kg/m\(^2\) – 25 kg/m\(^2\). The adjusted hazard ratios for a 5 kg/m\(^2\) increase in BMI are 1.27 (1.23 – 1.31, 95% CI) for coronary heart disease and 1.18 (1.14 – 1.22, 95% CI) for stroke. When further adjusted for elevated blood pressure, cholesterol and glucose, the hazard ratios fell to 1.15 (1.12 – 1.18, 95% CI) for heart disease and 1.04 (1.01 – 1.08, 95% CI) for stroke. Blood pressure was the prime risk factor, accounting for 31% (28% - 35%, 95% CI) of excess risk for coronary heart disease due to increases in BMI and 65% (56% - 75%, 95% CI) of the same for cerebrovascular disease. The principal conclusion is that efforts to reduce glucose, total cholesterol and high blood pressure might address approximately one half of the excess risk of coronary heart disease and up to three quarters of the excess risk of stroke attributed to high BMI. These results were not discriminatory between Asian and western cohorts.

A 2008 review made some further observations to add to the discussion, predicting that more than half of all cases of cardiovascular disease worldwide will be found in India by 2025\(^18\). The important determinants of cardiovascular diseases in India identified here were tobacco use, obesity, a high waist to hip ratio, high blood pressure, dyslipidaemia, diabetes, a low fruit and vegetable diet and sedentary lifestyles. It was suggested that perhaps genetic factors and a predisposition to diabetes may have a hand in the increase in cardiovascular diseases. Therefore it is clear that these risk factors identified globally to be responsible for the majority of morbidity and mortality are also responsible for the same health outcomes in India. Considering that 70% of India’s 1.2 billion population is rural, it is reasonable to focus efforts on early intervention in areas such as rural Andhra Pradesh to negate a potentially massive rise in cardiovascular risk factors.

Increased fasting glucose levels and diabetes features prominently in both disability and mortality in the Global Burden of Disease study and so would be a justified target for interventions both primarily as a risk factor for cardiovascular diseases but also as a disease process in itself. Two studies outline the changes and current status of worldwide diabetes prevalence and the attempts at controlling incidence. ‘National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: a systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants’ provides estimated trends in mean fasting plasma glucose and diabetes prevalence for adults aged over 25 in 199 countries and territories\(^7\). This showed that in 2008 the global age-standardised mean fasting plasma glucose was 5.5 mM (5.37 mM – 5.63 mM, 95% CI) for men and 5.42 mM (5.29 mM – 5.54 mM, 95% CI) for women which is a result of an average 0.07 mM and 0.09 mM increase respectively over every decade since data collection commenced. Additionally, the age-standardised diabetes prevalence was 9.8% (8.6% - 11.2%, 95% CI) in men and 9.2% (8.0% - 10.5%, 95% CI) in women. This is a significant increase from the 8.3% (6.5% - 10.4%, 95% CI) and 7.5% (5.8% - 9.6%, 95% CI) respectively seen in 1980. Oceania, in particular, had the most significant rise in mean fasting plasma glucose and diabetes prevalence but was only the leader in a closely following group of South Asia, Latin America, the Caribbean region, Central Asia, North Africa and the Middle East. These trends are distinct from other risk factors for cardiovascular disease such as systolic blood pressure which has actually decreased overall in most regions. Notably, there has been an increase in diabetes prevalence by about 7% each decade from 1980 and 70% of this increase can be attributed to ageing\(^19\). Rising adiposity rates are also powering this epidemic and estimate that this would account for much of the remaining 30% although there is a significant inter-regional heterogeneity to consider when determining the cause for an increase in diabetes in a particular area\(^20\).
One set of studies are presented in Table 1-1 with the full selection of diabetes prevalence studies based on specific rural regions of India. Another study based on India generally did not focus on specific regions, the Prevalence of Diabetes in India Study (PODIS) examined the general population through a random sampled multistage cross-sectional survey\textsuperscript{21, 22}. This survey was taken from 1999 to 2002 and was based on clusters of local areas and included 18,363 subjects with 7,746 of these from rural areas. It is difficult to discern if this was a case of simple random or systematic sampling after the clustering stage. The authors also quoted difficulties in accurate diagnosis since they could not rely on laboratory venous plasma levels of glucose in many of the clusters due to the remote locations. These practical issues underline the need for improved epidemiological techniques for measuring such analytes and are further discussed in chapter 3. The results from PODIS showed the total Indian diabetes prevalence was 4.3% (4.0% - 4.6%, 95% CI) with 5.6% (5.1% - 6.0%, 95% CI) urban and 2.7% (2.3% - 3.1%, 95% CI) in rural areas. The corresponding rates of impaired glucose tolerance were 5.2% (4.9% - 5.5%, 95% CI), 6.3% (5.8% - 6.8%, 95% CI) and 3.7% (3.3% - 4.1%, 95% CI). Based on these results the authors advocated for a diabetes action or prevention program carefully focussing on the rural segment where more than 70% of the population lives. These people are also the most vulnerable to the insidious effects of chronic disease.

**Table 1-1. Previous Rural Indian studies on diabetes prevalence.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Region</th>
<th>Sample Size</th>
<th>Diabetes Prevalence (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramachandran 1989\textsuperscript{23}</td>
<td>Tamil Nadu</td>
<td>1038</td>
<td>2.4%</td>
</tr>
<tr>
<td>Patandin 1994\textsuperscript{24}</td>
<td>Tamil Nadu</td>
<td>467</td>
<td>4.9%</td>
</tr>
<tr>
<td>Wander 1994\textsuperscript{25}</td>
<td>Punjab</td>
<td>809</td>
<td>4.6%</td>
</tr>
<tr>
<td>Singh 1998\textsuperscript{26}</td>
<td>Uttar Pradesh</td>
<td>1769</td>
<td>2.8%</td>
</tr>
<tr>
<td>Raghupathy 2002\textsuperscript{27}</td>
<td>Tamil Nadu</td>
<td>1221</td>
<td>2.1%</td>
</tr>
<tr>
<td>Balagopal 2003\textsuperscript{28}</td>
<td>Tamil Nadu</td>
<td>703</td>
<td>5.1%</td>
</tr>
<tr>
<td>Ramachandran 2003\textsuperscript{29}</td>
<td>Tamil Nadu</td>
<td>1213</td>
<td>6.36% (4.89-7.83)</td>
</tr>
<tr>
<td>Nirmalan 2004\textsuperscript{30}</td>
<td>Tamil Nadu</td>
<td>5150</td>
<td>4.4% (3.8-5.0)</td>
</tr>
<tr>
<td>Chow 2004\textsuperscript{31}</td>
<td>Andhra Pradesh</td>
<td>345</td>
<td>3.7% (1.8-5.5)</td>
</tr>
<tr>
<td>Chow 2005\textsuperscript{14}</td>
<td>Andhra Pradesh</td>
<td>4535</td>
<td>13.2% (12.1-14.3)</td>
</tr>
<tr>
<td>Ramachandran 2006\textsuperscript{32}</td>
<td>Tamil Nadu</td>
<td>2584</td>
<td>9.2% (8.0-10.5)</td>
</tr>
<tr>
<td>Kokiwar 2007\textsuperscript{23}</td>
<td>Nagpur</td>
<td>924</td>
<td>3.67%</td>
</tr>
<tr>
<td>Vijayakumar 2007\textsuperscript{34}</td>
<td>Kerala</td>
<td>1645</td>
<td>12.5%</td>
</tr>
<tr>
<td>Khatib 2008\textsuperscript{35}</td>
<td>Maharashtra</td>
<td>306</td>
<td>8.4%</td>
</tr>
<tr>
<td>Jonas 2010\textsuperscript{36}</td>
<td>Central India</td>
<td>2414</td>
<td>5.6% (5.1-6.1)</td>
</tr>
<tr>
<td>Vaz 2011\textsuperscript{37}</td>
<td>Goa</td>
<td>1266</td>
<td>10.3%</td>
</tr>
<tr>
<td>Madaan 2014\textsuperscript{38}</td>
<td>Haryana</td>
<td>4497</td>
<td>18.43%</td>
</tr>
</tbody>
</table>

Table 1-1 is similar to a previous table presented in “Review of the epidemiology of diabetes in India”\textsuperscript{39} although here it is updated to 2014 and includes the study year rather than the publication year where available. Included are studies surveying populations using fasting blood sugar, oral glucose tolerance test or glycosylated haemoglobin (HbA1c) as their diagnostic method. Also included are confidence intervals where possible, this is in addition to the previous publication. This table is also restricted to studies involving only specific rural populations, excluding urban studies or rural/urban mixed groups. The regional disparities in diabetes prevalence are evident, underlining the need for area specific studies if a registry system is not otherwise available. Only the studies in Tamil Nadu by Ramachandran et al. involved follow up of a series of results from the same area, their results are shown in figure 1-1. Otherwise, there is little information about the changing diabetes prevalence in each of these specific areas. By revisiting this rural region of Andhra Pradesh, valuable insight can be gained into the changes in diabetes in an area which has already shown an alarmingly high diabetes rate. The progress of the diabetes epidemic in rural areas can be elucidated further and the trend for diabetes prevalence can be identified as either stable, accelerating or potentially decelerating.
The directly relevant preceding studies in Andhra Pradesh were done by Chow et al. using the data gathered from APRHI in 2005. This involved a series of publications regarding cardiovascular risk factors in the Godavari region of rural Andhra Pradesh. It is these data that is being used as a baseline for comparisons with the GC13 and SMARTHealth results in this study. In an initial pilot study for the project, data were collected from two villages in June 2004 using a simple random sample after stratification by age and sex groups. The data collection was based on a questionnaire, physical measurements and a fasting venous blood sample. The authors concluded from the high quality of the data obtained that it was feasible to continue and conduct larger studies in the region. ‘Significant lipid, adiposity and metabolic abnormalities amongst 4535 Indians from a developing region of rural Andhra Pradesh’ showed a prevalence of metabolic syndrome of 24.6% (21.7% - 27.5%, 95% CI). When the specific Asian definition of metabolic syndrome was used, this increased to 30.2% (27.1% - 33.3%, 95% CI). In this paper this was defined as a waist circumference 12 cm (men) and 8 cm (Women) lower than the general definitions of an increased waist circumference. ‘Fatal and nonfatal cardiovascular disease and the use of therapies for secondary prevention in a rural region of India’ found a combined prevalence of coronary heart disease and stroke at 6.6% (5.8% - 7.4%, 95% CI). The mean age of cardiovascular disease incidence was 54 years (52 – 55, 95% CI). ‘The prevalence and management of diabetes in rural India’ was also based on the same data and showed a diabetes prevalence of 13.2% (12.1% – 14.3%, 95% CI) with a further 15.5% (14.2% – 16.8%, 95% CI) with impaired fasting glucose. These results from the APRHI study provide a background understanding of the already high rates of cardiovascular disease risk factors in this region.

A recent large study of 14 277 urban and rural adults examined the state of lipid levels in four Indian regions, determining that dyslipidaemia is extreme in India and calls for urgent lifestyle interventions for prevention and management of cardiovascular disease. Almost 80% of the general population showed abnormalities in at least one of total cholesterol, low density lipoprotein (LDL), high density lipoprotein (HDL) or triglycerides with no urban to rural differences. Specifically the highest rates of high LDL cholesterol (15.8%) and high total cholesterol (18.3%) were found in Tamil Nadu, the neighbouring state to Andhra Pradesh. The equivalent preceding study for blood lipids in rural Andhra Pradesh was also by Chow et al. It was concluded that the metabolic risk factors highly prevalent in this population are focal in the foundation of cardiovascular disease in the region.
Approaches to risk factor measurement

One of the difficulties when collecting data for rural, resource poor regions is attaining biological specimens. Unfortunately this is necessary when researching cardiovascular risk factors since questioning the population on prior diagnoses of diabetes or hypercholesterolaemia will significantly underestimate prevalence due to lack of awareness and undiagnosed cases. Currently, to assess diabetes and hypercholesterolaemia, venous blood samples via venepuncture are required and fasting blood glucose and cholesterol levels are measured. This presents challenges of collector training, safety, samples storage and participant willingness. A possible alternative to this is using dried blood spots (DBS). This is a spot of blood collected on paper which can be used for the same purposes as a full venepuncture sample. DBS was first used by Robert Guthrie to screen newborns for phenylketonuria\textsuperscript{41} and now has become more widespread to include large proteins such as insulin and antibodies, DNA sequences and various markers of metabolic diseases\textsuperscript{42}. The primary issue with DBS measurements has been sensitivity to detect minute amounts of the analyte since the amount of blood available is small. There has been progress with the increasing sensitivity of biochemical analysis techniques such as high performance liquid chromatography or immunoturbidimetry. Collecting the DBS itself only involves disinfection, pricking the site with a lancet, collecting the blood drop on filter paper, drying, extraction and then biochemical analysis\textsuperscript{42}. This requires far less of both the blood collector and the patient when compared to a traditional venepuncture to obtain larger blood samples. There are many other perceived benefits to using DBS over venepuncture samples. There are lower costs and more ease with storing the samples. It is unusual for the analysis centre to be in close proximity to the field where the specimens are collected and so a period of storage and transportation are usually necessary. Many more DBS samples can be stored with minimal disruption as there are no glass containers or liquid samples, there is a lower chance of losing samples to breakage during transportation. DBS also has a lower infectious risk with no liquid samples and with the breach of skin with a sharp being uncoupled from the collection of the sample, therefore having lower risk of needle-stick injuries. All of these advantages present DBS as an attractive option for biological specimen collection in such settings as rural villages in Andhra Pradesh.

Another adjustment to consider is the use of HbA\textsubscript{1c} rather than fasting or random blood glucose for diagnosing diabetes. This is more convenient as it avoids having to ensure that participants are fasted and so is beneficial for both specimen collectors and the participants. The HbA\textsubscript{1c} is representative of the glycaemic status of the participant over the previous three months and so is less affected by day to day variation. Therefore it is also more useful than random blood glucose measurements which will vary depending on the timing of the last meal.

The existing diagnostic criteria for diabetes using HbA\textsubscript{1c} do not account for DBS, they are based on venepuncture blood samples of HbA\textsubscript{1c}. These current definitions of diabetes and prediabetes are based primarily on studies examining the rate of diabetes related complications from increasing blood glucose or HbA\textsubscript{1c} levels\textsuperscript{43}. Table 1-2 provides the currently accepted definitions. It is also reasonable to group both diabetes and prediabetes as dysglycaemia in general. This is useful when working with studies which employ different diagnostic methods as none of the methods are completely concordant with each other.

\begin{table}
\centering
\caption{Definitions of Diabetes and Prediabetes from 'Diagnosis and Classification of Diabetes Mellitus'\textsuperscript{43}.}
\begin{tabular}{lll}
\hline
 & Fasting Blood Glucose & 2- hour or Random Blood Glucose & HbA\textsubscript{1c} \\
\hline
Prediabetes & 5.6 mM – 6.9 mM & 7.8 mM to 11 mM & 5.7% - 6.4% \\
Diabetes & > 7 mM & > 11.1 mM & > 6.5% \\
\hline
\end{tabular}
\end{table}

Unfortunately the data regarding the use of DBS for cardiometabolic analytes such as HbA\textsubscript{1c} and blood lipids was sparse. There was no consensus on the appropriate use of the technique for these specific analytes even
though DBS has been studied and validated in many other fields. Much of the current progress with DBS use has been presented in a review paper which did not have adequate conclusions on the use of DBS for the analytes of interest here. The 2005 APRHI and 2014 SMARTHealth surveys used venous blood sampling whereas the 2010 GC13 survey used dried blood spots. It was therefore necessary to collate these data and critically analyse the current literature in the form of a systematic review and meta-analysis before the results of GC13 were to be correctly analysed and published. This review aimed to identify the relationship between DBS and venous samples of cardiometabolic blood analytes and possibly provide a robust conversion equation to account for any systematic differences between the measures. Also the current literature on DBS for these analytes makes use of a variety of biochemical assays and storage conditions. The effects of these variables also need to be ascertained. The results of this study can have a significant effect on the field due to expedition of future large surveys such as in GC13.

In addition to specimen collection technique and diagnostic criteria for the disease of interest, the choice of survey methodology is important. Particularly in rural regions of the less developed countries, there may be difficulties in achieving robust data to adequately assess the epidemiology of diseases. The most ideal approach would be a full sample of every member of a target population. This is obviously impractical. Simple random sampling and systematic random sampling are reasonable alternatives but are effected by biases depending on the accessibility of villages and availability of participants. For example, during data collection, certain demographics may be away from home due to work etc. Post-stratification to a population standard can negate some of these biases by weighting the sample estimates to adjust a sample demographic to the population demographic. This requires up to date demographic data for the general population, such as a census. Simple random sampling could be conducted by random house visits or phone calls to members of a census list. Systematic random sampling would be to visit every nth household in a selection of villages. Both of these unfortunately are very resource intensive, therefore limiting the potential sample size. For a more practical approach, clustered sampling can be used. This involves selecting groups of households within villages. The potential bias here is that the standard error estimates may be artificially reduced due to similarities within clusters. For example, a group of households in close proximity may have more similar diets than a random selection of households would. To counter this, the standard error estimates are adjusted using various estimation techniques such as Taylor Linearization which provide more appropriate estimates for standard error. Post-stratification weighting can also be combined with a cluster sampling approach to mitigate the same demographic related biases as were discussed with simple random and systematic random sampling.

Discussion
Currently there are few studies examining the state of cardiovascular risk factors in the rural Indian population. It is also noted that there have been only eleven previous studies of coronary heart disease in rural India even though the current limited data shows coronary heart disease and stroke have increased incidence in both urban and rural India. For the rural population of 876 million this is not sufficient. This information is essential in characterizing the state of cardiovascular disease and its risk factors in India as the majority of the population is considered rural. The implication is that any change in the disease burden of the rural population will have a magnified effect on India and its economy in general. Of course it is more efficient to target the problem in its initial stages but before effective and informed preventative measures can be contemplated, the full extent of the problem must be assessed. This needs updated investigations and monitoring of the changing trends in disease burden which is part of the premise for this project. Appropriate choice of specimen collection technique and survey method can make this complex task more manageable.

Unfortunately, of the currently available studies there are some significant regional disparities in the precise nature of the cardiovascular risk factor profiles. This heterogeneity will also have implications on health policy, particularly in deciding on the most cost-effective interventions for a particular population. The path to identifying the accurate risk factor profiles of populations affected by the rise of cardiovascular disease is further observational studies. Studies of generally large areas such as states or countries can inform health policy to a certain extent. Further optimization of public health primary prevention policies and initiatives will be possible with detailed study of a region and its particularities. This is especially necessary when referring to
resource poor locations which are economically constrained to fulfil health recommendations and cannot afford to focus on risk factor reductions that may have a minimal effect for that population. Rather, it would be desirable to target the top cardiovascular disease risk factors responsible for the greatest portion of the total morbidity and mortality plaguing a particular society. These gaps in the data need to be adequately addressed to facilitate the required optimal interventions.

As it stands, India has made progress in effective cardiovascular disease monitoring with the implementation of a web based National Stroke Registry in 2013\textsuperscript{49,50}. This will provide data on stroke incidence and prevalence and can influence primary, secondary and tertiary prevention. The caveat to this is that to adequately assess primary prevention strategies, data on both cardiovascular events and cardiovascular risk factors is necessary. In the absence of national registries collecting regular data on cardiovascular risk factors, the options for more focused regional cardiovascular risk factor study are left to individual surveys of selected regions. These studies in Chapters 4 and 5 can serve as varied examples of cost-effective and efficient approaches to data collection in resource poor, rural but populous regions at risk of cardiovascular disease.
Chapter 2 - The Three Surveys

Three large scale cross-sectional studies, APRHI (2005), GC13 (2010) and SMARTHealth (2014) were analysed. In particular, the population diabetes prevalence from the three surveys were compared. Total cholesterol levels and low density lipoprotein (LDL) levels were also compared between APRHI and GC13. The methodology of the APRHI survey has been described previously\textsuperscript{13, 14, 46, 51, 52}.

APRHI

Setting and Survey Sampling

This was based on 20 villages in West and East Godavari. The participants were prestratified into age and sex strata and then sampled using simple random sampling within each age and sex stratum in each village. The full survey is found in Appendix 2. Those aged 30 and above were included.

Specimen Collection

APRHI used fingerprick blood glucose samples and the majority of participants were fasting. The fasting or random blood glucose criteria, and previous history of diabetes or diabetic medication use were used to diagnose participants as having diabetes.

Weighting of Data

Due to potential bias introduced by the simple random sampling method, post-stratification weights were used in APRHI to ensure the sample demographics were adequately representative of the larger population demographics. These post-stratification weights were based on age/sex demographic data collected on the larger background population in the form of a census in 2004. Without these weights the demographic of the sample could be significantly different from the true demographic of the villages.

GC13

Setting and Survey Sampling

In GC13 the study population included 14 villages in the Godavari region of rural Andhra Pradesh, India. These data were collected as a part of the Population Health Metrics Research Consortium Project in Andhra Pradesh, India under the Gates Grand Challenge-13 initiative. The aim of the overall project was to develop robust methods to measure mortality rates where vital registration systems are incomplete. The hypothesis was that the methods for estimating mortality rates in the population can be developed with suitable techniques surveying a subset of the population. This involved a census and household surveys. The census covered 45 villages and 180,000 people in total and the subsequent household survey covered a small subpopulation of this. Data were collected from those aged 15 and above. The data presented in this study is derived from this household survey collected in 2010. The surveys offered to each individual participant were divided into separate modules composed of household demographics, household mortality, coverage and verbal autopsy. This study focused further on the data available from the coverage module which included disease specific questions and physical measurements. The relevant sections of the survey are found in Appendix 3.
Specimen Collection
In GC13 dried blood spots (DBS) were collected for each participant involved at the final stage of selection. The participant’s finger was pricked with a lancet and after allowing a first drop to pass, the second drop was collected on a piece of filter paper. These samples were dried for two to three hours and placed in a plastic bag for storage at 4°C for 10 days maximum. They were transported to the Nizams Institute for Medical Sciences in Hyderabad, Andhra Pradesh. The glycosylated haemoglobin (HbA1c) was eluted from the filter paper specimens and analysed by immunoturbidimetry using the Randox Latex Agglutination Inhibition Assay, Randox Laboratories Ltd.

Post-Analysis Adjustment
The numerical results from the biochemical analysis of DBS for HbA1c were adjusted using the regression equation derived in Chapter 3. This allowed correction for factors that alter HbA1c measurements due to the use of DBS over a venepuncture sample. The diagnostic criteria for diabetes and prediabetes have only been
validated for venepuncture samples of HbA1c and so appropriate adjustment of DBS samples was required as per the results in Chapter 3.

Weighting of Data
GC13 data would be expected to have some bias due to the survey structure. To account as such post-stratification weights and inverse probability of selection weights were used. These post-stratification weights were based on age/sex demographic data collected on the larger background population using data from the census in 2010. The inverse probability of selection weights work to counter the increased likelihood of selecting an individual who belongs in a village with fewer households and/or a household with fewer people. Without these weights the results would be biased towards those who live in smaller villages/households. This is a result of the clustered survey structure.

SMARTHealth
Setting and Survey Sampling
The survey conducted in 2014 was much larger than both APRHI and GC13 with 62,254 adults. The West Godavari district is divided into multiple primary health care centres (PHC), 18 of which were selected as representative of the region. Within each PHC, three villages were selected by simple random sampling. A census list was compiled for the 54 villages and every household was visited with participants aged between 40 and 85 included in the survey. There was no sampling at the individual level. The probability of selection was equal for each individual. The aim in this survey was to gather simple census data than detailed surveys of each member in a limited sample.

Specimen Collection
SMARTHealth used blood glucose levels from fingerprick point of care samples. The majority of participants were non-fasting with > 6 hours of no oral intake considered to be fasting. The fasting or random blood glucose criteria, and previous history of diabetes were used to diagnose participants as having diabetes.

Weighting of the Data
SMARTHealth was a considerably large survey of 54 villages in West Godavari. Due to the extent of this survey with 62,254 participants, it was considered that no particular weighting of the data was required as it could be almost considered a population census in itself.

Discussion
The GC13 survey was structured as a stratified clustered sampling method. This is in contrast to a theoretically ideal simple random sampling survey which is logistically more difficult to perform and less easy to achieve a reasonable sample size. The full survey structure is demonstrated in figure 2-1. The use of such sampling methods has been well studied and, if accounted for appropriately, then should yield results similar to a corresponding simple random sample. There are three levels of statistical consideration required for this survey. The initial stratification into villages, the multistage clustered sampling and the final post-stratification adjustments for age and sex demographics. Sampling from within strata independently acts to reduce the sampling error by increasing the representation from all the different groups, in this case the villages. This allows for a weighted mean which can have a lower variance than the arithmetic mean derived from simple random sampling which may under-sample certain villages by chance.

The two main issues encountered with a clustered survey structure are the probability of selection of an individual and the calculation of the variance of the sample statistics. The probability of selecting an individual in a simple random sample is equal between each participant. In contrast, the probability of selecting an individual in such a multistage clustered survey will depend on the proportion of clusters sampled and the proportion of participants sampled within that participant’s cluster. This can reliably be accounted for by weighting the data from each participant appropriately as is explicitly demonstrated in chapter 4. The resulting weights are known as inverse probability of selection weights. The second issue is regarding variance which will affect the significance level of any results found. The clustered survey will have a higher variance, once the survey structure is correctly accounted for, when compared to an equivalent sample size in a simple random sample. This is known as having a design effect greater than 1. This dilated variance can be lessened by
increasing the sample size in a clustered sampling to achieve the same variance as that of a simple random sample. For example, assuming a design effect of approximately 2, the variance of a clustered survey will be equal to the variance of a simple random sample which has half the sample size. The design effect itself is a function of cluster size and intra-cluster correlation, Deff = 1 + (Cluster size – 1) * (Intra-cluster correlation)\textsuperscript{35}. Therefore the optimal clustered survey will have a minimal cluster size and maximum intra-cluster variation. Of these, only the cluster size is in the control of the investigators and here it is the size of the household for the primary sampling units and then the number of participants selected within each household for the secondary sampling units. In this survey the number of participants selected within each household was only one and so the cluster size at the second stage was at a minimum. Unfortunately the cluster size of the households is not within the investigator’s control in this case.

The third statistical adjustment is the post-stratification weighting after the simple random selection of participants for DBS. Here, the final data from each participant is weighted according to the discrepancy between the sample age and sex proportions and the population age and sex proportions. Of course for this to be feasible the population demographics are required and so the census data for the Godavari region was used in this case. This adjustment aims to ensure correct representation of the age and sex dependence of the data according to the population age and sex distribution. This is in case the sample has a different age and sex distribution to the overall population, this can then be corrected for to ensure that the data is not skewed due to the random selection of participants for dried blood spots. This post-stratification weighting was also used in the APRHI analysis. APRHI did not need inverse probability of selection weights since it only employed simple random sampling with no clusters. The variances of the means and proportions calculated were appropriately estimated using the Taylor Linearization Method through the STATA ‘svy’ command. Further methodological detail is found in chapter 4.

**Table 2-1. Summary of the survey methodology of all three studies.**

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Region</th>
<th>Villages</th>
<th>Survey Structure</th>
<th>Data Weights</th>
<th>Specimen Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>APRHI 2005</td>
<td>4535</td>
<td>East and West Godavari</td>
<td>20</td>
<td>Stratified Simple Random Sampling</td>
<td>Post-stratification weights</td>
<td>Fingerprick Blood Glucose</td>
</tr>
<tr>
<td>GC13 2010</td>
<td>2991</td>
<td>West Godavari</td>
<td>14</td>
<td>Stratified Clustered Random Sampling</td>
<td>Post-stratification weights and inverse probability of selection weights</td>
<td>DBS HbA1c</td>
</tr>
<tr>
<td>SMARTHealth 2014</td>
<td>62 254</td>
<td>West Godavari</td>
<td>54</td>
<td>Full Sampling</td>
<td>None</td>
<td>Fingerprick Blood Glucose</td>
</tr>
</tbody>
</table>
Chapter 3. Comparability of HbA1c and lipids measured with dried blood spots versus venous samples: a systematic review and meta-analysis

Publication Details: Eshan T Affan, Devarsetty Praveen, Clara K Chow, and Bruce C Neal. BMC Clin Pathol. 2014; 14: 21

Author Contributions: Eshan Affan was the principal investigator, designed the study, organized the literature review, analysed and presented the results and formed the main body of the text. Devarsetty Praveen was a secondary reviewer of the abstracts isolated for studies to be included in the meta-analysis. Clara Chow reviewed and provided guidance for the structure of the study. Bruce Neal oversaw the design of the study, edited and reviewed the publication.

INTRODUCTION
Cardiovascular diseases are increasing particularly rapidly in developing country settings with diabetes a key determinant of risk. Documenting the role of dysglycaemia and other metabolic risk factors can be challenging in these countries because the infrastructure and resources required to conduct research are limited. For example, assays of glycosylated haemoglobin (HbA1c) and blood lipids are usually done on venous blood samples which can be difficult to collect, transport and store. The use of dried blood spot (DBS) sampling is one possible solution. DBS involves pricking the participant’s finger with a lancet and collecting drops of blood on a piece of filter paper. Samples are then dried and placed in sealed plastic bags for transportation and storage. Compared to venous samples, collecting DBS requires minimal training of staff, is cheaper, is safer, provides for simpler transportation and is more acceptable to study participants.

DBS samples are now widely used for measuring serum antibodies, human immunodeficiency virus (HIV) loads and blood hormone levels with good data to define the comparability of results between analyses based upon DBS and standard venous samples. The absence of comparable data to define the associations for HbA1c and blood lipids means that DBS samples are not widely used in studies making assessment of cardiovascular risks. The objective of this project was to synthesise the available evidence describing the comparability of findings for assays of HbA1c and blood lipids based upon DBS samples compared to standard venous samples.

MATERIALS AND METHODS
This project was a systematic review and meta-analysis done to define the association of findings for HbA1c and blood lipids for analyses based upon standard venous samples compared to DBS samples. This was a secondary analysis of existing published data and no ethics review was therefore required.

Search strategy
The Cochrane, Embase and Medline databases were searched electronically during July 2012 using combinations of the terms “dried blood spot”, “dried blood”, “DBS”, “filter paper”, “triglycerides”, “triacylglycerides”, “HbA1c”, “glycosylated haemoglobin”, “glycated haemoglobin”, “cholesterol”, “high density lipoprotein”, “HDL”, “low density lipoprotein” or “LDL”. Additional studies were identified by a manual examination of the reference lists of all studies identified as eligible.

Eligibility criteria
Studies were eligible for inclusion if they directly compared values generated from analyses based on DBS samples to analyses based on venous samples. To be included, a study had to report in the form of a regression equation an association for one or more of the specified outcomes. There was no restriction on the type of study population.

Data extraction
Two independent observers reviewed the abstracts for eligibility and extracted standardised data into a data collection sheet for eligible studies. The data sought from each study were based upon a comparable prior
systematic review done in the HIV field and included: date of publication, study size, participant characteristics and sample storage conditions. For each risk factor reported upon we sought to identify the laboratory extraction method, biochemical assay method and regression coefficient. Where available we also noted data describing the stability of the DBS samples.

Outcomes
The outcomes studied in this overview were HbA1c, total cholesterol, high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol and triglycerides.

Statistical analysis
The characteristics of included studies were summarised in tabular form (Table 3-1). The linear regression coefficients from each study were pooled separately for each risk factor using a weighted least squares approach to estimate an overall coefficient. The same method was used to estimate a combined intercept. This gave a relationship of the form: DBS = bVenous + a where ‘b’ is the combined coefficient and ‘a’ is the combined intercept. The synthesis of the parameters was done as follows:

\[
b = \frac{\sum w_i b_i}{\sum w_i} \text{ and } a = \frac{\sum w_i a_i}{\sum w_i}
\]

Where weight \( w_i \) = the number of participants in study \( i \), and \( b_i \) and \( a_i \) represent the coefficient and intercept, respectively, for the regression line in study \( i \).

Heterogeneity of the individual study estimates contributing to each meta-analysis was assessed using the Cochran’s Q and \( I^2 \) statistics. Subsidiary analyses were done to explore the impact of assay method for the outcome of HbA1c.

RESULTS
There were 705 records identified by the electronic search for which abstracts were reviewed. Six further studies were found by the manual search of reference lists for included studies (Figure 3-1). One final study was found during the review process. Sixteen studies were ultimately included in the meta-analysis, 12 of which reported necessary data for HbA1c, 1 for triglycerides, 2 for both triglycerides and total cholesterol and 1 for HbA1c, total cholesterol and HDL (Table 3-1, Table 3-2). One other study of HbA1c was excluded because it did not provide a regression equation and one other study of triglycerides was excluded because it did not provide the sample size. There were no studies reporting data for LDL-cholesterol identified.
The total numbers of participants providing data were 1425 for HbA1c, 773 for triglycerides and 1093 for total cholesterol. Study sizes ranged from 30 to 613 participants. The assay methods varied for HbA1c which included immunoturbidimetric, high performance liquid chromatography (HPLC) and affinity chromatography assays but all studies measuring triglycerides, total cholesterol and HDL used colorimetry.

**Figure 3-1 Flow chart detailing identification of studies**

**HbA1c**

For HbA1c, the summary regression (DBS = 0.9858V + 0.3809) (Figure 3-2) showed close agreement between analyses based upon the venous and DBS sampling methods. There was, however, evidence of heterogeneity between the contributing regression lines for the intercepts (Cochran’s Q-test p<0.001 and I²=98%) but not the slopes (p= 0.833 and I²=0%). Subsidiary analyses by assay method showed that the heterogeneity was partially attributable to different results for the two studies that used affinity chromatography. Funnel plots did not provide clear evidence of publication bias (Figures 3-5 and 3-6).

**Blood lipids**

The summary regression line for total cholesterol (DBS = 0.6807Venous + 1.151) (Figure 3-2) indicates a requirement for moderate adjustment of values based upon analyses of DBS samples to obtain estimates equivalent to standard analyses based upon venous samples. The regression lines for the two studies contributing to this meta-analysis were directly comparable in terms of both slope and intercept although both were derived from studies done at the same investigational centre. For triglycerides, the summary regression for the three contributing studies showed a close association between the results obtained for the two
methods (DBS = 0.9557Venous + 0.1427) (Figure 3-2) without any evidence of heterogeneity between the three results. Only one data set was available for HDL.

Storage
Data about the circumstances and duration of storage of DBS samples were inconsistently reported with few data to describe whether the analysis findings were affected by extended storage duration or different storage temperatures. From the limited data available it was concluded that DBS samples collected for assay of HbA1c and intended for HPLC analysis can be stored for 5 days at room temperature or for up to 3 years at -70°C.65,67. If analysis by immunoturbidimetry is planned, the data variously suggest that samples can be stored safely at room temperature for up to 44 days, at 4°C for up to 15 days and storage at -80°C for up to 3 months.76

For total cholesterol samples were reported as stable for up to 1 month at room temperature82,89, and up to 3 months at 4°C89, and for triglycerides up to 1 month at room temperature and up to 2 months at 4°C82.

![Figure 3-2. Individual (solid) and summary (dotted) regression lines showing the associations between results for analyses based upon dried blood spot (DBS) compared to venous (V) samples for haemoglobin A1c (HbA1c) analysed by any method (A), HbA1c analysed by specific methods (B to D), triglycerides (E) and total cholesterol (F). - Thickness of line increases with sample size. Line length was defined as ±1 standard deviation (SD) of the study (or overall) mean. Where the mean or SD of a study was not available the average for that analysis was used.](image-url)
DISCUSSION
These analyses identified clear associations between assay results based upon blood samples collected using traditional venous approaches and blood samples collected using DBS techniques, for both HbA1c and selected blood lipids. The data provide a strong rationale for the further investigation of DBS sample collection techniques although also serve to highlight a number of areas that require further exploration before the method is considered mainstream in this field. If, however, standards and calibrations can be agreed, as has been achieved in other fields of research 77,83,97, the DBS method does appear to have significant potential to address the logistical challenges of venous sampling for studies of metabolic risks in resource poor settings73.

The differences between the intercepts of the regression lines obtained for the various analytic methods used for assay of HbA1c require careful consideration in terms of their implications. If the variation is due to the analytic method selected then it will be necessary to recommend a standard approach for each analyte of interest. However, while the analytic method is the obvious explanation for the observed variation it is not possible to exclude alternative causes on the basis of the available data. For example, other aspects of the preparation of the DBS samples such as transportation and extraction were not standardised across the different analytic methodologies and might also be a cause of the differences noted.

The incomplete and summary nature of the data available for analysis placed significant constraints upon the extent to which the results could be explored in this overview. In particular, measures of variance of the data were unavailable for most contributing studies, requiring that weighting be done by sample size alone74 with consequent limitations upon the methodologies that could be used to present uncertainty intervals around both the individual studies and the summary estimates. For example, we identified a possible relationship between the regression parameters and the mean HbA1c of the contributing studies suggesting that both the intercept and the slope might change when HbA1c rises above 8% (Figures 3-3 and 3-4). This implies a non-linear association of venous with DBS sample results that might require a more nuanced explanation than the simple linear regressions provided here65. Removing the studies with high average HbA1c levels from the meta-analysis resulted in a regression line approaching parity (DBS = 0.9553V + 0.2566) and with a reduced heterogeneity for both the slopes ($I^2=0\%$) and the intercepts ($I^2=92\%$). However, whether this simply reflects a chance finding in the data, or a true variation of the association by mean HbA1c level is still uncertain.

Likewise, several studies used the same patients for two rounds of analysis 78,87 and there would therefore have been some correlation between the findings for each. This would not be expected to substantially change the parameter estimates obtained but certainly would increase the uncertainty around them. Our inability to create robust uncertainty intervals around our estimates is the primary weakness of this piece of work. Unfortunately the lipid analyses were even less rigorous, with so few studies, confidence in the meta-analysis is limited. There is much potential for further work in this area to generate reliable regression analyses for use in the field.

On a more positive note, the association between results based on venous and DBS samples appeared to be consistent at the levels of HbA1c at which diabetes mellitus is diagnosed (HbA1c > 6.5%) 72. This implies that DBS samples could already be used for determining the presence or absence of diabetes with reasonable certainty, although measures of the extent to which blood glucose is controlled amongst those with diabetes would be less reliable.

Most of the studies reported some information about DBS sample preparation, transport and storage but the data were provided in diverse formats and were substantively incomplete. While it appears likely that DBS samples are stable for adequate periods of time this is an area that requires systematic evaluation and the development of standardised recommendations prior to widespread roll out of the methodology.

The establishment of World Health Organization “25 by 25” target for the prevention of non-communicable diseases 93 has added urgency to the need for data about the metabolic determinants of cardiovascular risk. With more than three quarters of all chronic disease now occurring in developing country settings, the introduction of low cost research techniques that will provide the data required to inform government decision making is a priority 94. DBS sample collection methods appear to have great potential for the
evaluation of cardiometabolic risk factors at the population level\textsuperscript{65, 69, 85} enabling data collection at scale in areas previously unstudied\textsuperscript{71}. There remain, however, important advances to be made in defining standard methodologies and adjustments before the DBS sampling method is confirmed as a sound proxy for traditional venepuncture samples for these types of blood analytes.
### Table 3-1. Characteristics and findings of included studies

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Diabetes</th>
<th>Population source</th>
<th>Mean (SD) of DBS values</th>
<th>Regression equation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HbA1c</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anjali 2007*</td>
<td>78</td>
<td>Yes</td>
<td>-</td>
<td>9.45 (±1.86)</td>
<td>DBS = 0.95V + 1.4</td>
</tr>
<tr>
<td>Buxton 2009*</td>
<td>115</td>
<td>-</td>
<td>Hospital</td>
<td>-</td>
<td>DBS = 0.85V + 0.81</td>
</tr>
<tr>
<td>Egier 2011*</td>
<td>85</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>DBS = 0.933V + 0.4</td>
</tr>
<tr>
<td>Fokkema 2009*</td>
<td>93</td>
<td>-</td>
<td>Community</td>
<td>-</td>
<td>DBS = 1.006V - 0.92</td>
</tr>
<tr>
<td>Fokkema 2009*</td>
<td>88</td>
<td>-</td>
<td>Community</td>
<td>-</td>
<td>DBS = 0.994V + 0.057</td>
</tr>
<tr>
<td>Fokkema. 2009*</td>
<td>73</td>
<td>-</td>
<td>Community</td>
<td>-</td>
<td>DBS = 0.987V - 0.011</td>
</tr>
<tr>
<td>Gay 1990*</td>
<td>58</td>
<td>Yes</td>
<td>Community</td>
<td>10.8 (±2)</td>
<td>DBS = 0.8V + 1.8</td>
</tr>
<tr>
<td>Jeppsson 1996</td>
<td>41</td>
<td>-</td>
<td>-</td>
<td>7.49</td>
<td>DBS = 0.99V + 0.16</td>
</tr>
<tr>
<td>Jones 2010*</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>6.74</td>
<td>DBS = 0.984V + 0.189</td>
</tr>
<tr>
<td>Jones 2010*</td>
<td>40</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>DBS = 0.998V - 0.204</td>
</tr>
<tr>
<td>Lacher et al 2013*</td>
<td>386</td>
<td>-</td>
<td>Community</td>
<td>5.92 (±1.2)</td>
<td>DBS = 0.94V + 0.37</td>
</tr>
<tr>
<td>Lakshmy 2009*</td>
<td>30</td>
<td>-</td>
<td>Community</td>
<td>5.94 (±1.58)</td>
<td>DBS = 0.9886V + 0.0018</td>
</tr>
<tr>
<td>Little 1986*</td>
<td>78</td>
<td>Yes/No</td>
<td>-</td>
<td>10.2</td>
<td>DBS = 1.09V + 2.17</td>
</tr>
<tr>
<td>Lomeo 2008*</td>
<td>97</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>DBS = 0.877V + 1.09</td>
</tr>
<tr>
<td>Tabatabaei-Malazy 2011*</td>
<td>33</td>
<td>Yes</td>
<td>Community</td>
<td>8.8 (±1.6)</td>
<td>DBS = 1.20V - 0.635</td>
</tr>
<tr>
<td>Tabatabaei-Malazy 2011*</td>
<td>33</td>
<td>Yes</td>
<td>Community</td>
<td>8.9 (±1.7)</td>
<td>DBS = 1.25V - 1.09</td>
</tr>
<tr>
<td>Wikblad 1998*</td>
<td>145</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>DBS = 1.03V - 0.405</td>
</tr>
<tr>
<td><strong>Triglycerides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lakshmy 2010*</td>
<td>85</td>
<td>-</td>
<td>Community</td>
<td>1.6 (±0.6)</td>
<td>DBS = 1.028V - 0.1690</td>
</tr>
<tr>
<td>Lakshmy 2012*</td>
<td>613</td>
<td>-</td>
<td>Community</td>
<td>1.16 to 1.87 (±0.45 to 0.79)</td>
<td>DBS = 0.9549V + 0.1875</td>
</tr>
<tr>
<td>Quraishi 2006*</td>
<td>75</td>
<td>-</td>
<td>Community</td>
<td>1.297 (±0.53)</td>
<td>DBS = 0.88V + 0.13</td>
</tr>
<tr>
<td><strong>Total Cholesterol</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lacher 2013*</td>
<td>395</td>
<td>-</td>
<td>Community</td>
<td>3.76 (±0.87)</td>
<td>DBS = 0.52V + 1.08</td>
</tr>
<tr>
<td>Lakshmy 2010*</td>
<td>85</td>
<td>-</td>
<td>Community</td>
<td>5 (±1)</td>
<td>DBS = 0.727V + 1.170</td>
</tr>
<tr>
<td>Lakshmy 2012*</td>
<td>613</td>
<td>-</td>
<td>Community</td>
<td>-</td>
<td>DBS = 0.7779V + 1.1943</td>
</tr>
<tr>
<td><strong>HDL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lacher 2013*</td>
<td>395</td>
<td></td>
<td>Community</td>
<td>1.41 (±0.42)</td>
<td>DBS = 0.7V + 0.46</td>
</tr>
</tbody>
</table>

*Some studies provided multiple estimates and are repeated in the table. SD=Standard deviation, DBS=dried blood spot, V=venous
### Table 3-2. Methods of sample analysis in the different studies

<table>
<thead>
<tr>
<th>Dried blood spot data collection details</th>
<th>Assay methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drying</td>
<td>Transportation</td>
</tr>
<tr>
<td>Anjali 2007&lt;sup&gt;75&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Buxton 2009&lt;sup&gt;95&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Egier 2011&lt;sup&gt;65&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Fokkema 2009&lt;sup&gt;78&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Fokkema 2009&lt;sup&gt;78&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Fokkema. 2009&lt;sup&gt;78&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Gay 1990&lt;sup&gt;79&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Jeppsson 1996&lt;sup&gt;77&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Jones 2010&lt;sup&gt;66&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Jones 2010&lt;sup&gt;66&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Lacher 2013&lt;sup&gt;80&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Lakshmy 2009&lt;sup&gt;81&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Little 1986&lt;sup&gt;84&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Lomeo 2008&lt;sup&gt;85&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>Tabatabaei-Malazy 2011&lt;sup&gt;87&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Tabatabaei-Malazy 2011&lt;sup&gt;87&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Wikblad 1998&lt;sup&gt;86&lt;/sup&gt;</td>
<td>-</td>
</tr>
</tbody>
</table>

**Triglycerides**

| Lakshmy 2010<sup>82</sup> | - | Packed in ice | Methanol | Col. | Col. |
| Lakshmy 2012<sup>83</sup> | Yes | Packed in ice | Methanol | Col. | Col. |
| Quraishi 2006<sup>89</sup> | Yes | - | Methanol | Col. | Col. |

**Total Cholesterol**

| Lacher 2013<sup>80</sup> | Yes | Packed in dry ice | Deionized water | Col. | Col. |
| Lakshmy 2010<sup>62</sup> | - | Packed in ice | Methanol | Col. | Col. |
| Lakshmy 2012<sup>83</sup> | Yes | Packed in ice | Methanol | Col. | Col. |

**HDL**

| Lacher 2013<sup>80</sup> | Yes | Packed in dry ice | Deionized water | Col. | Col. |

*Some studies involved multiple samples and are repeated in the table.*
Figure 3-3. The mean HbA1c values from the studies are compared with the regression intercept the study generated. The marker size is proportional to study size.

Figure 3-4. The mean HbA1c values from the studies are compared with the regression coefficient the study generated. The marker size is proportional to study size.
Figure 3-5. Funnel plot of the HbA1c regression intercepts.

Figure 3-6. Funnel plot of the HbA1c regression coefficients.


Author Contributions: Eshan Affan was the primary author in this publication and designed the comparative study with supervision from Bruce Neal and Clara Chow. Bruce Neal also critically reviewed and provided multiple edits of the publication. Eshan Affan also formed the main body of the text.

Andhra Pradesh Rural Health Initiative – This was a previously conducted survey with results published earlier by Clara Chow and Bruce Neal. Eshan Affan restructured the collected data and represented specific subsections to make a valid comparison with the other two surveys.

Gates Grand Challenge #13 Survey – Eshan Affan managed the data after collection. Developed and implemented the method of analysing the particular survey structure with Devarsetty Praveen. Eshan Affan conducted the analysis of the results and presented the diabetes and blood lipid prevalence. Jason Wu provided guidance on the methodology and statistical analysis.

SMARTHealth survey – David Pereis, Anushka Patel and Devarsetty Praveen were responsible for the data collection. Eshan Affan managed the diabetes data after collection, conducted the analysis and presented the diabetes prevalence results. Anushka Patel, Clara Chow, David Pereis and Jason Wu provided critical analysis of the methodology of the comparative study.

INTRODUCTION

Although infectious diseases remain a significant cause of morbidity and mortality in developing countries, non-communicable diseases are increasingly prevalent and now represent about half the burden of disease in South Asia. India is developing rapidly and lifestyle changes are occurring throughout much of the country, with a projected 150% increase in diabetes mellitus prevalence to more than 80 million by 2030. These estimates are, however, based on limited data and have significant uncertainty about them. There are few studies in India that have made repeat assessments of diabetes prevalence in the same population, using similar methods and this is particularly true in rural areas where some 70% of the population lives. The available data suggest a substantial rise in diabetes prevalence in rural areas over the last few decades with one large survey done in 1991 reporting a rate of just 0.4% compared to a 2005 estimate of 13.2% in a study done in rural Andhra Pradesh. In that 2005 Andhra Pradesh Rural Health Initiative (APRHI) study a further 15.5% of the population were classified as having impaired fasting glucose such that 28.7% of the population had some form of dysglycaemia. In 2010 the same Godavari area of Andhra Pradesh was resurveyed as part of the Gates Grand Challenge 13 (GC13) health metrics evaluation which included an assessment of diabetes prevalence. Most recently a further survey of the same population was done as a part of the Systematic Medical Appraisal Referral and Treatment Health (SMARThealth) initiative. SMARThealth is a project utilising low-cost smartphone technologies to provide high quality health care to rural India. In this paper we report and compare the estimated prevalence of dysglycaemia in the West Godavari district of rural Andhra Pradesh in 2005, 2010, and 2014.

METHODS

The APRHI project received approval from the ethics committees of the CARE Hospital, Hyderabad in India and the University of Sydney in Australia, the GC13 project from the ethics committee of the Gandhi Medical College and Hospital and the University of Sydney and the SMARThealth initiative from the ethics committees of the Centre for Chronic Disease Control, New Delhi and the University of Sydney in Australia. In every case survey participants were provided written and/or oral information about the respective study in their local language and written informed consent was obtained prior to participation.
Survey Characteristics

The 2005 APRHI survey - included 4535 adults aged 30 to 100 from 20 villages (total population 75 089) in the East and West Godavari regions of Andhra Pradesh. The 20 villages were selected to be representative of the area and used stratified random sampling. The population was divided into 8 groups on the basis of age (30-39, 40-49, 50-59, 60+) and sex with individuals sampled at random within these strata. This strategy was used to ensure similarly precise age and sex specific estimates with post-stratification weights applied to enable estimation of overall population means and prevalences. Each individual had a brief physical examination, completed a questionnaire, and a fasting plasma glucose (FPG) measurement made on a fingerprick blood glucose sample. The detailed methods of the APRHI study have been reported previously.13, 14, 46

GC13 survey - included 4024 adults aged 15 to 94 from 14 villages (total population 55 151) also in the West Godavari region of Andhra Pradesh. The 14 villages were selected to be representative of the area with stratified clustered sampling of individuals from the selected villages. The initial stratification was by village and then the primary sampling units in each village were randomly selected households from which an individual was randomly selected for inclusion and completion of the initial questionnaire stage. A further round of stratified random sampling of these individuals identified a subset from whom a dried blood spot (DBS) was collected for assessment of haemoglobin A1c (HbA1c). This sampling was done with predetermined post-stratification sampling weights to achieve similar sample sizes in each age and sex stratum of the population with a total of 851 DBS samples collected (Figure 4-2). The villages included in GC13 were from the same region as APRHI and 4 of the villages were common to both surveys.

The 2014 SMARTHealth survey - included 62 254 adults aged 40 to 85 years from 54 villages (total population 209 868) in the same West Godavari region of Andhra Pradesh. A total of 18 Primary Health Care Centres (PHCs), broadly representative of the West Godavari District of Andhra Pradesh were selected. From the region serviced by each PHC, three villages were randomly selected. A census list comprising age and sex of all residents was first enumerated from the 54 villages and each household was then visited. All eligible persons of age 40 years and above were identified and invited for an interview and clinical assessment. 62 194 participants had capillary blood glucose measurements made, with capture of information on fasting state at the time of measurement.

Data collection, blood sampling and analysis methods

2005 APRHI - Data collection was during February and March 2005 by trained staff using a structured questionnaire. Participants were asked about a previous diagnosis of diabetes by a medical practitioner or if they were currently taking an oral hypoglycaemic and/or insulin. Weight and height were measured with participants wearing light clothing and no shoes. A fasting finger-prick blood glucose measurement was sought for all participants using B-Braun (Germany) U.S.V. meters which were calibrated to give venous plasma glucose equivalents. Diabetes was classified as self-reported diabetes or a FPG ≥ 7 mM and impaired fasting glucose was designated as prediabetes within the range 5.6 mM < FPG ≤ 6.9 mM.4 For the very small proportion who were not fasting at the time of blood sampling, diabetes was classified as self-reported diabetes or a random plasma glucose (RPG) ≥ 11.1 mM and prediabetes was defined by the range 7.8 mM < RPG ≤11 mM.43

2010 GC13 - Data collection was done from July 2009 to February 2010 using laptops to guide the interviews and enable real-time data entry and data checking. Participants were asked whether they had a diagnosis of diabetes by a health practitioner and if they were currently taking ‘diabetic pills’. Weight and height were measured with shoes and heavy outer clothing removed. Blood collection was by a finger prick with the sample collected onto filter paper and air dried for two to three hours. Samples were then placed into a plastic bag and stored at 4°C for a maximum of 10 days before being sent by courier to Hyderabad for longer-term storage. DBS samples were analysed at the Nizams Institute for Medical Sciences in Hyderabad, Andhra Pradesh. Analysis was done by immunoturbidimetry using the Randox latex agglutination inhibition assay, Randox Laboratories Ltd. The DBS findings were minimally adjusted using a regression equation that defines the association between immunoturbidimetric analyses of HbA1c: DBS samples and fasting venous samples
tested using handheld glucose monitors. Diabetes was defined as self-reported diabetes or HbA1c > 6.5% and prediabetes as 5.7% < HbA1c ≤ 6.4%.

2014 SMARThealth - Data collection was done between February and May 2014. Participants were also asked about previously diagnosed diabetes. Capillary blood collected using the finger-prick method from participants in either a fasted or non-fasted state was assayed with the Abbott Freestyle Optium glucometer. Whether or not the participant was fasting at the time of sampling was recorded and the same diagnostic standards as APRHI were used for the classification of glycaemic status based upon FPG or RPG. There was no data on oral hypoglycaemic usage in 2014, although this is not expected to change the diabetes prevalence significantly as previously known diabetes was asked for.

In all three studies participants on oral hypoglycaemics, insulin or with a self-reported history of diabetes were classified as having diabetes and could not be classified as having pre-diabetes. The aggregate of diabetes and prediabetes defined dysglycaemia.

Statistical Analysis
Analyses were restricted to the 40-85 year age group common to all three surveys. Estimates of the prevalence of dysglycaemia were made for the 20 villages included in APRHI and for the 14 villages included in GC13 by weighting the data obtained from the survey participants. The weighting method used to obtain summary estimates for the 20 and 14 villages, respectively, varied between the two surveys reflecting the different approaches taken to sampling. For APRHI, which used simple stratified random sampling, post-stratification weighting was employed using data from a population census done by the Byrraju Foundation in 2004. For GC13 the stratified cluster sampling used to identify the households and the individual within each for inclusion, followed by secondary stratified sampling of the included individuals for measurement of HbA1c, required a more complex weighting methodology based upon the product of the probability weights and the post-stratification weights (Appendix 1). These weighting strategies provided appropriate age and sex standardised estimates as well as minimised sampling bias. The SMARThealth data required no weights since it was a full survey of the entire population with a substantially complete response of individuals in the 54 villages included. Four villages were common to APRHI and GC13, two were common to APRHI and SMARThealth, and two were common to GC13 and SMARThealth. None were included in all three studies.

The primary analyses based upon directly comparable assessment methods (using only participants with FPG samples) were made using data from 2005 and 2014. Secondary analyses included all participants with data aged 40-85 years in all three studies. The secondary analyses were done to explore the potential impact on the prevalence estimates of the different measurement methods upon which definitions of dysglycaemia, diabetes and prediabetes could be based. The statistical software package STATA 11.0 (StataCorp, College Station, TX, USA) was used for all analyses. The ‘svy’ command which supports the inclusion of survey weights was the basis of most analyses. Continuous variables are presented as mean and 95% confidence intervals and compared using the unpaired Student’s t-test or analysis of variance (ANOVA). Differences between proportions were tested for significance using the chi-squared test for homogeneity.

RESULTS
The primary analyses of dysglycaemia were based upon 3243 individuals from APRHI and 749 individuals from SMARThealth for whom fasting plasma glucose samples were available. The primary difference in population characteristics across the two surveys (Table 4-1) was a more than one unit rise in mean BMI driven by an approximate 1.7kg rise in mean weight. For the secondary analyses there were 3333 individuals aged 40 to 85 included from the APRHI study in 2005 with a response rate of 81%; 2200 individuals included from the GC13 survey done in 2010 with a response rate of 82%; and 62 254 participants included from the SMARThealth survey in 2014 with a response rate of 84%. The characteristics of the participants included in the secondary analyses were comparable to the characteristics of the populations used in the primary analyses with significant increases observed between 2005, 2010, and 2014 for both BMI (p<0.001) and weight (p<0.001) (Table 4-1).
In the primary analyses done for 2005 and 2014 amongst participants with fasting plasma glucose samples, the estimated prevalence of dysglycaemia was 53.7% (51.8 - 55.7, 95% CI) in 2005 and 62.0% (58.5 - 65.4, 95% CI) in 2014 (p<0.001) (Figure 4-1a). The corresponding prevalence estimates for diabetes were 17.4% (15.9 - 18.9, 95% CI) and 29.9% (26.6 - 33.2, 95% CI) (p<0.001) and for prediabetes were 36.3% (34.4 - 38.2, 95% CI) and 32.0% (28.7 - 35.4, 95% CI) (p=0.088). The proportions aware of their diabetes in each year were 49% and 31%. The mean body mass index (BMI) increased from 2005 to 2014 was 1.3 kg/m² in females (rising from 22.6 kg/m² to 23.9 kg/m²; p=0.002) and 0.8 kg/m² in males (21.8 kg/m² rising to 22.6 kg/m²; p=0.002). The increase in dysglycaemia over this period was entirely in women and driven by an approximate doubling in the prevalence of diabetes.

For the secondary analyses the estimated prevalence of dysglycaemia was 53.9% (52.0 - 55.9, 95% CI) in 2005, 50.5% (46.1 - 54.9, 95% CI) in 2010, and 41.3% (40.9 - 41.7, 95% CI) in 2014 with the data suggesting a decline across the three time points (p<0.001) (Figure 4-1b). The corresponding prevalence estimates for diabetes were 17.8% (16.3 - 19.3, 95% CI), 33.6% (29.4 - 37.9, 95% CI) and 18.1% (17.8 - 18.4, 95% CI) with the 2010 data based upon HbA1c assays providing a markedly higher estimate of diabetes for 2010 (p<0.001). For prediabetes the estimates of prevalence were 36.1% (34.3 - 38.0, 95% CI) in 2005, 17.2% (13.8 - 20.7, 95% CI) in 2010 and 23.2% (22.8 - 23.5, 95% CI) in 2014, again with substantial variation across years (p<0.001). The proportions of the population reporting use of an oral hypoglycaemic were 5.82% (4.94 - 6.71, 95% CI) in 2005 and 8.21% (6.89 - 9.53, 95% CI) in 2010.

**DISCUSSION**

These research findings leave little doubt that the prevalence of dysglycaemia is very high in this rural Indian population and indicate that the rates of both diabetes and pre-diabetes have been elevated for at least a decade. This bodes ill for the cardiovascular health of this community and the adverse effects are already apparent in recent measures of vascular disease burden.

More concerning is the possibility that similarly high, but as yet undocumented, rates of diabetes may be present in other parts of rural India. This raises the prospect that current projections for diabetes and diabetes-related disease burden in India may be significant under-estimates.

While the large magnitude of the diabetes problem in this population is clear from our data, whether it is improving or worsening, is less immediately apparent. The primary analyses based upon the subset with assessments made on directly comparable fasting blood samples suggest an increase in dysglycaemia from 2005 through 2014 which is consistent with the increase in BMI and weight observed over the same time period. By contrast, our secondary analyses that used all the data available suggest a decline in rates of dysglycaemia over the same period. These secondary estimates employed recognised, but different, methods of assessment at each time point and this appears to have confused the picture in terms of estimating change. It is well known that estimates of dysglycaemia prevalence based upon assays of RPG levels can be affected by factors such as the average time between sampling and the last meal. The use of prevalence assessments in 2010 based upon assays of HbA1c introduced further challenges to the interpretation of the data. It seems likely that the dysglycaemia prevalence rates based on HbA1c were a slight underestimate compared to what would have been obtained if fasting venous samples had been used, but the main issue arose with the use of HbA1c values to differentiate between diagnoses of diabetes and prediabetes. The very high rates of diabetes inferred for 2010 suggest that the recommended HbA1c cut point for differentiating between diabetes and pre-diabetes is unsuitable for this population. Of note, the use of DBS samples for HbA1c assay is well established and robust, and while a minor adjustment was made using a published regression equation, the impact of that adjustment on the primary results was minimal.

We believe that biases consequent upon the use of different assessment methods are the most likely explanation for the difference between the findings of the primary and secondary analyses we did. Restricting the primary comparison to the subset of participants assessed with directly comparable techniques should have provided a good approximation of truth by minimising the risk of confounding. While the quantity of participants upon which the primary analyses were based is much lesser than for the secondary analyses, the primary comparison still included fairly large numbers and provided reasonably precise estimates.
Furthermore, the characteristics of this subset were not markedly different from either the full 2014 SMARThealth sample or the 2005 APRHI sample (Tables 4-3 and 4-4) so the primary analysis findings should be generalizable to the broader population. Internal consistency in the findings for men and women within the primary data provide additional reassurance of the likely quality of these analyses - the increase in diabetes was observed in the female population for which a larger increase in body mass index was also observed.

The problems with estimating the prevalence of diabetes, pre-diabetes and dysglycaemia in this study raise important questions about the comparability of accepted methods for estimating diabetes prevalence from different types of samples using different assays, even when there are population-specific recommendations for each. A series of prior reports have identified similar challenges although the extent of the problem we observed has not previously been identified. The criteria used for the diagnosis of diabetes, pre-diabetes and dysglycaemia based upon different assay techniques in Indian populations are inconsistent and further work is required to clarify the cut points at which the correct prevalence estimates would be achieved.

Our conclusion of a high, and likely increasing, prevalence of dysglycaemia would not be unique to this rural Indian community – an almost three-fold increase (2.4% to 6.4%) in diabetes prevalence was reported for the 14 year period from 1989 to 2003 in neighbouring rural Tamil Nadu with a further escalation of prevalence to close to four times (9.2%) the 1989 value reported from a survey done in the same population in 2006. The rates of self-awareness of diabetes observed in our study compare favourably to the 31% reported in urban Indian settings. An elevated level of NGO activity in these villages in the period leading up to the initial APHRI survey and a series of subsequent initiatives in the area are likely to be a part of the explanation for the relatively high awareness.

Diabetes and prediabetes are responsible for considerable morbidity and mortality caused by macro-vascular and micro-vascular disease complications. Progressive urbanisation of the Indian population and unhealthy lifestyles are likely to further compound the consequences of dysglycaemia as is a likely genetic predisposition to type 2 diabetes and serious complications amongst South Asians. Our data shine a further spotlight on the huge issue of diabetes in the Indian population. Regardless of the challenges in interpreting the details of the findings reported here, it is clear that diabetes is, and will remain, a major health challenge for the country. Novel initiatives such as the SMARThealth project seeking to identify simple, low costs and scalable solutions to chronic disease management in rural India are urgently required. Robust evaluation of these types of initiatives will be vital and our findings highlight the need for the use of standardised methods to achieve this.
Table 4-1. Characteristics of surveyed populations in 2005, 2010, and 2014

Comparisons were made using ANOVA for continuous variables and a chi-squared test for homogenous proportions for categorical variables. Unpaired Student’s t-test for continuous variables used for comparing when only two studies included.

Values are weighted estimates reflecting the mean population values for the 20 villages surveyed in APRHI in 2005, the 14 villages surveyed in GC13 in 2010 and 54 villages from SMARTheath 2014.

‘-’ indicates data not collected.

<table>
<thead>
<tr>
<th></th>
<th>2005 % (95%CI)</th>
<th>2010 % (95%CI)</th>
<th>2014 % (95%CI)</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PARTICIPANTS INCLUDED IN PRIMARY ANALYSES BASED UPON THOSE WITH FASTING SAMPLES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>54·1 (53·7-54·6)</td>
<td>55·0 (54·2-55·8)</td>
<td>0·112</td>
<td></td>
</tr>
<tr>
<td>Males (%)</td>
<td>50·6 (48·6-52·6)</td>
<td>45·7 (42·1-49·2)</td>
<td>0·052</td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>54·4 (53·9-54·9)</td>
<td>56·1 (54·1-58·1)</td>
<td>0·019</td>
<td></td>
</tr>
<tr>
<td>Height (cm)</td>
<td>156·4 (156·0-156·7)</td>
<td>154·6 (153·3-156·0)</td>
<td>0·050</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22·2 (22·0-22·3)</td>
<td>23·3 (22·5-24·0)</td>
<td>&lt;0·001</td>
<td></td>
</tr>
<tr>
<td>Formal Schooling (%)</td>
<td>48·4 (46·5-50·4)</td>
<td>50·9 (47·3-54·5)</td>
<td>0·477</td>
<td></td>
</tr>
<tr>
<td>Smokers (%)</td>
<td>28·1 (26·3-29·9)</td>
<td>30·3 (27·0-33·6)</td>
<td>0·506</td>
<td></td>
</tr>
<tr>
<td>Previous MI (%)</td>
<td>1·73 (1·24-2·22)</td>
<td>3·47 (2·16-4·79)</td>
<td>&lt;0·001</td>
<td></td>
</tr>
<tr>
<td>Previous Stroke (%)</td>
<td>2·36 (1·81-2·90)</td>
<td>2·54 (1·41-3·67)</td>
<td>0·966</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2005 % (95%CI)</th>
<th>2010 % (95%CI)</th>
<th>2014 % (95%CI)</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PARTICIPANTS INCLUDED IN SECONDARY ANALYSES OF ALL SURVEYED</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>54·6 (53·6-54·5)</td>
<td>55·7 (52·2-56·3)</td>
<td>54·1 (54·0-54·2)</td>
<td>0·005</td>
</tr>
<tr>
<td>Males (%)</td>
<td>50·9 (49·0-52·9)</td>
<td>48·5 (46·2-50·8)</td>
<td>46·8 (46·4-47·2)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>54·4 (53·9-54·9)</td>
<td>55·9 (55·4-56·5)</td>
<td>58·8 (58·6-59·1)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>156·4 (156·1-156·8)</td>
<td>156·5 (156·1-156·9)</td>
<td>155·2 (155·0-155·4)</td>
<td>0·011</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22·2 (22·0-22·3)</td>
<td>22·8 (22·6-23·0)</td>
<td>24·3 (24·2-24·4)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Formal Schooling (%)</td>
<td>48·8 (46·9-50·8)</td>
<td>74·9 (72·9-76·9)</td>
<td>55·9 (55·5-56·3)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Smokers (%)</td>
<td>28·6 (26·8-30·3)</td>
<td>28·0 (25·9-30·1)</td>
<td>27·5 (27·2-27·9)</td>
<td>0·391</td>
</tr>
<tr>
<td>Previous MI (%)</td>
<td>1·70 (1·23-2·18)</td>
<td>3·43 (2·59-4·27)</td>
<td>2·29 (2·17-2·41)</td>
<td>0·001</td>
</tr>
<tr>
<td>Previous Stroke (%)</td>
<td>2·37 (1·83-2·92)</td>
<td>2·16 (1·44-2·88)</td>
<td>1·76 (1·65-1·86)</td>
<td>0·032</td>
</tr>
<tr>
<td>Previous Angina (%)</td>
<td>5·03 (4·19-5·87)</td>
<td>1·91 (1·29-2·53)</td>
<td>-</td>
<td>&lt;0·001</td>
</tr>
</tbody>
</table>
Table 4-2. Prevalence of diabetes and prediabetes in 2005, 2010, and 2014 and differences by age and by sex amongst all aged 40 to 85

<table>
<thead>
<tr>
<th></th>
<th>2005 % (95%CI)</th>
<th>2010 % (95%CI)</th>
<th>2014 % (95%CI)</th>
<th>P- Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRIMARY ANALYSES BASED UPON THOSE WITH DIRECTLY COMPARABLE FASTING SAMPLES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dysglycaemia (diabetes or prediabetes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>55·1 (52·2-57·9)</td>
<td>52·8 (48·0-57·7)</td>
<td>56·5 (50·4-62·7)</td>
<td>0·606</td>
</tr>
<tr>
<td>Women</td>
<td>52·3 (49·6-55·1)</td>
<td>72·8 (68·1-77·5)</td>
<td>42·1 (41·6-42·6)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>50·5 (47·1-54·0)</td>
<td>55·2 (49·4-60·9)</td>
<td>53·0 (46·9-59·1)</td>
<td>0·345</td>
</tr>
<tr>
<td>50-59</td>
<td>56·0 (52·4-59·6)</td>
<td>62·3 (55·6-69·1)</td>
<td>47·9 (47·3-48·6)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>60+</td>
<td>55·7 (52·6-58·8)</td>
<td>69·2 (63·6-74·9)</td>
<td>47·9 (47·3-48·6)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>18·9 (16·7-21·0)</td>
<td>23·6 (19·5-27·7)</td>
<td>19·3 (17·2 to 21·5)</td>
<td>0·115</td>
</tr>
<tr>
<td>Women</td>
<td>15·9 (13·9-17·9)</td>
<td>37·4 (32·2-42·6)</td>
<td>29·2 (24·8-33·7)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>13·4 (11·1-15·8)</td>
<td>23·4 (18·6-28·3)</td>
<td>31·7 (26·3-37·1)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>50-59</td>
<td>19·0 (16·2-21·8)</td>
<td>33·7 (27·1-40·3)</td>
<td>28·6 (22·3-35·0)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>60+</td>
<td>20·7 (18·1-23·2)</td>
<td>34·2 (28·4-40·0)</td>
<td>35·0 (29·2-40·8)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Pre-diabetes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>36·3 (33·6-39·0)</td>
<td>29·2 (24·8-33·7)</td>
<td>36·0 (33·4-38·7)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Women</td>
<td>36·4 (33·8-39·1)</td>
<td>35·4 (30·3-40·5)</td>
<td>28·6 (22·3-35·0)</td>
<td>0·999</td>
</tr>
<tr>
<td><strong>SECONDARY ANALYSES BASED UPON ALL SURVEYED</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dysglycaemia (diabetes or prediabetes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>55·3 (52·6-58·1)</td>
<td>56·5 (50·4-62·7)</td>
<td>40·3 (39·7-40·9)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Women</td>
<td>52·4 (49·7-55·2)</td>
<td>43·0 (36·6-49·3)</td>
<td>42·1 (41·6-42·6)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>50·5 (47·1-54·0)</td>
<td>48·5 (40·8-56·2)</td>
<td>33·6 (33·0-34·2)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>50-59</td>
<td>55·7 (52·1-59·3)</td>
<td>50·7 (41·4-59·9)</td>
<td>45·1 (44·4-45·9)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>60+</td>
<td>55·4 (52·4-58·5)</td>
<td>53·0 (46·9-59·1)</td>
<td>47·9 (47·3-48·6)</td>
<td>&lt;0·013</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>19·3 (17·2 to 21·5)</td>
<td>37·0 (31·0 to 43·0)</td>
<td>19·3 (17·2 to 21·5)</td>
<td>0·003</td>
</tr>
<tr>
<td>Women</td>
<td>16·2 (14·2 to 18·2)</td>
<td>29·6 (23·6 to 35·2)</td>
<td>21·8 (21·2-22·4)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>13·5 (11·0 to 16·0)</td>
<td>27·7 (20·7 to 34·7)</td>
<td>12·7 (12·3-13·1)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>50-59</td>
<td>19·1 (16·1 to 22·1)</td>
<td>34·2 (25·2 to 43·2)</td>
<td>21·8 (21·2-22·4)</td>
<td>0·003</td>
</tr>
<tr>
<td>60+</td>
<td>20·5 (17·8 to 23·2)</td>
<td>41·0 (34·9 to 47·1)</td>
<td>22·0 (21·4-22·5)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Pre-diabetes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>36·0 (33·4-38·7)</td>
<td>19·8 (14·7-24·8)</td>
<td>22·5 (22·0-23·0)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>Women</td>
<td>36·2 (33·6-38·9)</td>
<td>14·1 (9·46-18·6)</td>
<td>23·8 (23·3-24·2)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>37·0 (33·7 to 40·3)</td>
<td>20·8 (14·5 to 27·1)</td>
<td>20·9 (20·4-21·4)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>50-59</td>
<td>36·6 (33·2 to 40·1)</td>
<td>17·2 (10·2 to 24·2)</td>
<td>23·3 (22·7-24·0)</td>
<td>&lt;0·001</td>
</tr>
<tr>
<td>60+</td>
<td>34·9 (32·0 to 37·9)</td>
<td>12·5 (8·54 to 16·5)</td>
<td>25·9 (25·3-26·6)</td>
<td>&lt;0·001</td>
</tr>
</tbody>
</table>

Comparisons were made using Chi-Squared tests for homogeneity.
Figure 4-1. Prevalence of diabetes and pre-diabetes in 2005, 2010, and 2014
a) Fasting participants only (primary analyses)

b) All adults regardless of assessment method (secondary analyses)

‘Undiagnosed’ is equivalent to previously unknown cases of diabetes, diagnosed only during this study.
Figure 4-2. Survey structures for APRHI and GC13

GC13

\( i = \text{one of 14 strata (villages)} \)

\( M_i = \text{number of clusters (households) in the } i \text{-th stratum} \)

\( m_i = \text{number of clusters surveyed in } i \text{-th stratum} \)

\( j = \text{a particular cluster} \)

\( N_{ij} = \text{number of people aged above 15 in cluster } j \text{ of stratum } i \)

\( N_i = \text{number of people aged above 15 in stratum } i \)

\( n_{ij} = \text{number of people surveyed in cluster } j \text{ of stratum } i \)

\( n_i = \text{number of people surveyed in stratum } i \)

\( N_{ik} = \text{number of people aged above 15 in stratum } i \text{ in standard age/sex stratum } k \)

\( n_{ik} = \text{number of people surveyed in stratum } i \text{ in standard age/sex stratum } k \)

Where the 10 standard strata: 15-29, 30-39, 40-49, 50-59, 60+ for both males and females.

Probability weight, \( pw_{ij} = \frac{1}{P(\text{selecting household } j) \times P(\text{selecting individual})} = \frac{1}{m_i \times \frac{1}{N_{ij}}} \)

Therefore, \( pw_{ij} = \frac{M_i N_{ij}}{m_i} \)

Post-stratification weight, \( s_{ik} = \frac{N_{ik}}{N_i} \times \frac{N_i}{n_i} \)

Therefore final weight, \( w_{ijk} = pw_{ij} \times s_{ik} \)

APRHI

\( N_k = \text{number of people in standard age/sex stratum } k \)

\( n_k = \text{number of people surveyed in standard age/sex stratum } k \)

Where the 8 standard strata: 30-39, 40-49, 50-59, 60+ for both males and females.

Post-stratification weight, \( s_k = \frac{N_k}{n_k} \)

Detailed weighting structure of the studies.
Table 4-3. Characteristics of surveyed population in 2014 amongst all and fasting only aged 40 to 85 years of age

<table>
<thead>
<tr>
<th></th>
<th>2014 Full Sample</th>
<th>2014 Fasting</th>
<th>P- Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (95%CI)</td>
<td>% (95%CI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=62,254</td>
<td>n=749</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>54·1 (54·0-54·2)</td>
<td>55·0 (54·2-55·8)</td>
<td>0·860</td>
</tr>
<tr>
<td>Males (%)</td>
<td>46·8 (46·4-47·2)</td>
<td>45·7 (42·1-49·2)</td>
<td>0·841</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>58·8 (58·6-59·1)</td>
<td>56·1 (54·1-58·1)</td>
<td>0·054</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>155·2 (155·0-155·4)</td>
<td>154·6 (153·3-156·0)</td>
<td>0·520</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>24·3 (24·2-24·4)</td>
<td>23·3 (22·5-24·0)</td>
<td>0·031</td>
</tr>
<tr>
<td>Formal Schooling (%)</td>
<td>55·9 (55·5-56·3)</td>
<td>50·9 (47·3-54·5)</td>
<td>0·024</td>
</tr>
<tr>
<td>Smokers (%)</td>
<td>27·5 (27·2-27·9)</td>
<td>30·3 (27·0-33·6)</td>
<td>0·224</td>
</tr>
<tr>
<td>Previous MI (%)</td>
<td>2·29 (2·17-2·41)</td>
<td>3·47 (2·16-4·79)</td>
<td>0·085</td>
</tr>
<tr>
<td>Previous Stroke (%)</td>
<td>1·76 (1·65-1·86)</td>
<td>2·54 (1·41-3·67)</td>
<td>0·240</td>
</tr>
</tbody>
</table>

Comparisons were made using Unpaired Student’s t-test for continuous variables and a chi-squared test for homogenous proportions for categorical variables.
Values are weighted estimates reflecting the mean population values for the 54 villages from *SMARThealth* 2014.
Table 4. Characteristics of surveyed population in 2005 amongst all and fasting only aged 40 to 85 years of age

Comparisons were made using Unpaired Student's t-test for continuous variables and a chi-squared test for homogenous proportions for categorical variables.
Values are weighted estimates reflecting the mean population values for the 20 villages from APRHI 2005.

<table>
<thead>
<tr>
<th></th>
<th>2005 Full Sample % (95%CI)</th>
<th>2005 Fasting % (95%CI)</th>
<th>P- Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>54.1 (53.6-54.5)</td>
<td>54.1 (53.7-54.6)</td>
<td>0.999</td>
</tr>
<tr>
<td><strong>Males (%)</strong></td>
<td>50.9 (49.0-52.9)</td>
<td>50.6 (48.6-52.6)</td>
<td>0.970</td>
</tr>
<tr>
<td><strong>Weight (kg)</strong></td>
<td>54.4 (53.9-54.9)</td>
<td>54.4 (53.9-54.9)</td>
<td>0.999</td>
</tr>
<tr>
<td><strong>Height (cm)</strong></td>
<td>156.4 (156.1-156.8)</td>
<td>156.4 (156.0-156.7)</td>
<td>0.999</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>22.2 (22.0-22.3)</td>
<td>22.2 (22.0-22.3)</td>
<td>0.999</td>
</tr>
<tr>
<td><strong>Formal Schooling (%)</strong></td>
<td>48.8 (46.9-50.8)</td>
<td>48.4 (46.5-50.4)</td>
<td>0.942</td>
</tr>
<tr>
<td><strong>Smokers (%)</strong></td>
<td>28.6 (26.8-30.3)</td>
<td>28.1 (26.3-29.9)</td>
<td>0.909</td>
</tr>
<tr>
<td><strong>Previous MI (%)</strong></td>
<td>1.70 (1.23-2.18)</td>
<td>1.73 (1.24-2.22)</td>
<td>0.999</td>
</tr>
<tr>
<td><strong>Previous Stroke (%)</strong></td>
<td>2.37 (1.83-2.92)</td>
<td>2.36 (1.81-2.90)</td>
<td>0.999</td>
</tr>
</tbody>
</table>
Chapter 5. Blood Lipids in Rural Andhra Pradesh – 2005 and 2010

Introduction
As was discussed in chapter 1, dyslipidaemia is an important determinant of cardiovascular disease in India. On the background of limited data on temporal trends in cholesterol levels in rural India, this is a timely study focusing on the potential changes in these cardiovascular risk factors. This is pertinent in a geographic area which is likely to be at high risk for an epidemic of cardiovascular disease and the identification of increasing trends in risk factor prevalences will guide effective interventions and assess the severity of the problem.

This analysis was based on the lipids data extracted from the GC13 survey. A comparison with the APRHI study is included. The analytes measured were total cholesterol (TC) and low density lipoprotein (LDL) along with physical measurements of height, weight and body mass index (BMI).

Methods
The survey structure was as described in chapter 2. The participants aged 30 and above were included in this analysis. The finger pick dried blood spot (DBS) was collected on filter paper for the 2010 GC13 study. The samples were dried at room temperature, packaged in plastic and refrigerated at 2 – 4°C before transportation. Final biochemical assays for analysis were done at the Nizam’s Institute for Medical Sciences, Hyderabad. The analytes were eluted from the filter paper using a phosphate buffer solution. The assays were based on a Randox colorimetric end point assay (Randox Laboratories Ltd, UK). Participant’s characteristics such as weight, height and body mass index (BMI) were recorded. The 2005 APRHI study made use of fasting blood lipid samples collected via traditional venepuncture. Statistical analysis was done as previously described in chapters 2 and 4. Continuous variables were compared using the unpaired Student’s t test.

Results
The mean TC decreased from 4.6 mM (4.5 mM – 4.7 mM, 95% CI) to 3.4 mM (3.3 mM – 3.5 mM, 95% CI), p<0.001. The mean LDL decreased from 2.9 mM (2.8 mM – 3.0 mM, 95% CI) to 1.5 mM (1.4 mM – 1.6 mM, 95% CI), p<0.001. The significant fall in both blood lipids was consistent across all age and sex groups (Table 5-1). The weight and BMI had statistically significant increases from APRHI to GC13 but the average height was unchanged. When only the four common villages were included, the TC decreased from 4.8 mM (4.7 mM – 4.9 mM, 95% CI) to 3.3 mM (3.1 mM – 3.5 mM, 95% CI) and the LDL decreased from 3.0 mM (2.9 mM – 3.1 Mm, 95% CI) to 1.4 mM (1.3 mM – 1.5 mM, 95% CI). All the data appeared to be normally distributed except the GC13 TC and LDL results which showed right skewing (Figures 5-1 and 5-2).
**Figure 5-1.** LDL concentrations from each study. APRHI on the left, GC13 on the right. ‘Density’ on the vertical axis refers to the relative proportions of each LDL level on the horizontal axis.

**Figure 5-2.** TC concentrations from each study. APRHI on the left, GC13 on the right. ‘Density’ on the vertical axis refers to the relative proportions of each total cholesterol level on the horizontal axis.
Table 5-1. Mean levels of total cholesterol, LDL, weight and body mass index overall and by age and sex for APRHI (2005) and GC13 (2010).

<table>
<thead>
<tr>
<th></th>
<th>Total cholesterol (mmol/L)</th>
<th>LDL cholesterol (mmol/L)</th>
<th>Weight (kg)</th>
<th>BMI (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APRHI</td>
<td>GC13</td>
<td>APRHI</td>
<td>GC13</td>
</tr>
<tr>
<td></td>
<td>S.E.</td>
<td>S.E.</td>
<td>S.E.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Overall</td>
<td>4.6 0.04</td>
<td>3.4 0.05</td>
<td>2.9 0.03</td>
<td>1.5 0.03</td>
</tr>
<tr>
<td>30-39</td>
<td>4.3 0.08</td>
<td>3.3 0.09</td>
<td>2.7 0.06</td>
<td>1.5 0.06</td>
</tr>
<tr>
<td>40-49</td>
<td>4.8 0.09</td>
<td>3.4 0.11</td>
<td>3.0 0.07</td>
<td>1.6 0.07</td>
</tr>
<tr>
<td>50-59</td>
<td>5.0 0.07</td>
<td>3.7 0.12</td>
<td>3.2 0.06</td>
<td>1.6 0.08</td>
</tr>
<tr>
<td>60+</td>
<td>4.7 0.08</td>
<td>3.2 0.06</td>
<td>2.9 0.06</td>
<td>1.5 0.04</td>
</tr>
<tr>
<td>Men, all</td>
<td>4.5 0.06</td>
<td>3.5 0.08</td>
<td>2.8 0.05</td>
<td>1.6 0.05</td>
</tr>
<tr>
<td>30-39</td>
<td>4.3 0.12</td>
<td>3.5 0.15</td>
<td>2.7 0.1</td>
<td>1.5 0.11</td>
</tr>
<tr>
<td>40-49</td>
<td>4.6 0.15</td>
<td>3.6 0.16</td>
<td>2.9 0.1</td>
<td>1.7 0.11</td>
</tr>
<tr>
<td>50-59</td>
<td>4.8 0.11</td>
<td>3.7 0.19</td>
<td>3.1 0.09</td>
<td>1.8 0.11</td>
</tr>
<tr>
<td>60+</td>
<td>4.4 0.1</td>
<td>3.1 0.1</td>
<td>2.7 0.08</td>
<td>1.3 0.05</td>
</tr>
<tr>
<td>Women, all</td>
<td>4.8 0.06</td>
<td>3.2 0.06</td>
<td>3.0 0.04</td>
<td>1.5 0.03</td>
</tr>
<tr>
<td>30-39</td>
<td>4.4 0.11</td>
<td>3.2 0.11</td>
<td>2.8 0.08</td>
<td>1.5 0.06</td>
</tr>
<tr>
<td>40-49</td>
<td>5.0 0.1</td>
<td>3.1 0.12</td>
<td>3.2 0.09</td>
<td>1.4 0.08</td>
</tr>
<tr>
<td>50-59</td>
<td>5.2 0.09</td>
<td>3.4 0.15</td>
<td>3.3 0.07</td>
<td>1.6 0.1</td>
</tr>
<tr>
<td>60+</td>
<td>5.0 0.12</td>
<td>3.3 0.08</td>
<td>3.1 0.08</td>
<td>1.6 0.06</td>
</tr>
</tbody>
</table>
Discussion
These data suggest a large drop in mean blood lipids levels in the five years between 2005 and 2010 accompanied by a significant weight and BMI increase. This combination of changes and the magnitude of the blood lipid changes are implausible and almost certainly reflect systematic under-estimation of blood lipid values in the 2010 data.

For context, the normal range LDL in the US population is 2.3 mM to 3.4 mM with an average of 3.1 mM119. It is biologically and epidemiologically challenging to explain the fall in mean LDL observed here (2.9 Mm to 1.5 mM) which would take the mean value in Andhra Pradesh from within the US normal range to a value less than half of the US mean in the space of only five years. By way of further context, a study in rural college students of Andhra Pradesh aged 18-22 done in 2011 showed a mean LDL of 1.86 mM and TC of 3.34 mM117. While markedly lower than the levels observed in the US, these values are still significantly greater than those recorded for the general adult Andhra Pradesh population in the 2010 survey. Students of this age would be anticipated to have lower, not higher, mean lipid levels compared to the general population. Similarly in a recent study representative of most of India, the rural subgroup recorded a mean LDL of 2.30 mM and TC of 3.98 mM91 further highlighting the peculiarity of the GC13 values.

The GC13 lipid measurements were based on assays done on DBS and as shown in chapter 4 these appear to be unreliable of lipid assays and there is a strong likelihood that the biochemical analysis was biased and incorrect because of the use of DBS. The right skewing of only the blood analyte data in GC13 for TC and LDL also raises concerns about the integrity of the biochemical analysis technique or storage conditions of the GC13 samples. All of the other data showed the anticipated approximate normal distribution. A possible explanation would be that the higher LDL and TC values were not correctly measured resulting in both the lower than expected LDL and TC means and also the right skewing of the distributions. This could occur from degradation of the molecules while in storage, lack of complete elution from the filter paper before analysis or a problem with the colorimetry assay.

A similar survey structure and statistical analysis was used for both the blood analytes and the physical measurements but only the distributions of the analyte data are skewed. It is unlikely, therefore, that it is a statistical analysis error leading to the unusual LDL and TC results as the physical measurement data are normally distributed, as expected. The only additional step in the analysis of the blood analytes was the use of post-stratification weights after the final step of selection of DBS participants. This weighting procedure was re-examined and found to be robust. Also if the weights were the issue then there would be a systematic error in all the LDL and TC measurements which would not lead to the skewed distribution but rather a normal distribution with a shifted mean.

A further weakness in the temporal comparisons of this study was the slightly different sampling strategy used for the 2005 and 2010 populations. The subgroup analysis of the common villages does, however, provide results consistent with those observed in the total dataset. The survey methods and statistical analyses are based upon established approaches45 and are unlikely to have introduced substantive inaccuracies. The use of the Student’s t test is justified despite some non-normal distributions of the data, since the sample sizes are large and the variances are similar118.

In conclusion, the changes in lipids observed here are almost certainly a consequence of artefact and are very unlikely to reflect true changes in population lipid levels in this region. The observed increase in BMI and the rising rates of diabetes are both signs of adversely evolving cardiovascular risk factors and it is likely that LDL and TC levels in this population are either static or deteriorating. For further such studies it is recommended that traditional venous blood sampling be used at least until DBS for LDL and TC sampling has been further characterised.
Chapter 6. Discussion and Conclusions

These data provide new insight into the evolving state of cardiovascular risk factors in rural Andhra Pradesh and highlight the need for standardised approaches to monitoring of changes over time. The data show that the prevalence of diabetes mellitus and dysglycaemia is high and most likely increasing but fail to establish a clear pattern for blood lipids. That said, with body weight and body mass index (BMI) both increasing, it is more likely that blood lipid levels are deteriorating than improving.

The observed 72% increase in diabetes mellitus prevalence in the Godavari region of rural Andhra Pradesh between 2005 and 2014 is large. For comparison, a recent set of predictive figures on diabetes suggested that by 2030 there will likely be a 69% increase in the numbers of adults with diabetes in developing countries and a 20% increase in that of developed countries\(^9\). The increase in diabetes in Andhra Pradesh defined by our data is far greater in magnitude since it has occurred in just a decade. If true and if widespread throughout India, the predictions made that infer a rise in India’s 2010 population of 50.8 million with diabetes to 87 million by 2030\(^9\) would be a gross under-estimate. Of note, these predictions are significantly higher than previous predictions made just 5 years earlier\(^11\). In this previous estimate, it was thought that by 2030 the global diabetes prevalence would reach 4.4%, it has been shown that in 2010 that global prevalence has already crossed 6.4%. In 2015 there is an 8.6% prevalence with over 65 million Indians with diabetes\(^12\). This underlines the challenge of diabetes prevalence at either a national or an international level.

The primary analysis in Chapter 4, based on only fasting blood glucose levels, revealed an increasing prevalence of dysglycaemia in rural Andhra Pradesh. This was in parallel with BMI and was driven by the women in the population. There was a greater increase in BMI for women than men but this alone would not account for the majority of dysglycaemia and diabetes prevalence changes being in women. There have been suggestions that the women of rural India often place their own health needs after those of the men and the rest of the family and these data could in part be a manifestation of such. It has been noted that in South Asia women often have limited access to healthcare facilities and education due to negative social structures and customs\(^12\). A recent review identifies that women across India have higher prevalences of obesity. The suggested predisposing factors for obesity in Indian women included sedentary behaviour, imbalanced diets, sequential post-partum weight gain and cultural reasons for decreasing physical activity post-partum\(^12\). These factors are potential explanations for the comparatively higher increase in BMI and dysglycaemia seen in women compared to men.

On viewing the secondary analysis in Chapter 4, the prevalence of diabetes in the population in the Godavari region of Andhra Pradesh appears to have undergone a significant increase when considering the APRHI and GC13 results in isolation. Adding in the SMARTHealth data it seems that the dysglycaemia prevalence is decreasing (Figure 4-1b). This presents a confusing picture not apparent in the primary analysis and raises the issue of diagnostic standards since GC13 used glycosylated haemoglobin (HbA1c) but APRHI and SMARTHealth used plasma glucose levels. Both have been standardized with the diagnostic values determined by the rate of incidence in microvascular diseases, for example, diabetic retinopathy incidence\(^4\). An HbA1c over 6.5%, fasting plasma glucose over 7 mM and random plasma glucose over 11.1 mM all lead to a significantly increased risk of developing microvascular diabetic disease\(^4\), therefore it is expected that these standards would return similar population diabetes prevalences. Unfortunately this is not the case as there have been variations due to ethnicity or presence of haemoglobin subtypes in some populations\(^12\). Since the adoption of diagnosis by HbA1c concerns have been raised about the sensitivity and a correction to the diagnostic cut off has been suggested\(^12\). Even so, it has also been reported that groups diagnosed by either method are both still at increased risk of microvascular diabetic disease\(^4\),\(^12\). Therefore, the different diabetes prevalence identified by HbA1c in these studies may indicate a separate but overlapping population at risk of diabetic complications when compared to the group diagnosed using fasting plasma glucose. For this reason the extremely high diabetes prevalence seen in GC13 may not be completely a matter of overestimation, it may indicate a larger population at risk of cardiovascular disease\(^12\). This is consistent with previous studies showing the progressive rise in cardiovascular risk factors in India. In Chapter 1, the trend for diabetes prevalence in rural Tamil Nadu was noted to be potentially exponentially increasing in that rural region (Figure 1-1), even though the absolute prevalence did not reach the same value found here. The most recent study in that region was in 2006...
with a prevalence of 9.2% (8.0% – 10.5%, 95% CI). By the current stage that number could be approaching the prevalence levels shown here in rural Andhra Pradesh, assuming that there were no interventions.

The trend in the Godavari region cannot be interpreted with only the two data points on diabetes prevalence particularly with such uncertainty over the meaning of the prevalence diagnosed by two different methods. In this case the SMARTHealth study permits a greater insight into the problem. The results confirm that GC13 had either identified a different at risk population or overestimated diabetes prevalence due to a bias caused by the use of HbA1c-based criteria. Another approach to analysis was to combine the diabetes and prediabetes prevalence as a composite metric – dysglycaemia, and view that aggregate as an indication of increased cardiovascular risk. The changes in the level of dysglycaemia over time, rather than diabetes or prediabetes individually, showed a more plausible variation (Figure 4-1). This indicates that the group that GC13 identified as having diabetes included some of the population that was diagnosed as prediabetic according to plasma glucose levels in APRHI and SMARTHealth. Therefore, the diabetes prevalence was significantly higher but prediabetes prevalence was much lower in GC13 compared to APRHI and SMARTHealth (Figure 4-1b). Essentially, it appears that the diagnostic cut off for diabetes should be higher when using HbA1c for this population. Nevertheless, the changes in dysglycaemia are relevant to changes in cardiovascular risk profiles. The increase shown in the primary analysis suggests that diabetes rates are still on the rise. The contrasting decrease in dysglycaemia found in the secondary analysis was likely a consequence of relying mostly on non-fasting blood glucose levels in SMARTHealth, as was discussed in Chapter 4.

The blood lipids data has raised some interesting questions about the quality of analysis and biochemical methods. The decrease seen in blood lipids is unfeasibly rapid. On further inspection, and when viewed alongside the diabetes data, the decrease in blood lipids from 2005 to 2010 looks to be inconsistent with the increase in diabetes and BMI from 2005 to 2014 (figure 4-1a). A likely scenario is that there was a biochemical analytic error of sorts as expounded by the discussion in chapter 5. It would be preferable to repeat the study on blood lipids in the region using either traditional venepuncture based measurements or using dried blood spots (DBS) once it has been adequately characterised for these analytes. As seen in Chapter 3, the data is currently sparse and uniform methods for storage, elution and subsequent correction of the data need to be determined before further use of DBS for lipids.

The difficulty in assessing blood lipid levels in this study highlights the issues in describing blood analyte levels in developing regions. Chapter 3 examines the suitability of DBS as a possibility. It has been shown that DBS is a feasible option for detecting diabetes prevalence rates. It has the potential to become the standard method in studies involving large-scale population sampling. This is true particularly in resource poor regions where it has the greatest comparative advantage over traditional venepuncture sampling. It requires minimal training and is generally easy to store and safe to handle. DBS is shown to be reliable in diagnosing diabetes in a population sample where the individual inaccuracies will be averaged out over the mean. Specifically, the proportion of the sample who have an HbA1c above 6.5% can be adequately assessed using DBS and therefore population diabetes prevalence can be inferred, depending of course on the reliability of the survey structure and analysis.

For studies of cardiometabolic risk factors it would be advantageous to sample multiple blood analytes with a single DBS. For example, if a population can be sampled for HbA1c, total cholesterol, low density lipoprotein (LDL), high density lipoprotein (HDL) and triglycerides simultaneously using only a single drop of blood per person, then sampling would be seen as a more feasible task for potential investigators. This may encourage both more studies to be contemplated and more participants in each study. To achieve this utility from DBS, further studies of LDL, HDL, total cholesterol and triglycerides are required to characterize the relationship of DBS to venepuncture samples. In the background of changing disease burdens and risk factors this is of particular importance as it will allow more frequent sampling and follow up of specific regions. This will impact health policy decisions at a later stage where more informed decisions can be made to target the cardiovascular risk factors that pose the greatest threat to particular regions. The GC13 survey is a preliminary model of how this can be achieved relatively efficiently with DBS, only the technique requires some optimisation.
The value in characterizing the relationship between DBS and venous HbA1c is apparent in chapter 4. When presented with such a significant increase in diabetes prevalence from APRHI to GC13, one of the questions raised was about the use of different methods in the two surveys. The results of the meta-analysis were necessary to show that the unrealistic rise in diabetes over five years was not likely to be due to any issues with using DBS in GC13. This is particularly evident since the GC13 results were standardized using the regression equation derived in chapter 3 so as to be more comparable with the data in APRHI. This allowed attention to be diverted towards investigating other potential explanations for the results in GC13 which can be found in Appendix 1. Finally, the likely explanation for the excessive diabetes prevalence in GC13 was attributed to a non-robust diagnostic cut-off for diabetes in this population. Unfortunately there was no data available to assess how LDL measurements by DBS should be approached but total cholesterol levels seen in GC13 were adjusted for the use of DBS with a regression equation from the systematic review.

In placing this work in context we find that there has been a recent swing in the focus of cardiovascular epidemiology worldwide and increasing amounts of the latest literature has focused towards the rising epidemic of cardiovascular diseases in the developing world. These latest studies (Chapters 4 and 5) in rural Andhra Pradesh will provide further knowledge on how cardiovascular risk factors evolve in these environments and will highlight any differences from the progression of the same diseases in urban India and the developed world more broadly. In 2011, 'Chronic diseases and injuries in India' revisited the issues presented in 'Responding to the threat of chronic diseases in India', another review published 6 years earlier. At the moment there are a wide array of possible primary and secondary prevention strategies for cardiovascular diseases but the access to prevention is limited, especially in the more rural and poorer regions. In addition, a large proportion of current care for chronic diseases in India is in the private medical sector and this becomes prohibitively expensive for much of the population who are at risk. The authors advocate for a strengthening of social and public policy charters to allow the efficient implementation of appropriate interventions.

The World Health Survey of 2003 is highlighted on in the 'Chronic diseases and injuries in India' publication. This 2003 survey had a stratified multistage cluster design and was done in the Indian states of Assam, Karnataka, Maharashtra, Rajasthan, Uttar Pradesh and West Bengal. The purpose of the study was to assess chronic disease symptoms, previous diagnosis rates and economic status of the participants. A distinct pattern of an initial uptake of harmful health behaviours in the early stages of socioeconomic advancement was identified. Even though socioeconomic development generally tends to be associated with healthier behavioural changes, a steep climb in economic growth has led to a reduction of physical activity and increasing rates of obesity and diabetes. Current levels of health literacy and public awareness are inadequate and the result is the more affluent class are slowly adopting healthier lifestyles while the poorer groups are somewhat neglected. The fact that over 20% of the Indian population are smokers is further evidence of the need to improve public awareness of the well understood health hazards which are assumed as standard knowledge.

The authors of 'Responding to the threat of chronic diseases in India' also provide specific estimates for the impact of chronic diseases based on the trends and data observed to 2011. It is estimated that chronic diseases will move to become accountable for almost 75% of all mortality in India by the year 2030. The number of years of life lost due to coronary heart disease in those younger than 60 is projected to increase from the 7.1 million observed in 2004 to 17.9 million in 2030 - this is greater than that expected in China, Russia and the USA combined. Recommendations for prevention include population and individual focused efforts such as tobacco control, reduced dietary salt intake, blood pressure lowering medications, cholesterol lowering medications and combination drug treatment. With the common combination drug treatments such as an antihypertensive and a statin, scaling up production to provide for 50% of the total high risk individuals in India would avert 5.8 million deaths over 10 years. The cost would be less than 1 USD per person per year for all the high risk group if the treatment is targeted to those who would receive the most benefit, by effective screening of risk factors. This includes the cost of medications, health service delivery, laboratory testing and program administrative costs relevant to India, all summated and divided per head. In the public policy space, an economic modelling study has predicted that taxation on sugar sweetened beverages would mitigate rising obesity and type 2 diabetes with particularly large effects in young rural men rather than the expected urban beneficiaries. This provides another perspective on what interventions might be most effective.
Further pertinent predictions of cardiovascular disease risk are presented in ‘Distribution of 10-year and lifetime predicted risk for cardiovascular disease in the Indian Sentinel Surveillance Study population (cross-sectional survey results)’. This article is based on a nationwide risk factor surveillance study with 10,054 disease-free adult Indians who had their cardiovascular disease risk stratified into groups. The main findings are that high short term coronary heart disease risk was prevalent in more than one fifth of the population at 23.5% (22.7% - 24.4%, 95% CI). Also almost half, 48.2% (47.1% - 49.3%, 95% CI), of the participants who were classified as low short term risk were found to have a high predicted total lifetime risk. The fraction of the participants who had optimal levels of cardiovascular risk factors was 15.3% (14.6% - 16.0%, 95% CI). This was further divided into 20.6% (18.7% - 22.6%, 95% CI) of individuals from the highest educational group having optimal risk factors and only 8.8% (7.7% - 10.5%, 95% CI) of the lowest educational group having the same. This provides evidence of the socioeconomic status based health disparities within India. There is also a review of the progress made on previous recommendations made in 2005 for cardiovascular disease and diabetes. It is worth reproducing and reviewing the relevant list of recommendations and their current progress as presented by the authors of this review:

- **Tobacco control**: Banned on films and TV programmes. Ban on specific public places.
- **Production and supply of healthy foods**: Nil
- **Regulation of unhealthy foods**: Nil
- **Urban planning to promote physical activity**: Nil
- **Community empowerment through health promotion programmes**: National Rural Health Mission aims to integrate health promotion activities of the National Programme for prevention and control of diabetes, CVD and stroke into its overall goals
- **Health system strengthening aimed at early detection of high risk individuals**: Guidelines developed. Medical officer’s manual on prevention and control of diabetes, CVD and stroke. Health workers guide with a flip chart for community awareness. India-specific physical activity guidelines
- **Effective secondary prevention of chronic disease**: Involvement of medical colleges and private practitioners in setting up pilot special clinics.
- **Cost effective and lifesaving acute care**: Nil

From this there is some cause for optimism but there are many areas of public health policy that are yet to be harnessed to optimally tackle the challenge of the cardiovascular disease epidemic in India, of which a substantial component is the rising diabetes levels seen in the rural regions.

It must also be considered that there are other, less modifiable factors that may augment or assist the rapid rises of cardiovascular diseases and diabetes in India. Many have proposed that South Asians have a higher propensity for diabetes and cardiovascular disease for reasons other than demographics and urbanization alone. One study, ‘Serum cholesterol and coronary artery disease in populations with low cholesterol levels: The Indian paradox’, finds that in Indian populations serum total cholesterol level is directly related to coronary heart disease prevalence even amongst those with a low cholesterol (<5.18 mM). This result indicates that we may need to consider lipid levels as a continuous spectrum of risk rather than only at specific cut-off points. It is possible then that there may be an advantage to lowering total cholesterol levels to below the normal range considered optimum.

There are also questions in the literature of increased central obesity conferring a greater risk for Indian populations than for other groups. There is a hypothesis proposed in ‘Why might South Asians be so susceptible to central obesity and its atherogenic consequences? The adipose tissue overflow hypothesis’ as to the metabolic disturbances that may be affecting this population. The authors propose that the fat compartment of superficial subcutaneous adipose tissue has a lower total capacity in South Asians compared to Caucasians. Therefore excess fat is further deposited in the more atherogenic and metabolically active deep subcutaneous tissue and visceral adipose tissue compartments. This is presumed to lead to accelerated vascular disease as has been observed with disease occurring in younger age groups amongst the Indian population.
Genetic studies have also suggested some differences in the susceptibility of Indians to diabetes mellitus. The authors of ‘Genetic predisposition to type 2 diabetes among Asian Indians’ determine that there are certain genes predisposing Indians to developing diabetes while there are also other genes which are normally protective against insulin resistance in Caucasians but do not appear to be protective for the Indian population\textsuperscript{115}. More specifically, a recent publication has identified a new type 2 diabetes mellitus associated gene locus at 2q21\textsuperscript{136}. This was a two stage genome wide association study of diabetes in 12,535 Indians. The pertinent results showed that both the previously known loci and the newly distinguished 2q21 together explained 7.65\% of the variance in the risk of diabetes in Indians. This new evidence for a true genetic predisposition of the Indian population towards diabetes can partially explain the high diabetes rates seen in many parts of India. This is notwithstanding that India is not the only area with rapid urbanization and demographic changes but still appears to be one of the most severely affected by the diabetes and cardiovascular disease epidemic.

The overall conclusions of my work are that the DBS technique is very promising in its application to resource poor regions when surveying cardiovascular risk factor blood analytes but needs some further work to determine the best methodology for each analyte. The rural Andhra Pradesh population followed up in these surveys have shown an increasing dysglycaemia prevalence but changes in blood lipid levels were difficult to determine with any certainty. The proportion of the population with an HbA1c over 6.5\% is very concerning for the future development of diabetic complications and cardiovascular disease. More data is required on blood lipid levels in this population to ensure adequate surveillance of evolving cardiovascular risk factors in a region likely undergoing significant changes in cardiovascular disease. The increasing dysglycaemia prevalence requires urgent attention, particularly in the context of global and regional shifts in the burden of disease from non-communicable diseases.
References


56. Organization WH. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus Screening for Type 2 diabetes. 2011.


113. Singh Z. Aging: the triumph of humanity—are we prepared to face the challenge? *Indian journal of public health* 2012; 56(3): 189-95.


128. Reddy AA. Growth, Structural Change and Wage Rates in Rural India.
Appendix 1. Assessing Potential Sources of Bias in GC13

Introduction
In light of the diabetes prevalence results in GC13, further queries can be made regarding the nature of any errors or possible overestimation of diabetes in the GC13 data. As was shown in chapter 3, error is unlikely to be due to the use of DBS. In chapter 4 it was proposed that GC13 had likely overestimated diabetes and underestimated prediabetes. The evidence was seen in the variation amongst the diabetes and prediabetes aggregates which showed a more understandable trend than what appeared to be a severe spike in diabetes prevalence as seen in GC13. To provide further insights into the comparison of APRHI and GC13 and possible explanations for the results, a propensity score matched analysis\textsuperscript{137} was undertaken. It was hypothesized that once the participants in APRHI and GC13 were matched on propensity scores, any differences in target population or sampling errors would be minimized therefore allowing diabetes prevalence to be assessed with little residual confounding effects. The purpose of this was to search for any bias in the GC13 data that could have been from confounding factors, eg. Sample BMI distribution or exercise habits. The results from each interviewer involved in the study were also examined, searching for any clear bias amongst individual interviewers due to poor technique.

Methods
Potential effects on diabetes prevalence due to differences in background covariates in the APRHI and GC13 samples were explored using a propensity score matching method. This involved assigning scores to each participant depending on their propensity to be in either study based on their background covariates; age, sex, body mass index (BMI), smoking status, previous school education, controlling weight, physical exercise, reducing dietary salt and fat or past history of angina, stroke or myocardial infarction. These scores were derived using a logistic regression of study, either APRHI or GC13, on the background covariates. Then the diabetes prevalence was compared only within those with similarly matched propensity scores thereby reducing potential confounding error. The matching procedure was single nearest neighbour matching without replacement and with set calipers. This involved matching a participant in APRHI to only the nearest participant in GC13. This must also be within a minimum propensity score difference, the caliper. Therefore the matched pairs never had a difference in propensity score greater than the chosen caliper value. The caliper value used was 0.2 of the standard deviation of the logit propensity score as this was shown as optimal\textsuperscript{138}.

The GC13 results were separated into mean HbA1c by interviewer and questionnaire diabetes prevalence by interviewer. This was to detect potential sample collection error and incorrect recording of previously known diabetes respectively. Analysis of variance (ANOVA) was conducted on the mean HbA1c by interviewer data. All statistical analysis was using the statistical software package STATA 11.0 (StataCorp, College Station, TX, USA) with the PSMATCH2 module applied to conduct the propensity score matched analysis.
Results

The propensity score matching uncovered a persistent bias in the GC13 data compared to APRHI even after accounting for the other measured covariates (Figure A1-1). As expected, the propensity score matched samples had little difference in the background covariates measured but it did not reduce the difference in diabetes prevalence between APRHI and GC13. This difference increased after matching process.

Figure A1-1. Propensity score matched analysis. Bias in diabetes prevalence and other covariates between APRHI and GC13. Top panel shows the bias between APRHI and GC13 for each covariate. The propensity score matching process has reduced the bias and the matched subgroup now have similar levels of background covariates. The bottom panel demonstrates the effect on the bias in diabetes prevalence, between APRHI and GC13, before and after matching for the other covariates. The bias or difference in diabetes prevalence is not significantly affected when we use the propensity score matched subgroup.
No particular interviewer could clearly be found responsible for spurious results due to either incorrect blood spot collection or data entry (Figure A1-2). An ANOVA of the data, divided into groups by interviewer ID, confirmed the differences between interviewers were non-significant, p = 0.881.

A subgroup analysis of the 4 common villages showed a non-significant increase in diabetes prevalence but prediabetes decreases consistent with the general data (Table A1-1).

![Graph](image)

**Figure A1-2. Results by interviewer. Assessing for potential data collection errors in GC13.**
### Table A1-1. Prevalence of diabetes and prediabetes in 2005 and 2010 by age and sex for the four villages common to both APRHI and GC13.

<table>
<thead>
<tr>
<th></th>
<th>2005 % (95%CI)</th>
<th>2010 % (95%CI)</th>
<th>Absolute Difference % (95%CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DIABETES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>17.6 (14.5-20.7)</td>
<td>21.9 (14.9-29.0)</td>
<td>4.30 (-3.34 to 11.9)</td>
<td>p=0.270</td>
</tr>
<tr>
<td>Men</td>
<td>18.9 (14.4-23.3)</td>
<td>19.3 (10.6-28.0)</td>
<td>0.40 (-9.32 to 10.1)</td>
<td>p=0.936</td>
</tr>
<tr>
<td>Women</td>
<td>16.3 (12.1-20.6)</td>
<td>25.2 (13.9-36.5)</td>
<td>8.90 (-3.12 to 20.9)</td>
<td>p=0.147</td>
</tr>
<tr>
<td><strong>PREDIABETES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>35.0 (31.1-38.8)</td>
<td>15.6 (9.61-21.6)</td>
<td>-19.4 (-26.5 to -12.3)</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Men</td>
<td>34.8 (29.5-40.2)</td>
<td>16.3 (8.16-24.4)</td>
<td>-18.5 (-28.2 to -8.83)</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Women</td>
<td>35.1 (29.7-40.6)</td>
<td>14.8 (5.84-23.8)</td>
<td>-20.3 (-30.8 to -9.83)</td>
<td>p&lt;0.001</td>
</tr>
</tbody>
</table>

#### Discussion

In effect, the propensity score matching allows an observational study to simulate some characteristics of randomized studies. It minimizes the effect of confounding factors amongst treatment and intervention groups, in this case the APRHI and GC13 studies. These results of the propensity score analysis argue against any of the measured covariates being a major contributor to the increased diabetes seen in GC13, although this does not rule out any unmeasured confounders. Considering the difference between the diabetes prevalence values were not reduced by matching on the sample background characteristics it is seems likely that there is another cause for such a sharp increase in diabetes prevalence from 2005 to 2010. The suggestion in Chapter 4 that the diagnostic cut-off definitions based upon fasting plasma glucose and HbA1c are not equivalent for this population seems to be the remaining explanation.
Appendix 2. APRHI Survey Questionnaire

Relevant questions for this work are included below:

THANK YOU FOR AGREEING TO TAKE PART IN THIS SURVEY TO FIND OUT ABOUT THE RISKS OF HEART ATTACK, STROKE AND OTHER CHRONIC DISEASES FOR RESIDENTS OF RURAL ANDHRA PRADESH. MOST OF THE QUESTIONS I WILL ASK WILL BE ABOUT YOU, AND A FEW WILL BE ABOUT YOUR HOUSEHOLD. AT THE END, THERE WILL BE A BRIEF PHYSICAL EXAMINATION.

1. How old are you? [___]

2. Can you read and write?
   1. Yes [___]
   2. No [___]

3. What is the highest level of education that you have completed? **MARK ONE BOX ONLY.**
   1. No formal schooling [___]
   2. Primary school [___]
   3. Secondary school [___]
   4. Higher education (eg. Diploma/Technical/University studies) [___]

4. Which of these best describes your usual main occupation? **READ OUT AND MARK ONE BOX ONLY.**
   1. Unemployed/retired [___]
   2. Housewife [___]
   3. Skilled manual worker [___]
   4. Unskilled manual worker [___]
   5. Owner of business or farm [___]
   6. Office worker/non professional [___]
   7. Professional [___]
   8. Student [___]
5. How many people live in your HOUSEHOLD? [___]

6. What is the combined MONTHLY income of everyone in your HOUSEHOLD, including yourself? [________] Rs

7. In an average MONTH, how much do you think your HOUSEHOLD spends on
   i. Medication (including prescription tablets and other medicines) [________] Rs
   ii. Other health care (including doctor consultations, tests, procedures and travel) [________] Rs
8. Which of these best describes your quality of life? *MARK ONE BOX*
   1. Excellent  
   2. Very Good  
   3. Good  
   4. Fair  
   5. Poor  

9. Which of these best describes your health? *MARK ONE BOX*
   1. Excellent  
   2. Very Good  
   3. Good  
   4. Fair  
   5. Poor  

**I WILL NOW ASK YOU A FEW QUESTIONS ABOUT WHAT YOUR DOCTOR HAS TOLD YOU ABOUT YOUR HEALTH**

10. Have you ever been told by a doctor that you have had any of the following?

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Heart attack</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Angina</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Stroke</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Peripheral vascular disease (disease of the arteries)</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Diabetes (sugar)</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>Hypertension (high blood pressure)</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>High cholesterol</td>
<td>Yes</td>
</tr>
</tbody>
</table>

11. Have you ever been told by a doctor that you have any of the following chronic medical problems?

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Heart failure</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Rheumatic heart disease or valve disease</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Anaemia</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Tooth or gum disease</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Cataract</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>Thyroid disease</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>Depression</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>Chronic lung disease (NOT asthma)</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>Asthma</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>Pneumonia</td>
<td>Yes</td>
</tr>
<tr>
<td>11</td>
<td>Tuberculosis</td>
<td>Yes</td>
</tr>
<tr>
<td>12</td>
<td>Malaria</td>
<td>Yes</td>
</tr>
<tr>
<td>13</td>
<td>HIV/AIDS</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>Cancer</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>Epilepsy</td>
<td>Yes</td>
</tr>
<tr>
<td>16</td>
<td>Stomach ulcer</td>
<td>Yes</td>
</tr>
<tr>
<td>17</td>
<td>Malnutrition</td>
<td>Yes</td>
</tr>
</tbody>
</table>
I AM NOW GOING TO ASK YOU ABOUT FACTORS RELATED TO YOUR FAMILY AND LIFESTYLE

12. Have any of your close relatives (mother, father, brothers, sisters) had a heart attack before the age of 60 years?

   1. Yes □
   2. No □

13. Have any of your close relatives (mother, father, brothers, sisters) had a stroke before the age of 60 years?

   1. Yes □
   2. No □

14. Have any of your close relatives (mother, father, brothers, sister) been diagnosed with diabetes?

   1. Yes □
   2. No □

15. On average, how much physical activity do you do each day during working hours? MARK ONE BOX

   1. LOTS □ Eg. Lifting heavy weights, construction work, labourers, running
   2. MEDIUM □ Eg. Bike riding, rickshaw drivers, carrying buckets of water or loads of laundry to and from wells multiple times a day, walking long distances up and down hills
   3. LIGHT □ Eg. Walking on the level, standing all day working at a shop, housework such as cooking, cleaning in the house
   4. ALMOST □ Eg. Seated at a desk, driving a car, watching television, reading, resting

16. On average, how much physical activity do you do each day after working hours? MARK ONE BOX

   1. LOTS □ Eg. Lifting heavy weights, construction work, labourers, running

   □
2. **MEDIUM**  Eg. Bike riding, rickshaw drivers, carrying buckets of water or loads of laundry to and from wells multiple times a day, walking long distances up and down hills

3. **LIGHT ACTIVITY**  Eg. Walking on the level, standing all day working at a shop, housework such as cooking, cleaning in the house

4. **ALMOST NONE**  Eg. Seated at a desk, driving a car, watching television, reading, resting
17. Have you ever smoked regularly? (i.e. on most days for at least a year)
   1. Yes □
   2. No □

   Do you currently smoke?
   1. Yes □
   2. No □

1. How many years have you smoked for? □□
2. How many of each do you smoke per day?
   Cigarettes □□
   Cigars □□

18. Do you use chewing tobacco regularly? (i.e. on most days for at least a year)
   1. Yes □
   2. No □

19. How many people smoke in your household? (include yourself if you are a current smoker) □□ people

20. On average, how many hours a day can you smell tobacco smoke at home (either your own or others) □□ hours □□ minutes

21. On average, how many hours a day can you smell tobacco smoke at work (either your own or others) □□ hours □□ minutes
23. On average, how many days in a week do you eat fruit? ___ days

24. On average, how many days in a week do you eat green leafy vegetables? ___ days

I AM NOW GOING TO ASK YOU ABOUT TREATMENTS YOU RECEIVE:

25. No [ ] Yes [ ] Have you ever had your BLOOD PRESSURE checked?
   Yes [ ] No [ ] Was your BP checked by the Byrraju Foundation?
   Yes [ ] No [ ] Was your BP checked in the last 12 months?
   Yes [ ] No [ ] Are you taking BP-lowering tablets?
   Yes [ ] No [ ] Are these tablets from the Byrraju foundation?

26. No [ ] Yes [ ] Have you ever had your BLOOD SUGAR checked?
   Yes [ ] No [ ] Was your sugar checked by the Byrraju Foundation?
   Yes [ ] No [ ] Was your sugar checked in the last 12 months?
   Yes [ ] No [ ] Are you taking sugar-lowering tablets?
   Yes [ ] No [ ] Are these tablets from the Byrraju foundation?
27. No [ ] Yes [ ] Have you ever had your CHOLESTEROL checked?

Yes | No | Was your cholesterol checked in the last 12 months?
Yes [ ] No [ ] Are you taking cholesterol lowering tablets?

28. Yes [ ] No [ ] Do you take aspirin?
29. Are you taking any medications regularly?
   1. Yes □ —
   2. No □

A. What medications are you taking? (Please write the exact name of the medication from the medicine packet or a doctor’s script.)

<table>
<thead>
<tr>
<th>Code</th>
<th>Medication Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>
30. Have you ever had any pain or discomfort or any pressure or a feeling of heaviness in your chest?

1. Yes □ □
2. No □ □

1. Do you get it when you walk uphill or hurry?
   Yes □ □
   No □ □

2. Do you get it when you walk at an ordinary pace on the level?
   Yes □ □
   No □ □

3. What do you do if you get it while you are walking?
   Stop or slow down
   Carry on

4. If you stand still what happens to it?
   Relieved
   Not relieved

5. How soon?
   10 minutes or less
   More than 10 minutes

6. Will you show me where it was? **MARK THE AREA(S) ON THE DIAGRAM**

7. Have you ever had a severe chest pain across the front of your chest lasting for 30 minutes or more?
   Yes □ □
   No □ □
31. Do you get more breathless than people your own age when you go walking either on the level or up a hill?

1. Yes
2. No
Sex (M or F)

I AM NOW GOING TO ASK YOU SOME QUESTIONS ABOUT DISEASE PREVENTION

32. Which of the following actions may prevent a person getting a heart attack or stroke?

<table>
<thead>
<tr>
<th></th>
<th>Lose weight</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Unsure</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Quit smoking</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Increase exercise</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Eat more fish</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Drink less alcohol</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Reduce fat in meals</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Reduce salt in meals</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Eat more fresh fruit</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>Eat more green leafy vegetables</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

33. In the last 12 months, have you done any of the following to improve your health?

<table>
<thead>
<tr>
<th></th>
<th>Lose weight</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Unsure</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Quit smoking</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Increase exercise</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Eat more fish</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Drink less alcohol</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Reduce fat in meals</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Reduce salt in meals</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Eat more fresh fruit</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>Eat more green leafy vegetables</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

34. During the past 12 months have you been hurt in any way which required you to seek medical attention, or to stay away from work or school for at least one day? This includes you being hurt by an object or by a person (eg
pushed or hit), or from a fall, road accident, burn, electric shock, poisoning, animal bite or drowning. Include injuries that needed medical attention but you did not receive treatment due to cost.

1. Yes ☐
2. No ☐

If YES, complete the supplementary injury questionnaire AFTER you finish the main questionnaire.

THANK YOU VERY MUCH FOR YOUR TIME, YOU NOW NEED TO GO TO THE EXAMINATION ROOM

35. Interviewer Sign off
Name (BLOCK LETTERS) ______________________
Signature ________________________________
DOCUMENTATION OF PHYSICAL EXAMINATION

36. Height in centimetres

37. Weight in kilograms

38. Waist circumference in centimetres

39. Hip circumference in centimetres

40. Blood pressure in millimetres of mercury and heart rate

<table>
<thead>
<tr>
<th>Time BP taken</th>
<th>Systolic</th>
<th>Diastolic</th>
<th>Heart rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

42. Blood sample taken

43. Splitting sample

44. Urine dipstick

<table>
<thead>
<tr>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood</td>
</tr>
<tr>
<td>Glucose</td>
</tr>
<tr>
<td>Protein</td>
</tr>
</tbody>
</table>

45. ECG

46. Researcher - sign off for physical examination
Appendix 3. GC13 Survey Questionnaire

Relevant questions for this work are included below:

**Demographics Module**

Includes: Introduction, Consent, and Demographic Questions; Mortality Module; Coverage Module

Q2000. [INTERVIEWER: RECORD RESPONDENT’S SEX AS OBSERVED]

1. Male
2. Female

Q2001. Just to check, could you tell me if you are 15 years or older?

1. Yes
2. No [GO BACK TO RANDOM SELECTION PROCESS TO GET CORRECT ELIGIBLE RESPONDENT]

**Demographics of the Respondent**

Thank you for agreeing to participate in the survey. I would like to start by asking you some background questions.

Q2002. In what year were you born? PROBE WITH EVENT CALENDAR

___ ___ ___ ___ Year

PROBING PROCESS TO BE FOLLOWED WHEN YEAR OF BIRTH IS NOT KNOWN:

**How old do you think you are?**

Try to get the best guess to start the probing process. Based on the response (for example – 70 years), locate the event(s) within that timeframe (70 years before the current year) in the event calendar. Check with the respondent if she is aware of that event (if not, identify another event or combination of events within that timeframe). Then, using that event ask –

**Were you born before or after (EVENT)?**

Based on the response, ask –

IF BORN BEFORE THE EVENT, ASK - How old were you at (EVENT)?

IF BORN AFTER THE EVENT, ASK - How many years after (EVENT) were you born?

Arrive at the best estimate for the year of birth by narrowing down the time before or after the (EVENT). Based on this information, record the current age.

Q2003. What is your current marital status?

1. Never married [GO TO Q2006]
2. Currently married
3. Separated
4. Divorced
5. Widowed
6. Cohabiting/living together [GO TO Q2006]
Q2004. In what year did you first marry? 
   ___ ___ ___ ____ [ENTER 00 IF LESS THAN 1 YEAR]

Q2005. [ASK EVER MARRIED WOMEN 15-49 ONLY] Have you ever had a pregnancy? 
   1. Yes 
   2. No

Q2006. Have you ever attended school? 
   1. Yes 
   2. No [GO TO Q3000]

Q2007. What is the highest level of school you attended? 
   1. Primary 
   2. Secondary 
   3. Higher

Q2008. What is the highest (grade/form/year) you completed at that level? 
   Grade/form/year: ____

Coverage Module

Now I’d like to ask you some questions about your health. I’d like to remind you that the information you provide is totally confidential and will not be disclosed to anyone.

GENERAL HEALTH

Q5000. In general, how would you rate your health today? Is it... [READ ALL OPTIONS]

   1. Excellent 
   2. Very Good 
   3. Good 
   4. Fair 
   5. Poor

VISIT TO A HEALTH PROVIDER

Q5001. When was the last time you visited a health provider [or insert a locally specific term for a health provider in the “formal”, possibly a trained, licensed practitioner]? 
   1. Less than 1 year ago 
   2. 1-5 years ago 
   3. More than 5 years ago 
   4. Never [DO NOT ASK Q5002-Q5004, Q5007-Q5009, Q5012-Q5014, Q5021-Q5022, Q5024-Q5025, Q5027-Q5028, Q5030-Q5033, Q5045-Q5048]
BLOOD PRESSURE
Now I would like to ask you some questions focused on health conditions that are common around the world.

Q5002. Have you ever had your blood pressure measured by a health provider?
   1. Yes
   2. No [GO TO 5005]

Q5003. In what year was your blood pressure last measured by a health provider?
     Year ___ ___ ___ ___

Q5004. The last time your blood pressure was measured, were you told it was … [READ OPTIONS]
   1. Normal
   2. Higher than normal
   3. Lower than normal
   4. Or were you not told anything?

Q5005. Are you currently taking western medicine to control your blood pressure or prevent high blood pressure?
     ‘Western’ refers to all western allopathic medicines that the respondent is taking. Do not include homeopathic or ayurvedic medicines.
     1. Yes
     2. No

Q5006. Are you currently taking any herbal or traditional remedy to control your blood pressure or prevent high blood pressure?
     ‘Herbal or traditional remedy’ refers to all non-western medicine, including homeopathic, ayurvedic, and any other local traditional medicines.
     1. Yes
     2. No

CHOLESTEROL

Q5007. Have you ever had your blood cholesterol checked by a health provider?
     ‘Cholesterol’ is a substance found among the fats in the bloodstream that is an important part of a healthy body, but can be dangerous to health if levels are too high.
     1. Yes
     2. No [GO TO Q5010]

Q5008. In what year was your cholesterol last checked by a health provider?
     Year ___ ___ ___ ___

Q5009. The last time your cholesterol was checked, were you told it was … [READ OPTIONS]
     1. Normal
     2. Higher than normal
     3. Lower than normal
     4. Or were you not told anything?
Q5010. Are you currently taking western medicine to help control your cholesterol or prevent high cholesterol?
   1. Yes
   2. No

Q5011. Taking any herbal or traditional remedy?
   ‘Herbal or traditional remedy’ refers to all non-western medicine, including homeopathic, ayurvedic, and any other local traditional medicines.
   1. Yes
   2. No

DIABETES

Q5012. Have you ever had a blood test or urine test for diabetes?
   1. Yes
   2. No [GO TO Q5015]

Q5013. In what year were you last tested for diabetes?
   Year ___ ___ ___ ___

Q5014. Have you ever been told by a health provider that you had diabetes, pre-diabetes or borderline diabetes, or diabetes but only at a time when you were pregnant?
   1. No
   2. Diabetes
   3. Pre-diabetes or borderline diabetes
   4. [Women only] Diabetes, but only at a time when you were pregnant

Q5015. Are you currently taking diabetic pills?
   1. Yes
   2. No

Q5016. Are you currently taking insulin?
   1. Yes
   2. No
HEALTH BEHAVIOR QUESTIONS

Q5017a. Are you currently controlling your weight or losing weight?
   1. Yes
   2. No [GOTO Q5018a]

Q5017b. Why are you currently controlling your weight? Is it because...
   [INTERVIEWER: CHECK ALL THAT APPLY]
   1. Your health provider has recommended you to do so, or
   2. You think it’s good for you

Q5018a. Are you currently increasing your physical activity or exercise?
   1. Yes
   2. No [GOTO Q5019a]

Q5018b. Why are you currently increasing your physical activity or exercise? Is it because...
   [INTERVIEWER: CHECK ALL THAT APPLY]
   1. Your health provider has recommended you to do so, or
   2. You think it’s good for you

Q5019a. Are you currently trying to cut down on salt? For example, have you cut down on how much [insert site-specific salty foods] you eat?
   1. Yes
   2. No [GOTO Q5020a]

Q5019b. Why are you currently trying to cut down on salt? Is it because...
   [INTERVIEWER: CHECK ALL THAT APPLY]
   1. Your health provider has recommended you to do so, or
   2. You think it’s good for you

Q5020a. Are you currently trying to eat fewer high fat or high cholesterol foods? For example, you have cut down on how much [insert site-specific high fat/high cholesterol foods] you eat?
   1. Yes
   2. No [GOTO Q5021]

Q5020b. Why are you currently trying to eat fewer high fat or high cholesterol foods? Is it because...
   [CHECK ALL THAT APPLY]
   1. Your health provider has recommended you to do so, or
   2. You think it’s good for you
ANGINA

Q5027. Have you ever been told by a health provider that you have angina (or angina pectoris)?

‘Angina’ is a condition that is characterized by chest pain originating from the heart muscle.

1. Yes
2. No [GO TO Q5029]

Q5028. In what year were you told by a health provider that you have angina?

Year ___ ___ ___

Q5029. Are you currently taking western medication for angina?

1. Yes
2. No

OTHER CHRONIC CONDITIONS

Q5030. Have you ever been told by a health provider that you had a heart attack?

‘Heart attack’ refers to a myocardial infarction. This is a health condition where the blood supply to part of the heart is interrupted, so that the heart does not get enough oxygen and can become damaged.

1. Yes
2. No

Q5031. Have you ever been told by a health provider that you had a stroke?

‘Stroke’ refers to a disturbance in the blood supply to the brain.

1. Yes
2. No

TOBACCO

Q5053. Have you smoked at least 100 cigarettes or beedis [term can be translated to site-specific terminology for hand-rolled cigarettes throughout this section of questionnaire] in your entire lifetime?

1. Yes
2. No

Q5054. Do you currently smoke cigarettes/beedis?

1. Yes
2. No

[IF RESPONDENT ANSWERED “NO” TO Q5053 and Q5054, GO TO Q5057]

Q5055. How old were you when you first began smoking?

1. Age: ____
2. Don’t know
Q5056. How many cigarettes do(did) you smoke a day?

Enter number of cigarettes: ___ ___

Q5057. [ASK ONLY IF Q5054=1] Do you use any other kind of tobacco? [Add/modify site-specific types of tobacco.]

[CHECK ALL THAT APPLY]
1. Pipe
2. Chewing tobacco
3. Local form of tobacco
4. Other

SYMPTOM QUESTIONS

During the last 12 months, have you experienced any of the following:

Q5064. A feeling of tightness in your chest?
1. Yes
2. No

Q5065. Waking up with a feeling of tightness in your chest in the morning or any other time?
1. Yes
2. No

Q5067. Pain or discomfort in your chest when you walk uphill or hurry?
1. Yes
2. No
3. Never walks uphill or hurries

Q5068. Pain or discomfort in your chest when you walk at an ordinary pace on level ground?
1. Yes
2. No

[IF NO TO BOTH Q5067 and Q5068, GO TO Q5072]

Q5069. What do you do if you get the pain or discomfort when you are walking?
1. Stop or slow down
2. Carry on after taking a pain relieving medicine that dissolves in your mouth
3. Carry on

Q5070. If you stand still, what happens to the pain or discomfort? Is it...
1. Relieved
2. Not relieved

Q5071. Will you show me where you usually experience the pain or discomfort?
INTERVIEWER: SHOW FIGURE 2, PICTURE OF CHEST
(RECORD ALL AREAS OF BODY MENTIONED OR SHOWED)

1. 
2. 
3. 
4. 
5. 
6. 
7. 
8.
In your lifetime:

Q5074. Have you ever had sudden painless weakness on one side of your body?
   1. Yes
   2. No

Q5075. Have you ever had sudden numbness or a dead feeling on one side of your body?
   1. Yes
   2. No

Q5076. Have you ever had sudden painless loss of vision in one or both eyes?
   1. Yes
   2. No

Q5077. Have you ever suddenly lost one half of your vision?
   1. Yes
   2. No

Q5078. Have you ever suddenly lost the ability to understand what people are saying?
   1. Yes
   2. No

Q5079. Have you ever suddenly lost the ability to express yourself verbally or in writing?
   1. Yes
   2. No

INVENTORY OF MEDICINE

We are interested in knowing about the availability and use of certain medicines and drugs. Remember that whatever information you give me is confidential and will only be used for research purposes.

Q5080. Do you have any medicines in the house that a health provider has prescribed for you or given to you?
   1. Yes
   2. No [IF NO, GO TO CLOSE-ENDED QUESTIONS TO THE INTERVIEWER]

Q5081. May I see what medicines you personally have been using in the last 2 weeks?
   1. Yes
   2. No, not using any [GO TO CLOSE-ENDED QUESTIONS TO INTERVIEWER]
   3. Refuse [GO TO CLOSE-ENDED QUESTIONS TO INTERVIEWER]
Q5082. [INTERVIEWER: ASK THE RESPONDENT IF YOU CAN SEE THE MEDICATION AND RECORD NAME OF PRESCRIPTION OR ANY OTHER RELEVANT INFORMATION FROM THE PRESCRIPTION OR BOTTLE/BOX LABEL. IF THE RESPONDENT REFUSES TO ALLOW YOU TO SEE THE MEDICATION, SKIP THAT MEDICATION(S)]

1. Name of prescription from bottle/label:__________________ [we will need site to code]
   1. How frequently do you use this medicine?
      ___times per day OR
      ___times per week OR
      ___when needed/symptoms occur

2. Name of prescription from bottle/label:__________________ [we will need site to code]
   2. How frequently do you use this medicine?
      ___times per day OR
      ___times per week OR
      ___when needed/symptoms occur

3. Name of prescription from bottle/label:__________________ [we will need site to code]
   3. How frequently do you use this medicine?
      ___times per day OR
      ___times per week OR
      ___when needed/symptoms occur

4. Name of prescription from bottle/label:__________________ [we will need site to code]
   4. How frequently do you use this medicine?
      ___times per day OR
      ___times per week OR
      ___when needed/symptoms occur

5. Name of prescription from bottle/label:__________________ [we will need site to code]
   5. How frequently do you use this medicine?
      ___times per day OR
      ___times per week OR
      ___when needed/symptoms occur

[Close-ended question to the interviewer: DO NOT READ OUT LOUD] Is there anything the research team should be aware of about this respondent that may have affected the quality of data?

   1. No significant problems
   2. Respondent seemed to have difficulty understanding
   3. Respondent seemed distracted
   4. Respondent seemed tired
5. Respondent seemed drunk
6. Respondent seemed rushed
7. Other, specify: ______________________________________________________

PROGRAM CHECK:

✓ IS THE RESPONDENT A FEMALE?
✓ IS THE RESPONDENT 15-49 YEARS OLD?
✓ HAS THE RESPONDENT EVER BEEN MARRIED?
✓ HAS THE RESPONDENT EVER HAD A PREGNANCY?

IF YES TO ALL ABOVE, GO TO MATERNAL AND CHILD HEALTH MODULE.

OTHERWISE, PROCEED TO PHYSICAL MEASUREMENTS

PHYSICAL MEASUREMENTS

INTERVIEWER: PLEASE TAKE HEIGHT, WEIGHT, AND BLOOD PRESSURE MEASUREMENTS AND RECORD BELOW.

PM1. Height [ENTER 9999 FOR REFUSED]
Enter height in centimeters: ___ ___ ___.__ cm

PM2. Weight [ENTER 9999 FOR REFUSED]
Enter weight in kilograms: ___ ___ ___. ___ kg

PM3a. Blood pressure, 1st measurement [ENTER 999 FOR REFUSED]
Systolic: ___ ___ ___
Diastolic: ___ ___ ___

PM3b. Blood pressure, 2nd measurement [ENTER 999 FOR REFUSED]
Systolic: ___ ___ ___
Diastolic: ___ ___ ___

PM3c. Blood pressure, 3rd measurement [ENTER 999 FOR REFUSED]
Systolic: ___ ___ ___
Diastolic: ___ ___ ___

IF RESPONDENT NOT SELECTED FOR DBS, THAT COMPLETES THE INTERVIEW

IF RESPONDENT IS SELECTED FOR DBS, PROCEED

PM4. How many hours ago did you last eat or drink anything other than water?
   ____ hours

PM5. Dried blood spots:
1. Dried blood spots taken, enter barcode number for DBS: XXXXX- ___ ___ ___ ___
2. Refused
Appendix 4. *SMARTHealth 2014 Survey Questionnaire*

Relevant questions for this work is included below:

Page 1 of the app:

*Welcome page:* 5 digits unique combination of alphabets and numeric username and password will be provided to log in as Interviewer.

**Page 1:**

- **PHC ID:** Two digits numeric ID, will be provided to interviewers. [Dropdown]
- **Village ID:** Two digits numeric ID, will be provided to interviewers. [Dropdown]
- **Locality:** Select the locality from a pre-populated list. [Dropdown]

- **Household ID:** Nine digits numeric ID (1 digit uniform +2 digits PHC id+2 digits Village id+ 4 digits HH id), will be provided to interviewers based on the listing sheet.

To click search with above details

(Pre filled respondents names will come here, except Village option we can put other options here)

**Q1: Could you please tell me your house number?**

_Interviewers will confirm names displayed under the House hold, if given names are not available there then interviewer will choose the House hold disposition code:_

**Household disposition code**

a. Household moved away permanently from the village and a new family is staying [option to enter details of eligible individuals in the new household]

b. New household [option to enter details of eligible individuals in the new household]

c. Household refused to participate [close the interview]

**Q2: List of all members (name, age and gender) aged 40 and above in the household?**

_The interviewer selects a particular member to enter his details. Upon end of the questionnaire, he/she will be directed back to this page if there is another member whose details have to be collected._

[The head of the household needs to be highlighted either with a star at the end]

[Another 2 columns with interview status and respondent disposition code for each respondent]

**Respondent disposition code** for interview not done

a. Respondent not at home during the visit [option to click on this respondent on next visit]

b. Respondent moved away permanently [name is greyed out]

c. Respondent not eligible [name is greyed out]

d. Respondent is not alive [name is greyed out]

e. Respondent incapacitated [name is greyed out]

f. Respondent refuses to participate [option to click on this respondent on next visit]

If the household member gives consent and agrees to participate in the study Interviewer need to click **Start New Risk Assessment** to initiate the screening.
Page 2 of the app:

**Participant ID:** eleven digits numeric ID, first digit 1 is uniform to all IDs, second 2 digits – village id, next four digits - household, last two digits individual order in the household – [automatically generated from the household number]

**Consent Number:** 5 digits alphanumeric ID, type in the consent number on the top of the consent form

**Q3:** Just to check, could you please tell us your name?

[Interviewers to first record his/her surname and then the given name]

**Sex** (male/female) to appear on this screen

**Q4:** What is your date of Birth? (Date /Month /Year)?

If date of birth is not known, what is your age?

[Interviewer will use “I” button to check Event calendar for males and age probing methods for females to get accurate age in case the respondents do not remember their date of birth or age.]

Age is collected again during the baseline study to confirm the age we have recorded during the listing. At this point, if it is found that the respondent is less than 40 years old, the interviewer will thank and terminate the Interviewer

**Q5:** What is your current marital status? [Responses in drop down box]

1. Single
2. Married
3. Separated
4. Divorced
5. Widowed

**Q6:** What is the highest level of formal education you have obtained? [Responses in drop down box]

1. No School.
2. Primary (till 5th class)
3. Upper Primary (till 7th class)
4. Secondary School (up to 10th class)
5. Higher Secondary (Intermediate)
6. Graduate & Higher

**Q7:** What is your present occupation? [Responses in drop down box]

1. Agriculture Labourer
2. Manual Labourer
3. Skilled worker
4. Farmer
5. Business
6. Housewife
7. Retired
8. Government Employee
9. Others

Page 3 of the app:

**Past History**

Interviewer: Have you ever been told by a health provider that you have or have had any of following
"I" button will be used by interviewer to make uniform explanation about the diseases/condition, interviewer will read the message in “I” button as it is to explain to the respondent.

Q8: Have you ever been told by a health provider that you have or have had heart attack or angina?
   1. Yes [go to q 8b]
      0. No

8b: When were you told - __________ years ago [drop down box – 1,2,3,4,5,>5 year ago, don’t know]

Q9: Have you ever been told by a doctor that you have or have had Stroke?
   1. Yes [go to q 9b]
      0. No

9b. When were you told __________ years ago [drop down box – 1,2,3,4,5,>5 year ago, don’t know]

Q10. Have you ever been told by a doctor that you have or have had Diabetes?
     1. Yes [go to q 10b]
        0. No

10b. When was it done - __________ years ago [drop down box – 1,2,3,4,5,>5 year ago, don’t know]

Q11: Have you ever been told by a doctor that you have or have had Hypertension (High BP)
     1. Yes [go to q 11b]
        0. No

11b. When were you told - __________ years ago [drop down box – 1,2,3,4,5,>5 year ago, don’t know]

Q12: Have you ever had a surgery done in legs to restore blood flow by redirecting blood around the blocked or narrowed artery? Peripheral vascular diseases
     1. Yes [go to q 11b]
        0. No

12b. When were you told - __________ years ago [drop down box – 1,2,3,4,5,>5 year ago, don’t know]

Family History
Interviewer: Have any one of your close relatives (mother, father, brother, sister) had ever suffered from any of the following

Q13: Have any one of your close relatives (mother, father, brother, sister) had ever suffered from heart attack?
     1. Yes
        0. No

Q14: Have any one of your close relatives (mother, father, brother, sister) had ever suffered from stroke?
     1. Yes
        0. No

Q15: Have any one of your close relatives (mother, father, brother, sisters) had suffered from Diabetes?
     1. Yes
        0. No
Smoking Status
Q16: Have you ever used tobacco products in any form
   1. Smoking cigarette/Beedi/Chutta
   2. Chewing tobacco like Khaini/Ghutka/Zarda
   3. Both
   4. No [Go to next section]

If smoking cigarette/Beedi/Chutta:
Q16a: Do you currently smoke cigarettes/Beedi/Chutta (at least 1 per day)?
   1. Yes [GO TO Q16c]
   0. No [GO TO Q16b]

Q16b: Did you quit smoking cigarettes/Beedi/Chutta in last 12 months?
   1. Yes [GO TO Q18c]
   0. No [Go to Q19a]

Q16c: Age when you started smoking? _____ years.

If chewing tobacco:
Q17a. Do you currently chew tobacco?
   1. Yes [GO TO Q17c]
   0. No [GO TO Q17b]

Q17b: Did you quit chewing tobacco in last 12 months?
   1. Yes [GO TO Q17c]
   0. No [Go to next section]

Q17c: Age when you started chewing tobacco? _____ Years.

If the respondent using both:
Q18. Do you currently smoke and chew tobacco both?
   Interviewer will get both Questions 16b, c and 17 b,c to fill in.

Page 4
CVD Risk Factor Measurements
Interviewer: Now I am going to measure your Blood Pressure 3 times with this instrument. Please be seated comfortably, [BP VALUES ARE AUTOMATICALLY TRANSFERRED INTO APPLICATION]

“1” info button explains about how to check blood pressure and using A&D BP machine.

Blood Pressure
1st Measurement SBP___ /DBP___mmHg HR___bpm
2nd Measurement SBP___ /DBP___mmHg HR___bpm
3rd Measurement SBP___ / DBP___mmHg HR___bpm

Cuff size used:
1. Standard
2. Large

**Blood Glucose:**

Now I am going to collect one drop of blood from your finger by prick with sterilized lancet to measure your blood glucose level.

Blood Glucose Values _________ mg/dl

Q 19. How many hours ago did you last eat or drink anything other than water _______ hrs

Select from the drop down list:

1. 0-2 hr.
2. 2-4 hr.
3. 4-6 hr.
4. >6 hours.

Time of day blood specimen taken (24 hour clock) – *Automatically records time*

Referral to Doctor recommended due to presence of the following

- SBP > 180 or DBP > 100 mm/hg
- FBG > ___ or RBS > ________ mg/dl

**PROFILING (PROGRAM CHECK):**

- HAS THE RESPONDENT CONFIRMED OF AN ESTABLISHED CVD (EITHER CORONARY HEART DISEASE, STROKE/ TRANSIENT ISCHEMIC ATTACK OR PERIPHERAL VASCULAR DISEASE)?
- IS THE RESPONDENT’S TEN-YEAR CVD RISK ≥30%?
- IS THE RESPONDENT’S TEN-YEAR CVD RISK ≥ 20% AND A SYSTOLIC BLOOD PRESSURE (SBP) ≥ 140mmHg?

IF YES TO ANY OF THE ABOVE, GO TO THE MODULE FOR FURTHER QUESTIONS.

IF NOT:

END THE SURVEY HERE

[Thank you for answering all the questions. We appreciate your participation in this study and will give you the results of the measurements right away. Please feel free to ask any questions that you might have]

Page 6 of the app: Further Questions:

IF RESPONDENT IS SELECTED FOR FURTHER QUESTIONS:

[Thank you for answering all the questions. As I mentioned before, in this study, we are also trying to collect more information about your health. Therefore, we would like to ask you some more questions about your]
health. We will also measure your height and weight. I would like to remind you that the information you provide is totally confidential and will not be disclosed to anyone.

**Medicine history:**

**Q22a: Are you currently on any of following medication?**

"I" button gives commonly available different type of BP Lowering medicines just to get an idea for Interviewer

May I see what medicines you have been using? [Interviewer to see the medication and record name of the medication from the bottle/strip label]

1. **BP Lowering medication**
   1. Yes
   0. No
2. **Lipid Lowering medication (Statin)**
   1. Yes
   0. No
3. **Anti-plate let Therapy**
   1. Yes
   0. No

If Yes there will be pop up question comes on

**22b. Record Medication use:**

- Information available
- Currently not available
- Refused to show

If Information is available Input Medication box will be prepopulated to enter the details of Medication as follows:

**Q22b: Could you please tell me the number of tablets/medicines you are currently consuming in a day?**

[All the medications from Q22a should appear below]

1. Brand Name _____ number of pills in a day_____strength____
   Press to Save medication details or Cancel to delete.
2. Brand Name _____ number of pills in a day_____strength____

**Q23: Do you take herbal or AYUSH medicine?**

1. Yes
0. No

**Page 7 of the app:**

**Q24: Physical activity :( We modified the shorter version of IPAQ)**

1. **How many days in the past week have you done vigorous physical activity?**

"I" Button: explains more about what is vigorous activity

- Number of Days per week [if >=1, go to q 1b. If no go to q 2] [range check – 1 to 7]
- No vigorous activity
- Don’t know/not sure
- Refused

1b. Time spent on one of those days: _____ Hours ______ mins

2. **How many days in the past week have you done moderate physical activity?**
2b. Time spent on one of those days: _____ Hours _____ mins

3. How many days in the past week have you walked (with the duration being at least 10 minutes or more per walk)

3b. Time spent on one of those days: _____ Hours _____ mins

Q27. Physical Measurements:
“i” button will explain record height and weight measures in the given blanks.

Height _____ cms
Weight _____ kgs

Page 12 of the app:
If the patient is identified as high risk, an alert appears -

Referral to a doctor is recommended.

If the patient doesn’t fall under high risk then the following message appears -

Thank you for answering all the questions. We appreciate your participation in this study and will give you the results of the measurements right away. Please feel free to ask any questions that you might have.

CLICK OK TO SAVE THE CASE.